

Федеральное государственное бюджетное учреждение науки  
Институт философии Российской академии наук

**СОЦИОГУМАНИТАРНЫЕ АСПЕКТЫ  
ЦИФРОВЫХ ТРАНСФОРМАЦИЙ  
И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

Под редакцией *В.Е. Лепского, А.Н. Райкова*

Когито-Центр  
Москва – 2022

УДК 001.8  
ББК 87.25  
С 69

### Рецензенты

Академик РАН, доктор философских наук  
*Владислав Александрович Лекторский*

Член-корреспондент РАН, доктор экономических наук  
*Владимир Викторович Иванов*

**С 69** Социогуманитарные аспекты цифровых трансформаций и искусственного интеллекта / Под ред. В.Е. Лепского, А.Н. Райкова. – М.: Когито-Центр, 2022. – 308 с.

Авторский коллектив:

*Агеев А.И., Дубровский Д.И., Ефимов А.Р., Журенков Д.А., Ильин Н.И.,  
Лепский В.Е., Малахова Е.В., Матвеев Ф.М., Пойкин А.Е., Райков А.Н.,  
Савельев А.М., Славин А.Б., Славин Б.Б.*

**ISBN 978-5-89353-665-2**

УДК 001.8  
ББК 87.25

Прогресс в развитии цифровых технологий, включая искусственный интеллект, несомненен. Однако, методологический базис этого развития до сих пор носит преимущественно техногенный характер. Как результат, например, системы искусственного интеллекта так и не вышли за рамки «слабого» и «узкого», они остаются однозадачными, не дают вразумительного объяснения своим выводам, слабо учитывают специфику поведения субъектных сред.

Вместе с тем, уже явно происходит свертывание идеи техногенной цивилизации, которая характеризуется проявлением в ее базовой ценности научно-технического прогресса и ценности науки как основы управления социальными процессами. Как следствие, развитие цифровых технологий, включая искусственный интеллект, происходит без всеобъемлющего учета социально-гуманитарных ценностей и этических аспектов регулирования жизнедеятельности людей.

Настоящая монография для умаления подобных ограничений в развитии цифровых технологий, включая искусственный интеллект, обосновывает необходимость использования субъектно-ориентированного подхода, который формируется в рамках постнеклассической кибернетики саморазвивающихся полисубъектных (рефлексивно-активных) сред гибридной реальности, основные положения которой опираются на представления об обществе как социально-гуманитарной системе.

Монография предназначена для специалистов из гуманитарных, естественнонаучных и технических областей знания, ориентированных на проблемы управления и развития, для студентов и аспирантов, а также для широкой аудитории практиков управления, в том числе связанных с разработкой и использованием цифровых технологий, включая искусственный интеллект.

*Работа выполнена при финансовой поддержке РНФ, проект № 21-18-00184  
«Социогуманитарные основания критериев оценки инноваций, использующих  
цифровые технологии и искусственный интеллект»*

# СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ .....	5
<b>1. ФЕТИШ ЦИФРОВЫХ ТРАНСФОРМАЦИЙ И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В XXI ВЕКЕ</b>	
1.1. Введение	10
1.2. Фетиш искусственного интеллекта	12
1.3. Искусственный интеллект: туманность определений в неопределенности реалий	23
1.4. Противоречие технического и социогуманитарного представления искусственного интеллекта	32
1.5. Непредсказуемые угрозы от злонамеренного использования сильного искусственного интеллекта	42
1.6. Выводы	51
<b>2. ЦИФРОВЫЕ ТРАНСФОРМАЦИИ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ (ПРЕОДОЛЕНИЕ ОГРАНИЧЕНИЙ ТЕХНОГЕННОЙ ЦИВИЛИЗАЦИИ)</b>	
2.1. Введение	54
2.2. Системные основания для преодоления ограничений техногенной цивилизации	56
2.3. Социогуманитарные вопросы развития искусственного интеллекта в условиях глобального кризиса земной цивилизации	69
2.4. Ценностные ориентиры разработки искусственного интеллекта в США и Китае	79
2.5. Парадигмальные основания внедрения цифровых инноваций в современное образование	85
2.6. Образование будущего в контексте развития научной рациональности	99
2.7. Выводы	107
<b>3. СОЦИОГУМАНИТАРНЫЙ АНАЛИЗ ЧЕЛОВЕКО-ОРИЕНТИРОВАННЫХ ПРЕДСТАВЛЕНИЙ ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ (ЭТИКА ИИ, ДОВЕРИТЕЛЬНЫЙ ИИ, ОБЪЯСНИТЕЛЬНЫЙ ИИ И ДР.)</b>	
3.1. Введение	110
3.2. Субъектно-ориентированный подход к представлениям элементов и систем искусственного интеллекта (цифровой двойник, цифровой субъект и цифровой метасубъект)	113
3.3. Машины как люди	122
3.4. Субъектность объяснимого искусственного интеллекта	137
3.5. Этические дилеммы систем искусственного интеллекта	147
3.6. Этические аспекты развития искусственного интеллекта в Китае	157
3.7. Проблема моделирования этических свойств интеллектуальных роботов	166
3.8. Выводы	172

4. СОЦИОГУМАНИТАРНЫЕ ОСНОВАНИЯ КРИТЕРИЕВ ОЦЕНКИ ИННОВАЦИЙ, ИСПОЛЬЗУЮЩИХ ЦИФРОВЫЕ ТЕХНОЛОГИИ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ, В УПРАВЛЕНИИ СОЦИАЛЬНЫМИ СИСТЕМАМИ	
4.1. Введение	175
4.2. Социогуманитарные основания оценки инноваций в сфере искусственного интеллекта	176
4.3. Общий искусственный интеллект в саморазвивающихся рефлексивно-активных средах	189
4.4. Нейротехнологии в системах искусственного интеллекта и применения в сфере управления	201
4.5. Социотехническая оценка инновационных цифровых управленческих информационных систем	213
4.6. Выводы	229
5. СОЦИОГУМАНИТАРНЫЕ ОСНОВАНИЯ КРИТЕРИЕВ ОЦЕНКИ ИННОВАЦИЙ, ИСПОЛЬЗУЮЩИХ ЦИФРОВЫЕ ТЕХНОЛОГИИ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ОРГАНИЗАЦИИ ЦИФРОВЫХ ПЛАТФОРМ, ТЕХНОЛОГИЧЕСКИХ ДОЛИН, ЭКОСИСТЕМ	
5.1. Введение	232
5.2. Критерии оценки инноваций с использованием цифровых технологий: анализ международного опыта стандартизации	233
5.3. Особенности управления инновациями в области цифровых технологий и искусственного интеллекта	244
5.4. Архитектура Общего искусственного интеллекта (социогуманитарные аспекты)	251
5.5. Построение эталонной модели социогуманитарных оснований критериев оценки инноваций в виде модели результативности в соглашениях архитектурного подхода к проектированию цифровых платформ	261
5.6. Социогуманитарные основания критериев оценки цифровых медиа платформ	269
5.7. Выводы	279
ЗАКЛЮЧЕНИЕ	282
ЛИТЕРАТУРА	287
АННОТАЦИЯ И КЛЮЧЕВЫЕ СЛОВА	305
TITLE, ABSTRACT, KEYWORDS	306
СВЕДЕНИЯ ОБ АВТОРАХ	307

## ПРЕДИСЛОВИЕ

---

Прогресс цифровых технологий, включая искусственный интеллект (ИИ), демонстрируемый почти во всех областях человеческой деятельности, от изучения атома и предоставления услуг до управления огромными промышленными комплексами и исследования загадок космоса, в последние годы несомненен. Цифровые экономики ведущих стран мира более, чем на 40% делают вклад в ВВП. Сквозные цифровые технологии, включая ИИ охватывают большинство направлений производства и научно-исследовательских работ. Без ИИ, машинного обучения и анализа больших данных сейчас немыслимо достижение высоких показателей производительности труда, обеспечения первенства на глобальном конкурентоспособном рынке. ИИ преодолевает все более высокие барьеры по решению сложных задач обеспечения жизнедеятельности человека, управления сложными процессами, диагностики и лечения болезней людей и животных.

Успехи ИИ оказались столь значительными, что многие полагают возможным, что он сможет скоро заменить человека почти во всех областях деятельности. ИИ стал фетишем XXI века, который одни стали превозносить как наше будущее, а другие начали вести с ним борьбу, призывать к ограничению его развития. Вместе с тем, от жизненных реалий никуда не уйти, сейчас на рынке побеждает тот, кто делает и внедряет более быстрые, смелые и необычные инновации. Для этого с применением цифровых технологий, включая ИИ, формируются специализированные и гибкие технологические среды поддержки процессов синтеза прорывных решений.

Вполне очевидно, что развитие цифровой экономики – это безальтернативный стратегический выбор для того, чтобы воспользоваться новым витком научно-технической революции и новыми возможностями для промышленных преобразований. Цифровая экономика сейчас является важным показателем мощи любой продвинутой страны и важным двигателем построения современной экономической системы. Совместное продвижение технологических, модельных, форматных и системных инноваций, а также эффективное использование ИИ и средств анализа данных придают мощный импульс цифровому развитию экономики и общества.

Однако, яркий успех цифровых технологий зачастую закрывает обратную сторону медали. Масштабы цифровой экономики в мире быстро растут, но проблемы несбалансированного и нерегулярного развития становятся все более заметными. Методологический, научный и технологический базис развития цифровых технологий пока несет на себе вериги атавизмов прошлого века. Скажем, системы ИИ так пока и

не вышли из шор «слабого» и «узкого» ИИ; в искусственных нейронных сетях многократно увеличилось число слоев, усложнились процессы обработки данных, растет скорость вычислений, однако эти системы так и остались в основном однозадачными, они не дают хотя бы сносные объяснения своим выводам, хорошо решают только круг задач типа распознавания, анализа и синтеза голоса и текста. Продвижение перспективных версий развития ИИ в виде «сильного» и «общего» ИИ, как можно заметить, пока имеют в своей основе эволюционный характер модернизации, методы почти не отрываются от традиционного логического инструментализма, пока слабо учитываются особенности нелокальных и когнитивных аспектов сознания человека, поведения социума и космологической взаимосвязи явлений.

Наращение глобального кризиса потребительской цивилизации ведет к разжиганию все более масштабных экономических, политических, всевозможных социальных конфликтов, силовой борьбе за ресурсы, за сферы влияния, за передел всей мировой структуры социально-экономической и политической самоорганизации. Рыночная гонка с интенцией победы в конкурентоспособной борьбе зачастую оставляет на втором плане социогуманитарные аспекты внедрения цифровых технологий, то есть, собственно, того, чему эти технологии должны служить. Эта гонка при внедрении цифровых технологий, включая ИИ, заставляет игнорировать социальные ценности, забывать об уникальности человека и духовности человеческого общества. Идут ломка традиционной этики, строится некая новая, цифровая, этика; при этом деградирует мышление, растут риски злонамеренного использования ИИ, с новой силой поднимаются извечные вопросы обеспечения справедливости. Решение этических вопросов применения ИИ – очень сложное дело, которое невозможно выполнить лишь принятием кодексов этики и провозглашением благородных целей. Необходимо учесть, что этические принципы есть лишь точки в бесконечном пространстве морально-этических решений.

Растет риск угрозы будущему развитию человечества, развитию патологий озлобления человека, становления его более агрессивным. Как результат, ведущими учеными и политическими деятелями мира вопрос развития ИИ ставится иногда в один ряд с наиболее опасными проявлениями человеческой деятельности, такими как политический и религиозный экстремизм. Человечество подходит к точке бифуркации своего развития.

В этой бифуркационной точке особого внимания заслуживает то, что сейчас техническая сторона вопроса экономического развития стран, регионов и промышленного производства становится не единственным фактором успеха. В контексте упомянутого выше свертывания идеи

техногенной цивилизации впервые в истории развития человечества технологии покусились на, казалось бы, недоступное для технологий в человеческой жизни – на когнитивные способности человека, на распознавание и экспликацию его эмоций и бессознательных состояний, которые невозможно описать с помощью математической формулы.

В этой бифуркационной точке особого внимания заслуживает то, что сейчас техническая сторона вопроса экономического развития стран, регионов и промышленного производства становится не единственным фактором успеха. В контексте упомянутого выше свертывания идеи техногенной цивилизации впервые в истории развития человечества технологии покусились на, казалось бы, недоступное для технологий в человеческой жизни – на когнитивные способности человека, на распознавание и экспликацию его эмоций и бессознательных состояний, которые невозможно описать с помощью математической формулы.

Появился новый риск: если технологии смогут заместить хотя бы частично способности человека мыслить, не значит ли это, что близок момент сингулярности, когда технологии смогут получить равносильную субъектность по трансформации общества. Вместе с тем, в настоящее время замещение человека системами ИИ происходит повсеместно, в нарастающих и даже угрожающих, по мнению отдельных экспертов, масштабах. При разработке таких систем неизбежно допускается некоторая степень неопределенности, которая сходна со свободой воли человека: этические самоограничения, знания о нормах поведения, имитация саморегуляции поведения, способность к эмпатии и др.

В современном социуме формирование целостности личности происходит в сопровождении построения цифровых проекций (двойников) процессов и объектов динамического взаимодействия людей и надличностных систем. При этом в личности воплощаются когнитивные проекции социума, индивидуализированные: биологически, информационно, когнитивно, социально. Это приводит к неизбежности комплексирования различных цифровых моделей поддержки лояльной к правовым нормам, управленческим ключевым установкам и глубинным регуляторам жизнедеятельности социальной среды. Современный уровень цифровых и интеллектуальных технологий позволяет осуществлять зондаж и воздействие с учетом психосемантических качеств личности расширенного характера.

В приведенном контексте, когда фетиш цифровых технологий, включая ИИ нарастет, а риски внедрения этих технологий явно отличаются от нуля, ищутся альтернативы техногенной цивилизации с усилением акцента на социально-гуманитарные факторы развития человека и общества. Например, в качестве варианта пост-техногенной

цивилизации предлагается модель социогуманитарной цивилизации. Ключевой проблемой для перехода к такой цивилизации является проблема становления субъектности развития человечества. Эта тема также рассматривается в предлагаемой монографии.

В монографии на конкретном уровне идет поиск эффективного и одновременно социально-значимого подхода к построению цифровых систем, включая систем ИИ, прежде всего, для решения социально-гуманитарных задач, которые решаются в гибридной (человеко-машинной) реальности, характеризующейся высоким уровнем субъектности. Это реальность, в которой пока отсутствуют способы репрезентации субъектов и событий классическими формализованными методами.

В монографии не остаются без внимания вопросы определения критериев оценки социальной значимости развития ИИ и процессов цифровизации, что потребовало основательного рассмотрения прежде всего самого понятия социальной значимости в условиях цифровизации экономики. При этом рассматриваются различные социальные аспекты жизнедеятельности, включая существующие неклассические и постнеклассические образовательные парадигмы и педагогические технологии, в том числе с оценкой того, насколько они в принципе позволяют включать в себя те или иные цифровые технологии, и как последние влияют на процесс обучения и социализацию учащихся в различных возрастах.

Для решения конфликтных вопросов и упреждения угрожающих проблем, связанных с экспансией цифровых технологий, включая ИИ, в настоящей монографии прорабатывается новый механизм цифрового социально-экономического развития с диверсифицированным участием и синергией взаимодействия различных экономических и социальных субъектов. Этот механизм опирается на субъектно-ориентированный подход, который формируется в рамках постнеклассической кибернетики саморазвивающихся полисубъектных сред. Становление субъектно-ориентированного подхода в настоящее время определяется в рамках представлений о научной рациональности: классика, неклассика, постнеклассика. Важность субъектно-ориентированного подхода нарастает по мере движения от деятельностного к субъектно-деятельностному и, наконец, к субъектно-ориентированному подходу.

Основные положения этого подхода опираются на представления об обществе как социальной системе; идеи деятельностного и субъектно-деятельностного подходов; концептуальные модели рефлексивно-активных процессов; идеи устойчивого развития и управления сложностью; опыта разработки больших информационных систем для поддержки управления страной, включая создание системы



распределенных ситуационных центров для информационно-аналитической и технологической поддержки решения вопросов национальной безопасности.

В рамках предлагаемого подхода детерминированную ясность приобретает построение концепция цифрового двойника (виртуальная интерактивная копия реального физического объекта или процесса), которая, сейчас реализуется в ограничениях парадигмы кибернетики первого порядка, где этот двойник, независимо от сложности, остается формализованным и пассивным по отношению к субъекту управления. Субъектно-ориентированный подход показывает путь деятельного устранения этого ограничения.

Предлагаемый в монографии подход показывает направление для построения объяснимого ИИ с ориентацией на потребности человека, на его когнитивную атрибуцию и с учетом внешнего окружения. Делается попытка показать, что тема построения объяснимого ИИ выходит далеко за рамки вопроса каузальности, выявляемой статистическими методами, охватывает феномен нелокальности семантической интерпретации моделей ИИ, и, как следствие, гибридную и субъективную реальность.

Субъектно-ориентированный подход обозначает необходимое условие для обеспечения включенности общества в процессы управления и развития через формирование двухконтурной структуры организации процессов государственного управления и развития. Первый контур может трактоваться как исполнительный механизм управления мировыми процессами, а второй – формируется с доминирующим началом общественных структур, в котором интегрируются представители администраций, бизнеса и общества.

Предлагаемый в монографии субъектно-ориентированный подход в реальной социально-экономической практике государственного управления и человеческой жизнедеятельности создает необходимые условия для цифрового обеспечения формирования слаженной системы государственного управления и единой, справедливой, упорядоченной конкуренции, зрелой и завершенной современной рыночной системы цифровой экономики, опирающейся на преимущества использования инновационных сред, создаваемой в гармоничном сочетании поведения субъекта, включая коллективного, и инструментов цифровой экономики.

Авторы благодарят коллег за участие в обсуждении проблем, представленных в данной монографии: В.И. Аршинова, В.Г. Буданова, С.В. Гарбука, В.В. Иванова, В.А. Лекторского, М.В. Локтионова, Т.А. Нестик, Д.А. Новикова.

*От редакторов*

*В.Е. Лепский, доктор психологических наук, профессор*

*А.Н. Райков, доктор технических наук, профессор*

# 1. ФЕТИШ ЦИФРОВЫХ ТРАНСФОРМАЦИЙ И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В XXI ВЕКЕ

---

## 1.1. ВВЕДЕНИЕ

В данном разделе представлены основания, позволяющие констатировать фетиш искусственного интеллекта (ИИ). Выделяются принципиальные отличия ИИ от всех предшествующих технологических инноваций, связанные прежде всего с внедрением в когнитивную сферу человека и принципиально новыми неконтролируемыми последствиями для общества. Представлены аргументы, что лидеры глобалистского проекта являются главными интересантами и заказчиками фетиша ИИ. Это отчетливо проявляется в работах философов, приближенных к гигантским ИТ-корпорациям, и в мегапроектах этих корпораций. Предлагается к рассмотрению проблем использовать возможности ИИ для преодоления нарастающих международных конфликтов и в целом мирового кризиса. В центре внимания оказывается проблема субъектности, решение которой с позиций антропоморфного подхода к ИИ чревато серьезными негативными последствиями. При наделении субъектностью ИИ неявно снимается ответственность с человека, который использует эту технологию, а также разрушается сложившаяся законодательная практика. Предлагается представление ИИ как агента, наделенного набором инвариантных упрощенных качеств, которыми обладают естественные субъекты. К таким качествам можно отнести: способность к когнитивным размышлениям, целеустремленности, своего рода рефлексивности, наличие эмоций и чувств, коммуникативность и упрощенные элементы социальности.

Проводится анализ, выделяются приоритетные направления и ставятся острые вопросы в области развития систем ИИ в условиях превращения традиционных методов управления социумом в единую систему с мощными цифровыми и коммуникационными возможностями. Оценивается фактор трансформации экономических и общественных отношений, которые все больше обладают субъектным качеством в переустройстве мира.

Быстрый рост производительности вычислительной техники, успехи в области использования нейронных сетей, средств анализа больших данных, технологий ИИ повышают их востребованность. Однако любой прогресс имеет и свои ограничения. Так, с увеличением внимания к ИИ во власти и обществе усиливаются требования ограничить использование инструментов ИИ, например, распознавания,

которые нарушают права личности и несут риски принятия ошибочных решений.

Не обходит развитие ИИ и поиск ответов на извечные вопросы обеспечения справедливости, этики, взаимодействия систем ИИ с субъектами. Все большее внимание уделяется проблеме создания перспективных систем ИИ, которые призваны решать одновременно множество разных междисциплинарных задачи. На повестку встает вопрос угроза наступления сингулярного момента, когда ИИ может стать сильнее человека.

Все более актуальным становится решение этических вопросов применения ИИ. При этом особое внимание уделяется проблеме субъектности в области разработок этических основ применения ИИ.

В этом контексте рассматривается тема, связанная с фетишизацией ИИ, которому приписывается грандиозное будущее. Международный и национальный стандарты ИИ, задающие его базовые определения и области применения, однако они находятся в фазе начальной разработки. В этой работе активно участвует Технический комитет ИСО/МЭК 42 «Искусственный интеллект», состоящий из 5 рабочих групп.

Рассматриваются подходы к построению систем ИИ, которые погружены в гибридную (человеко-машинную) реальность, характеризуемую высоким уровнем субъектности. В ней почти отсутствуют формализованные способы репрезентации явлений.

Рассматриваются риски, связанные с решениями в области ИИ и ошибках в данных, ломкой естественной этики, деградацией аналитических способностей человека. Исследуются вопросы возможности злонамеренного использования ИИ.

## 1.2. ФЕТИШ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

По мере того, как экономика и социальные коммуникации оцифровываются, возрастает и роль цифровых технологий. Технологические инновации всегда привлекали к себе внимание, так было и во времена появления паровых машин, и в период электрификации, и в эпоху расцвета электроники. Появление новых технологий приводило к трансформации экономических и общественных отношений, и поэтому часто технологиям приписывали субъектные возможности по переустройству мира. Однако время все расставляет на свои места, и предсказания подобные известному «Со временем, телевидение перевернет жизнь всего человечества. Ничего не будет: ни кино, ни театра, ни книг, ни газет – одно сплошное телевидение», как правило, не оправдываются. Технологии находят свое место в общей экосистеме научного прогресса и не ломают привычных устоев. Поэтому всегда необходим элемент скепсиса при появлении новых революционных технологий.

Такой подход можно было бы применить и к современным цифровым технологиям, и не обсуждать особенно долго возможные трансформации, к которым они приводят. Однако у цифровых технологий есть особенность, которая их существенно отличает от всех других технологий. Все новшества, которые появлялись в доцифровую эпоху, были призваны повысить производительность труда человека, либо облегчить и сделать более комфортной его жизнь. Они преобразовывали окружающую среду, делая ее более удобной для человека, но при этом не затрагивали саму сущность человека, его когнитивные возможности. Конечно, совершенствование орудий труда вело к тому, что человечество накапливало новые знания, но эти знания по-прежнему находились в головах людей. Книги были лишь инструментом передачи знаний от человека к человеку, но без человека они не имеют смысла.

Цифровая эпоха дала возможность накапливать и обрабатывать знания в цифровом виде. В начале это, как и любые новшества, лишь облегчало использование знаний человеком – стало возможным читать книги и статьи с электронных носителей, искать в них информацию. Но по мере развития технологий обработки данных оказалось, что аналитические программы способны так анализировать данные (так называемый *data mining*), чтобы находить в них то, чего невозможно напрямую обнаружить. Впервые технологии покусились на самое святое – на когнитивные возможности человека. Если технологии смогут заместить хотя бы частично способности человека мыслить, не значит ли это, что технологии смогут получить ту самую субъектность по

трансформации общества, которой они раньше никогда не обладали? Этот вопрос сегодня в той или иной мере будоражит умы многих ученых.

### **1.2.1. Нарастающая волна интереса к ИИ и предвестники шторма**

Благодаря увеличению производительности вычислительной техники, успехам в области разработки алгоритмов использования искусственных нейронных сетей (в первую очередь машинного, и в особенности глубокого обучения) и появлению инструментов для работы с большими данными технологии ИИ перешли из разряда перспективных в разряд самых востребованных технологий. Надо сразу сказать, что существующие технологии ИИ касаются в основном задач распознавания, предсказаний и имитации человеческой деятельности, и не претендуют на создание реального аналога человеческого интеллекта.

Текущие успехи ограничены решением проблем определенных классов – ИИ все еще нельзя доверить самостоятельное принятие сложных решений, от которых зависит жизнь человека. Однако даже в таком виде ИИ стал широко использоваться в предиктивной аналитике, скоринге, распознавании лиц и игровых приложениях. В 2021 крупнейшие мировые компании вложили более 250 млрд. долларов США в исследования и разработки в области ИИ<sup>1</sup>. Прогнозируется, что мировой рынок ИИ технологий составит 554,3 млрд долл. США к 2024 г. при среднегодовом темпе роста 17,5%<sup>2</sup>.

Одновременно с ростом успеха ИИ началась череда принятия стратегий его развития на национальных уровнях. В 2016 году в США Национальным советом по науке и технологиям были подготовлены доклады «Подготовка к будущему искусственного интеллекта», «Национальный стратегический план исследований и разработок в области искусственного интеллекта» и «Искусственный интеллект, автоматизация и экономика», определившие стратегию развития ИИ в стране. В июле 2017 года Госсовет Китая опубликовал «План развития искусственного интеллекта нового поколения», рассчитанный на создание индустрии искусственного интеллекта и превращения Китая в ведущую державу в области ИИ к 2030 году. На ежегодной встрече Всемирного экономического форума 2018 года Премьер-министр

---

<sup>1</sup> <https://t.me/abulaphia/3277>

<sup>2</sup> Прорывные инновации 2022 – Прорывные инновации Прорывные инновации: человек 2.0: доклад к XXIII Ясинской (Апрельской) международной научной конференции по проблемам развития экономики и общества, Москва, 4–8 апреля 2022 г. / С.А. Гавриш, Л. М. Гохберг, Д.Е. Грибкова и др.; под ред. Л.М. Гохберга, А.Р. Ефимова, Ю.В. Мильшиной; Нац. исслед. ун-т «Высшая школа экономики», ПАО Сбербанк. – М.: НИУ ВШЭ, 2022. – 56 с.

Великобритании Тереза Мэй объявила, что собирается сделать Великобританию мировым лидером в области искусственного интеллекта. В 2019 году в России президентом В.В. Путиным была утверждена «Национальная стратегия развития искусственного интеллекта на период до 2030 года»<sup>1</sup>, а позже в рамках программы «Цифровая экономика» выделен проект по развитию ИИ. В самые последние годы набирает оборот новый мировой тренд – разработка Общего искусственного интеллекта, который по своим функциям должен приблизиться к решению задач, специфичных для естественного интеллекта. Об этом подробнее будет сказано ниже.

Одновременно с блицкригом ИИ в бизнесе и на государственной ниве начались протесты и требования ограничить использование инструментов распознавания, которые нарушают права личности и несут риски принятия ошибочных решений. Так европейские страны, уже при подписании Декларации о сотрудничестве в области искусственного интеллекта в апреле 2018 года делали акцент на необходимость учета социальных, экономических, этических и юридических вопросов. В частности, Европейская комиссия учредила Группу экспертов, которые опубликовали руководящие принципы этики ИИ в апреле 2019 года. В сентябре 2021 года Совет по правам человека ООН выпустил доклад с рекомендациями «для государств и предприятий относительно разработки и внедрения гарантий для предотвращения и сведения к минимуму вредных последствий и содействия полному использованию преимуществ, которые может предоставить искусственный интеллект». Руководитель Совета (Верховный комиссар ООН по правам человека) Мишель Бачелет, комментируя доклад, вообще призвала ввести мораторий на системы ИИ, которые угрожают правам человека, до тех пор, пока правительства не смогут установить гарантии.

В докладе Совета по правам человека ООН, в частности говорится, что «Процессы принятия решений во многих системах ИИ непрозрачны. Сложность информационной среды, алгоритмов и моделей, лежащих в основе разработки и функционирования систем искусственного интеллекта, а также преднамеренная секретность государственных и частных субъектов являются факторами, которые подрывают значимые способы понимания общественностью вопросов влияния систем искусственного интеллекта на права человека». Интересно, что именно машинное обучение вызывает у защитников прав человека наибольшие опасения, поскольку результат таких вычислений непредсказуем:

---

<sup>1</sup> Указ Президента Российской Федерации от 10.10.2019 N 490 О развитии искусственного интеллекта в Российской Федерации (вместе с Национальной стратегией развития искусственного интеллекта в Российской Федерации на период до 2030 года)

«Системы машинного обучения добавляют важный элемент непрозрачности; они могут быть способны выявлять закономерности и разрабатывать рецепты, которые трудно или невозможно объяснить. Это часто называют проблемой “черного ящика”. Непрозрачность затрудняет тщательное изучение системы искусственного интеллекта и может стать препятствием для эффективной подотчетности в тех случаях, когда системы искусственного интеллекта наносят ущерб»<sup>1</sup>. Кроме того, вызывает большую озабоченность то, что любое машинное обучение очень сильно зависит от тех данных, на которых происходит само обучение. Если эти данные не проверять на соблюдение элементарных этических норм или хуже того, специально подобрать их, чтобы они им противоречили, то системы машинного обучения неизбежно будут выдавать заведомо неэтичные рекомендации.

### 1.2.2. Главные интересанты и заказчики фетиша ИИ

Недавно в печати появилась статья Ясона Габриэля «К теории справедливости для искусственного интеллекта», который работает штатным философом в компании, принадлежащей Google – DeepMind<sup>2</sup>. Статья посвящена разработке гуманистических принципов использования технологий ИИ. Проанализируем философско-методологические основания позиции автора, их влияния на представления о потенциальных последствиях развития ИИ, а также, что особенно важно, о главных интересантах предложенного подхода.

В первую очередь, обращает на себя внимание идеологическая установка автора статьи, опирающегося на труды Джона Ролза, посвященные его «Теории справедливости»<sup>3</sup>. Суть этой установки изображается в следующем виде: люди заинтересованы в увеличении своей и уменьшении общей доли выгоды. Это ярко выраженная позиция идеологии социального либерализма. Для представления роли и места ИИ автор предлагает понимать основную структуру общества как совокупность социотехнических систем, функционирование которых складывается под все возрастающим влиянием ИИ. Такое представление ИИ может быть охарактеризовано как яркое проявление технократического редукционизма в организации социальных

---

<sup>1</sup> Human Rights Council (2021). A/HRC/48/31: The right to privacy in the digital age - Report of the United Nations High Commissioner for Human Rights // <https://www.ohchr.org/>. 2021. URL: <https://www.ohchr.org/en/documents/thematic-reports/ahrc4831-right-privacy-digital-age-report-united-nations-high> (дата обращения: 18.05.2022)

<sup>2</sup> *Gabriel Iason*. Toward a Theory of Justice for Artificial Intelligence. *Daedalus*, Spring 2022, Vol. 151, No. 2, AI & Society (Spring 2022), pp. 218-231.

<sup>3</sup> *Rawls John*. A Theory of Justice, rev. ed. Cambridge, Mass.: Harvard University Press, 1999.

процессов. Здесь в качестве следствия неявно предлагается нарастающее под влиянием ИИ снижение роли государства и общества. Интересантами таких результатов являются лидеры глобалистского проекта. Но кто именно? Это надо конкретизировать.

Легко увидеть, что лидерами глобалистского подхода и яркими выразителями и защитниками идеологии западного либерализма выступают в первых рядах именно владельцы гигантских ИТ-корпораций, которые подобно спруту охватывают мировое коммуникативное пространство. Владельцы таких компаний как Facebook, Twitter, Amazon, Google и др. и их наемные теоретики, приписывая системам ИИ *качество субъектности*, пытаются создать впечатление, что они наделены некими «божественными» функциями. Ведь они же не просто разрабатывают новые программные продукты, а создают инструменты, которые способны якобы устанавливать справедливость и защищать права человека. Такие инструменты – утопия, но прокламирование веры в них, имитация такой веры является хорошей почвой для защиты глобалистских идей, а в то же время и сокровенных интересов крупнейших мировых производителей ИТ-технологий. Манипулируя массовым сознанием, используя все формы социальных коммуникаций, все средства ИИ, они заявляют, что будут для всех нас «сеять разумное, доброе, вечное». Не случайно же основатель Facebook Марк Цукерберг решил на основе своих ИТ-продуктов создавать *Метавселенную*, в которой, по его мнению, будет выстроена новая, более правильная и благополучная жизнь. Однако крайне трудно допустить, что Марк Цукерберг действительно озабочен столь высокими гуманистическими устремлениями. Судя по всей его бизнес-деятельности и его опыту манипуляции массовым сознанием в целях достижения максимальной прибыли, сказки о «правильной Метавселенной» служат лишь маскировкой для осуществления тех же целей.

Глобалистский подход, опирающийся на идеологию западного либерализма с его псевдогуманистическими клише, убедительно показал свою несостоятельность в условиях пандемии COVID-19<sup>1</sup> и тем самым продемонстрировал свою негативную роль в деле оценки и разработки способов преодоления нарастающих угроз для человечества. В том числе угроз, связанных с цифровыми трансформациями и развитием ИИ.

---

<sup>1</sup> *Лепский В.Е.* Рефлексия пандемии COVID-19: субъектно-ориентированный подход // Экономические стратегии. 2020. № 8 (174). С. 66–71.



### **1.2.3. Перспективы ИИ в контексте глобального кризиса мировой цивилизации**

Именно этот озаглавленный контекст рассмотрения процесса развития ИИ и его социальной значимости, отодвигаемый часто на дальний план (особенно в публикациях того типа, которые рассматривались выше), приобретает сейчас первостепенное, судьбоносное значение для человечества. Неуклонное нарастание глобального кризиса нашей потребительской цивилизации ведет к разжиганию все более масштабных экономических, политических, всевозможных социальных конфликтов, бескомпромиссной борьбы за ресурсы, за сферы влияния, в конечном итоге, за передел всей мировой структуры социально-экономической и политической самоорганизации. И в этой связи ИИ становится инструментом борьбы в противостоянии различных сил.

В этом отношении, как уже отмечалось выше, особое внимание привлекает задача создания Общего искусственного интеллекта, решение которой стало в последние годы предметом конкуренции между крупнейшими бигтехами, а в более широком плане – между государствами-лидерами в области ИИ, в числе которых находятся наши стратегические противники. Это обязывает нас максимально сконцентрировать усилия в этом направлении и добиваться опережения конкурентов. Недавно в России вышла первая книга, посвященная специфической проблематике Общего искусственного интеллекта<sup>1</sup>. В ней подробно проанализированы основные теоретические и методологические вопросы разработки Общего искусственного интеллекта и необходимые для этого научные подходы, обозначены те его специфические функции, которые должны быть созданы и, что особенно интересно, те предполагаемые масштабные перемены, которые он способен произвести в нарождающемся новом мироустройстве.

Авторы выделяют две главные способности Общего искусственного интеллекта, которые характерны для естественного интеллекта. В отличие от «узкого» ИИ, решающего одну определенную задачу, он должен быть интегральным, т.е. способным решать одновременно разные задачи. И он должен обрести качество автономности, т.е. способность самостоятельно эффективно действовать в широком диапазоне сред. Все это качественно повышает деятельные возможности систем ИИ, их использование для военной техники и военных действий, для решения задач производства, управления, планирования, организации экономических процессов, оптимизации самых

---

<sup>1</sup> Сильный искусственный интеллект: На подступах к сверхразуму / Александр Ведяхин и др. М.: Интеллектуальная Литература, 2021. 232 с.

разнообразных сфер общественной жизни и научных исследований, что чрезвычайно важно для нашей страны в нынешних условиях. И это столь же важно для осмысления и осуществления тех грядущих исторических изменений в развитии земной цивилизации, которые связаны с крушением принципа и практики монополярного мира.

Вместе с тем развитие Общего искусственного интеллекта ставит новые сложные теоретические вопросы, касающиеся его взаимодействий с естественным интеллектом, создания и перспектив гибридного интеллекта, возможного состязания с естественным интеллектом, возможных рисков и угроз для человека и общества. Достижение высокой степени автономности Общего искусственного интеллекта создает вероятность появления таких видов и способов его «самодетельности», которые могут представлять опасность для человека и общества, потребуют разработки новых методов обеспечения безопасности. Здесь перед нами окажется новый аспект все той же проблемы *субъектности* систем ИИ, сохраняющей свою высокую актуальность.

#### **1.2.4. ИИ – это субъект или агент, контролируемый субъектами естественного интеллекта**

Обратимся еще раз к истолкованиям теоретических вопросов о субъектности систем ИИ по отношению к реальным человеческим субъектам. Подход автора упомянутой выше статьи Ясона Габриэля<sup>1</sup> заключается в том, чтобы перенести на ИИ те же принципы, которые устанавливаются для людей. Автор пишет, что «ИИ все больше формирует элементы базовой структуры общества», и, «следовательно, его проектирование, разработка и развертывание потенциально взаимодействуют с принципами правосудия». По мнению автора «ИИ взаимодействует с поведением людей, принимающих решения, и формирует характер этих практик, включая распределение выгод и бремени среди населения».

Получается, что у автора технологии как бы «оживают»: «Для нашей цели – пишет он – важно учесть, что в современных обществах фоновая справедливость все больше осуществляется алгоритмически». Здесь вводится понятие фоновой (видимо, массовой) справедливости, и неявно предполагается, что эту справедливость осуществляет алгоритм. И далее: «Делая оценки или прогнозы на основе прошлого выбора человека и предоставляя решения или рекомендации, которые затем формируют

---

<sup>1</sup> *Gabriel Iason. Toward a Theory of Justice for Artificial Intelligence. Daedalus, Spring 2022, Vol. 151, No. 2, AI & Society (Spring 2022), pp. 218-231.*

набор возможностей, доступных этому человеку в будущем, эти системы сильно влияют на разворачивающуюся взаимосвязь между индивидуальным выбором и коллективными результатами». Как видим, опять неявно предполагается, что системы что-то «разумно» делают и что-то предоставляют. Автор неявно придал субъектность технологиям ИИ, и требует распространить законы социальных отношений на использование ИИ, которое должно «поддерживать основные свободы граждан, способствовать справедливому равенству возможностей и приносить наибольшую пользу тем, кто находится в наихудшем положении».

Такого рода антропоморфный подход к искусственному интеллекту чреват серьезными последствиями. Во-первых, перенося субъектность на ИИ, неявно снимается ответственность с человека, который использует эту технологию, что нивелирует законодательную практику. Любая технология несовершенна, и не может сама по себе принимать решение. Ошибка в работе детектора лжи никогда не будет равна нулю, также как и при использовании ИИ. Задача людей состоит в том, чтобы учесть ограничения технологий, а не пытаться просто их поставить в какие-то заданные рамки. Не случайно в судебной практике при всех возможностях криминалистики окончательное решение принимают люди. Системы ИИ ничем не отличаются от других технологий, поэтому технологии не должны снимать ответственности с человека, а следовательно и не должны обладать человеческими характеристиками – т.е. быть справедливыми, гуманными и т.д.

Успехи ИИ оказались настолько значительными, что многие решили, что эта технология может заменить человека, только надо ее поставить в определенные рамки. Фактически ИИ стал фетишем XXI века, который одни стали превозносить как наше будущее, а другие к таким прогнозам относятся очень осторожно. На самом деле ИИ, хотя и может распознать то, что не удастся человеческому взгляду, или выявить корреляцию, которую не может найти человек, вместе с тем остается далеким от реальных когнитивных возможностей человека. Не исключено, что в будущем удастся создать ИИ, интегрированный в социальную среду, так называемый сильный ИИ, но пока мы далеки от этого, и придавать ему субъектность не только неправильно, но опасно.

### **1.2.5. Этические вопросы развития ИИ**

Вопросы этики использования ИИ широко обсуждаются сейчас не только общественными деятелями, политиками, но и учеными. Большое внимание привлекла ситуация с подготовленной к изданию статьи об этике ИИ сотрудником компании Google Тимнит Гебру, одним из

ведущих мировых экспертов по проблемам необъективности алгоритмов и извлечения данных (data mining). Научные и коммерческие позиции принципиально разошлись, и, как следствие, исследовательница покинула компанию Google.

Дискуссии ведутся и среди российских ученых. Наряду с конструктивными предложениями, отдельные авторитетные ученые предлагают весьма сомнительные идеи, призывают к политизации науки и технологий, наделяя технологии свойствами «патриотизма». Так 23 ноября 2021 года состоялось заседание Президиума Российской Академии наук, на которой академики обсуждали в том числе и этическую сторону ИИ, и возможности для ИИ быть доверенным. Один из выступающих заявил: «С моей точки зрения, искусственный интеллект должен быть не только доверенным, о чем сегодня говорилось, но и патриотичным, то есть он должен в первую очередь работать на интересы страны, а не против нее»<sup>1</sup>.

Сегодня многие протестуют против технологий распознавания лиц (в некоторых городах США такое распознавание запрещено, Европарламент тоже предложил запретить технологии распознавания лиц). Однако вредно не само распознавание, а использование его в противоправных целях. Человек не скрывает свои болезни перед врачом потому, что доверяет ему и рассчитывает на помощь. Необходимо регулировать законы применения технологий, в том числе и ИИ, и серьезно наказывать, если они были использованы во вред человеку или незаконно. Россия сегодня оказалась во многом слабо защищенной перед мошенниками, которые используют средства коммуникаций для обмана доверчивых граждан. Все, что сегодня власть может сделать – это предупредить население о новых способах мошенничества. При этом власть вполне эффективно, в том числе и с использованием ИИ, борется с политическими противниками. Этические проблемы ИИ должны решаться не путем ограничения технологий, а в первую очередь путем ограничения действий тех людей, кто их неправомерно использует.

Решение этических вопросов применения ИИ – комплексное дело, которое невозможно выполнить лишь принятием декларативных кодексов Этики ИИ. Необходимо учесть, что даже наши этические принципы, какие бы мы не взяли – от библейских заповедей до кодекса строителя коммунизма – есть лишь точки в бесконечном пространстве морально-этических решений, в котором мы движемся ежедневно. Дискуссии об этике ИИ только начинают разворачиваться и сообщества

---

<sup>1</sup> Славин Б. Может ли искусственный интеллект быть справедливым // БИТ. 2021. №10 (113). С. 32-35.

философов и инженеров должны тесно сотрудничать для выработки ответов на новые вызовы времени.

Необходимо подчеркнуть, что сейчас крайне важно рассматривать этические проблемы применительно к развитию ИИ в более широком концептуальном плане – под углом реальных особенностей функционирования этических норм в социальной жизнедеятельности, реального состояния нравственности массового сознания, индивидуальных, групповых, институциональных субъектов. Ведь сплошь и рядом, всегда, на всех этапах истории человечества ясно наблюдался разрыв между знанием этических норм и их исполнением. Вспомним древнеримскую поговорку: «Вижу лучшее и одобряю, но следую худшему». Слишком уж часто интерес оказывался выше нравственных установлений, а обман подавлял правду и вил себе уютные гнезда в самых высоких этических наставлениях. Говорить о нравственном прогрессе в развитии человечества нет достаточных оснований – обстоятельные материалы на этот счет представлены в соответствующих философских исследованиях, см., например<sup>1</sup>. Все эти обстоятельства необходимо учитывать, когда мы рассуждаем на тему «Этики ИИ», причем как в отношении создателей систем ИИ, так и в отношении пользователей ими.

При попытках же моделирования принципов этики и воплощения их в работе ИИ дело осложняется тем, что совокупность этических норм не может быть упорядочена в виде четкой иерархической структуры, допускающей альтернативный выбор. Здесь выбор практически всегда может быть сделан лишь при рассмотрении и оценке конкретных условий. Поэтому указанное моделирование возможно лишь в специально определенных частных случаях.

Вместе с тем проблема субъектности в области разработок ИИ остается заслуживающей пристального внимания. Ведь так или иначе способность системы ИИ решать сложные задачи связывается нами с описаниями некоторых функций естественного интеллекта. Если ИИ нецелесообразно представлять как субъекта аналогичного человеку, то как тогда понимать и определять ИИ, которому передаются возможности принятия решений в определенных ситуациях и который побеждает чемпиона мира по шахматам или игры в Го. Если ИИ нецелесообразно представлять как субъекта аналогичного человеку, то как тогда понимать и определять ИИ, которому передаются возможности принятия решений в определенных ситуациях и который побеждает чемпиона мира по шахматам или игры в Го. Наиболее адекватным подходом, на наш

---

<sup>1</sup> Дубровский Д.И. О нравственном прогрессе и нравственном регрессе (К проблематике развития морального сознания // Философские науки, № 11, 2007. С. 81–102.

взгляд, может быть представление ИИ как агента, наделенного набором инвариантных упрощенных качеств, которыми обладают естественные субъекты. К таким качествам можно отнести: подобие целеустремленности, своего рода рефлексивность, коммуникативность и упрощенные элементы социальности. Такое представление ИИ как агента (псевдосубъекта) согласуется с принципом распределенного управления в биологии и психологии, который был назван принципом двойного субъекта<sup>1</sup>. Это представление ИИ, в сочетании с системами принципов и онтологий, задаваемых в концепции постнеклассической кибернетики саморазвивающихся сред, позволяет использовать ИИ как средство социальных инноваций, при сохранении контроля над технологиями ИИ, а также ставить и решать проблему интеграции образований искусственного и естественного интеллекта при сохранении базовых качеств носителей естественного интеллекта<sup>2</sup>.

### 1.2.6. Выводы

ИИ переживает фазу бурного роста. Однако, масштабные цифры эффектов от внедрения не должны вводить нас в заблуждение. Именно поэтому нам следует очень внимательно относиться к возможным когнитивным искажениям при исследованиях возникающих феноменов, что требует более широкого диалога и перехода от создания этических кодексов поведения ИИ к созданию следующего поколения ИИ, действующего в полисубъектной рефлексивно-активной среде вместе с человеком и для человека.

---

<sup>1</sup> *Лепский В.Е.* Концепция субъектно-ориентированной компьютеризации управленческой деятельности. М.: Институт психологии РАН, 1998. – 204с.

<sup>2</sup> *Lepskiy V.* Evolution of Cybernetics: Philosophical and Methodological Analysis // *Kybernetes*. 2018. Vol. 47. No. 2. P. 249–261. *Лепский В.Е.* Искусственный интеллект в субъектных парадигмах управления // *Философские науки*. 2021. Т. 64. № 1. С. 88–101.

## **1.3. ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ТУМАННОСТЬ ОПРЕДЕЛЕНИЙ В НЕОПРЕДЕЛЕННОСТИ РЕАЛИЙ**

### **1.3.1. Введение**

Системы ИИ (СИИ) справедливо признаются ансамблем наиболее всеобъемлющих технологий современности, которые окажут и уже оказывают существенное влияние на развитие систем управления на всех уровнях и направлениях деятельности человека. СИИ инкрустируются едва ли не во все новейшие цифровые решения, колоссальные инвестиции в разработку и внедрение СИИ делаются в ведущих государствах мира.

СИИ являются ключевыми среди всех сквозных цифровых технологий. Цифровая трансформация ведет к экспоненциальному образованию автономных агентов искусственной и смешанной генеалогии, их сообществ различной устойчивости, сложных, противоречивых и динамичных конструкций информационной и нормативной среды. При этом возрастает риск деструктивных действий формально «естественных агентов» социальных процессов, в том числе и в случаях воздействия на их сознание «искусственных агентов». Все это таит множество вызовов для принятия управленческих решений для общественной динамики в целом. На множественном и усиливающемся кризисном фоне возможно быстрое манифестирование как предвидимых угроз, так и событий из семейства «черных лебедей», «серых лебедей» или «розовых фламинго».

Следует учитывать масштаб уже наступившего «информационного потопа», в который ежесекундно вливаются новые реки данных от пользовательской активности человека и электронных устройств и систем. Принципиально новые феномены порождают динамика «естественного сознания» и множества новых виртуальных пространств, что создает проблемы идентификации активных и пассивных субъектов социальной деятельности, предотвращения целенаправленных и случайных деструктивных действий. Необходимо создавать новые модели анализа и прогнозирования обстановки, учитывающие непрерывное изменение состояния объектов и субъектов, их кластеризации или распада на базе информационных взаимосвязей и их рефлексивных эффектов.

Философские аспекты данной проблематики охватывают вопросы определения сущности ИИ, коэволюции искусственных и естественных систем, их идентичности и коллективности, перспектив и последствий внедрения СИИ, и особенно нейротехнологий, в социальную реальность.

### 1.3.2. Запутанность понимания ИИ

В настоящее время не существует общепринятого определения ИИ. Международные и национальные стандарты ИИ, задающие его базовые определения и области применения, находятся в фазе интенсивной разработки, хотя уже введено в действие множество стандартов, охватывающих различные аспекты тематики.

В структуре Технического комитета ИСО/МЭК 42 «Искусственный интеллект» из 5 рабочих групп, действовавших в 2019–2022 гг., первая – «Основополагающие принципы ИИ», была нацелена на решение базовых вопросов ИИ (включая определения и этику ИИ). Сам российский ТК 164 «Искусственный интеллект» в РФ образован в июле 2019 г., по структуре зеркально с ТК 42 «ИИ» ИСО и МЭК [Приказ Росстандарта 2020]. Всего по проекту базового стандарта ИИ было высказано более 1000 комментариев уполномоченными представителями более 70 стран. Большая часть из них была так или иначе учтена. В разработке в настоящее время находится более 30 профильных стандартов.

Все это означает, что еще некоторое время будут существовать риски неточного, несогласованного (недоопределенного) толкования СИИ и как следствие – недобросовестного или некомпетентного отнесения к ним продукции, которая не соответствует критериям ИИ.

Применяемые для конкретных целей определения делаются либо дедуктивно, либо индуктивно, либо функционально, либо через родовидовое отличие, либо через отрицание и пр.

Представляется полезным при определении сущности ИИ оттолкнуться от понимания естественного интеллекта.

Латинский термин *intellectus* означает «ум, рассудок, разум; мыслительные способности человека». В словаре Брокгауза и Ефрона статья «интеллект» отсылает к весьма обширной статье «умъ». Среди трех групп психических явлений: 1) восприятия и их интеллектуальная переработка; 2) изменение эмоционального равновесия; 3) волевые импульсы, только содержание первой группы более века назад относилось к собственно «уму», хотя постоянно делались попытки либо расширить, либо сузить даже это определение.

В понимании «ума» более ста лет назад подчеркивалось выдающееся влияние памяти, внимания и «утомляемость личности» на течение интеллектуальной жизни. Важнейшая функция умственной деятельности, как полагалось, состояла в сочетании вновь воспринимаемых явлений с накопленными воспоминаниями и опытом и выработке «целесообразной и планомерной реакции». Считалось, и не без оснований, что сложные последовательности реакций на разные



ряды впечатлений рутинизируются, становясь инстинктами и протекают рефлекторно, «не проникая в сознание».

Однако история ИИ как нового направления в науке начинается в середине XX в. К этому времени сложилось множество предпосылок в математике, кибернетике, гносеологии, нейрофизиологии, психологии. В науке сформировалось множество вычислительных традиций (теория алгоритмов, первые компьютеры). Алан Тьюринг в 1950 г. в статье «Вычислительная техника и интеллект» обсуждал критерии, позволяющие считать машину интеллектуальной. Важно отметить, что в английском языке словосочетание *artificial intelligence* не имеет такой антропоморфной окраски, которую оно приобрело в русском переводе. Но одним из ключевых вопросов при обсуждении ИИ в то время стала способность компьютеров мыслить и осознавать себя как отдельную личность. Вопрос об этике и самоидентификации СИИ вновь стал актуальным в наши дни, поскольку стала просматриваться и техническая возможность оперирования СИИ в пространстве целеполагания и расширения классов и контуров используемых данных, в частности – в Интернете вещей.

В настоящее время под «интеллектом» понимается «общая познавательная способность, которая проявляется в том, как человек воспринимает, понимает, объясняет и прогнозирует происходящее, какие решения он принимает и насколько эффективно он действует (прежде всего в новых, сложных и необычных ситуациях)».

Палитра пониманий интеллекта включает:

- 1) рассмотрение специфики организации его как «базы знаний»: объем, разнообразие, актуальность, компетентность, которая рассматривается в качестве критерия развитости интеллекта;
- 2) интеллект как система мыслительных операций – анализ, синтез и обобщение, при этом скорость переработки данных выступает критерием развития интеллекта;
- 3) механизмы ментального самоуправления, форма организации ментального опыта и т.п.

Интеллект также определяют как «способность приходить к решению при помощи вычислений».

Говард Гарднер в 1980-е годы выделял семь сторон интеллекта и, соответственно, их различную выраженность и пропорции у людей: лингвистическая интеллектуальность, логико-математические составляющие, оцениваемые тестом IQ, музыкальные способности, способность к пространственному видению, кинестетические способности и др.

На сегодня самое общее современное определение интеллекта описывает «способность к процессу познания и эффективному решению

проблем, в частности при овладении новым кругом жизненных задач». Нельзя не заметить, что данное понимание недалеко оторвалось от определения «ума» в словаре Брокгауза и Ефрона более века назад. Тем не менее определение интеллекта варьирует в зависимости от сферы применения. Релевантным может быть определение интеллекта как «социально полезной адаптации». Кроме того, представляется уместным использовать понятие «домен»: принципиальное отличие и преимущество человека до сих пор заключается именно в способности оперировать знаниями и опытом из разных «доменов», развивая способности, создавая новые знания, навыки и даже новые «домены».

Эпицентром современной дискуссии об определениях и границах СИИ стали алгоритмы как своего рода понятийный водораздел. В математике и кибернетике класс задач определенного типа считается решенным, когда для их решения установлен алгоритм. Изучение и нахождение алгоритмов является естественной целью человека при решении разнообразных проблем. В определенном смысле рост культуры можно описать как накопление запаса алгоритмов (стереотипов, рутин, навыков, эвристик) решения тех или иных проблем. Равно как бескультурье можно трактовать как либо утрату или уничтожение подобных навыков, либо как изначальное их отсутствие. Задачи, для решения которых алгоритм еще не найден и требуется усилие, изобретательность и проницательность человеческого ума относят к категории интеллектуальных. Соответственно, в строгом смысле этот класс задач и следует относить к объему понятия СИИ. Стремление к компромиссу в определении по критерию «интеллектуальности задач» в зависимости от нахождения алгоритма привело к разделению СИИ на «слабые» и «сильные». Так, в работе<sup>1</sup> даны определения:

- Сильный искусственный интеллект – система, способная моделировать человеческие чувства, намерения и мышление путем обработки символов, физических полей и других видов материи и энергии для решения сложных междисциплинарных задач с глубоким пониманием того, что она делает.
- Слабый искусственный интеллект – система, способная решать интеллектуальные задачи, обрабатывая только символы, не понимая, что она делает, но быстрее, чем это сделали бы люди.

В формирующихся международных стандартах (проект ISO/IEC 22989) также выделяется два вида ИИ: слабый ИИ и сильный ИИ. В

---

<sup>1</sup> Raikov A. (2021). Cognitive Semantics of Artificial Intelligence: A New Perspective. Springer Singapore, Topics: Computational Intelligence XVII, 128 p. <https://doi.org/10.1007/978-981-33-6750-0>.

слабом ИИ система может обрабатывать только символы, не понимая, что она делает. В сильном ИИ система тоже обрабатывает символы, но уже понимает, что она делает, потому что она оснащена разумом, подобно людям. Правда, в указанном стандарте отмечается, что обозначения «слабый ИИ» и «сильный ИИ» в основном важны для философов и неактуальны для практиков ИИ.

В упомянутой выше Национальной стратегии развития искусственного интеллекта в России на период до 2030 года дано прагматичное определение ИИ как «комплекса технологических решений, позволяющих имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека».

В стандарте ISO CD 22989 ИИ понимается как «дисциплина, которая изучает инженерные системы, способные приобретать, обрабатывать и применять знания и опыт». При этом как научная дисциплина ИИ включает в себя несколько подходов и методов: от машинного обучения и машинного мышления, включающего прогнозирование, планирование, представление знаний и рассуждение, поиск и оптимизацию; робототехники до интеграции всех других методов в киберфизические системы.

Анализ определений в национальных и международных документах по ИИ показывает, что предпочтение отдается «рабочим» определениям. Однако есть основания рассчитывать, что на уровне ИСО/МЭК будут в ближайшее время сформулированы относительно общепринятые представления об ИИ и утверждены в качестве стандартов.

В проекте национального стандарта, посвященного классификации систем ИИ 2020, подготовленного АО «ВНИИС» с учетом предложений РГ 01 ТК 164 ИИ, указано, что он разрабатывается в целях установления принципов классификации систем ИИ и повышения эффективности их использования для решения прикладных задач как в автономном режиме, так и во взаимодействии с человеком-оператором.

Классификация ИИ позволяет сравнивать различные решения в СИИ по таким параметрам, как вид деятельности, структура знаний, функции контура управления, безопасность, конфиденциальность, степень автоматизации, методы обработки информации, интеграция/интероперабельность, комплексность системы, архитектура, специализация.

Таким образом, для прагматических целей смыслы и значения понятия ИИ (точнее – СИИ) можно считать определенными. Неопределенности являются предметом как дальнейших научных

изысканий, так и осмысления практических проблем, возникающих в ходе цифровой трансформации, имеющей серьезные социальные проекции.

### **1.3.3. Статус жизненного цикла систем ИИ**

Имеются обоснованные ожидания, что массовое внедрение технологий ИИ обеспечит к 2025 г. удвоение темпов роста ВВП ведущих государств мира и увеличение мирового ВВП на 15 трлн долл. Внедрение технологий ИИ даст значительный экономический и социальный эффект в промышленности, энергетике, сельском хозяйстве, финансовом секторе, образовании, городской инфраструктуре, в области безопасности и противодействия терроризму, на транспорте, в обороне и во многих других сферах.

В 2017 г. первые пять стран приняли национальные программы в сфере ИИ, в 2022 г. таких стран уже больше сорока. При этом в мире есть лишь пять стран, которые создали такие важные собственные элементы цифровой экосистемы, как поисковик, социальные сети, развитые школы криптографии, инфраструктуру кибербезопасности, систему подготовки кадров в области математики, логики, программирования, ИТ. К этим странам относится и Россия.

Однако в истории развития СИИ присутствовали как моменты бурного развития, привлекающего значительные инвестиции, так и этапы утраты интереса к феномену ИИ, разочарования в ожиданиях. В 60-х годах XX в. ставились и решались достаточно сложные задачи в науке и технике, управлении, военном деле. Но часть из них не была удовлетворительно решена из-за отсутствия требуемых вычислительных мощностей, производительности компьютеров, а отчасти из-за социальных причин. В настоящее время наблюдается новый всплеск интереса к СИИ, зачастую в новых формулировках, создающих иллюзии новизны. В любом случае на сегодняшний момент можно констатировать фазу завершения этапа стартового развития систем, связанных со «слабым» ИИ. Созданы социальные сети, электронные торговые площадки, системы видеомониторинга, военные и транспортные устройства, центры сбора и обработки «больших данных», где используются СИИ, возникло общественное принятие СИИ как «нормальной» социальной данности, облегчающей жизнь при заказе такси, уборке квартиры, поиске полезной информации, получении госуслуг, обучении и т.д.

Выделяются следующие приоритетные направления исследований в области ИИ: нейронные сети; глубокое обучение; теория управления, умный контроль; компьютерное зрение; технологии поиска и

оптимизации; анализ речи; генерация естественного языка и речи; компьютерная логика и рассуждение, когнитивные вычисления; вероятностные методы выбора в условиях неопределенности; классификаторы и методы статистического обучения; технологии взаимодействия систем с искусственным интеллектом; нейроинтерфейсы; чтение сигналов мозга, нейроинформатика; электроэнцефалография. Эти направления представлены в деятельности научно-исследовательских организаций России практически повсеместно, коррелируя с размещением основных научных центров страны. Однако в практике работы предприятий промышленного сектора России данная тематика представлена в меньшей степени. Это отчетливо подтвердилось в проведенном нами исследовании<sup>1</sup>. Оно, в частности, выявило, что одной из двух наиболее существенных угроз, сопровождающих цифровизацию, является «деградация естественного интеллекта». Готовность ее парировать является низкой. Прежде всего потому, что возникновение такой угрозы является результатом сочетания множества процессов эволюции социума. В связи с этим исследования СИИ должны, очевидно, сопровождаться приоритетными усилиями в понимании особенностей эволюции «естественного» сознания, коэволюции и гибридизации искусственных и естественных систем.

#### **1.3.4. Проблема субъектности личности и надличностные системы**

При разработке СИИ неизбежно допускается некоторая степень неопределенности, которая сходна со свободой воли человека (встроенные этические самоограничения, знания о нормах поведения людей, имитация процесса саморегуляции своего поведения, способность к эмпатии, механизм прогнозирования рисков и последствий собственных действий, возможность исправления собственной ошибки, в том числе исключение действий, связанных с особым риском для человечества).

Среди наиболее острых вопросов развития ИИ сегодня – нейротехнологии, где наблюдается стремительный прогресс НИР, ОКР, практических применений. Нейротехнологии – это класс технологий, которые используют или помогают понять работу мозга, мыслительные

---

<sup>1</sup> Агеев А.И., Асанова Е.А., Глибенко О.В. и др. К «цифре» готов? Оценка адаптивности высокотехнологичного комплекса России к реалиям цифровой экономики / Под ред. А.И. Агеева. М.: ИНЭС, 2018. 60 с.

процессы, высшую нервную деятельность, в том числе технологии по усилению, улучшению работы мозга и психической деятельности.

Формирование целостности личности в современном социуме происходит в процессе динамического взаимодействия людей и надличностных систем. В личности воплощаются когнитивные проекции социума, индивидуализированные, во-первых, биологически (тело); во-вторых, информационно (первичная и вторичная социализация с освоением видов грамотности, коммуникативные связи и базы данных); в-третьих, когнитивно (знания, чувствования и понимания); в-четвертых, социально (как члена агрегированной группы, которых может быть множество по разным критериям). Отсюда – неизбежность комплексирования различных моделей поддержки лояльной к правовым нормам, управленческим ключевым установкам (в том числе и цифровым, анонимным) и глубинным регуляторам жизнедеятельности («матрикс») социальной среды. При этом сбои в работе киберфизических систем могут быть вызваны дистанционным способом. Современный уровень технологий позволяет осуществлять зондаж и воздействие с учетом психосемантических качеств личности расширенного характера (официальной и реальной политической ориентации, качества ее профессиональной подготовки, культурного уровня, интересов, волевых качеств, внутренней мотивации и т.п.).

Новейшие поколения массовых технических устройств дают возможность заинтересованной стороне, опираясь на автоматизированные системы сбора, накопления, обработки и использования данных, не только идентифицировать и определять геолокацию обладателя устройства, его эмоциональное отношение к содержанию электронных сообщений, но и выявлять характеристики окружающего оборудования. Получение информации о состоянии личности через «гаджеты здоровья» дают возможность при необходимости прогнозировать переходы состояний личности в нормальных и в чрезвычайных условиях. Выявление особенностей личности по характеристикам просматриваемых ею информационных программ, активности в социальных сетях, выбору компьютерных игр и т.п. (портфель данных индивидуального электронного контента) позволяет сформировать когнитивно-рефлексивную модель личности. На основе такой модели возможно нейропрограммирование мировоззренческих и ситуативных ориентиров и актов поведения личности и групп людей.

Сведение полученных данных из всех возможных форм электронного контента в пакет информации о психосемантической субъектности личности позволяет с высокой степенью достоверности обнаружить ее поведенческие доминанты и скрываемые качества, а

также принадлежность к социопатической группе. Объектом управления человеком становится его же цифровой двойник, через воздействие на параметры которого можно корректировать его реальное поведение, мышление, интерпретации событий и процессов. При этом сам двойник непрерывно актуализируется по мере «электронной активности» самого человека. Применение СИИ позволяет осуществлять всю необходимую предиктивную аналитику цифрового двойника и его прототипа, правда, пока преимущественно в парадигме Кибернетики 1.0.

По сути, в настоящее время традиционные методы управления социумом превращаются в единую систему с новыми коммуникационными интерфейсами, нейро- и биоинтерфейсами. Для разработки методов прогнозирования интеллектуальной динамики поведенческой активности накоплен значительный исследовательский опыт<sup>1</sup>. Имеется множество перспективных концепций создания многофункциональной информационной мониторинговой системы как платформы прогнозирования (с обратной связью) явных и неявных глубинных процессов и тенденций в социуме, техносфере и природной среде.

### **1.3.5. Выводы**

Окружающая нас социальная реальность насыщается явным и латентным присутствием цифровых технологий, включая СИИ. Это создает растущий поток новых явлений в массовом сознании, наибольшую опасность из которых представляет манипулятивное использование новых зависимостей человеческого мышления и поведения от информационных, виртуализированных и других киберфизических систем. Риски дестабилизации социума вследствие как техногенных катастроф, вызванных «человеческим фактором», так и деструктивного воздействия с применением цифровых технологий растут практически по экспоненте. Новейший опыт показывает феноменальные возможности современных СМИ и Интернета по влиянию на массовое сознание, политический выбор, повседневность. Активное введение в информационный и научный оборот терминов «информационная», «гибридная», «когнитивная», «ментальная», «цивилизационная» война, «управление смыслами» и иных тому подобных отражает наиболее опасные аспекты обсуждаемых здесь проблем.

---

<sup>1</sup> Агеев А.И., Логинов Е.Л., Шкута А.А. Конвергентный мониторинг и программирование личности как инструмент оперирования интеллектуальной динамикой поведения больших групп людей // Экономические стратегии. 2018. № 2 (152). С. 70–87.

## 1.4. ПРОТИВОРЕЧИЕ ТЕХНИЧЕСКОГО И СОЦИОГУМАНИТАРНОГО ПРЕДСТАВЛЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

### 1.4.1. Актуальность разрешения противоречия

Цель этого пункта состоит в поиске эффективного подхода к построению систем ИИ для социогуманитарных задач, которые решаются в гибридной (человеко-машинной) реальности и характеризуются высоким уровнем субъектности. Это реальность, в которой отсутствуют традиционные формализованные способы репрезентации классическими методами ИИ. Например, сравнительно новой задачей является описание процессов интеграции ИИ и человека для коллективного и междисциплинарного синтеза научных открытий.

Об этом много говорят и пишут, постоянно идет поиск адекватной методологии для построения такого ИИ, поскольку феномен социально-гуманитарной динамики противоречит традиционным технологическим характеристикам ИИ, и, по всей видимости, такой поиск будет идти еще долго. Настоящий пункт монографии предлагает методику, которая охватывает формализованную и неформализованную части социально-гуманитарных явлений<sup>1</sup>. Используется комбинация холонического и конвергентного подходов, что обеспечивает целенаправленность и устойчивость коллективного управления и принятия решений. Холонический подход обеспечивает интерпретацию соотношения между двумя реальностями, между макро- и микроуровнем, между формальными системами и непредставимыми формально, искусственно разыгрывающими и «живыми» объектами. Конвергентный подход опирается на методы решение обратных задач в топологических пространствах, контролируемой термодинамики, когнитивной семантики моделей ИИ – с учетом включения наблюдателей, которые рефлексивно и когнитивно влияют на реальность.

Во внимание принимается, что коллективное принятие решений в социально-гуманитарной среде принципиально не может быть представлено формальными средствами. Для учета этого императива предлагается идея построения *уравнения искусственной жизни*, которое объединяет формализуемые и неформализуемые части пространства решений, передающие и объединяющие холонические подходы, термодинамическую теорию, когнитивное моделирование и решение обратных задач.

---

<sup>1</sup> Raikov, A.N. and Pirani, M. (2022), "Contradiction of modern and social-humanitarian artificial intelligence", *Kybernetes*, Vol. 51 No. 13, pp. 186-198. <https://doi.org/10.1108/K-01-2022-0057>



Такой взгляд на развитие ИИ направлен на исследование новых возможностей продвинутого (прежде всего, общего и сильного) ИИ, чтобы социально-гуманитарные процессы целенаправленно и устойчиво достигали своих целей, даже если цели эти изначально неизвестны.

Очевидно, что любой подход имеет ограничения. Среднесрочные ограничения, скорее всего, будут состоять в построении когнитивной семантики моделей ИИ, которая не может быть представлена полностью формализованно, и необходимо искать способы ее косвенной репрезентации. Заметим, что разрешение намечающейся проблематизации должно иметь весомые практические последствия, поскольку социально-гуманитарные объекты и события охватывают все явления, связанные с индивидуальным и коллективным поведением людей. Это повлияет на работу экспертов в сетях, развитие сетевых сообществ, деятельность в экстремальных условиях спасательных отрядов, синтез коллективных открытий, функционирование профессиональных сообществ, устойчивое развитие общества и окружающей среды.

#### **1.4.2. Социально-гуманитарная реальность**

Социально-гуманитарные объекты и события охватывают все явления, связанные с индивидуальным и коллективным поведением людей в таких образованиях, как международные организации, правительственные ведомства, компании и другие организованные группы. За основу построения таких групп берутся гуманитарные принципы. Например: гуманность, нейтралитет, беспристрастность и оперативная независимость. Приоритетами гуманитарной деятельности являются, прежде всего, безопасность жизни и уважение к людям<sup>1</sup>.

Все аспекты социогуманитарной реальности не могут быть полностью формализованы с помощью компьютерных моделей и современного ИИ, которые оперируют логикой и нейронными сетями. Современный ИИ может распознавать, предсказывать и отвечать на простые вопросы. Но мышление, понимание, объяснение, постановка и решение проблем, принятие правильных и в то же время беспричинных решений зависит от наблюдателя. Примечательно, что преобладание эффекта наблюдателя сегодня стало очевидным даже в физических процессах, не требующих непосредственного участия человека. Кроме того, человек обладает интуицией, может медитировать, впадать в транс и ощущать себя в медитативном состоянии. Это аспекты построения

---

<sup>1</sup> *Raikov, A.N. and Pirani, M. (2022), "Contradiction of modern and social-humanitarian artificial intelligence", Kybernetes, Vol. 51 No. 13, pp. 186-198. <https://doi.org/10.1108/K-01-2022-0057>*

возможного будущего ИИ<sup>1</sup>, который необходим для того, чтобы помочь человеческому обществу решать нарастающие по сложности социальные проблемы и преодолевать трудности.

Чтобы построить ИИ будущего, необходимо изменить существующую сейчас его парадигму, явно включить в систему с ИИ наблюдателя (агента), который рефлексивно и когнитивно влияет на ситуацию, может тем самым привести ее к «коллапсу». Включение агента в когнитивные процессы требует подключения методов решения обратных задач, учета новой семантики моделей ИИ, возможно, основанной на нелокальных эффектах, реализуемых на квантово-оптическом и релятивистском уровнях. Мощность систем ИИ с такими семантиками может оказаться на 30-50 порядков выше, чем у традиционных детерминированных систем ИИ. Эта семантическая сила ИИ может быть достигнута через использование социально-гуманитарного феномена, характеризующегося одновременно коллективным сознанием и бессознательным. Любая попытка формализации этого феномена искажает его эмоционально-трансцендентную целостность, делает процесс моделирования реальности вырожденным, создает коллапс целого, то есть исключает учет таких явлений как:

- скрытые мысли, эмоции и намерения человека;
- отсутствие сходства между языком и сознанием;
- неформализуемое взаимодействие между субъектами;
- способность к взрывной мыслительной деятельности;
- захват сознания через поведение соседних агентов и др.

Обзор литературы по этой тематике показывает, что используемые для учета этих явлений методы ИИ обычно сталкиваются со следующими трудностями:

- ошибочная диагностика состояния и поведения лиц в сообществе;
- невозможность устранения информационных помех, возникающих при диагностике поведения групп людей;
- необходимость использования различных, зачастую плохо совместимых, средств и методов для оценки поведения различных субъектов, объектов и явлений;

---

<sup>1</sup> Raikov A. (2021). Cognitive Semantics of Artificial Intelligence: A New Perspective. Springer Singapore, Topics: Computational Intelligence XVII, 128 p. <https://doi.org/10.1007/978-981-33-6750-0>

- ограниченная возможность использования визуальной диагностики для реализации функций дорогостоящих датчиков;
- неполнота знаний о глубинном строении организма человека во всей его сложности в контексте внешней визуальной диагностики, ненадежность и ошибки данных («темные» и «грязные» данные»);
- невозможность комплексной интеграции работы различных нейронных сетей на основе алгоритмического подхода и др.

Хотя эффективность современного ИИ заключается в том, что он отказывается от попыток имитировать неформализованное человеческое мышление и эмоции<sup>1</sup>, социально-гуманитарный ИИ по-прежнему требует усилий в направлении создания сильного ИИ. Компьютеры и робототехника должны быть наделены способностью расширять свои возможности за пределы любой конкретной логики и формальной системы, охватывая многие из них быстро и беспрепятственно, чтобы объяснить феномены понимания, трансцендентального сознания и пр.

В рамках этого видения кибернетика третьего порядка является наиболее многообещающей и в настоящее время наиболее зрелой методологической структурой<sup>2,3,4</sup>, весомым вкладом которой является переход от классической к постнеклассической научной рациональности, где реальность нестатична, а погружена в саморазвивающуюся рефлексивно-активную среду, состоящую из явных или неявных сущностей и их явных или неявных отношений. Строится новая трансдисциплинарная связь между всеми социальными и природными сущностями, с использованием новых подходов, которые глубже проникают в чувственные и эмоциональные слои человеческого сознания и коллективного бессознательного. Новое видение гибридной реальности (HyR) и ее структура в ее обновленном понимании<sup>5</sup> – это область, в которой точка контакта между людьми и машинами происходит как динамическая нестабильная система, в которой люди и технология ИИ сосуществуют и влияют друг на друга.

---

<sup>1</sup> *Esposito, E.* (2017), "Artificial communication? The production of contingency by algorithms", *Zeitschrift für Soziologie*, Vol. 46 No. 2, pp. 249–265. doi: 10.1515/zfsoz-2017-1014

<sup>2</sup> *Lepskiy, V.* (2018), "Evolution of cybernetics: philosophical and methodological analysis", *Kybernetes*, Vol. 47 No. 2, pp. 249–261. doi: 10.1108/K-03-2017-0120

<sup>3</sup> *Espejo, R. and Lepskiy, V.* (2020), "An agenda for ontological cybernetics and social responsibility", *Kybernetes*, Vol. 50 No. 3, pp. 694–710. doi: 10.1108/K-06-2020-0390

<sup>4</sup> *Umpleby, S. A. and Medvedeva, T. A. and Lepskiy, V.* (2019), "Recent developments in cybernetics, from cognition to social systems", *Cybernetics and Systems*, Vol. 50 No. 4, pp. 367–382. doi: 10.1080/01969722.2019.1574326

<sup>5</sup> *Perko, I.* (2020), "Hybrid reality development - can social responsibility concepts provide guidance? ", *Kybernetes*, Vol. 50 No. 3, pp. 676–693. doi: 10.1108/K-01-2020-0061

### 1.4.3. Концепция холона

Как отмечено выше, холон интерпретирует соотношение между двумя реальностями, является их посредником. В этом контексте концепция холона<sup>1</sup>, которая введена уже давно для того, чтобы примирить атомистический и холистический (целостный) подходы, получает новую жизнь и важную роль в технологических разработках. Холон может быть связан с любой из сущностей вертикальной рекурсивной связи, которые смешиваются, чтобы интегрировать уровни иерархии организованной системы систем. Например, роль, которая была отведена «марковским одеялам»<sup>2</sup>, может представлять собой пример холонов, которые могут (в так называемых Р-системах) выполнять функцию мембран между различными уровнями биологической организации. Холон – это любая сущность (абстрактная или естественная), которая одновременно охватывает два соседних уровня иерархии с соответствующими объектами реальности, благодаря так называемому эффекту Януса<sup>3</sup>: каждый элемент иерархии видит автономное целое и вершину, обращенную вверх. В современной промышленной и, по аналогии, социально-гуманитарной, инженерии холоны часто воплощают точку контакта между физическим и кибер-физическими системами.

Наиболее примечательной и практически используемой концепцией, реализуемой холоном, является концепция холархии. Холархия – это временное целенаправленное группирование холонов в иерархические отношения. Концепция холархии особенно полезна при использовании в качестве функциональной структуры для выражения аутопоэзиса (аутопоэза)<sup>4</sup> и эмерджентной самоорганизации в системах.

Открытость в детализации и количестве уровней холархии требует эпистемического инварианта. Архитектура многоагентной системы реляционной модели (RMAS)<sup>5</sup> определяет реляционную модель как такой инвариант-кандидат. Агенты, реализующие холоны в RMAS,

---

<sup>1</sup> Koestler, A. (1970), "Beyond atomism and holism—the concept of the holon", *Perspectives in Biology and Medicine*, Vol. 13 No. 2, pp. 131-154, doi: 10.1353/pbm.1970.0023 Koestler, A. (1970), "Beyond atomism and holism—the concept of the holon", *Perspectives in Biology and Medicine*, Vol. 13 No. 2, pp. 131-154, doi: 10.1353/pbm.1970.0023

<sup>2</sup> Palacios, E.R., Razi, A., Parr, T., Kirchhoff, M. and Friston, K. (2020), "On Markov blankets and hierarchical self-organisation", *Journal of Theoretical Biology*, Vol. 486, p. 110089, doi: 10.1016/j.jtbi.2019.110089

<sup>3</sup> Raikov, A.N. and Pirani, M. (2022), "Contradiction of modern and social-humanitarian artificial intelligence", *Kybernetes*, Vol. 51 No. 13, pp. 186-198. <https://doi.org/10.1108/K-01-2022-0057>

<sup>4</sup> Чайковский Ю.В. Аутопоэз. Опыт пособия тем, кто хочет понять эволюции живого. М.: Товарищество научных изданий КМК. 2018. 560 с.

<sup>5</sup> Pirani, M., Bonci, A. and Longhi, S. (2022), "Towards a formal model of computation for RMAS", *Procedia Computer Science*, Vol. 200, pp. 865-877, doi: 10.1016/j.procs.2022.01.284

являются кандидатами на роль амфибий во множестве реальностей, совместно созданных в свете кибернетики третьего порядка<sup>1</sup>, в которой может быть сформирована адекватная структура для проактивного объяснения в среде НуР. Такая холархическая структура для понимания НуР и функционирования в этой среде системы ИИ очень желательна, поскольку в настоящее время существует пробел в теории систем и инженерии, когда речь идет о точке контакта между людьми и технологиями.

В таком контексте семантика социально-гуманитарных моделей ИИ не может быть всесторонне описана формализуемым способом. Однако эта семантика может быть охвачена косвенными подходами, основанными на учете субъективной реальности, когнитивных семантик, квантовых и релятивистских эффектов<sup>2</sup>.

#### 1.4.4. Конвергентная интерпретация

Неформализуемые и нелокальные средства семантической интерпретации моделей ИИ используются в упомянутом выше конвергентном подходе, который позволяет допустить ситуацию, когда цель может быть очень плохо определенной, или неизвестной вовсе. Процесс решения задачи в таком случае будет иметь неустойчивый характер, однако применение специального структурирования информации создает необходимые условия для того, чтобы сделать процесс решения более целенаправленным и устойчивым. В процессе решения обратной задачи группа людей делает оценки промежуточных этапов и вносит в процесс качественную информацию, подсказанную, возможно, мыслями и чувствами, обеспечивая тем самым устойчивую целенаправленность (конвергенцию) решения. Правила структурирования информации специальным конвергентным способом (конвергентные правила) таковы:

- дерево целей должно ранжироваться по уровням и важности;
- пространства интерпретаций семантики модели ИИ должны быть сепарабельны (по Хаусдорфу);
- пространство ресурсов должно быть покрыто конечным числом классов (компактное пространство);
- лица, принимающие решения, могут вводить в процесс информацию, лежащую за пределами логики и формализма.

---

<sup>1</sup> *Lepskiy, V.* (2018), "Evolution of cybernetics: philosophical and methodological analysis", *Kybernetes*, Vol. 47 No. 2, pp. 249–261. doi: 10.1108/K-03-2017-0120

<sup>2</sup> *Raikov A.* (2021). *Cognitive Semantics of Artificial Intelligence: A New Perspective*. Springer Singapore, Topics: Computational Intelligence XVII, 128 p. <https://doi.org/10.1007/978-981-33-6750-0>

Поскольку нейроны мозга и клетки тела состоят из атомов, квантовые эффекты также могут влиять на процессы мышления и принятия решений в НуР. В этом случае квантовые уравнения, такие как уравнение Шредингера и принцип неопределенности Гейзенберга, возможно, могут подсказать, как лучше сделать описания состояний и поведения мыслительного процесса. Таким образом, неформализуемая когнитивная семантика моделей ИИ, основанная на квантовых и корпускулярно-волновых аспектах теории поля, которая не может быть полностью описана с использованием классического логического и нейросетевого подходов к ИИ, поддается косвенной репрезентации.

Для создания НуР необходимо построение пространства доверия между мотивированными людьми, которые принимают коллективные решения. Требуется два основных столпа для поддержки коллективного решения. Первый, математический, формирует упомянутые выше конвергентные правила. Второй, термодинамический, предлагает правила обеспечения устойчивости коллективного процесса принятия решений, а именно тонкий баланс между формализуемыми и хаотическими компонентами события обеспечивает устойчивость процесса принятия решения. При этом атомарная и волновая природа нейронов человеческого мозга и клеток тела также влияет на процессы коллективного мышления и принятия решений.

Перечисленные компоненты формируют когнитивную (неформализуемую) семантику искусственного сильного интеллекта (ASI). В отличие от слабого ИИ в ASI агент может давать значимые результаты для наблюдателя, даже не понимая смысла своих операций. ASI может принимать правильные решения только погружаясь в НуР. Известный пример такого слабого процессора ИИ был дан в знаменитом примере китайской комнаты. Сущность ASI должна проявлять понимание как сознательное Я. Обнаружение, моделирование и контроль возникновения такого сознания у искусственного существа, независимо от того, появится оно когда-либо или нет<sup>1</sup>, является пока далеко не решенной проблемой. Такие возможности ASI можно описать только интегрируя специальным образом множество различных дисциплин, см. на эту тему работу<sup>2</sup>.

Если электромагнитная волна характеризует когнитивную семантику, то тогда она квантуется, что является имманентным

---

<sup>1</sup> *Fjelland, R.* (2020), "Why general artificial intelligence will not be realized", *Humanities and Social Sciences Communications*, Vol. 7 No. 1, pp. 1-9, doi: 10.1057/s41599-020-0494-4.

<sup>2</sup> *A. Raikov*, "Convergent Ontologization of Collective Scientific Discoveries," 2021 14th International Conference Management of large-scale system development (MLSD), 2021, pp. 1-5, doi: <https://doi.org/10.1109/MLSD52249.2021.9600184>

свойством квантовых частиц. Такая модель и аналогии могут быть отчасти полезны и для объяснения человеческого понимания. Прежде чем слова смогут выразить мысль, процесс мышления выглядит хаотичным. Цель мышления – создать целенаправленный порядок в хаосе. В то же время, если этот порядок строится только на основе имеющихся знаний, то ожидаемого прозрения новой идеи не произойдет, поскольку мысль будет ограничена ретроспективными логическими стереотипами. Когда работают только правила дедукции и индукции вывода, а также абдукции, то процесс может потерять целенаправленность.

Кроме того, логическая формула не может полностью описать мысли и эмоции, потому что они сразу же подчиняются символической схеме, то есть мысли и эмоции претерпевают коллапс. Вербализация всегда обречена на пренебрежение потенциалом эмоций и мыслей. Однако некоторые фундаментальные законы могут помочь косвенно представить хаотическую мысль. Например, хаотическая особенность мышления может быть представлена соответствующей динамической системой, но имеющей некоторое многомерное пространственное состояние, которое можно представить, например, средствами из теории групп. Два события связаны друг с другом, образуя связанные подгруппы Лоренца: одно вращение переходит в другое посредством непрерывного изменения. При такой интерпретации состояние мыслящей системы задается набором положений и векторов событий, например, в системе обобщенных координат и скоростей пространства Минковского или векторов и их производных в бесконечномерном гильбертовом пространстве. Состояния этих пространств с течением времени переходят одно в другое. Эти изменения могут быть определены как формализованными, так и неформализованными способами.

#### **1.4.5. Уравнение Гибридной Реальности**

Агент в системе НуR пытается использовать модели реальностей, чтобы кумулятивно обучаться и достигать своей цели. Однако модели, основанные на математике, справедливы лишь для репрезентации бифуркации конвергенции и дивергенции, организации и дезорганизации<sup>1</sup>. Действительно, стабильность является предпосылкой существования и представления реальности. Агенты должны выйти за пределы любого математического представления, чтобы материализовать целенаправленную конструкцию организации как

---

<sup>1</sup> *Bogdanov, A. (1984), Essays in Tektology: The General Science of Organization, 2nd ed., Intersystems Publications, Seaside, CA, Transl. Gorelik, C.*

естественного «инстинктивного» встроенного механизма, создающего устойчивое состояние знания и бытия как некую собственную форму<sup>1</sup>.

В то время как символические системы ИИ или математические системы обрабатывают моментальные снимки или символические представления вещей, собственная форма – это вещностная сущность, имеющая модельную ассоциацию в математическом контексте. Такая сущность, погруженная в свою среду, делает субъектные запросы к среде с помощью органов чувств и исполнительных механизмов. Соответственно, субъект получает ответы на свои запросы в виде информации и испытывает связанное с этим изменение внутреннего состояния, что является актом познания. Сущность немного узнает что-либо новое, когда повторные запросы дают одни и те же ответы. Правда, в этом случае агент находит (или ощущает) себя в стабильном состоянии бытия. Однако, это может привести к состоянию «скуки», которое может вызвать смену парадигмы состояния. Если же некоторые ответы не приходят постоянно или не соответствуют ожиданиям, сущность испытывает «раздражение»<sup>2</sup>. Раздражение запускает новую процедуру познания, чтобы достичь другого стабильного состояния бытия.

Примером собственной формы является автономный агент, пытающийся контролировать положение дел; она/он/оно должны испытывать равновесие между неформализуемым в целом бытием (т.е. эффективным действием в окружающей среде) и формализуемым знанием.

При использовании теории категорий с двумя реальностями (субъектной и объектной, соответственно: *Mir № 1* и *Mir № 2*) формируется структурирующий функтор, преобразующий любую символическую систему (вычислительный процесс, конечный автомат и т. д.) в динамический системный процесс с участием субъекта. В этом направлении вычисление воплощается в «законах физики». Наоборот, забывающий функтор определяет абстракцию, которая представляет собой подъем из физического мира в некоторую логическую или символическую систему. Эти функториальные отношения при использовании в двух направлениях создают возможность трансформации: сущности *Mira № 1* могут видеть и использовать абстракции *Mira № 2*; сущности *Mira № 2* могут видеть и использовать

---

<sup>1</sup> Von Foerster, H. (2003), “Objects: tokens for (eigen-) behaviors”, in *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer-Verlag, New York, Berlin, Heidelberg, pp. 261-271.

<sup>2</sup> Füllsack, M. (2014), “The circular conditions of second-order science sporadically illustrated with agent-based experiments at the roots of observation”, *Constructivist Foundations*, Vol. 10 No. 1, pp. 46-54, available at: <http://constructivist.info/10/1/046>



абстракции *Мира № 1*. Такая трансформация в работе<sup>1</sup> называется *Уравнением Гибридной Реальности*. Оно введено для того, чтобы выразить огромное влияние подобных трансформаций на гибридную реальность, поскольку его решение может оказать видимое воздействие на науку, а его последствия – на внедрение технологий. Любая реализация, которая достигает решения такого уравнения, становится организованной и стабильной сущностью.

#### **1.4.6. Выводы**

На основе изложенного в этом разделе, можно выделить следующие основания для построения искомым социогуманитарных критериев оценки инноваций, использующих цифровые технологии, включая ИИ:

- Все аспекты социогуманитарной реальности не могут быть формализованы с помощью компьютерных моделей и традиционного ИИ, который оперирует логикой и нейронными сетями.
- Мышление, понимание, объяснение, постановка и решение проблем, принятие правильных и в то же время беспричинных решений зависит от наблюдателя, который должен быть непосредственно учтен в работе системы ИИ.
- Необходимо использование концепции холона, чтобы примирить атомистический и холистический (целостный) подходы.

---

<sup>1</sup> Raikov, A.N. and Pirani, M. (2022), "Contradiction of modern and social-humanitarian artificial intelligence", *Kybernetes*, Vol. 51 No. 13, pp. 186-198. <https://doi.org/10.1108/K-01-2022-0057>

## 1.5. НЕПРЕДСКАЗУЕМЫЕ УГРОЗЫ ОТ ЗЛОНАМЕРЕННОГО ИСПОЛЬЗОВАНИЯ СИЛЬНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

### 1.5.1. Актуальность роста рисков

Основные способы и подходы, которыми ИИ может негативно повлиять на социальную стабильность, международные отношения и национальную безопасность, исходят не столько из футурологического супер-разума или «дронов-убийц», а из-за рисков, связанных с доверием людей к результатам работы компьютерного ИИ, который делает выбор и дает советы. Внедрение ИИ порождает риски решений и ошибки в данных, ломку естественной этики, деградацию аналитических способностей человека, возможность неправильного восприятия и недопонимания.

Заметные риски в системах ИИ возникают из-за того, что для обучения им требуются большие объемы данных. Но данные могут быть «грязными» или «темными», и они несут в себе ретроспективные стереотипы. Обмен данными сталкивается с политическими проблемами, потому что государства могут не желать делиться конфиденциальными данными из-за страха раскрыть слишком много информации, важной с точки зрения решения вопросов национальной безопасности.

Компьютеры и ИИ хорошо работают в контролируемых средах, например, при решении задач распознавания ситуаций, образов, болезней и пр. Но в иных условиях ИИ может быть даже опасен. Функционирование ИИ в реальных условиях характеризуется высоким уровнем неопределенности, необъяснимости, «туманности». В литературе можно найти выделение трех уровней, на которых отсутствие явного погружения в социальную среду ИИ предполагает опасности и сбой социального развития<sup>1</sup>:

- 1) Между новыми и устаревшими системами ИИ;
- 2) Между системами ИИ, операторами и лицами, принимающими решения;
- 3) Между организациями, использующими различные системы ИИ.

Чрезмерная уверенность в возможностях ИИ может привести к тому, что на эти возможности будут рассчитывать в ситуациях, которые слишком сложны для машины. Более того, в гибридных (человеко-машинных) системах избыточная уверенность в правильности работы

---

<sup>1</sup>[https://tnsr.org/roundtable/policy-roundtable-artificial-intelligence-and-international-security/#\\_ftn12](https://tnsr.org/roundtable/policy-roundtable-artificial-intelligence-and-international-security/#_ftn12)

системы ИИ может привести к ситуации, когда люди в процессах мышления будут переносить суждения на алгоритмы, в которых скрыты недостатки.

Уже сейчас ИИ приносит в нашу жизнь социально-экономические риски, к которым, как известно, можно отнести:

- Возможность манипулирования общественным мнением. ИИ знают своего пользователя и могут эффективно предлагать ему целевую рекламу, социально-политическую пропаганду.
- Возникновение нового цифрового класса и цифровой власти. Лица, получившие доступ к управлению системами ИИ в социально значимых областях (персональные данные, трудоустройство, образование, медицина, кредитование и др.) получают в свои руки новый вид власти.
- Вмешательство в частную жизнь с помощью ИИ. Установка в городах камер с распознаванием лиц, сведение вместе и анализ персональных данных из различных источников и т.п., нарушают конституционные права граждан. Массовую слежку могут оправдать соображения безопасности жизни, однако ситуацию, при которой все граждане рассматриваются как потенциальные преступники нельзя назвать адекватной.
- Рыночное внедрение ИИ порождает риски потери миллионов рабочих мест за счет того, что часть работы будет выполнять ИИ.
- Большинство данных граждан, собираемых и анализируемых цифровыми платформами и системами ИИ, сейчас находятся в «серой зоне», то есть эти процессы сбора и обработки данных не урегулированы в законодательстве.
- Отсутствует правовой статус у систем ИИ, принимающих автономные решения и имитирующих когнитивные функции человека.

Развитие перечисленных рисков в ближайшем будущем может привести к серьёзному социальному напряжению. Вместе с тем угроза от повышения возможностей ИИ многократно возрастает в связи с его парадигмальным развитием. Смена парадигмы ИИ в направлении от классической к неклассической и пост-неклассической версии ИИ может привести к скачкообразному наращиванию рисков использования ИИ.

Искусственный интеллект свои 70 лет жизни развивался волнообразно. Вера в его возможности то росла, то падала, а иногда совсем пропадала, например, в начале 1990-х. При этом, однако, увеличиваются финансовые расходы, расширяется его сфера применения. Одновременно растут и риски его использования для

общества, увеличивается глубина его проникновения в личную жизнь и сознание человека.

Вместе с тем в прошлом веке всегда оставалась вера в то, что мышление можно представить с помощью логики и искусственных нейронных сетей. Для этого строились семантические сети, развивались неклассические логики, создавались системы дедуктивного и индуктивного вывода, плодились нечеткие базы знаний и пр. Результат зачастую оставался неожиданным – компьютер и система ИИ нас не понимает, на вопросы он дает совсем неправильные ответы.

### **1.5.2. Угрозы злонамеренного использования ИИ**

Даже в рамках сложившихся классических стереотипов ИИ становится опасным. Например, растут риски злонамеренного использования ИИ (Malicious use of AI, MUAI) государственными и негосударственными структурами для дестабилизации психологической стабильности общества в виде увеличения активности по нейтрализации угроз социуму. Так, в своих работах Евгений Пашенцев (см. ссылки ниже) показывает, что некоторые негативные факторы и последствия развития ИИ могут сильно угрожать международной психологической безопасности. Автор выделил три уровня угроз:

1) MUAI-угрозы связаны с намеренно искаженной интерпретацией обстоятельств и последствий развития AI в пользу антисоциальных групп.

2) MUAI нацелен в первую очередь не на управление целевой аудитории в психологической сфере, а для совершения других злонамеренных действий, например, разрушение критической инфраструктуры.

3) MUAI нацелен в первую очередь на нанесение ущерба в психологической сфере.

MUAI приобретает большое значение в целенаправленной психологической дестабилизации политических, социальных и экономических систем. Целенаправленные операции при поддержке систем ИИ уже являются фактом реальности. Этот факт предъявляет новые требования к обеспечению национальной и международной безопасности<sup>1</sup>. Различными авторами отмечается, что раздельный анализ вредоносного психологического воздействия через дипфейки, боты, предиктивную аналитику и т.д. не позволяет учесть синергию и

---

<sup>1</sup> Telley, Ch. The Influence Machine: Automated Information Operations as a Strategic Defeat Mechanism; The Institute of Land Warfare, Association of the United States Army, Arlington, VA, USA, 2018

эмерджентность такого воздействия на социум. Отсутствие всестороннего анализа объясняется новизной темы. Технологии и системы ИИ входят в гражданскую и военную сферы; быстро распространяется на рынке и социальной среде; по мере повышения их качества и доступности минует порог зрелости ИИ, когда они созреют для использования в МУАИ антисоциальными субъектами.

Авторы работ<sup>1,2</sup> считают маловероятным, что общество может быть сохранено без радикальных прогрессивных преобразований, которые помогли бы устранить угрозу МУАИ. Неизбежная широкая доступность средств глобального воздействия на общество на основе новых цифровых технологий и ИИ для небольших групп и отдельных лиц ставит вопрос о прекращении существования благоприятных условий для антисоциальных групповых интересов. Поэтому ИИ рассматривается не только как набор конкретных преимуществ и угроз, но так же, как практический и необходимый инструмент для достижения нового баланса в человеческом обществе.

Опасной сферой использования ИИ остается этика: идет ее ломка, свертывание. Этика, по результатам исследования А.И. Агеева, характеризуется следующими аспектами и вопросами:

- Типами онтологий;
- Древним наследием;
- Этическими нововведениями Средневековья и позже;
- Антинормативным поворотом;
- Результатами долгой эволюции человечества;
- Постмодерном;
- Этическим разнообразием;
- Что такое хороший экшен? Ценность человеческой жизни? Что такое справедливость? Что такое хорошая жизнь?

В результате же воздействия цифровой среды происходит:

- Взлом психики и сознания;
- Растут риски гибридизации кибернетических и естественных сфер;
- Эволюция морали ведет к падению ее всеобщих принципов.

---

<sup>1</sup> Experts on the Malicious Use of Artificial Intelligence: Challenges for Political Stability and International Psychological Security. Report by ICSPSC. Coordinated by Evgeny Pashentsev. Moscow. 2020.

<sup>2</sup> Darya Bazarkina, Kaleria Kramar. The Malicious Use of Artificial Intelligence and International Psychological Security. ASRIE Analytica Geopolitical Report. 2020. Vol. 1. Rome: ASRIE. URL: [http://www.asrie.org/wp-content/uploads/2020/02/Geopolitical-Report-Volume-1\\_2020\\_-Russian-researchers-and-MUAI.pdf](http://www.asrie.org/wp-content/uploads/2020/02/Geopolitical-Report-Volume-1_2020_-Russian-researchers-and-MUAI.pdf)

### 1.5.3. Аспекты злонамеренной этики

В системы ИИ при моральном разнообразии закладываются столь же разнообразные этики. Внедрением всеобщих этических принципов развития ИИ это не остановить. Кодексы этики ИИ делаются и их делать нужно, но в силу цифровой деградации мышления и цифрового искажения самой этики, они пока слабо воспринимаются и бессильны против бизнес-стимулов рынка и интенций внедрения милитаристских ИИ-приложений.

В настоящее время ключевое различие между искусственным и естественным интеллектом можно интерпретировать как разницу между решением одной задачи и множеством задач в пространствах с различиями в структурных и динамических аспектах, а также отсутствие способности систем ИИ давать объяснение своим выводам.

Способность естественного интеллекта оперировать множеством концепций из разных областей, комбинировать их и создавать новые знания, причем, зачастую нелогичным образом, беспричинно, явно превосходит возможности систем искусственного интеллекта. Однако вопрос о совместной эволюции искусственного и естественного интеллектов, и возникновении комбинированных форм интеллекта (например, гибридный ИИ) уже перерос из теоретической в практическую плоскость.

При этом, следует заметить, естественный интеллект в цифровой среде деградирует, чему способствует следующее:

- Развитие «клипового мышления», иждивенческим образом (можно не думать) заменяющего мыслительные процессы, восходящие к логике и традиционной грамотности.
- Замена непрерывного естественного сигнала передачи данных дискретным, цифровым, что накапливает ошибки в передаваемой информации<sup>1</sup>.
- Рост интеллектуальной и психической зависимости от электронных устройств (передача функций памяти гаджетам).
- Стирание границ между реальностью и иллюзией, что облегчает манипулирование массовым сознанием в коммерческих и политических целях.
- Заимствование ценностей и потребностей из цифровых шаблонов, предлагаемых компьютерными программами, играми, социальными сетями, что происходит почти подсознательно.

---

<sup>1</sup> Райков А.Н. Ловушки для искусственного интеллекта // Экономические стратегии. – 2016. № 6.- С. 172-179.

С другой стороны, инструментарий ИИ постоянно развивается, правда, развивается экстенсивно: растет число слоев в нейронных сетях, увеличивается скорость вычислений, множится число параметров для обучения нейронных сетей. Такое эволюционное развитие идет уже много лет, с момента зарождения феномена ИИ в эпоху появления компьютеров.

#### **1.5.4. Потребность парадигмального прорыва**

Возможен парадигмальный прорыв в развитии ИИ. Так, теоретически и практически ИИ пока не захватил релятивистскую, некаузальную, квантовую, термодинамическую природу процессов мышления. Возможно, именно поэтому человечество переживает сейчас кризис не только в экономике, финансах, доверии, но и в физике, математике. Та же физика вот уже много лет ограничила свою сферу исследования. Размеры наблюдаемых объектов, к которым относится и человек, измеряются от  $10^{-33}$  м (длина Планка) до  $10^{27}$  м (радиус Вселенной). Наблюдаемые времена жизни Вселенной находятся на шкале от  $10^{-43}$  с до  $10^{20}$ . Однако наука пока не может эти границы преодолеть. В электронных библиотеках много на эту тему скопилось знаний. По-видимому, требуется новая парадигма ИИ, которая поможет человеку эти ограничения преодолеть, в том числе, за счет поддержки междисциплинарной креативной работы.

Более продвинутый ИИ (общий, сильный, гибридный и пр.), по всей видимости, сможет помочь человеку это сделать. Однако, если новый ИИ получит способность делать подобно человеку творческие прорывы, породить инсайты, тогда вопрос о злонамеренном использовании ИИ встанет с новой силой, правда риски последствий от того, в чьи руки он попадет, будут несоизмеримо выше.

Сильный ИИ (ASI) принципиально отличается от современного Слабого ИИ (AWI) и его многозадачного варианта – Общего ИИ (AGI). ASI – это пока футурологический проект. Такой ИИ носит гибридный характер, то есть система ASI включает человека и машину, которые слаженно функционируют особым сходящимся и целенаправленным образом. Компьютеры для ASI будут построены не только на цифровой основе, но и, скорее всего, на основе оптической. Задачи, которые современные суперкомпьютеры решают месяцами, оптический процессор будет решать почти мгновенно. Это произойдет за счет отказа от представления обрабатываемых данных в дискретной (цифровой) форме. Цифровой (современный) компьютер будет выполнять определенные операционно-управленческие функции.

Модели ASI будут учитывать когнитивную семантику, которая не может быть формализована напрямую, но может быть учтена путем представления моделей ИИ с учетом атомарного уровня человеческого тела и мозга, используя нелокальные квантовые, релятивистские и волновые эффекты. Он также будет обрабатывать символы как AWI и AGI. ASI будет (с участием человека) действительно понимать, что делает; он будет погружен в реальный мир и обладать субъективностью; у него будет способность принимать решения в рамках определенной этической системы. В результате ASI будет принимать и беспричинные решения, которые могут нести непредсказуемые угрозы в развитии социума, которые социум не может спрогнозировать. Отдельные детали построения ASI можно посмотреть в работе<sup>1</sup>.

Поэтому непредсказуемые угрозы из-за злонамеренного использования ASI, обладающего субъективностью и «свободой воли», будут являться наиболее опасными в информационных войнах в сфере психологической безопасности и международных отношений.

Основные способы и подходы, которыми ИИ может негативно повлиять на социальную стабильность, международные отношения и национальную безопасность, исходят, как отмечено выше, из-за рисков, связанных с доверием людей к результатам работы компьютерного ИИ, который делает выбор и дает советы. Внедрение ИИ порождает риски решений и ошибки в данных, ломку естественной этики, деградацию аналитических способностей человека, возможность неправильного восприятия и недопонимания.

### **1.5.5. ИИ в неконтролируемых средах**

Компьютеры и ИИ хорошо работают в контролируемых средах, например, при решении задач распознавания. Вместе с тем чрезмерная уверенность в возможностях ИИ может привести к тому, что на эти возможности будут рассчитывать в ситуациях, которые слишком сложны для машины.

Несмотря на то, что уже получены обнадеживающие результаты внедрения ИИ для решения очень сложных проблем, которые до сих пор были неразрешимы, в литературе отмечается наличие двух факторов, ограничивающих его потенциал:

- очень низкий уровень лидерства и понимания изменений в процессе реагирования на вызовы Индустрии 4.0;

---

<sup>1</sup> Raikov A. (2021). Cognitive Semantics of Artificial Intelligence: A New Perspective. Springer Singapore, Topics: Computational Intelligence XVII, 128 p. <https://doi.org/10.1007/978-981-33-6750-0>



- отсутствие достаточно последовательного, позитивного и общего нарратива, в котором описываются возможности и проблемы Индустрии 4.0.

Такой потенциал необходим для того, чтобы обрести легитимность и доверие. В то время как непонимание руководством организаций изменений приводят к неадекватной институциональной структуре для управления распространением инноваций, отсутствие адекватного и последовательного нарратива также вызывает недоверие к преимуществам Индустрии 4.0. Как отмечают некоторые авторы<sup>1,2</sup>, в отношении ИИ наблюдается далеко ненормальное сопротивление изменениям<sup>3</sup> и что в обществе возникает амбивалентное ощущение, проявляющееся в виде поляризации по всему спектру юридических, технических, академических и государственных практиков. С одной стороны, общество обдумывает ожидаемые результаты и новые пути экономического роста и занятости, которые могут быть открыты с помощью ИИ; с другой – его негативное влияние на отрасли, рынки, организации и граждан, от моделей работы и занятости до неприкосновенности частной жизни, цифрового неравенства, фрагментации общества, безопасности, прозрачности и способов, которыми правительство может защитить своих граждан и его идентичность.

Как правило, риски цифровых инноваций пропорциональны величине выгод. ИИ не является исключением: исследование выявило очень длинный список рисков, связанных с ИИ<sup>4</sup>. Среди них литература выделяет непрозрачность как в структуре, так и в механизмах обучения. В экстремальных ситуациях ИИ проявляет так называемый «эффект черного ящика» (ЭЧЯ), то есть сложность понимания, а скорее, просто непонимания, внутреннего механизма работы систем ИИ. ЭЧЯ вызывает серьезную озабоченность не только потому, что предполагает передачу неконтролируемым и интеллектуальным машинам управления процессами, влияющими на человеческую жизнь без контроля со стороны человека, но и потому, что она связана с нарушениями прав

---

<sup>1</sup> Johnson, D. G. (2015). Technology with No Human Responsibility? *Journal of Business Ethics* 127(4): 707-715

<sup>2</sup> Martin, K. (2019). Ethical Implications and Accountability of Algorithms. *Journal of Business Ethics* 160: 835–850.

<sup>3</sup> Pardo del Val, M. & Martinez Fuentes, C. (2003). Resistance to change: A literature review and empirical study. *Management Decision* 41: 148-155.

Pashiri, R. T.; Rostami, Y.; Mahrami, M. (20) Spam detection through feature selection using artificial neural network and sine–cosine algorithm, *Mathematical Sciences*, 14: 193–199.

<sup>4</sup> Floridi, L., Cowls, J., Beltrametti, M. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations, *Minds & Machines* 28:689–707

человека<sup>1</sup>, ущербом на неприкосновенность частной жизни, несправедливость и сохранение предубеждений, коренящихся в социальном неравенстве<sup>2</sup> или дискриминации<sup>3</sup>.

### 1.5.6. Выводы

На основе изложенного в этом разделе, можно выделить следующие основания для построения искомым социально гуманитарных критериев:

- Внедрение ИИ порождает риски решений и ошибки в данных, ломку естественной этики, деградацию аналитических способностей человека, возможность неправильного восприятия и недопонимания.
- Вопросы злонамеренного использования ИИ приобретают все большее значение в целенаправленной психологической дестабилизации политических, социальных и экономических систем.
- Риски цифровых инноваций, включая ИИ, могут быть пропорциональны величине выгод; сейчас имеется очень длинный список рисков, связанных с ИИ<sup>4</sup>.

---

<sup>1</sup> Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016) The ethics of algorithms: Mapping the debate, *Big Data & Society*, 3(2), 1–21.

<sup>2</sup> Jacobson, N.C.; Bentley, K.H. et al (2020) Ethical dilemmas posed by mobile health and machine learning in psychiatry research, *Bulletin of the World Health Organization*, 98(4): 270–276.

<sup>3</sup> Benthall S. & Haynes, B.D. (2019). Racial categories in machine learning. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, Atlanta.

<sup>4</sup> Floridi, L., Cowls, J., Beltrametti, M. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations, *Minds & Machines* 28:689–707

## 1.6. ВЫВОДЫ

1. В этой главе выделяются приоритетные направления исследований в области ИИ: нейронные сети; глубокое обучение; теория управления, компьютерное зрение; поиск и оптимизация; распознавание речи; генерация языка и речи; когнитивные вычисления; вероятностные методы выбора в условиях неопределенности; классификация и статистическое обучение; взаимодействие с системами ИИ; чтение сигналов мозга, нейроинформатика и др. Среди наиболее острых вопросов развития ИИ сегодня – нейротехнологии, которые помогают понять работу мозга, мыслительные процессы, высшую нервную деятельность.

2. Традиционные методы управления социумом превращаются в единую систему с новыми коммуникационными возможностями. Уже имеется множество перспективных концепций создания многофункциональной информационной мониторинговой системы как платформы прогнозирования явных и неявных глубинных процессов и тенденций в социуме, техносфере и природной среде.

3. Устойчиво убеждение, что новые технологии приводят к трансформации экономических и общественных отношений, и что они все больше обладают качеством субъектов в процессах переустройства мира. Но это далеко не так. Культурные традиции, человеческие качества и способности всегда много сильнее, особенно в когнитивных областях. Технологии, конечно, находят свое место в развитии общей экосистемы научного прогресса, ускоряют этот прогресс, но не ломают кардинально привычных устоев. Без человека цифровые технологии не имеют смысла.

4. С ростом производительности вычислительной техники, успехами в области усложнения и использования нейронных сетей для глубокого обучения, методов и средств анализа больших данных технологии ИИ стали исключительно востребованными. Многие страны мира приняли национальные стратегии развития ИИ. Однако пока они относятся к классу слабого и узкого ИИ, поскольку могут решать хорошо только отдельные задачи, включая задачи распознавания, простых предсказаний и упрощенной имитации человеческой деятельности.

5. Одновременно с увеличением внимания к ИИ во власти и обществе усиливаются требования ограничить использование инструментов ИИ, например, распознавания лиц, которые нарушают права личности и несут риски принятия ошибочных решений. В

развитии ИИ усиливается акцент на необходимости повышения прозрачности работы систем ИИ, учета личностных, социальных, экономических, этических, юридических и иных вопросов.

6. В контексте развития ИИ ставятся извечные вопросы обеспечения справедливости. Например, обращает на себя внимание позиция идеологии социального либерализма, в которой для представления роли и места ИИ предлагается понимать структуру общества как совокупность социотехнических систем, функционирование которых складывается под все возрастающим влиянием ИИ, что можно охарактеризовать как упрощенное понимание организации социальных процессов в виде проявление технократического редукционизма.

7. В контексте рассмотрения вопросов справедливости апологеты либерализма проталкивают идею, что системы ИИ, как бы становятся «живыми», все больше и решительней формирует элементы базовой структуры общества, и что тема развивается при тесном взаимодействии с принципами правосудия, с поведением и мыслями людей, что, в свою очередь, формирует характер человеческих практик, включая распределение выгод и бремени среди населения. Однако, это не так, перенося субъектность на ИИ, скрытым образом снимается ответственность с человека, что нивелирует императив законодательной практики.

8. Все большее внимание уделяется проблеме создания систем Общего и Сильного искусственного интеллекта, идея которого родилась в самом конце прошлого века. Именно такие системы призваны решать одновременно множество разных междисциплинарных задачи. В этом контексте утверждается угроза наступления в скором времени момента сингулярности, когда ИИ станет сильнее человека во всех областях. Вместе с тем, при более внимательном рассмотрении оказывается, что важность разрешения все той же проблемы, а именно субъектности систем ИИ, еще более возрастает.

9. Все более актуальным становится решение этических вопросов применения ИИ. Это не просто комплексное и междисциплинарное дело, и его невозможно выполнить лишь принятием декларативных кодексов этики применения ИИ. Дискуссии об этике ИИ уже ведутся и довольно интенсивно, что предполагает потребность сходимости различных направлений обсуждения в единое конструктивное русло. При этом особого и стержневого внимания

заслуживает проблема субъектности в области разработок этических основ применения ИИ.

10. Начало этого века ознаменовалось фетишем, а скорее – фантомом, ИИ, которому приписывается грандиозное будущее. Однако не обходится и без протестов. Вместе с тем ИИ в традиционном виде или в виде Общего ИИ пока остается «узким», и по парадигме своего развития он не выходит за рамки сложившихся еще в прошлом веке подходов и методов. Его возможности остаются пока далекими от реальных когнитивных возможностей человека.

11. Косвенным признаком фетиша ИИ является отсутствие общепринятого определения ИИ. Международный и национальный стандарты ИИ, задающие его базовые определения и области применения, находятся в фазе начальной разработки, хотя уже введено в действие множество стандартов, охватывающих различные аспекты тематики. Пока применяемые для конкретных целей определения делаются дедуктивно, индуктивно, функционально, либо через родовидовое отличие, либо через отрицание. В этой работе активно участвует Технический комитет ИСО/МЭК 42 «Искусственный интеллект», состоящий из 5 рабочих групп.

12. Идет поиск эффективных подходов к построению систем ИИ, которые погружены в гибридную (человеко-машинную) реальность, характеризуемую высоким уровнем субъектности. В ней почти отсутствуют формализованные способы репрезентации явлений. Например, сравнительно новой задачей является описание процессов интеграции ИИ и человека для коллективного и междисциплинарного синтеза научных открытий. В рамках этого видения кибернетика третьего порядка и исследование вопросов развития ИИ в контексте поведения рефлексивно-активных сред является наиболее многообещающей и в настоящее время наиболее зрелой методологической структурой.

## 2. ЦИФРОВЫЕ ТРАНСФОРМАЦИИ И ИИ (ПРЕОДОЛЕНИЕ ОГРАНИЧЕНИЙ ТЕХНОГЕННОЙ ЦИВИЛИЗАЦИИ)

---

### 2.1. ВВЕДЕНИЕ

В последние годы все более четко проявляется негативная роль использования модели техногенной цивилизации, ориентированной на унификацию локальных цивилизаций, в организации различных сфер жизнедеятельности человечества и особенно в связи с процессами цифровой трансформации и использования ИИ. В первую очередь это связано с ценностно-целевыми ограничениями техногенной цивилизации, в игнорировании социальных ценностей и доминировании ценностей научно-технического прогресса. Это создает благоприятные условия для относительно скрытного использования модели техногенной цивилизации как инструмента рефлексивного управления человечеством в интересах глобалистского проекта.

Одним из признаков заката доминирования модели техногенной цивилизации является практически неконтролируемый со стороны общества бум цифровой трансформации и ИИ, порождающий новые глобальные угрозы для человечества. Среди этих угроз унификация цивилизаций и культур, становление новых жестоких форм колониальной политики, расчеловечивание. Человечество оказалось не готовым к адекватному реагированию на новации техногенной цивилизации. Как следствие справедлив диагноз «бессубъектность развития человечества».

В настоящее время интенсивно ищутся альтернативы техногенной цивилизации, одной из которых могла бы стать модель цивилизации, ориентированной на сохранение многоцивилизационного мира при обеспечении его гармоничного развития. Исходные философско-методологические основания для поиска новой модели цивилизации сформулированы В.С. Степиным, который утверждал, что речь должна больше идти о придании научно-техническому прогрессу существенно большего гуманистического измерения. Этот облик помещает в центр внимания субъектов познания и развития, предполагает анализ их ценностно-целевых структур в неразрывной связи с культурой. При этом необходимо выявить единый процесс эволюции человечества и биосферы.

Модель посттехногенной цивилизации представляется как модель саморазвивающихся полисубъектных (рефлексивно-активных) сред. Если техногенную цивилизацию представить как парадигму, то

становление посттехногенной цивилизации можно представить как разработку мета-парадигмы, например социотехногенной цивилизации, в которой техногенная цивилизация будет представлена как частная парадигма.

При поиске модели пост-техногенной цивилизации в центре внимания оказываются базовые цивилизационные ценности. В этой главе предложены социогуманитарные критерии оценки новаций в техногенной цивилизации и показано, что человечество не готово к адекватному реагированию на новации техногенной цивилизации. Обоснован вывод, что бессубъектность развития – диагноз состояния человечества в техногенной цивилизации. В качестве альтернативы для техногенной цивилизации предлагается модель социогуманитарной цивилизации.

В главе предлагается упреждающий переход к VII социогуманитарному технологическому укладу. Россия может стать лидером развития этого уклада, чтобы не рассчитывать на безнадёжные шансы догнать лидеров VI технологического уклада. Предлагается социогуманитарная цивилизация в виде целостного социального организма, предполагающего определенный тип культуры. Задается вектор возможности России стать инициатором разработки такой модели цивилизации.

Рассматриваются вопросы включенности общества в процессы управления и развития, необходимым условием чего является формирование двухконтурной структуры организации процессов развития.

Весомую роль в развитии нового типа цивилизации начинает играть внедрение цифровых технологий, включая ИИ, в том числе Общий ИИ. Для разработки такого ИИ требуются новые парадигмальные, теоретические и методологические подходы, отвечающие принципам постнеклассической эпистемологии с учетом опыта философского анализа междисциплинарных проблем и результатов когнитивной науки в социогуманитарном антропологическом контексте.

Особое место в главе уделено неклассической концепции системы образования, которая характеризуется субъект-субъектным типом взаимодействия. Учащийся в ней рассматривается как равноправный участник образовательного процесса. В постнеклассической парадигме образования в центре внимания оказываются также субъект-метасубъектные отношения.

## 2.2. СИСТЕМНЫЕ ОСНОВАНИЯ ДЛЯ ПРЕОДОЛЕНИЯ ОГРАНИЧЕНИЙ ТЕХНОГЕННОЙ ЦИВИЛИЗАЦИИ

Современная эпоха характеризуется свертыванием проекта глобализации и закатом однополярного мира. Как следствие, это привело к кризису обслуживающей этот проект цивилизационной модели, которую многие исследователи называют «техногенной цивилизацией»<sup>1</sup>. В цивилизационном контексте техногенная цивилизация ориентирована на унификацию локальных цивилизаций, свертывание многоцивилизационного мира<sup>2</sup>.

Альтернативой техногенной цивилизации могла бы стать пост-техногенная цивилизация, ориентированная на сохранение многоцивилизационного мира, на гармонию цивилизаций в планетарном развитии. В настоящее время делаются первые шаги в поиске такой цивилизационной модели<sup>3</sup>.

Суть техногенной цивилизации проявляется в ее базовой ценности научно-технического прогресса и ценности науки как основы управления социальными процессами. Важно отметить, что при этом фактически игнорируются социальные ценности и этические аспекты регулирования жизнедеятельности.

Как следствие, ограниченность такого подхода проявляется в нарастающих темпах угроз для человека, общества, экологии и собственно ее собственного порождения техносферы.

Угрозы человеку проявляются в разрушение сложившегося образа жизни и экономики стран, не входящих в «круг избранных».

Угрозы обществу проявляются: в нарастании вероятности возникновения глобальных войн, в распространение массовой культуры, общества потребления и культа денег, в перехвате управления через мировые финансово-экономические институты и цифровые платформы и др.

Экологические угрозы связаны с неконтролируемыми процессами новаций в техногенной цивилизации, с доминированием принципа

---

<sup>1</sup> *Степин В.С.* XXI век – радикальная трансформация типа цивилизационного развития // Глобальный мир: системные сдвиги, вызовы и контуры будущего: XVII Международные Лихачевские научные чтения, 18–20 мая 2017 г. СПб.: СПбГУП, 2017. С. 185–188.

<sup>2</sup> *Смирнов А.В.* Проект многоцивилизационного мира как основание идеи многополярности: концепция всечеловеческого сегодня // Национальная философия в глобальном мире: тезисы Первого белорусского философского конгресса. Минск: Беларуская навука, 2017. С. 350-352.

<sup>3</sup> *Лепский В.Е.* Системные основания для перехода от техногенной цивилизации к социогуманитарной цивилизации // Проблемы цивилизационного развития. 2019. № 1. С.33-48.



максимизации прибыли, с игнорированием «мировыми лидерами» международных соглашений по защите окружающей среды и др.

Угрозы техносфере связаны с игнорированием социальных ценностей, с формирующимся барьером между субъектной и виртуальной (цифровой) реальностями, с использованием разработок техносферы в интересах манипулирования обществом и рядом других аспектов<sup>1</sup>.

Критический анализ отдельных аспектов техногенной цивилизации содержится в докладах авторитетных сообществ: Римский клуб, ЮНЕСКО и др. Предлагаются разнообразные концепции и модели: концепция устойчивого развития, модель «Общество 5.0» и др. Эти инициативы крайне важны и полезны, однако следует заметить недостаточно высокий уровень системной проработки этих подходов. В них доминируют отдельные аспекты: экологический, цифровой и др.

Исходные философско-методологические основания для системного подхода к решению этой проблемы были сформулированы В.С. Степиным<sup>2</sup>. Он утверждал, что речь не идет об отказе от научно-технического прогресса, а о придании ему существенно большего гуманистического измерения. Степин утверждал также, что этот тезис достаточно обоснован и принят в отечественной философской науке. При этом новый облик науки понимается в контексте постнекласической научной рациональности, включающей в себя в явном виде гуманистические ориентиры и ценности.

В данном разделе предпринята попытка анализа философско-методологических оснований системного подхода в критике техногенной цивилизации и поиске образа пост-техногенной цивилизации, которую мы предлагаем называть – «социогуманитарная цивилизация».

### **2.2.1. Философско-методологический анализ кризиса техногенной цивилизации**

Большой вклад в осмысление философских аспектов кризиса техногенной цивилизации и поиску путей дальнейшего цивилизационного развития внес В.С. Степин. Он жестко фиксирует свою позицию, согласно которой «современная цивилизация вошла в стадию неустойчивости, кризисных состояний и нестабильности»<sup>3</sup>. Эти

---

<sup>1</sup> Иванов В.В., Малинецкий Г.Г. Философские основания гуманитарно-технологической революции. *Философские науки*. 2019, 62(4). С. 76-95.

<sup>2</sup> Степин В.С. Научное познание и ценности техногенной цивилизации // *Вопросы философии*. 1989. №10. С. 3-18.

<sup>3</sup> Степин В.С. XXI век – радикальная трансформация типа цивилизационного развития // *Глобальный мир: системные сдвиги, вызовы и контуры будущего: XVII Международные*

исследования опирались на разработанную ранее Степиным систему представлений о трех этапах научной рациональности: классическом, неклассическом и постнеклассическом<sup>1</sup>. Фактически, это три научные революции и научные парадигмы в понимании Т. Куна. Важно отметить, что каждый последующий тип рациональности, имея свою специфику, одновременно включает предыдущий тип, задавая соответствующую рамочную конструкцию. В таком понимании развитие научной рациональности представляется как единая система философии науки. Постнеклассическая научная рациональность, помещая в центр внимания субъектов научного познания и развития, предполагает анализ их ценностно-целевых структур в неразрывной связи с культурой, влияющей на их картины мира и индивидуальные особенности. Развитие науки рассматривается во взаимодействии с развитием социальной среды, что задает основания поиска путей трансформации техногенной цивилизации в новые формы цивилизационной организованности. Этот переход требует принципиально новых социальных инноваций, в поиске и оценке которых ведущая роль будет отведена постнеклассической научной рациональности.

Это утверждение опирается на четыре важнейших философско-методологических аргумента в пользу постнеклассической научной рациональности. Во-первых, формирование методологических оснований для целостного представления и конвергенции субъектов, средств и объектов в познавательных процессах и разнообразных проявлениях деятельностной активности. Во-вторых, интеграция интернальных и экстернальных механизмов научного и инновационного развития. В-третьих, введение этических регуляторов в научную деятельность. В-четвертых, представление постнеклассической рациональности как рамочной конструкции, объединяющей все виды научной рациональности.

Основания кризиса техногенной цивилизации связаны в первую очередь с ограничениями ее базовых ценностей. Это ценность научно-технического прогресса и ценность науки как основы управления социальными процессами. «Эти ценности не подвергались сомнению до последней трети XX столетия, пока техногенная цивилизация не столкнулась с глобальными проблемами, порожденными предшествующим научно-техническим развитием»<sup>2</sup>.

---

Лихачевские научные чтения, 18–20 мая 2017 г. СПб.: СПбГУП, 2017. С. 185–188.

<sup>1</sup> *Степин В.С.* Теоретическое знание. М.: Прогресс-Традиция, 2003. 744с.

<sup>2</sup> *Степин В.С.* Научное познание и ценности техногенной цивилизации // Вопросы философии. 1989. №10. С. 3-18.

Особое значение в поиске новых цивилизационных моделей на основе постнеклассической научной рациональности приобретает концентрация внимания на социальных ценностях и целях. По мнению Степина, цивилизационное развитие будет производиться с учетом исторического опыта на основе и в гармонии четырех взаимозависимых базовых ценностей: сохранения и развития человека, человечества, биосферы и техносферы (включая цифровую реальность). Эти ценности оказались за пределами внимания господствующей в современном мире техногенной цивилизации, поскольку с ней органично связаны рыночные отношения и общество потребления. Неразрывно связан с ценностным аспектом образа пост-техногенной цивилизации аспект моральных основ, которые должны быть органично связаны и определять специфику модели пост-техногенной цивилизации<sup>1</sup>.

Принципиально важна ведущая ориентация Степина на системность и междисциплинарность в осмыслении кризиса техногенной цивилизации. Анализируя культурно-генетический код и ценностно-смысловое ядро техногенной цивилизации, Степин выделяет важнейший для нее принцип стимулирования опережающего роста потребления и распространения рыночных отношений на все новые сферы человеческой жизнедеятельности. Как следствие это вызывает обострение экологического и антропологического кризисов. Степин делает вывод, что нужен переход к принципиально новому типу цивилизационного развития. В центре его исследований оказываются механизмы переходов в сложных системах. Он описывает стадии перехода систем в новое состояние, связывая этот переход с идеями синергетики. Принципиально важно, что в завершающей стадии ведущей становится целевая причинность, которую можно представить как целевую детерминацию, тогда процессы самоорганизации общества рассматриваются как базирующиеся на проектной идентификации. В системно-генетическом аспекте выдвигается нестандартный тезис о влиянии будущего на настоящее и даже прошлое.

Снова возникает ключевой вопрос о ценностях, задающих ориентиры перехода к новому типу цивилизационного развития. Степин считает, что новые ценности извне не придут, они должны начать формироваться в недрах техногенной культуры, и важно отыскать их точки роста. Точки роста новых ценностей Степин связывал также с предложенной им постнеклассической научной рациональностью, которая ориентирована на рассмотрение развивающихся

---

<sup>1</sup> Гусейнов А.А. Великие пророки и мыслители. Нравственные учения от Моисея до наших дней. М.: Вече, 2009. 496 с.

человекообразных систем. В контексте этой рациональности в центре внимания оказывается совместное рассмотрение внутринаучных и внеаучных (социальных) этических регуляторов.

Системность подхода проявляется в органичной связи рассматриваемых аспектов фазового перехода систем с политической организацией мирового сообщества, переходящего от однополярного к многополярному миру. Степин ставит нерешенные актуальные проблемы отношений «общество – власть» в XXI в. Для этих отношений еще не найдены новые точки роста ценностей, но проблематизация кризиса техногенной цивилизации является началом поиска такого рода ценностей.

В контексте прогнозов цивилизационного развития Степин обоснованно приходит к выводу<sup>1</sup>, который согласуется с мнением К. Леви-Стросса, что XXI век будет веком социальных наук, или этого века не будет.

К аналогичному выводу приходит и Н.Н. Моисеев. Одна из важнейших идей Моисеева связана с органичной связью человечества и биосферы. Опираясь на труды В.И. Вернадского о ноосферной организации жизнедеятельности и понимая ограниченность постановки проблемы устойчивого развития, он разработал *концепцию коэволюции*. Понятие коэволюции как со-развития (совместного, сонаправленного развития) определяет способ существования человека в мире<sup>2</sup>.

Суть концепции состоит в представлении единого процесса эволюции человека (человечества) и биосферы. Фактически эта концепция задает условия выживания человечества и ценностные ориентиры для преодоления кризиса техногенной цивилизации. Как математик с большим опытом моделирования биосферы он прекрасно понимал ее нелинейную специфику и потенциально возможные труднопредсказуемые скачкообразные изменения ее состояний. Это обосновывает его призывы бережно относиться к биосфере и устанавливать с ней не потребительские, а партнерские отношения.

Рассмотренные идеи сближают двух творцов, Моисеева и Степина, не только цивилизационным масштабom их мышления, но и философско-методологическим представлением биосферы как саморазвивающейся системы, тесно связанной с саморазвивающимися системами общества и культуры. Фактически оба исследователя стоят на

---

<sup>1</sup> Степин В.С. Социальные системы и методология прогнозирования их будущих состояний // Контуры будущего в контексте мирового культурного развития: XVIII Международные Лихачевские научные чтения, 17–19 мая 2018 г. СПб.: СПбГУП, 2018. С. 178-181.

<sup>2</sup> Моисеев Н.Н. Универсум. Информация. Общество. М.: Устойчивый мир, 2001.

позиции преодоления ограничений антропоцентризма, когда в центре внимания был исключительно человек. Сохранение биосферы является также необходимым условием существования человека.

В мировом сообществе идет активный поиск новых способов организации жизнедеятельности, новых форм цивилизационного развития и работы Степина и Моисеева создали философский фундамент для этого поиска.

### **2.2.2. Кризис экономической детерминации развития в техногенной цивилизации**

В последние годы все отчетливее проявляются тенденции, позволяющие сделать вывод, что экономика утратила позицию доминирования в развитии и конструировании будущего. По мнению Р.С. Гринберга<sup>1</sup> для экономики устройство и функционирование окружающего мира становится все менее понятным, доминирующими становятся неэкономические механизмы. Все больше внимания уделяют механизмам общественного участия в процессах управления и развития и их влиянию на экономические механизмы<sup>2</sup>.

Становится все более обоснованной точка зрения, что экономический детерминизм развития порождает также нарастание угроз для человечества при переходе к очередным технологическим укладам. Технологические инновации не проверяются на готовность человечества к их внедрению, на потенциальные последствия для человечества<sup>3</sup>.

Нарастающей угрозой для всех технологических укладов становится отстраненность технологий от этического осмысления. Эта точка зрения подвергнута ревизии. В рамках постнеклассической научной рациональности ставится проблема возврата науки в лоно этической регуляции, однако в отношении технологий делаются только первые шаги в их регламентации со стороны этики<sup>4</sup>.

Для гармонизации мирового сообщества необходим поиск адекватных сложившейся ситуации форм жизнедеятельности, под которые должны быть созданы адекватные им экономические механизмы.

---

<sup>1</sup> Гринберг Р.С. Глобализация, трансформация, кризис – что дальше? М.: Магистр, 2011. С. 9.

<sup>2</sup> Богомолов О.Т. Крупный неэкономический резерв модернизации // Экономические стратегии. 2013, №1. С.14-23.

<sup>3</sup> Лепский В.Е. Проблемы становления субъектности человечества и модели развития // Развитие и экономика. 2011, сентябрь. С.95-101.

<sup>4</sup> Степин В.С. Теоретическое знание. М.: Прогресс-Традиция, 2003. 744с.

Преодолеть кризис экономической детерминации развития возможно за счет социального контроля над техногенной «возгонкой» технократического уклада. Для этого предлагается упреждающий переход к *VII социогуманитарному технологическому укладу*. Если Россия станет лидером этого уклада, то у нее появляются шансы не гнаться за лидерами VI технологического уклада, которых догнать едва ли удастся, и в настоящее время разрыв только увеличивается, а перегнать их, не догоняя<sup>1</sup>.

Философско-методологические основания VII социогуманитарного уклада базируются, на наш взгляд, на современных представлениях философии науки (постнеклассический подход). Нами разработаны концептуальные основы и социогуманитарные модели саморазвивающихся рефлексивно-активных сред, в которых создаются условия для гармоничного развития всех субъектов, обеспечение рефлексивных процессов на основе новых информационных платформ, адекватные механизмы сборки субъектов развития, механизмы конвергенции представительской и прямой демократии и др.

В VII социогуманитарном технологическом укладе в центре внимания окажется проблема регулирования техногенной возгонки по технологическим укладам через переход в VII социогуманитарный технологический уклад<sup>2</sup>. На основе социальных ценностей будет осуществляться управление предыдущими и будущими техногенными укладами в интересах развития человечества.

### **2.2.3. Социогуманитарные критерии оценки новаций в техногенной цивилизации**

Анализ культуры междисциплинарных исследований процессов организации различных видов человеческой деятельности позволяет выделить четыре базовых вида критериев их оценки: продуктивность, безопасность, удовлетворенность и развитие субъектов и самих видов деятельности.

---

<sup>1</sup> Иванов В.В., Малинецкий Г.Г. Философские основания гуманитарно-технологической революции. *Философские науки*. 2019, 62(4). С. 76-95.

<sup>2</sup> Лепский В.Е. Седьмой социогуманитарный технологический уклад — контуры будущего человечества / Глобальный мир: системные сдвиги, вызовы и контуры будущего : XVII Международные Лихачевские научные чтения, 18–20 мая 2017 г. СПб. : СПбГУП, 2017. С. 357-360. Лепский В.Е. Седьмой социогуманитарный технологический уклад – адекватный ответ технологическим вызовам XXI века / Философия в диалоге культур: материалы Всемирного дня философии. М.: Прогресс-Традиция, 2010. С. 1010-1021.

Эти критерии можно использовать и для оценки новаций в техногенной цивилизации. На их основе можно выделить следующие приоритетные направления реагирования на новации.

*Эффективное реагирование* человечества на позитивные аспекты технологических новаций. Готовность общества к разработке и внедрению новаций, к восприятию инновационных предложений и др.

*Контролирующее реагирование* человечества на потенциальные угрозы от внедрения создаваемых технологий.

*Справедливое реагирование* человечества на возможности технологических новаций. Возможность использовать технологические новации в интересах всего человечества (проблема качества жизни и др.), а не только и не столько в интересах узкой группы лиц, обладающих финансовыми и административными ресурсами.

*Развивающее реагирование* человечества на технологические новации. Готовность и возможность человечеству решать задачи стратегического целеполагания, разрабатывать проект своего развития, формировать видение будущего, и с этих позиций оценить создаваемые технологические новации. Осуществлять соотнесение последствий влияния технологических новаций на проект развития, оценивать готовность человечества к использованию новаций, наличие необходимых ресурсов на достижение позитивных результатов и нейтрализацию негативных и др.

Способно ли человечество оценивать технологические новации в соответствие с рассмотренными критериями? Способно ли человечество адекватно оценивать вызовы, угрозы и риски от технологических новаций с учетом базовых ценностей развития? Ответ очевиден – человечество пока не готово к адекватной оценке технологических новаций XXI века!

#### **2.2.4. Бессубъектность развития – диагноз состояния человечества в техногенной цивилизации**

Мировое сообщество постоянно отстает в принятии решений по адекватному реагированию на возникающие глобальные угрозы (ядерную, экологическую, демографическую, информационную, террористическую, финансово-экономическую), во многих случаях вообще отсутствуют механизмы предотвращения и нейтрализации глобальных угроз. Это приводит к нецелесообразному гигантскому расходованию ресурсов, формированию нестабильности в экономических и политических процессах, к возникновению непреодолимых последствий.

Мировое сообщество не располагает действенными механизмами регулирования своей жизнедеятельности, формирования образа будущего и его воплощения в жизнь. Отсутствуют адекватные механизмы рефлексии, позволяющие осознать свое состояние, вызовы, угрозы, риски, свои потенциалы и ресурсы, разработать действия, обеспечивающие поддержание процессов жизнедеятельности и развития.

В мировое сообществе отсутствуют действенные механизмы обеспечения социальной ответственности перед самим мировым сообществом как целостным образованием, перед населением планеты, перед государствами и разнообразными социальными образованиями, перед Природой.

«Бессубъектность развития» – болезнь мирового сообщества<sup>1</sup>. Бессубъектность развития человечества не позволяет поставить на повестку дня актуальную проблему преодоления ограничений техногенной цивилизации, разработки модели пост-техногенной цивилизации. Эту проблему в основном ставят представители философии и социогуманитарных областей знания. Эта проблема должна быть главной для научной дипломатии в XXI веке.

В центре внимания становления субъектности развития человечества, на наш взгляд, должны быть две первоочередные проблемы:

- разработка на основе методов постнеклассической научной рациональности моделей организации управления и развития человечества, ориентированных на обеспечение его целостности и субъектности;
- формирование культуры международной и национальной безопасности, ориентированной на преодоление бессубъектности развития человечества.

### **2.2.5. Мир как саморазвивающаяся полисубъектная среда в социогуманитарной цивилизации**

Как вариант пост-техногенной цивилизации нами предлагается социогуманитарная цивилизация<sup>2</sup>. При этом термин «цивилизация» будем использовать в смысле ее рассмотрения как некоторого

---

<sup>1</sup> Лепский В.Е. Проблемы становления субъектности человечества и модели развития // Развитие и экономика. 2011, сентябрь. С.95-101.

<sup>2</sup> Лепский В.Е. Гармония культур в саморазвивающихся рефлексивно-активных средах (От техногенной к социогуманитарной цивилизации) / Контуры будущего в контексте мирового культурного развития: XVIII Международные Лихачевские научные чтения, 17-19 мая 2018 г. СПб.: СПбГУП, 2018. С. 431-433.



целостного социального организма, предполагающего определенный тип культуры<sup>1</sup>.

Россия могла бы выступить инициатором разработки модели социогуманитарной цивилизации и механизмов перехода от техногенной к социогуманитарной цивилизации. Она имеет обоснованные аргументы претендовать на эту роль: за счет исторического опыта интеграции народов с разнообразными культурами; философско-методологических разработок, ориентированных на планетарное мировоззрение, на преодоление эгоизма «свой – чужой», что могло бы позволить организовать гармонию субъектов развития, преодолеть рыночный эгоизм, неэквивалентный обмен и принцип максимизации прибыли.

В соответствие с логикой постнеклассической научной рациональности мировое сообщество в пост-техногенной цивилизации может быть представлено как саморазвивающаяся полисубъектная (рефлексивно-активная) среда<sup>2</sup>.

В настоящее время ведутся интенсивные разработки постнеклассической кибернетики саморазвивающихся полисубъектных (рефлексивно-активных) сред. В соответствие с эволюцией кибернетики, соотношенной с эволюцией типов научной рациональности, предлагаемая постнеклассическая кибернетика названа нами – кибернетикой третьего порядка<sup>3</sup>. *Кибернетика третьего порядка* органично вписывается в линию развития кибернетики в целом и интегрирует все предшествующие виды кибернетики через систему онтологий. Кибернетике первого порядка («наблюдаемых систем» – Н. Винер) соответствует классическая научная рациональность. Кибернетике второго порядка («наблюдающих систем» – Фон Ферстер) – неклассическая научная рациональность. Кибернетике третьего порядка («саморазвивающихся рефлексивно-активных сред») – постнеклассическая научная рациональность.

Важно заметить, что опора на постнеклассическую научную рациональность, которая является рамочной парадигмой для классической и неклассической рациональности, позволила интегрировать в кибернетике третьего порядка предшествующие типы кибернетики.

Мы полагаем, что становление постнеклассической кибернетики третьего порядка могло бы заложить научный фундамент для перехода

---

<sup>1</sup> Степин В.С. Цивилизация и культура. СПб.: СПбГУП, 2011. 408 с.

<sup>2</sup> Лепский В.Е. Методологический и философский анализ развития проблематики управления. М.: Когито-Центр, 2019. 340 с.

<sup>3</sup> Lepskiy V. Evolution of cybernetics: philosophical and methodological analysis // *Kybernetes*. 2018. V.47, Iss.2. pp.249–261.

от техногенной к социогуманитарной цивилизации, что способствовало бы разработке и реализации новых подходов к решению проблем обеспечения международной безопасности и развития человечества.

Уточнение сходства и различия философских оснований видения кибернетики будущего на Западе и в России было проведено на мировом конгрессе WOSC2017 в январе 2017 г. (Италия, Рим), на XI международном симпозиуме «Рефлексивные процессы и управление» (Россия, Москва) в октябре 2017 г., на международной конференции IRDO2019 по проблеме социальной ответственности (Словения, Марибор) в июне 2019 г., на международной конференции IFAC TECIS2019 (Болгария, Сазополь) в сентябре 2019 г. Предложенные нами идеи кибернетики третьего порядка привлекли внимание зарубежных специалистов, одним из оснований для развития международного сотрудничества, стало понимание принципиально разных философских оснований на Западе и в России для развития управленческой науки социальными системами. Более того в понимании западных специалистов предложенная модель кибернетики третьего порядка не заостряет внимания на идеологии и разного рода «измов»<sup>1</sup>. Эти соображения явились стимулом для развертывания интенсивного сотрудничества исследователей в данной области, планирования проведения в Москве 16-18 сентября 2020 года конгресса WOSC2020 по системным исследованиям и кибернетике, устремленной в будущее человечества.

Наши установки для разработки кибернетики третьего порядка в первую очередь связаны с ориентацией на создание научного обеспечения социогуманитарной цивилизации. Возможно, кибернетика саморазвивающихся полисубъектных (рефлексивно-активных) сред позволит сделать первые шаги на пути перехода от техногенной к социогуманитарной цивилизации.

В настоящее время имеются основания для утверждения, что кибернетика третьего порядка является конкурентоспособной российской разработкой, позволяющей претендовать на лидерство в проблематике управления и развития социальными системами. Важно заметить, что это особенно важно, поскольку в нашей стране до 1954 года кибернетика была лженаукой, как следствие мы отстали в пятом

---

<sup>1</sup> *Stuart A. Umpleby, Tatiana A. Medvedeva & Vladimir Lepskiy* “Recent Developments in Cybernetics, from Cognition to Social Systems”, *Cybernetics and Systems*, 2019. Volume 50, Issue 4, pp. 367-382. *Stuart A. Umpleby* *Expanding Science and Advancing Reflexive Government: Two Current Projects in Cybernetics Journal on Policy and Complex Systems* • 2018. Volume 4, Number 2, pp. 171-184.

технологическом укладе. Не хотелось бы снова наступать на те же грабли и отстать в создании пост-техногенной цивилизации.

Высокая актуальность кибернетики третьего порядка для современной России является ее ведущая ориентация на создание действенных механизмов *включенности общества в процессы контроля, управления и развития*. Такой подход может обеспечить решение проблемы сложности в управлении страной, обеспечить согласование сложности системы управления и объекта управления. В настоящее время имеет место явное нарушение одного из важнейших принципов кибернетики (принцип Эшби). Этот подход также позволит более полно обеспечить соблюдение важнейшего положения Конституции Российской Федерации о народе как носителе властных полномочий и на современном научном уровне решать междисциплинарную задачу сборки субъектов развития<sup>1</sup>. Расширение возможностей общественного участия в механизмы развития страной позволит обеспечить эффективные процессы стратегического целеполагания в стране, которое в настоящее время в России практически отсутствует. Как следствие страна обслуживает в основном чужие стратегии, при недостаточном внимании к интересам общества<sup>2</sup>.

Одним из необходимых условий включенности общества в процессы управления и развития является формирование двухконтурной структуры организации процессов развития. Первый контур образуется из сложившихся структур и механизмов управления и фактически может трактоваться как «исполнительный механизм» управления мировыми процессами. Второй контур формируется с доминирующим началом общественных структур и механизмов, в котором интегрируются представители администраций, бизнеса и общества. Консолидация двух контуров осуществляется на основе системы онтологий, предусматривающих конвергенцию иерархических, сетевых и средовых механизмов в условиях гибридной реальности (субъектная, цифровая, физическая).

Построение двухконтурного механизма управления и развития мирового сообщества (социальных систем) – это актуальная и сложнейшая проблема. В первую очередь потребуется совершенствование и снятие ангажированности у существующих международных организаций, создание новых организаций и

---

<sup>1</sup> Лепский В.Е. Аналитика сборки субъектов развития. М.: «Когито-Центр», 2016. 130 с.

<sup>2</sup> Авдеева З.К., Барышников П.Ю., Журенков Д.А., Зацаринный А.А., Ильин Н.И., Колин К.К., Лепский В.Е., Малинецкий Г.Г., Райков А.Н., Савельев А.М., Сильвестров С.Н., Славин А.Б., Славин Б.Б. Стратегическое целеполагание в ситуационных центрах развития / Под ред. В.Е. Лепского, А.Н. Райкова. М.: Когито-Центр, 2018. 320 с.

механизмов интеграции в интересах становления субъектности развития человечества.

Миссия России в XXI веке может быть сформулирована как лидерство в обеспечении перехода мирового сообщества от Техногенной к Социогуманитарной цивилизации, в формировании субъектности развития человечества. Переход неизбежен, ибо в противном случае человечество будет нарастающими темпами приближаться к катастрофе.

### **2.2.6. Выводы**

Предложены системные основания для созидания пост-техногенной цивилизации, которые базируются на результатах философско-методологический анализ кризиса техногенной цивилизации. В центр внимания этого пункта монографии положены работы В.С. Степина и Н.Н. Моисеева. Ограничения техногенной цивилизации рассмотрены в контексте кризиса экономической детерминации развития и нарастания угроз в эволюции технологических укладов.

Сформулированы социогуманитарные критерии оценки новаций в техногенной цивилизации и показано, что человечество не готово к адекватному реагированию в рамках этой цивилизации. Обоснован вывод, что бессубъектность развития – диагноз состояния человечества в техногенной цивилизации.

В качестве варианта пост-техногенной цивилизации предлагается модель социогуманитарной цивилизации. Ключевой проблемой для перехода от техногенной к социогуманитарной цивилизации является проблема становления субъектности развития человечества, решение которой предполагает две первоочередные задачи. Во-первых, разработка современных моделей организации управления и развития человечества на основе постнеклассической научной рациональности. Во-вторых, формирование современной культуры международной и национальной безопасности, ориентированной на преодоление бессубъектности развития человечества. Предложено решать эти задачи на основе постнеклассической кибернетики саморазвивающихся полисубъектных (рефлексивно-активных) сред – кибернетике третьего порядка.

## **2.3. СОЦИОГУМАНИТАРНЫЕ ВОПРОСЫ РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В УСЛОВИЯХ ГЛОБАЛЬНОГО КРИЗИСА ЗЕМНОЙ ЦИВИЛИЗАЦИИ**

### **2.3.1. Введение**

Для основательной разработки и определения критериев социальной значимости развития ИИ необходимо более глубокое осмысление самого понятия *социальной значимости* технологических новаций в современных условиях, с учетом, прежде всего, двух фундаментальных факторов: нового этапа развития ИИ, связанного с разработкой общего искусственного интеллекта (ОИИ) и резкого обострения глобального кризиса мировой цивилизации, вызванного ее монополярной структурой, которую стремятся сохранить США со своими западными союзниками.

### **2.3.2. Социально значимые перспективы развития ИИ**

Крупные достижения последних лет в развитии информационных технологий и процессов цифровизации, охватившие практически все основные сферы социальной жизнедеятельности, стали предметом широкого обсуждения в философской и научной литературе, в том числе в плане анализа явлений негативного характера. Опасения общественности вызывают риски и угрозы, обусловленные исключительно быстрым возрастанием сложности систем ИИ, трудностями контроля за их функционированием. Это, в первую очередь, связано с результатами использования машинного обучения, в особенности глубокого обучения нейронных сетей, количество параметров которых исчисляется триллионами. Такой масштаб усложнения ведет к тому, что специалисты называют проблемой «Черного ящика», т.е. недостаточной обоснованности, «объяснимости» получаемых на выходе результатов работы систем ИИ<sup>1</sup>. А это снижает к ним доверие пользователей: ведь именно эти результаты должны служить в качестве инструментов и рекомендаций для решения всевозможных социальных задач (экономических, образовательных, коммуникативных, медицинских и т.д.). Отмечаются к тому же многие существенные издержки внедрения цифровизации в разных сферах

---

<sup>1</sup> Славин Б. Может ли искусственный интеллект быть справедливым // БИТ. 2021. №10 (113). С. 32-35.

социальной жизни. Крупные достижения ИИ порождают столь же крупные проблемы, касающиеся вопросов их социальной значимости.

Но гораздо более серьезные проблемы такого рода возникают сейчас, когда начался качественно новый этап развития ИИ, связанный с разработкой ОИИ, который по своим функциям приближается к способностям естественного интеллекта (ЕИ). Все существующие, традиционные системы ИИ являются «узкими», т.е. функционируют лишь в ограниченном наборе задач и условий, представляют отдельные виды ИИ, такие как системы компьютерного зрения, обработки естественного языка, символьной информации и т.п. Они не способны самообучаться в существенно новых условиях, требуют каждый раз специального программирования, не могут полноценно функционировать без участия человека-оператора

Что же касается ОИИ, то такая система ИИ по своим функциональным возможностям должна быть не «узкой», а *интегральной* (в смысле *общей*), но в то же время способной самостоятельно решать задачи в разных условиях внешней среды<sup>1</sup>. Разработка ОИИ стала мировым трендом, предметом возрастающей конкуренции между крупнейшими корпорациями и государствами-лидерами в области ИИ. Надо подчеркнуть, что ее результаты будут иметь особенно существенное значение (по сравнению с развитием «узкого» ИИ) для судеб мировой цивилизации в условиях дальнейшего нарастания глобального кризиса, имея в виду при этом как их мыслимое положительное, так и отрицательное значение.

Для успешной разработки ОИИ требуются новые теоретические и методологические подходы, отвечающие принципам постнеклассической эпистемологии<sup>2</sup> с учетом опыта философского анализа междисциплинарных проблем и результатов когнитивной науки, включая социогуманитарный контекст. Вместе с этим для моделирования необходимых ОИИ функций требуется исследование и описание специфических когнитивных архитектур естественных процессов мышления. Здесь необходимыми ресурсами для построения моделей ОИИ могут служить результаты психологических и феноменологических исследований ценностно-смысловых,

---

<sup>1</sup> Недавно вышла первая в нашей стране книга, посвященная обзору проблематики ОИИ, анализу состояния ее разработки, основных трудностей и возможных способов их преодоления. См.: Сильный искусственный интеллект. На подступах к сверхразуму / Александр Ведяхин, ред. и др. / М.: Интеллектуальная Литература 2021. 232 с.

<sup>2</sup> Лекторский В.А. Эпистемология классическая и неоклассическая. М., 2001.

динамических структур субъективной реальности и результаты новейших нейронаучных исследований сознания<sup>1</sup>.

Принципиальное значение для осмысления нового этапа развития ИИ имеют основательные социогуманитарные подходы, на основе которых оцениваются проблемы, результаты и перспективы разработки ОИИ. В этом плане первостепенное значение приобретает предложенная В.Е. Лепским концепция саморазвивающихся рефлексивно-активных сред (кибернетики третьего порядка)<sup>2</sup>, и рассмотрение на ее основе вопросов социогуманитарной эргономики, особенно, социогуманитарных критериев оценки инноваций цифровой реальности<sup>3</sup>.

Наряду с акцентом на достижении успехов в разработке ОИИ столь же важно рассматривать те новые широкомасштабные риски и угрозы социогуманитарного характера, которые они могут принести человечеству. Ведь система ИИ такого рода должна стать автономной и самообучающейся, а, следовательно, способной к «самодеятельности»; она сможет легко выходить из-под контроля человека, более того, как часто полагают, сможет генерировать собственные цели, конкурировать с ним и даже подчинить его себе. Подобные сюжеты дискутируются во многих публикациях. Некоторые авторы (как, например, Ник Бостром) говорят о *сингулярном рубеже*, – когда наделенный интеллектом искусственный агент догонит, а затем и превзойдет человека, причем для этого ему не обязательно приписывать обладание сознанием, достаточно одного когнитивного превосходства.

Такой образ мысли немалого числа западных философов и ученых является, однако, весьма «узким», основанным на очевидно несостоятельной абстрактной предпосылке разделенного (не взаимообусловленного) развития человека и ИИ. При этом, как ни странно, игнорируются повсеместные факты и фундаментальные возможности *гибридного интеллекта*, сочетающего в себе свойства ЕИ с ИИ<sup>4</sup>. Уже сейчас у нас такое сочетание проявляется в разной, но

---

<sup>1</sup> Одна из попыток анализа, обобщения и систематизации этих исследований содержится в моих работах (См.: Дубровский Д.И. Задача создания Общего искусственного интеллекта и проблема сознания // Философские науки. 2021. Т. 64. № 1. С. 13–44 ; Дубровский Д.И. Значение нейронаучных исследований сознания для разработки общего искусственного интеллекта (методологические вопросы) // Вопросы философии, 2022., № 2 , с. 83 -93.

<sup>2</sup> Лепский В.Е. Концепция субъектно-ориентированной компьютеризации управленческой деятельности. М.: Институт психологии РАН, 1998. – 204с. См. также: Лепский В.Е. Рефлексивно-активные среды инновационного развития. – М.: Когито-Центр. 2010. – 280 с

<sup>3</sup> Лепский В.Е. Искусственный интеллект в субъектных парадигмах управления. Филос. науки. 2021. 64(1). С. 88–101; Лепский В. Е. Социогуманитарные критерии оценки новаций цифровой реальности // SocioTime / Социальное время. 2018. – № 4(16). С. 16-26.

<sup>4</sup> Аршинов В.И., Гримов О.А., Чеклецов В.В. Киберанимизм: искусство быть живым в

достаточно определенной форме, как результат вовлеченности в процессы *антропотехнологической эволюции*, которая создает, по выражению Бруно Латура, все новые «*гибриды природы и культуры*». Идущая уже несколько десятилетий и набирающая темпы антропотехнологическая эволюция формирует быстро возрастающее множество новых интерфейсов между физическими, биологическими, техническими, социальными объектами и процессами, делокализуя фрактальные границы субъекта и создавая качественно новые диапазоны и формы его *коммуникативной* и практической деятельности, Все это способно привести к существенным, скорее, качественным, преобразованиям нашего интеллекта, творческих способностей, и, что особенно важно подчеркнуть, – к преобразованиям нашей *телесности*, а в силу этого и основного, комплекса наших потребностей, целей и видов деятельности. Но это как раз то, что необходимо для преодоления глобального кризиса и качественной трансформации социума.

Мы уже были свидетелями того, что буквально за несколько десятилетий в связи с появлением ИИ и развитых на его основе технологий, произошел переход от индустриального общества к информационному обществу, качественно изменилась вся система производства, потребления и коммуникаций (достаточно сравнить 1950-й год с 1990-м, не говоря уже о последних десятилетиях).

Подобная перспектива вероятна в случае создания и развития технологий, связанных с ОИИ. Особенности такого процесса важно учитывать в состоянии резкого обострения глобального кризиса нашей цивилизации. Выход из этого состояния предполагает существенное *преобразование мировой социальной самоорганизации*, прежде всего преодоления ее монополярной структуры, породившей фундаментальные противоречия, чреватые третьей мировой войной (об этом еще речь будет далее).

### **2.3.3. Антропологический контекст социальной значимости развития ИИ**

Рассуждая о социогуманитарных аспектах нового этапа развития ИИ в современных условиях, важно расширить рассмотрение этой проблематики в *антропологическом контексте*. Нашу цивилизацию справедливо называют «потребительской», поскольку главный, неуклонный вектор ее развития состоит в том, чтобы еще больше потреблять, чтобы еще больше производить, чтобы еще больше потреблять и т.д. и т.п. Из этого замкнутого круга не видно выхода.



Главным «двигателем» и охранителем такого «прогресса» выступает пресловутый «золотой миллиард», западные страны во главе с США, которые потребляют неизмеримо больше остальных и навязывают свою систему ценностей и свои интересы всему миру. Но чтобы еще больше производить, нужно производить всё новые и новые, более эффективные технологии, а для всего этого нужны новые и новые природные ресурсы, которые ограничены, и чтобы их обрести, необходимо политическое доминирование в мире, нужна власть над теми «нижестоящими» странами, которые ими располагают и, конечно, нужно, чтобы эти страны тоже потребляли больше продукции, производимой для них западным миром. Везде – круговая зависимость экономических, политических и иных интересов субъектов разного вида и ранга, интересов, лишенных, как правило, подлинного гуманистического экзистенциального смысла. Это ведет к нарастанию абсурда во всех сферах и на всех уровнях социальной жизнедеятельности, что представляет собой яркий симптом глобального кризиса нашей цивилизации.

Здесь надо еще раз обратить внимание на то принципиальное обстоятельство, что информационные технологии играют в общественном развитии двоякую роль. С одной стороны, они являются средством эффективного решения неоглядного множества насущных практических задач и способны при определенных условиях служить глобальным целям положительного характера. Но, с другой стороны, будучи интегральными, «сквозными», в том смысле, что являются одним из главных факторов совершенствования и использования большинства современных технологий, они определяют общее состояние и направленность развития производства, потребления, коммуникаций, в конечном итоге, основные, в том числе негативные, особенности нынешнего этапа мировой цивилизации. Тем самым выходит, что они ответственны за неуклонно нарастающий глобальный кризис

Мы часто слышим, что дело не в самих по себе технологиях, а в том, как их использует человек. Во многих конкретных случаях это справедливо. Но возникает вопрос: почему же абсолютное большинство человечества, многие сотни миллионов людей (независимо от национальной, государственной, религиозной принадлежности, от социального статуса, уровня благосостояния, образования и т.д.) охотно использует их в целях еще большего потребления и осуществления личных или институциональных интересов.

Получается, что ответственность лежит на самом человечестве. Это давно и предельно ясно для многих, стало уже неким общим местом, но, тем не менее, требует осмысления с опорой на более глубокое антропологическое основание, именуемое *«природой человека»*. Это

понятие весьма часто используется в философской и научной литературе. Оно обозначает комплекс устойчивых психических и телесных свойств социального индивида, которые обнаруживают инвариантность по отношению ко всем различным эпохам, этносам, социальным устройствам. А это однозначно указывает на их биологическую обусловленность. Взятые по отдельности и в комплексе эти свойства выражают потребности, способности, склонности, коммуникативные особенности, формы поведения и деятельности человека. За обозримый период истории природа человека не изменилась. Так позволяют думать обширные материалы о людях Древнего Египта и Месопотамии, труды древнегреческих и древнеримских историков и философов. Читая «Характеры» Теофраста, поражаешься удивительному сходству: прошло две с половиной тысячи лет, но перед нами – все те же столь знакомые каждому человеческие типы, формы общения и поведения. Поскольку природа человека биологически обусловлена, она не могла измениться за столь мизерный срок, по сравнению с процессом антропогенеза.

Когда мы размышляем о социогуманитарных проблемах, связанных с развитием ИИ, то часто недооцениваем социальную значимость биологических факторов, то обстоятельство, что наши фундаментальные ценности, такие как жизнь, здоровье, семья, дети, любовь и ряд других, имеют биологическую основу. Природа человека, наряду с альтруистическими свойствами, включает ряд явно негативных свойств, играющих весьма важную роль в социальных процессах. Если говорить кратко, то это – неумеренное потребительство, агрессивность к себе подобным и чрезмерное эгоистическое своеволие. Они в разной степени присущи как отдельным людям, так и коллективным и институциональным субъектам, выражают существенные черты массового сознания, которые определяют цели, формы и результаты деятельности массового человека. Но именно эта деятельность в итоге служит неуклонному нарастанию глобального кризиса земной цивилизации. И надо отдавать себе ясный отчет, что этот кризис является по своему существу *антропологическим кризисом*, несущем угрозу гибели земной цивилизации, *антропологической катастрофы*.

#### **2.3.4. Как реагирует на антропологический кризис западная философская антропология?**

Перед современной философской антропологией стоит неотложная задача гораздо более глубокого осмысления, в ряде отношений переосмысления, представлений классической философии о человеке и социуме, во всяком случае – тех ее метафизических конструкций, на

которые опирается *экзистенциальная проблематика*. Усиливается впечатление, что у классической философии нет ответа на судьбоносные мировоззренческие, экзистенциальные вызовы нашего времени. Скорее, это связано во многом с тем, что она опирается на неявную *предпосылку о безграничном существовании социума*. Смерть, исчезновение – удел каждого человека, но не социума. Если раньше возможность гибели цивилизации связывалась лишь с космическими, природными факторами, то теперь становится ясным, что наиболее вероятной причиной этого будут социальные факторы, сам человек. Вот почему этот судьбоносный вопрос – угроза гибели земной цивилизации, *антропологической катастрофы* – должен стать центральным в философской антропологии и в смежных с ней областях философского и научного знания.

Однако, как ни странно, обзор литературы показывает, что подавляющее большинство философов, занимающихся вопросами развития мировой цивилизации и философской антропологией, не слишком озабочены этой экзистенциальной угрозой. Они чаще всего обсуждают метафизические вопросы онтологии человека, причем в классическом ключе, и с этих позиций резко критикуют идеи трансгуманизма, видят именно в нем чуть ли не главную угрозу человечеству. При этом трансгуманистические концепции зачастую изображаются в упрощенном, иногда в окарикатуренном виде, а главное, игнорируется или затемняется их суть, – направленность на поиск путей противодействия глобальному кризису, спасение земной цивилизации от надвигающейся гибели.

Стоит привести в подтверждение пример, касающийся лично одного из авторов монографии (Д. Дубровского). Не так давно известный по этой тематике американский философ Янес-Фернандес (Школа Декстера Саутфилда, Бостон, США) выступил с развернутой критикой концептуальных подходов Д. Дубровского к указанной проблематике. Вот интересное название его статьи: «Давид И. Дубровский и Мераб Мамардашвили: Второе падение Адама и пришествие киберлевиафана»<sup>1</sup>. В этой большой статье (более пяти печатных листов) автор ставит задачу проанализировать «антропологические парадигмы Дубровского и Мамардашвили в их отношении к феноменам кибербытия и киберкультуры» и «разработать критический подход к кибернетической антропологии Дубровского» в отличие от онтологии

---

<sup>1</sup> *Yanes-Fernandez, Inti* . David I. Dubrovsky and Merab Mamardashvili : Adam's Second Fall and the Advent of the Cyber-Leviathan // CREATIVE-COMMONS-BY Forum Philosophicum 24 (2019) . 2, 301–41 ISSN 1426-1898 e-ISSN 2353-7043 DOI:10.35765/forphil.2019.2402.14

Мамардашвили», которая оказывается с ней «принципиально несовместимой» (см. с 301 указанной статьи; в дальнейшем сноски будут обозначаться в скобках номером страницы).

Надо признать, что автор упомянутой статьи внимательно прочел необходимые публикации, весьма подробно и в большинстве случаев адекватно излагает высказанные в них взгляды, четко формулирует свои возражения (все это, конечно, стоило ему немалого труда, о чем свидетельствует несколько десятков (!) цитирований работ критикуемого автора и сносок на них). Такой оппонент заслуживает уважения. Он сразу обозначил себя сторонником антропологической парадигмы Мамардашвили, и с этих позиций высказывал свои критические положения. Здесь нет возможности полемизировать с ним, этому следует посвятить отдельную статью. Остановимся здесь лишь на вопросах, которые были поставлены выше.

В статье Янеса-Фернадеса нет ни слова о глобальном кризисе мировой цивилизации. В одном месте, правда, при рассмотрении критикуемой им концепции упоминается «проблема сохранения земной цивилизации в ситуации приближения сингулярного рубежа». Но в дальнейшем автор излагает концепцию Мамардашвили и свои суждения без малейшей связи с этой проблемой. Он видит достоинство концепции и антропологической парадигмы «грузинского философа» Мамардашвили в интеграции двух основных классических традиций: греко-римской и иудео-христианской, что позволило ему развить свою философскую антропологическую концепцию с ее традиционным «принципом онтореальности», который опирается на понятия языка и истории как «атрибуты действительного бытия человека». По словам Янес-Фернадеса: «В онтологии Мамардашвили бесконечное и конечное бытие раскрываются и даруются друг другу языком в истории как органом человека. В этом смысле я понимаю ее как историко-трансцендентальную онтологию» (с. 335). Таким образом, «онто-реальные трансцендентальные атрибуты Бытия открывают для человечества путь к действительному сближению сущности и существования», создают «возможность реального синтеза между Абсолютом и конечным человеком в органе истории как наиболее подходящем месте обитания человека» (см. с. 334, 336). И в заключение еще одна цитата: «И Мамардашвили, и Дубровский производят метафизику, избегающую разговоров о Боге, но по совершенно разным причинам. Для Дубровского Бог – это ненужный товар... Мамардашвили, со своей стороны, акцентирует внимание на раскрытии Смысла в истории через язык, при условии, что это будет язык, открытый «евангельскому принципу» (с. 336).

Здесь приведены основные положения философско-антропологической концепции Мамардашвили, поддержанной Янес-Фернандесом. Они оба, как видим, целиком отрешены от вопросов антропологического кризиса, от судьбоносных проблем нашего времени. И эта «успокоенность», вытеснение из теоретических размышлений реальной возможности гибели земной цивилизации составляет характерную черту большинства западных философов, которые обсуждают проблемы и перспективы земной цивилизации. В связи с этим считаем важным отметить, что такую же позицию демонстрируют и некоторые российские философы, явно разделяющие принципы прозападной либеральной, глобалистской идеологии. Они принимали активное участие в обсуждавшемся в Институте философии проекте развития Российской цивилизации, заявляя, что она «является недоразвитой», что образцом для нее должна служить западная цивилизация, что столь же недоразвитым оказывается и сознание русского народа, в силу чего он и поддерживает Президента России В.В. Путина, что в России, тем не менее, все равно назревает революция, которая покончит... и т.д. и т.п.<sup>1</sup> Такие явно враждебные нам сентенции всячески «анализируются», «обосновываются», выдаются за некое «глубокомыслие» наших «продвинутых интеллектуалов», которые возвышаются над «примитивными по своему мышлению», «зомбированными патриотами». И удивительно, что все это не получает должного решительного отпора и осуждения.

Сейчас глобальный кризис, развивавшийся уже много лет, приближается к тому рубежу, за которым либо деградация и гибель нашей цивилизации, либо переход ее на новый этап развития. В этих условиях понятие социогуманитарной значимости технологий ИИ существенно изменяет свое содержание в плане системы приоритетов, главным из которых становится жизнь и здоровье людей, выживание человечества. Вместе с этим все же повышается вероятность того, что развитие ОИИ, разных форм гибридного интеллекта смогут, как уже отмечалось выше, качественно обогатить и ускорить антропотехнологическую эволюцию, представляющую собой форму и средство перехода к новому этапу жизни земной цивилизации.

---

<sup>1</sup> Проведен критический анализ подобных взглядов, высказанных в процессе обсуждения в Институте философии проекта развития Российской цивилизации. См.: Д.И. Дубровский. К проблемам цивилизационного развития России: Критический анализ концепции А. В. Рубцова // ЛИЧНОСТЬ. КУЛЬТУРА. ОБЩЕСТВО. 2021. Том XXIII (№№ 109-110). С. 106-125. См. также: Д.И. Дубровский. К вопросу о развитии российской цивилизации. Патриотизм и либерализм // Проблемы цивилизационного развития 2021. Т. 3. № 1. С. 232–240.

### 2.3.5. Выводы

В современных условиях вопросы определения критериев социальной значимости развития ИИ и процессов цифровизации требуют основательного рассмотрения прежде всего самого понятия социальной значимости, с учетом не только особенностей нового этапа развития ИИ, связанного с разработкой ОИИ, но также, в первую очередь, с предельным обострением глобального кризиса мировой цивилизации. В этом отношении указанные критерии должны быть разделены на две группы: 1) относящиеся к масштабам мировой цивилизации в целом и 2) относящиеся к жизненным интересам нашей страны. В первом плане важно рассматривать их в антропологическом контексте с учетом «природы человека», как стойкого комплекса биологически обусловленных свойств человека, определяющих его основные потребности, условия существования и связанные со способами преодоления глобального кризиса и выживания человечества. Во втором плане – это критерии для создания средств эффективного экономического, политического, информационного противостояния, противодействия блоку наших противников, но в то же время сохранения и развития нашей экономики, социального и продовольственного обеспечения населения, решения комплекса насущных задач развития ИИ.

## 2.4. ЦЕННОСТНЫЕ ОРИЕНТИРЫ РАЗРАБОТКИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В США И КИТАЕ

Искусственный интеллект (ИИ) как инструмент преобразования общества, его структуры, функционирования, вызывает ускорение парадигмальных сдвигов в науке и методологии, усложняя современные исследовательские подходы. На современном этапе развития технологии ИИ необходимо рассматривать с позиций постнеклассической научной рациональности<sup>1</sup> как сложную среду, объединяющую технологические, биологические, социальные и этические аспекты своего развития. Проблематика «человек-машина» расширяется на пространства «страна-машина», «общество-технология» – субъектами инновационной деятельности в сфере развития ИИ являются как отдельные индивиды, так и коллективы, страны, нации, что усложняет технологический дискурс культурной оценкой, этической, религиозной.

### 2.4.1. Этические концепции и принципы разработки ИИ

Социогуманитарное измерение научных исследований в области ИИ прежде всего подразумевает включение ценностных ориентаций и этических аспектов не только в процессы разработки технологии, но и в процессы кросс-культурных коммуникаций и взаимодействия стран.

Большая часть важнейших инициатив, связанных с проработкой этических принципов и подходов к разработке и развитию технологий ИИ, были выдвинуты различными международными ассоциациями, объединениями, институтами<sup>2</sup>. Многие из данных документов были исследованы Л. Флориди и Д. Каулзом<sup>3</sup>, и на их основе были выделены

---

<sup>1</sup> *Степин В.С.* Теоретическое знание: Структура, ист. эволюция. М.: Прогресс-Традиция, 2003. – 743 с.

<sup>2</sup> Среди них: «Асиломарские принципы ИИ», разработанные под эгидой Института будущего жизни в 2017 году, «Монреальская декларация ответственного ИИ», разработанная под эгидой Монреальского университета в 2017 году, Общие принципы, предложенные во второй версии книги “Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems («Этически обоснованный дизайн ИИ: взгляд на благополучие человека в автономных и интеллектуальных системах)”, как результата совместных усилий 250 экспертов под патронажем Института инженеров по электротехнике и электронике (IEEE), Этические принципы, предложенные в Докладе об искусственном интеллекте, робототехнике и «автономных» системах, опубликованном Европейской группой по этике в науке и новых технологиях Европейской комиссии в марте 2018 года, Встреча Министров ИКТ и промышленности стран «Большой семерки» (G7) в 2017 году и многие другие.

<sup>3</sup> *Floridi L., Cowls J.* A Unified Framework of Five Principles for AI in Society // Harvard Data Science Review. Issue 1(1). June 2019.

базовые этические принципы, которые наиболее часто упоминались и формулировались международным экспертным сообществом (Таблица 2-1). Данные принципы были дополнены исследователями – был введен принцип «обоснованности» в контексте понятности, объяснимости и подотчетности человеку-пользователю.

**Таблица 2-1**

**Общая матрица ценностных оснований технологий ИИ**

	<b>Принцип ценностной ориентации</b>	<b>Онтологические особенности</b>
1.	<i>«делай благо»</i>	Содействие благополучию индивидуальных и коллективных субъектов, сохранение достоинства и поддержание гармонии в масштабе всего мира.
2.	<i>«не навреди»</i>	Неприкосновенность жизни, безопасность и «осторожность в возможностях»
3.	<i>«уважение автономии субъекта»</i>	Право субъекта принимать решение, право полностью или частично делегировать принятие решения иным субъектам.
4.	<i>«справедливость»</i>	Содействие процветанию, солидарности, предотвращение несправедливости.
5.	<i>«обоснованность»</i>	Обеспечение работы других принципов через понятность и подотчетность.

Рассмотрим с позиций данных «универсальных» принципов (или близких к ним) стратегические документы в области развития ИИ таких стран, как США и Китай, которые рассматривают как преимущества развития данной технологии, так и угрозы, которые она может представлять.

### **2.4.2. Ценностные ориентиры разработки ИИ в США**

Политика США в вопросах этики и кросс-культурных коммуникаций в сфере ИИ неоднозначна. С одной стороны, США активный участник в международных глобальных площадках по вопросам ИИ. Страна активный участник Глобального партнерства по ИИ (GPAI) (в рамках инициативы ОЭСР) представляет собой объединение ведущих экспертов из науки, промышленности, гражданского общества, международных организаций и правительств. Организация предоставляет механизмы для междисциплинарных исследований и коммуникаций для преодоления разрыва между теорией и практикой в области ИИ. На глобальных площадках США отмечает важность «ориентированного на человека ИИ, который стимулирует инновации и рост в цифровой экономике»<sup>1</sup>, а также использование

<sup>1</sup> G7 ICT and Industry Ministerial Meeting. Making The Next Production Revolution Inclusive. Open And Secure. 2009. 24 p.



технологий ИИ в содействии социальной интеграции и культурному и языковому разнообразию.

Междисциплинарность исследований ИИ на государственном уровне рассматривается в том аспекте, что специалисты в области ИИ должны взаимодействовать со специалистами из других областей, таких как психология, экономика, теория управления, этика, лингвистика, математика, философия<sup>1</sup>. На университетском уровне – как межкультурное и междисциплинарное взаимодействие (например, междисциплинарный институт при Стэнфордском университете BUILDING A BRAIN, как продукт взаимодействия Stanford Bio-X и Wu Tsai Neurosciences Institute).

С другой стороны, в США также рассматривают технологии ИИ как приоритетные для усиления национальной экономики, общества. В «Исполнительном приказе» по ИИ<sup>2</sup> обозначена стратегия, в рамках которой обозначен ряд приоритетов страны. Один из них предусматривает содействие созданию международной среды, поддерживающей американские инновации в области ИИ и их ответственное использование. То есть, обозначается верховенство национальных интересов над глобальным общественным.

Такой подход «усиливается» приоритетами Национального стратегического плана НИОКР в области ИИ<sup>3</sup>, в котором фиксируется «мировое лидерство» США в области ИИ, разработка эффективных методов взаимодействия ИИ и человека для «эффективного дополнения и расширения человеческих возможностей». Планируется разработка систем ИИ, способных решать этические, правовые общественные проблемы (с помощью технологических инструментов), особо выделяется безопасность и надежность технологии ИИ, в т.ч. они должны вызывать доверие пользователя-человека.

Наиболее часто встречающимися базовыми принципами (матрицу которых мы позаимствуем у Флориди и Каулза) являются «делай благо», «справедливость» и «не навреди». Проблемой является направленность данных принципов – в стратегических документах США есть только субъекты данной страны, для которых будут соблюдаться данные принципы. Справедливость и блага для американских студентов, исследователей, безопасность для американцев, улучшение жизни

---

<sup>1</sup> Executive Office of the President. Maintaining American Leadership in Artificial Intelligence, E.O. 13859 of Feb 11, 2019. s.l.: Executive Office of the President, Presidential Document. Retrieved at: <https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence>

<sup>2</sup> Там же

<sup>3</sup> National Artificial Intelligence Research and Development Strategic Plan. 2019.

американских домохозяйств и т.д., что является довольно эгоистичным подходом к общецивилизационной проблеме.

### 2.4.3. Ценностные ориентиры разработки ИИ в Китае

В настоящее время Китай обладает самой разветвленной системой законодательных актов и государственных планов, декларирующих приоритеты развития ИИ-технологий, но это многообразие компенсируется достаточно размытым характером этих документов<sup>1</sup>.

В Дорожную карту научно-технического развития Китая до 2050 года<sup>2</sup> включен принцип «опора на внутренние силы и эффективная интеграция глобальных инновационных ресурсов в соответствии с открытостью внешнему миру». Национальная политика формирования инновационной среды отражается в новом проекте закона о научно-техническом прогрессе<sup>3</sup>. Влиятельными являются дискурс марксистской практики, построение «сообщества единой судьбы» и других культурных дискурсов, оказывающих влияние на международные научно-технические и культурные взаимодействия<sup>4</sup>.

Последовательно ведется работа на международных площадках – Китай активно участвует в глобальных дискуссиях ООН по вопросам этики в ИИ, в вопросах об ограничениях для смертоносных автономных систем (САС), участвует в создании гражданских стандартов применения ИИ на площадке Международной организации по стандартизации (ISO). Идеи открытости и глобализации (Open & Global) во многом отвечают позициям страны по вопросам кросс-культурного взаимодействия в сфере ИИ.

Во внутренней образовательной и научной политике Китай транслирует тенденции к трансдисциплинарному подходу. Осознается проблематика ИИ как сложного проекта, требующего взаимодействия различных институтов, культур, практик и дисциплин науки. Необходимость трансдисциплинарности для науки Китая отражается и на высшем уровне – в 2016 году генеральный секретарь КНР Си Цзиньпин заявил, что необходимо ускорить развитие новых научных дисциплин и междисциплинарных исследований, имеющих

---

<sup>1</sup> Савельев А.М., Журенков Д.А. Национальные стратегии развития систем искусственного интеллекта: опыт стран - лидеров и ситуация в России // Научный вестник ОПК России, №3. 2019. С. 75-82

<sup>2</sup> *Yongxiang Lu. Science & Technology in China: A Roadmap to 2050. Strategic General Report of the Chinese Academy of Sciences. Springer Berlin, Heidelberg. 2010. 138 p.*

<sup>3</sup> China mulls law revision to advance innovation in science, technology. 2021. Retrieved at: <https://goo.su/eUvHt>.

<sup>4</sup> *Pan Chung. New Age philosophy and social science discourse innovation (in Chinese). 2021. Retrieved at: <https://goo.su/2YPz0>.*

практическое значение<sup>1</sup>. Включение социальных и философских дисциплин в инновационное и технологическое развитие отражает новую сложность общественного развития и кросс-культурного взаимодействия. В Китае осознают, что необходим системный подход к преодолению противоречий между различными научными дисциплинами для решения практических научных задач. Уже есть ряд значимых исследований по тематике ИИ с включением культурных и ценностных характеристик, этических и эстетических исследований<sup>2</sup>.

Однако, наряду с коммуникационной активностью, научным взаимодействием и междисциплинарными исследованиями, в Китае область ИИ рассматривается и как сфера международной конкуренции<sup>3</sup>. ИИ как стратегическая технология будет определять конкурентоспособность, уровень влияния страны, её безопасность. Планы развития Китая в данной технологической сфере рассматриваются другими странами, например США<sup>4</sup>, как политика конфронтации и «мягкой» агрессии.

Несмотря на приоритет человека в стратегических документах Китая, часто «благо» рассматривается в контексте страны и всего населения Китая, а «уважение к автономии субъекта» вообще отсутствует в привычном нам понимании. Принцип «не навреди» является одним из приоритетов любого плана и стратегии Китая в области ИИ, но он рассматривается с позиций безопасности технологической (открытость кода и пр.). «Справедливость» рассматривается как законность, ответственность и благонамеренность использования технологии.

Стоит отметить, что в разрабатываемых данными странами стратегических документах этические принципы и прописанные ценностные ориентации пока остаются целевыми пожеланиями, без прописанных конкретных механизмов их практического внедрения и соблюдения. Решение проблемы кросс-культурного взаимодействия в контексте технологий ИИ предлагается искать в социогуманитарных подходах, например в изучении рефлексивных коммуникативных

---

<sup>1</sup> *Xi Jinping*. In philosophy and social science work talks on (in Chinese). 2016. Retrieved at: [http://www.xinhuanet.com/politics/2016-05/18/c\\_1118891128.htm](http://www.xinhuanet.com/politics/2016-05/18/c_1118891128.htm)

<sup>2</sup> *Songchun Zhu*. Intelligence needs to be driven by «mind,» achieving a dynamic balance between «mind» and «principle» (in Chinese). 2022. Retrieved at: <https://mp.weixin.qq.com/s/FMraladW0255C-yeU5alfA>

<sup>3</sup> The State Council of the People's Republic of China. The new generation artificial intelligence development plan (in Chinese). 2017. Retrieved at: [http://www.gov.cn/zhengce/content/2017-07/20/content\\_5211996.htm](http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm)

<sup>4</sup> Office of the United States trade representative executive Office of the President. Findings of the investigation into China's acts, policies, and practices related to technology transfer, intellectual property, and innovation under section 301 of the trade Act of 1974. 2018. 215 p.

платформ<sup>1</sup>. Эта модель как инструмент инновационной политики характерна для постнеклассической научной рациональности. Рефлексивная платформа – это позиция индивидуального и группового субъекта, оснащенная языковыми (коммуникативными) средствами для осознания, структурирования и реструктурирования реальности себя и своей деятельности, а также коммуникации с другими субъектами через согласование принимаемых ими реальностей. Включение рефлексивных коммуникативных платформ изменяет характер связей между инновационными субъектами на характерные полисубъектной среде. Это означает прозрачность, возможное взаимодействие с любым субъектом инновационной среды, взаимоизменение, трансдисциплинарность. Процесс рефлексивного взаимодействия субъектов инновационной политики, обеспечит непрерывный диалог между государством, обществом, заказчиком и инициатором исследований, исполнителем и добровольными инвесторами за рамками традиционной двусторонней схемы.

#### **2.4.4. Выводы**

Такие понятия, как «общественное благо», «социальная справедливость», «мультикультурализм», «свобода», «равенство» – во многом абстрактны. В различных культурах они воспринимаются по-разному, существует проблема восприятия, языка.

В условиях нарастающей конвергенции биологических и цифровых субъектов, вызванной практически повсеместным внедрением технологий ИИ, уже мало создать некий набор тех или иных ценностных и морально-этических правил для пользования технологией во благо. Необходимо формировать трансдисциплинарный подход, учитывающий и технологические и социогуманитарные аспекты, определять субъектов разработки и использования ИИ (включая сам ИИ), учитывать биологические, психологические, культурные, этические аспекты.

---

<sup>1</sup> Lepskiy, V., Zhurenkov, D., Saveliev, A. Science Diplomacy and Social Innovations // SocioTime, 2(18). 2019. pp. 54-62.

## 2.5. ПАРАДИГМАЛЬНЫЕ ОСНОВАНИЯ ВНЕДРЕНИЯ ЦИФРОВЫХ ИННОВАЦИЙ В СОВРЕМЕННОЕ ОБРАЗОВАНИЕ

Цифровизация образования – это комплексный процесс внедрения самых разнообразных цифровых технологий в работу существующих образовательных институтов – и процесс этот идет уже много лет. Однако настоящий бум массового интереса к цифровизации данной сферы ожидаемо пришелся на 2020 год<sup>1</sup> в связи с вынужденным переходом многих учебных заведений на удаленные форматы обучения<sup>2</sup>. Как следствие, цифровизация образования в массовом сознании оказалась связана именно с удаленным обучением и отчасти с различными онлайн-курсами, хотя всем этим она никак не ограничивается, включая в себя, в том числе, цифровые инструменты, которые можно применять и при очном формате обучения.

При этом многие работы, рассматривающие последствия ускоренной цифровизации, делают акцент на технических проблемах, с которыми столкнулись многие участники данного процесса<sup>3</sup>, в то время как более глубокие методологические вопросы нередко остаются «за кадром».

Мы, со своей стороны, в данном разделе хотим поставить следующий фундаментальный для современного образования вопрос: какие парадигмы в сфере образования могут с успехом использовать цифровые инновации, а какие к подобным инновациям резистентны. Соответственно, это могло бы помочь прояснить, в случае господства каких образовательных парадигм и связанных с ними педагогических технологий цифровые инновации способны сделать обучение более или менее эффективным – то есть, при каких обстоятельствах подобного рода инновации способны выполнить тот запрос, который формирует образовательная парадигма и воплощает прикладная педагогическая технология.

Мы полагаем, что оценка эффективности цифровизации невозможна без рассмотрения существующих образовательных парадигм и педагогических технологий и того, насколько они в принципе позволяют включать в себя те или иные цифровые инструменты, так как именно образовательная парадигма создает ту систему взаимоотношений, которая составляет образование как институт и способна принимать или

---

<sup>1</sup> Уроки «стресс-теста». Вузы в условиях пандемии и после нее // Аналитический доклад ректоров и рабочих групп университетов, Минобрнауки РФ, 2020. URL: [https://minobrnauki.gov.ru/ru/press-center/card/?id\\_4=2777](https://minobrnauki.gov.ru/ru/press-center/card/?id_4=2777)

<sup>2</sup> Назаров В. Л., Жердев Д. В., Авербух Н. В. Шоковая цифровизация образования: восприятие участников образовательного процесса // Образование и наука. 2021. Т. 23, № 1. С. 156–201. DOI: 10.17853/1994-5639-2021-1-156-201

<sup>3</sup> Там же. С. 168-174.

отторгать в зависимости от своих особенностей конкретные инструменты, в том числе, цифровые.

## 2.5.1. Методология исследования

### 2.5.1.1. Полипарадигмальность современного образования

Если говорить о самой сфере образования, то изнутри нее довольно сложно выделить теоретические конструкты, которые действительно заслуживали бы статуса парадигмы. Ряд авторов<sup>1</sup>, говоря об этой ситуации, называют данный этап развития науки об образовании полипарадигмальным или даже допарадигмальным. В качестве обоснования этой позиции упоминается очень широкий разброс взглядов, которые различные авторы склонны относить к образовательным парадигмам, причем, нередко, исключительно декларативно. Так, например, с точки зрения влияния образования на социум, могут выделять подходы к образованию, связанные с природосообразностью или культуросообразностью, а также ориентированные или на личность, или на общество и его запросы<sup>2</sup>. Все это может быть названо парадигмами, однако, когда речь заходит о том, что является содержанием пунктов этой классификации, оказывается, что наполнить их смыслом тоже возможно достаточно произвольно, что нередко и происходит.

Кроме того, подходы к образованию могут быть классифицированы еще по целому ряду факторов, при этом каждый такой подход может быть достаточно произвольно назван парадигмой<sup>3</sup>.

Все это заставляет для начала вспомнить классическое определение парадигмы, данное Т. Куном: парадигмы – это «признанные всеми научные достижения, которые в течение определенного времени дают научному сообществу модель постановки проблем и их решения»<sup>4</sup>. В

---

<sup>1</sup> Яцык В.З., Чернышенко Ю.К., Пискарева О.В. Парадигмы современного образования // Физическая культура, спорт - наука и практика. 2008. №2. URL: <https://cyberleninka.ru/article/n/paradigmy-sovremennogo-obrazovaniya-1>.

Герасимов Г. И. Парадигмальный плюрализм как проявление кризиса образования // Теория и практика общественного развития, №. 3-4, 2009, С. 7-21.

Володина О. В. Полипарадигмальные обоснования современного образования // Известия Волгоградского государственного педагогического университета, №. 7 (111), 2016, С. 4-11.

<sup>2</sup> Агапова, Н.Г. Парадигмальные ориентации и модели современного образования (системный анализ в контексте философии культуры): монография / Н.Г. Агапова; Ряз. гос. ун-т им. С.А. Есенина. – Рязань, 2008. – 364 с.

<sup>3</sup> Колесникова И.А. Педагогические цивилизации и их парадигмы // Педагогика. 1995. № 6. С. 84–89.

<sup>4</sup> Кун Т. Структура научных революций: Пер. с англ. / Т. Кун; Сост. В.Ю. Кузнецов. — М.: ООО «Издательство АСТ», 2003. С. 17.

сфере образования подобным образом воспринимаемых парадигм до сих пор, очевидно, не хватает.

При этом, бесспорно, в образовании есть ряд признанных и успешно, а главное, массово применяемых методик и педагогических технологий, которые, определенно, связаны с какими-то устойчивыми, пусть и иногда имплицитно существующими взглядами на образование. Вот эти имплицитно присутствующие взгляды мы хотим найти и вербализовать, так как именно они могут претендовать на тот статус, которым обладают парадигмы.

Для выявления этих взглядов мы предлагаем использовать теоретическую концепцию, внешнюю для сферы образования, чтобы рассмотреть последнюю как предмет.

Внешней, метатеоретической по отношению к образованию, может стать концепция типов научной рациональности, предложенная В.С. Степиным<sup>1</sup>. Это, соответственно, классический, неклассический и постнеклассический типы.

#### *2.5.1.2. Цифровые инструменты в образовании*

Если говорить о цифровизации образования с точки зрения используемых технологий, то можно встретить множество классификаций используемых инструментов с точки зрения особенностей самих технологий<sup>2</sup> и гораздо меньше – с точки зрения потребностей образования.

Обращаясь к последним, можно в наиболее общем виде разделить цифровые технологии, применяющиеся в образовании, на 1) те, что могут быть использованы непосредственно в учебных классах и аудиториях; 2) те, что могут применяться для дистанционного обучения; 3) и те, что могут быть использованы в смешанных форматах.

Для первого случая можно упомянуть, к примеру, VR и AR технологии<sup>3</sup>, которые при необходимости могут позволить преподавателю продемонстрировать учащимся ряд обучающих материалов не просто в форме видеозаписи, а в трехмерном формате, позволяющем также осуществлять обучающие операции параллельно с просмотром. Также упоминания заслуживают роботы-помощники и роботы-тьюторы, которые стали появляться относительно недавно, и влияние которых на эффективность обучения в стенах учебных

---

<sup>1</sup> *Степин В.С.* Классика, неклассика, постнеклассика: критерии различения // Постнеклассика: философия, наука, культура. Санкт-Петербург, 2009. С.249 – 295.

<sup>2</sup> *Maud Chassignola, Aleksandr Khoroshavin, Alexandra Klimova and Anna Bilyatdinova.* Artificial Intelligence trends in education: a narrative overview, *Procedia Computer Science* 136 (2018) 16–24.

<sup>3</sup> *Virtual reality & augmented reality (англ.)* – технологии виртуальной и дополненной реальности.

заведений активно изучается одновременно разработчиками и педагогами<sup>1</sup>.

Для второго случая, который как раз чаще всего упоминается в разговорах о цифровизации образования, может применяться множество инструментов<sup>2</sup>, таких, как продукты, представляемые в рамках национальных и коммерческих образовательных платформ для создания массовых открытых онлайн-курсов (далее – MOOC), разнообразные облачные сервисы для хранения и распространения контента, платформы для видеосвязи, иногда предлагающие специальные возможности именно для целей обучения, LMS<sup>3</sup>, используемые школами и вузами, и многое другое. Все это нацелено, во-первых, на создание возможностей для распространения образовательного контента и работы с ним онлайн. Во-вторых, сюда иногда добавляются и возможности для коммуникации преподавателя и учащихся в режиме реального времени с помощью инструментов видеосвязи, интерактивной доски и др. Наконец, также имеют место попытки (более или менее эффективные) контроля знаний учащихся посредством онлайн-инструментов<sup>4</sup>, таких, как тесты, текстовые ответы на вопросы, беседа с преподавателем по видеосвязи и др.

Для третьего случая (*blended learning* (англ.) – смешанных форматов обучения) возможно применение как технологий, указанных в первом случае, так и части тех, что указаны во втором, так как нередко, особенно в последние год-два можно наблюдать случаи частичного перехода учебных заведений, особенно высших, на удаленный формат: например, проведение части занятий одного учебного курса в удаленном, а части – в очном формате. При этом часто применяются LMS, видеосвязь, облачные сервисы.

---

<sup>1</sup> *Smakman, Matthijs, Vogt, Paul, Konijn, Elly A. (2021) Moral considerations on social robots in education: A multi-stakeholder perspective, Computers & Education, Volume 174, DOI: <https://doi.org/10.1016/j.compedu.2021.104317>.*

<sup>2</sup> *Alenezi, M. Deep Dive into Digital Transformation in Higher Education Institutions. Educ. Sci. 2021, 11, 770. <https://doi.org/10.3390/educsci11120770>*

Shanks JD (2020) Toward A Contemplative Technopedagogy Framework: A Discourse Analysis. *Front. Educ.* 5:553212. doi: 10.3389/feduc.2020.553212

Røe Y, Wojniusz S and Bjerke AH (2022) The Digital Transformation of Higher Education Teaching: Four Pedagogical Prescriptions to Move Active Learning Pedagogy Forward. *Front. Educ.* 6:784701. doi: 10.3389/feduc.2021.784701

<sup>3</sup> *Learning management systems (англ.) – системы управления обучением, включающие в себя инструменты, позволяющие размещать обучающий контент, частично или полностью автоматизировать процесс контроля и оценки знаний учащихся и др.*

<sup>4</sup> *Yang, C. C. Y., Chen, I. Y. L., & Ogata, H. (2021). Toward Precision Education: Educational Data Mining and Learning Analytics for Identifying Students' Learning Patterns with Ebook Systems. Educational Technology & Society, 24 (1), 152–163.*



Применение технологий ИИ<sup>1</sup> возможно при всех указанных вариантах обучения – начиная от роботов, которые могут использоваться в классах<sup>2</sup>, и до интеллектуальных помощников-тьюторов<sup>3</sup>, анализирующих скорость прогресса учащегося и в зависимости от этого предлагающих ему задания того или иного уровня сложности. Успешность обучения, как мы предполагаем и далее постараемся это обосновать, зависит не только от технической оснащенности и грамотности, позволяющей применить ту или иную цифровую технологию, но, возможно, в первую очередь от методологического и организационного соответствия цифровых инструментов применяемой парадигмальной модели обучения и связанной с этим педагогической технологии.

### 2.5.2. Классическая парадигма образования

Как уже было сказано, классическая парадигма возникает в образовании и науке практически одновременно. В интересующей нас сейчас образовательной сфере эта парадигма нашла отражение в комплексе подходов к обучению, который также получил название классической или традиционной педагогической технологии<sup>4</sup>. Одним из основоположников ее в сфере образования является Я.А. Коменский<sup>5</sup>. Его «Великая дидактика», по сути, заложила основы массового образования последующей индустриальной эпохи. В этой его работе указаны едва ли не все особенности школьного массового образования и значительной части высшего, которые стали доминировать уже в XX веке. Большинство современных образованных людей имели возможность почувствовать работу этих принципов на себе.

Во-первых, это связь образования с воспитанием, а точнее, с социализацией учащихся. Во-вторых, весьма четкое и понятное наполнение социальных ролей педагога и учащихся. В-третьих, регламентация того, как должен быть организован сам образовательный

---

<sup>1</sup> Goksel, N., & Bozkurt, A. (2019). Artificial Intelligence in Education: Current Insights and Future Perspectives. In S. Sisman-Ugur, & G. Kurubacak (Eds.), Handbook of Research on Learning in the Age of Transhumanism (pp. 224-236). Hershey, PA: IGI Global.

<sup>2</sup> Léna, P. (2021). Robotics in the Classroom: Hopes or Threats? In: von Braun, J., S. Archer, M., Reichberg, G.M., Sánchez Sorondo, M. (eds) Robotics, AI, and Humanity. Springer, Cham. [https://doi.org/10.1007/978-3-030-54173-6\\_9](https://doi.org/10.1007/978-3-030-54173-6_9)

<sup>3</sup> Alkhatlan Ali and Kalita Jugal. (2019) Intelligent Tutoring Systems: A Comprehensive Historical Survey with Recent Developments. International Journal of Computer Applications 181(43):1-20.

<sup>4</sup> Селевко Г.К. Энциклопедия образовательных технологий: В 2 т. Т. 1. М.: НИИ школьных технологий, 2006. 816 с-(Серия «Энциклопедия образовательных технологий».)

<sup>5</sup> Коменский Я. А. Великая дидактика / Ян Амос Коменский – М.: Книга по Требованию, 2012. – 321 с.

процесс. Все эти особенности позволили сделать образование всеобщим, на чем настаивал Коменский. При этом подобный педагогический подход, как мы увидим, не особенно подходит для онлайн-обучения.

Начать стоит с того, что классическая парадигма, воплощенная в соответствующей педагогической технологии, возникшая в Новое время и распространившаяся во многих странах частично в XIX, частично уже в XX веке, была, естественно, рассчитана только на очное взаимодействие. Поэтому, кстати, она вызывала ряд вопросов и сложностей, будучи использована для дистанционных (нецифровых) форматов обучения, которые существовали на протяжении значительной части XX века, например, в виде заочного обучения.

Эта парадигма рассматривала обучение, происходящее именно в стенах учебного заведения, как один из важнейших инструментов того, что значительно позднее было названо термином «социализация». Воспитание нравственных качеств учащихся, а также их умения взаимодействовать между собой и с педагогом, следовать установленному в учебном заведении порядку, выполнять качественно и в срок необходимые задания – все это становилось во многом обязанностью и ответственностью именно образовательных учреждений, а еще более точно – педагогов. Причем все перечисленное мыслилось важным и ценным как само по себе, так и в качестве условия для успешного обучения. Учащийся мог прийти в учебное заведение практически неподготовленным, а выйти оттуда полностью сформированной, обученной и разносторонне развитой личностью. Как уже было сказано, ответственность и обязанности по организации учебного процесса, в данном случае, в значительно большей мере принадлежат педагогу, а не ученику. От ученика требуется аккуратно выполнять указания и соблюдать дисциплину. Те, кто привык только к такой парадигме, могли бы возразить, что иначе и быть не может, и педагог всегда должен нести эту ответственность именно в подобном объеме, но далее мы увидим, что в других парадигмах могут заметно варьироваться как обязанности педагога, так и мера ответственности учащихся за организацию их же собственного обучения.

Из этого вытекает, в частности, то, что учитель/преподаватель в данной парадигме не только является преимущественным источником знаний, но также отвечает за то, насколько хорошо эти знания усвоены, и достаточно ли мотивированы учащиеся. Кроме этого, именно педагог может и должен осуществлять функции контроля как усвоения знаний, так и поведения учащихся в учебном заведении. Для того, чтобы все это выполнять, педагог должен обладать для учащихся практически непререкаемым авторитетом и пользоваться их безусловным уважением и доверием.

Таким образом, сама система обучения неминуемо сначала формирует, а затем предполагает определенные социальные роли – педагога и учащихся – опять же, с четким содержанием, которое должно быть известно как самим участникам учебного процесса, так и остальному сообществу, которое для собственного благополучия должно поддерживать и воспроизводить эту систему. Если же содержание этих социальных ролей меняется, то их взаимодействие может существенно нарушиться. Именно об этом процессе говорят те, кто высказывает недовольство падением авторитета учителей<sup>1</sup>. В этой системе падение авторитета педагога напрямую мешает ему выполнять весь тот, на самом деле, огромный объем функций, что ему определен. Также само возмущение этим уже говорит о том, что речь идет именно о классической системе, в реальности, применяющейся во многих учебных заведениях и по сей день.

С указанной образовательной парадигмой связаны, как мы видели, вытекающие из нее определенные инструменты обучения, которые складываются в так называемую классическую педагогическую технологию. Это всем известная классно-урочная (в вузах – лекционно-семинарская) система, которая получила едва ли не повсеместное распространение – в силу своего удобства, изначальной направленности на массовое обучение и, не в последнюю очередь, из-за наибольшей дешевизны по сравнению со всеми остальными моделями. Такая система (и это было заложено еще Коменским) позволяет силами одного педагога обучать большое количество учащихся – до нескольких десятков или даже сотен одновременно.

Теперь вновь о нашей основной проблеме – цифровизации образования. Как уже было указано, все инструменты этой в целом эффективной системы были созданы для ситуаций очного общения и завязаны на одну ключевую фигуру – педагога. Любое уменьшение возможностей педагога не только по передаче знаний, но и по влиянию на мотивацию учащихся, а также по контролю их знаний и поведения автоматически понижает эффективность обучения. Она падает прямо пропорционально ослаблению роли педагога. Поэтому те дистанционные форматы (в том числе, цифровые), которые эту роль уменьшали, в такой системе по определению не могли быть эффективны.

Однако помимо упомянутых выше вопросов организации обучения необходимо учитывать особенности участников учебного процесса. Здесь мы оставим в стороне характеристики педагогов и обратим

---

<sup>1</sup> Маанди В.И. Школа: начинаем обратный отсчёт // Народное образование, № 7, 2010, С. 45-49.

внимание на учащихся, а среди их многочисленных особенностей выделим ту, которая определяет черты образования как ступенчатого процесса – и это, конечно, возраст учащихся.

Очевидно, что обучение как подача информации для детей младшего школьного возраста отличается от обучения подростков в последних классах школы и, тем более, студентов в вузах. Важный момент здесь – влияние обучения на социализацию учащихся. Чем это влияние выше, тем важнее для учащегося очное общение как с педагогом, так и, нередко, с группой других учащихся, так как такой комплексный процесс, как социализация, предполагает усвоение не только информации, но и определенных паттернов коммуникации, поведения, интеллектуальной культуры. Это возможно только в ситуации личного общения.

Мы можем предположить, что по описанной выше причине классическая педагогическая технология обучения будет существовать еще достаточно долго, и удаленный формат с применением цифровых средств не будет вводиться на постоянной основе. Для этой технологии в современном социуме просто нет альтернативы, которая с такой же эффективностью массово обеспечивала бы одновременно обучение и напрямую связанную с ним социализацию.

При этом вовсе не обязательно, что такая модель исключает цифровизацию вообще. В ее рамках цифровые инструменты вполне могут использоваться и быть эффективными – но при условии применения их в стенах учебного заведения, на очных занятиях, где присутствуют и учащиеся, и педагог, способный направлять и контролировать процесс обучения.

### **2.5.3. Неклассическая парадигма образования**

Следующая парадигма, которую мы планируем рассмотреть, это неклассическая парадигма в образовании. Она, как уже упоминалось, характеризуется субъект-субъектным типом взаимодействия. При этом учащийся рассматривается уже не как только лишь реципиент педагогических усилий, но как практически равноправный участник образовательного процесса, постоянно, а не эпизодически, дающий педагогу обратную связь. Мотивация учащегося, в таком случае, зависит от него самого не в меньшей степени, чем от педагога. Уровень ответственности учащегося, таким образом, тоже повышается в сравнении с классической парадигмой. Сам учащийся уже движется к тому, чтобы определять интересующие его направления обучения. Такое обучение нередко называют гуманистическим или личностно-ориентированным.

В рамках неклассической парадигмы нельзя выделить какого-то одного магистрального подхода, как в классической парадигме. Гуманистические подходы к обучению достаточно многочисленны, и каждый имеет свои сильные стороны<sup>1</sup>. Не имея своей целью представить их всесторонний обзор, здесь мы ограничимся лишь несколькими широко известными примерами.

Всплеск интереса к неклассическим гуманистическим подходам к обучению пришелся на начало – первую половину XX века, но основания для них были заложены значительно раньше. Возможно, одним из первых авторов здесь был Джон Локк. В своих трудах «Мысли о воспитании» и «Мысли о том, что читать и изучать джентльмену» он как раз высказывает те положения, которые в дальнейшем легли в основу гуманистических направлений педагогики: *необходимость учитывать личностные особенности учащегося и воспитуемого, развивать его самостоятельность и ответственность, воспитывать полезные для него в будущем привычки, включая бытовые, приучать его к тому, что позднее назовут здоровым образом жизни*<sup>2</sup>. Все это, действительно, будучи применено в практике обучения и связанного с ним воспитания, должно было способствовать развитию гармоничной, самостоятельной, разносторонне развитой личности.

При этом Локк указывал на ту особенность, которая препятствовала широкому распространению его концепции обучения, а также, впоследствии, многих других гуманистических педагогических технологий, хотя сам философ явно проблемой это не считал. Неминуемая элитарность (как противоположность массовости) данных подходов происходила как из заявленного стремления к разностороннему развитию, так и из необходимости постоянной обратной связи от учащихся. Локк прямо указывал, что его интересует, в первую очередь, «выяснение надлежащих методов воспитания молодого джентльмена»<sup>3</sup>. Данная образовательная модель неминуемо должна была включать очень большие затраты на обучение каждого ученика, что делало невозможным ее массовое внедрение.

Попытки создания других гуманистических технологий обучения предпринимались и в XIX веке – здесь можно вспомнить, в частности, работы Песталоцци и Ушинского, писавших о развивающем обучении. Их наработки в некоторой небольшой части вошли в массовое образование, но полностью использовать эти безусловно полезные идеи

---

<sup>1</sup> Селевко Г.К. Энциклопедия образовательных технологий: В 2 т. Т. 1. М.: НИИ школьных технологий, 2006. 816 с-(Серия «Энциклопедия образовательных технологий».)

<sup>2</sup> Локк Дж. Мысли о воспитании / Локк Дж. Сочинения: В 3 т. Т.3. М.: Мысль, 1988. 668 с.

<sup>3</sup> Локк Дж. Мысли о воспитании / Локк Дж. Сочинения: В 3 т. Т.3. М.: Мысль, 1988. С. 414.

не получалось по все той же причине – они требовали едва ли не индивидуального подхода к каждому ученику.

Тем не менее, новые гуманистические педагогические технологии, нередко носящие экспериментальный характер, продолжают возникать и по сей день, но при этом не становятся массовыми и зачастую уже не ставят перед собой подобной цели.

Если говорить о цифровизации образования применительно к таким педагогическим технологиям, то многое, видимо, будет выглядеть здесь сходно с ситуацией с классическим подходом, но некоторые важные различия также возможны. Уровень ответственности педагога и учащегося за мотивацию к обучению, контроль выполнения заданий и др. может варьироваться как в зависимости от конкретных подходов, так и от возраста и уровня подготовки учащихся. Если вновь взять параметр возраста как один из ведущих, то здесь смешанные форматы обучения могут быть возможны для более старших и подготовленных учащихся, уже завершающих или завершивших первичную социализацию и развивших те навыки и личностные качества, которые предусмотрены конкретной педагогической технологией. Если же первичная социализация учащихся еще далека от завершения, то применение цифровых инструментов в обучении для них было бы возможно при условии параллельного личного общения с педагогом и между учениками, когда речь идет о группе.

#### **2.5.4. Постнеклассическая парадигма образования**

Если говорить о третьей парадигме, включающей в себя субъект-метасубъектные отношения, то в образовании она пока что не представлена какими-то устоявшимися педагогическими технологиями. И вместе с тем, мы полагаем, что именно такой подход может быть наиболее продуктивен, когда речь идет о цифровизации образования, особенно об удаленных онлайн-форматах.

Для того, чтобы обосновать этот тезис, необходимо сначала прояснить, что именно в сфере образования может пониматься под субъект-метасубъектными отношениями, каковы их характеристики – тогда можно будет сделать вывод сначала о возможных здесь педагогических технологиях, а затем об их комплементарности цифровым форматам обучения.

Если мы спросим, кто здесь является субъектом, ответ будет не всегда однозначен. Конечно, субъектом, как в предыдущих моделях, может быть педагог (учитель, преподаватель). В таком случае метасубъектом для него будут выступать все те отдельные учащиеся и их группы, которым может быть адресован его курс, и количество,

интересы и подготовка которых могут быть педагогу заранее не известны. В некоторых ситуациях он может общаться со всеми или частью из них, например, по видеосвязи, в других случаях такого общения может и вовсе не быть.

Кроме того, субъектом может быть учащийся, который самостоятельно выбирает направления собственного обучения, понимая, что оно не прекращается по выходу его из учебного заведения, и может продолжаться неограниченно долго. Более того, объем и содержание полученных знаний и навыков может также быть очень вариабелен, и такой субъект-учащийся понимает, что не всегда в состоянии прогнозировать, что именно он будет изучать в дальнейшем, и у кого будет всему этому учиться. В таком случае для него преподаватели и образовательные организации будут своего рода метасубъектом.

Взаимоотношения субъекта и метасубъекта в сфере образования могут строиться, таким образом, как и в субъект-субъектных отношениях, через взаимную коммуникацию и учет целей и интересов другой стороны. Но эта вторая сторона – метасубъект – может быть постоянно изменчивой, что значительно усложняет коммуникацию. Субъект здесь не может просто «отвечать» на требования другого субъекта, т.к. заранее может о них и не знать или лишь предполагать их с некоторой долей вероятности. Таким образом, для субъекта становится необходимым развитие тех качеств, которые позволят ему успешно коммуницировать (обучать или обучаться) с неограниченным и заранее не предсказуемым числом других индивидуальных и коллективных субъектов. Среди таких качеств можно выделить, в самую первую очередь, повышение ответственности субъекта, который теперь сам способен и даже вынужден определять предпочтительные форматы учебной коммуникации, ее объем и направленность вне зависимости от принимаемой им социальной роли.

То есть, преподаватель в данной парадигме не знает заранее (или даже вообще может не узнать), кого ему предстоит обучать, сколько будет учащихся, с какими интересами и ожиданиями они обращаются к его учебному курсу<sup>1</sup>. В доцифровую эпоху, вероятно, подобное могло иметь место в ситуации открытых лекториев и всегда требовало от лектора высочайшего уровня профессионализма.

Если говорить об учащемся, то он, опять же, не всегда способен заранее спрогнозировать, какие знания и навыки ему потребуются в постоянно меняющейся среде, и куда он за ними обратится. Он получает

---

<sup>1</sup> *Abduvakhidov, A. M., Mannapova, E. T. & Akhmetshin, E. M. (2021). Digital Development of Education and Universities: Global Challenges of the Digital Economy. International Journal of Instruction, 14(1), 743-760. <https://doi.org/10.29333/iji.2021.14145a>*

возможность, например, комбинировать очный или смешанный формат обучения в вузе<sup>1</sup> с онлайн-изучением практически любого желаемого количества курсов от других образовательных (и иных) организаций, сознательно выстраивая собственное образование. От автодидактики как самообразования предыдущих эпох современная ситуация отличается разительно – в первую очередь тем, что предполагает не просто знакомство с литературой, но полноценное прохождение учебного курса.

Итак, если попробовать обратиться к вопросу о том, какая именно педагогическая технология могла бы соответствовать этой возникающей образовательной парадигме, то здесь можно исходить из данных, имеющих место для предыдущих парадигм. Если в классической парадигме мы имеем преимущественно однонаправленное взаимодействие, где наибольшая доля ответственности, обязанностей и авторитета принадлежит педагогу, в неклассической парадигме – взаимно-направленную коммуникацию, где доля ответственности и обязанностей ученика возрастает, но все равно остается меньшей, чем у педагога, который определяет основные цели и средства обучения и воспитания, то в постнеклассической парадигме это, опять же, взаимно-направленное взаимодействие, но ответственность здесь распределяется в относительно равных долях, потому что каждый из участников коммуникации может рассматриваться как субъект (центральный в схеме) для коммуникации по отношению к сформированному из различных субъектов метасубъекту.

Если в первом случае мы имеем четкое расписание, строго определенные учебные планы и количество предметов, классно-урочную (или лекционно-семинарскую) систему обучения, во втором случае все может быть достаточно вариабельно в зависимости от взглядов педагога на то, как лучше развить гармоничную личность, то в третьем случае сам учащийся может определять как набор предметов, так и время, которое готов отвести на их изучение, так как последнее может быть более или менее углубленным.

Педагоги оказываются такими же субъектами образовательного процесса, которые в этой системе могут быть не привязаны к какому-то конкретному расписанию, и не всегда связаны неким не ими созданным учебным планом. Расписание и учебные планы, созданные для педагогов кем-то, кроме них – скорее, наследие именно классического подхода,

---

<sup>1</sup> Семенова Т. В., Вилкова К. А. Типы интеграции массовых открытых онлайн-курсов в учебный процесс университетов // Университетское управление: практика и анализ Том 21, № 6, 2017, с. 114-126.



который мы здесь намеренно не оцениваем, так как каждый из подходов имеет свою область эффективного применения.

Тем не менее, из сказанного уже очевидно, что именно постнеклассический подход к обучению в наибольшей мере соответствует удаленным форматам цифрового образования, в особенности там, где речь идет об обучении взрослых людей, уже прошедших первичную социализацию, не нуждающихся во внешней мотивации для работы над заданиями, не стремящихся к формально высокой оценке знаний в отрыве от их получения и не зависящих от мнения и даже наличия референтной группы других учащихся. Вероятно даже, что только для таких учащихся удаленное цифровое обучение может быть эффективным.

### **2.5.5. Выводы**

Итак, резюмируя все сказанное, мы можем констатировать, что успехи или неудачи в применении тех или иных цифровых инструментов в образовании напрямую связаны не только с техническим оснащением учебных заведений и цифровой грамотностью участников учебного процесса, но не в меньшей степени с тем, в рамках какой образовательной парадигмы указанные инструменты применяются.

Так, классическая парадигма образования подразумевает коммуникацию, преимущественно направленную от преподавателя к ученику. Ответственность за все аспекты образовательного процесса тоже лежит преимущественно на преподавателе. Обучение неразрывно связано с первичной социализацией учащихся и даже прямо ее подразумевает. Такого рода обучение, вероятно, в силу вышеуказанной особенности, до сих пор является предпочтительным вариантом для массового общедоступного обязательного среднего образования, возможно, также для среднего специального и отчасти для высшего, по крайней мере, младших курсов вузов. Этот подход может быть успешно реализован только при регулярном очном взаимодействии в стенах учебного заведения. Касаясь цифровизации, здесь могут применяться те цифровые технологии, которые можно использовать на занятиях в очном и, отчасти, смешанном форматах.

Неклассическая парадигма предполагает, что коммуникация взаимно направлена, а обратная связь от учеников чрезвычайно важна. При этом ученик постепенно становится все более ответственным за процесс и результаты обучения наряду с педагогом. Такого рода обучение также предполагает одновременную с ним социализацию учащихся, поэтому может применяться для учеников того же возраста и уровня подготовки, что и классическая парадигма. В связи с

необходимостью социализации очный формат обучения также здесь более предпочтителен, однако постепенно может дополняться смешанными форматами при условии, что педагог может получать постоянную обратную связь от учащихся, а их мотивация уже была должным образом сформирована до этого. Что касается цифровых инструментов, то здесь также предпочтительнее те, что могут применяться непосредственно в классе/аудитории, однако на последних ступенях обучения для более старших учащихся могут частично дополняться теми, что подходят для смешанных форматов обучения.

Постнеклассическая парадигма основывается на коммуникации каждого из участников учебного процесса как субъекта с потенциально неограниченным количеством других участников, которые представляют собой своеобразный метасубъект. При этом именно субъект оказывается в очень значительной мере ответственен за само наличие коммуникации и ее параметры. Поскольку субъектом в такой системе может выступать как преподаватель, так и учащийся, подобный подход подразумевает, что все участники учебного процесса прошли социализацию, то есть лучше всего подходит для обучения высоко мотивированных взрослых, а не детей или подростков. В данном случае могут применяться, вероятно, все существующие цифровые инструменты образования, соответствующие очным, смешанным или удаленным форматам – в зависимости от конкретных задач, возникающих при освоении учебного материала. В целом, мы полагаем, что цифровые инновации в сфере образования в наибольшей степени отвечают содержанию именно постнеклассической парадигмы. Вполне вероятно, что именно при помощи подобных инновационных инструментов как раз и будут окончательно сформированы педагогические технологии для этой лишь совсем недавно появившейся парадигмы образования.

## **2.6. ОБРАЗОВАНИЕ БУДУЩЕГО В КОНТЕКСТЕ РАЗВИТИЯ НАУЧНОЙ РАЦИОНАЛЬНОСТИ**

Новые технологические, социальные и экономические возможности современного общества (цифровизация общественных процессов, автоматизация производств, рассмотрение вопроса о безусловном базовом доходе в развитых государствах) при адекватном подходе могут стимулировать потребности человека в свободе, счастье, творческой активности, что связано с повышением внимания к социогуманитарным аспектам образования. В первую очередь это должно быть связано с разработкой соответствующих критериев оценки образовательных инноваций с использованием инструментов современной техносферы (виртуализация образовательного процесса, новые методы и навыки работы с информацией, структурирование знаний, проблематика различий поколений и их опыта взаимодействия с новыми технологиями и т.д.).

Формирование образовательного пространства как среды, создающей новые знания, новые культурные формы, этические установки, ценности требует не только рациональности, осознанности, рефлексивности, эмпатии (как от педагогов, так и от учеников), но и определенных философско-методологических подходов к осмыслению и практике современного образовательного процесса.

Наращение темпов социальной динамики и становление экономики знаний обосновывает важность субъекта, неотделимости знаний, культуры и ценностей от субъекта. Образование будущего, ориентированное на человека, предполагает динамичные коммуникационные площадки, саморазвивающуюся образовательную среду, пространства доверия, взаимоуважения, культуры. Исходя из данной предпосылки, постулируется гипотеза о социогуманитарном подходе в качестве базового подхода для образования будущего, технологии которого предстоит сформулировать, разработать, органично связать с научно-техническим прогрессом, развитием техносферы, виртуальных реальностей как части физической реальности.

### **2.6.1. Современное представление о научной рациональности как философско-методологическая основа образования будущего**

Рассмотрим процесс образования с позиций субъекта образования (как педагога, так и обучающегося), развития культуры организации знания, соответствующего учебного процесса<sup>1</sup> и проблематики скрытого

---

<sup>1</sup> *Rubtsov V. Learning in children: organization and development of cooperative actions. M. J. Hall, Trans. Nova Science. 1991.*

знания, которое неотрывно связано с субъектом знания. Заявленный контекст исследования образования неотрывно связан со сменой научных парадигм, развитием философии науки и научной рациональности<sup>1</sup> – от парадигмы классической научно рациональности к неклассической, а затем к постнеклассической научной рациональности.

Представление о знании в контексте классической научной парадигмы формировалось в рамках деятельностного подхода и с позиции отношения «субъект – объект». Знание определялось истинным без связи с инструментами получения знания и субъектами, породившими знание. Ограниченность данного подхода была осознана ближе к концу 20 века, однако часть образовательных практик, характерных для данной научной рациональности имеет место и в настоящее время. Основной подход – функциональный, т.е. важно передать необходимые знания ученику, проконтролировать запоминание через определенную повторяемость образовательных «сигналов», ученик в такой системе является объектом обучения.

Представление о знаниях в контексте неклассической научной рациональности в значительной степени формировалось под влиянием парадигмы «субъект – субъект» и субъектно-деятельностного подхода. Осуществлялся переход к философскому конструктивизму – субъект (в данном случае ученик) активно осваивает, творит, конструирует окружающий мир. Знания оказываются неразрывными с субъектами, их порождающими<sup>2</sup>. В основание познавательной деятельности закладываются феномены личных мотивов, мировоззрения, переживаний, субъективных реальностей, а сетевая связь частных «субъектных» картин мира образует общую «неклассическую» картину мира. Такого рода представления о знаниях позволяли формулировать соответствующие требования к образованию: усиление роли непосредственного общения преподавателей с учениками; рассмотрение коммуникаций как процессов создания нового знания и культуры, понимания знания на новом рациональном и чувственном уровнях; «прикосновения» к неявным (скрытым) знаниям; приоритет самоорганизации учеников, их сетевых взаимодействий; усиление роли гуманитарных предметов.

---

<sup>1</sup> *Степин В.С.* Теоретическое знание: Структура, ист. эволюция. М.: Прогресс-Традиция, 2003. 743 с.

<sup>2</sup> *Turchin V. and Joslyn C.* The Cybernetic Manifesto. *Kybernetes*. 1990. 01.19, Nos.2 and 3. P. 63-64, 52-55.

**Таблица 2-2**

**Сравнительная таблица подходов к образовательной деятельности  
в контексте парадигм научной рациональности**

<b>Параметры сравнения</b>	<b>Классическая научная рациональность</b>	<b>Неклассическая научная рациональность</b>	<b>Постнеклассическая научная рациональность</b>
<i>Цель образовательной деятельности</i>	Достижение образовательных целей, успеваемость	Обучение обучению, саморазвитие	Творческая активность, саморазвитие личности и развитие общества
<i>Функции образования</i>	Контроль – исполнение	Организационно-стимулирующая, самоорганизационная	Социогуманитарная, коммуникационная, саморазвивающаяся
<i>Адаптивность к изменениям</i>	Закрытые системы, ретроспективность	Открытые системы, гомеостатичность	Открытые рефлексивно-активные среды, устремленность в будущее
<i>Система коммуникаций</i>	Вертикальные каналы коммуникации (сверху-вниз)	Вертикальные и горизонтальные связи (сети)	Средовые коммуникации (общее коммуникационное целевое пространство)
<i>Роль педагога</i>	Выполнение функциональных обязанностей, постановка задач, без привлечения учащихся	Взаимодействие и обсуждение образовательного процесса с учащимися, расширение коммуникативного пространства	Формирование образовательного процесса с учетом целей обеспечения жизнедеятельности, безопасности и развития общества
<i>Роль ученика</i>	Исполнительская	Самоуправление	Коммуникации и гармонизация личных и общественных интересов и благ
<i>Доминирующие методы и подходы</i>	Операциональный, функциональный подходы, «ориентация на знание»  Повторяемость, внешние стимуляции, отрицательная обратная связь	Деятельностный, смысловой подходы, «ориентация на понимание»  Динамические средства предоставления информации, внутренняя стимуляция, положительная обратная связь	Рефлексивно-активный, стратегический подходы, «ориентация на смыслы»  Коммуникационная и рефлексивная активности, стратегические цели, этическая среда
<i>Базовые парадигмы взаимодействия</i>	Субъект-объект	Субъект-субъект	Субъект-полисубъектная среда

Постнеклассический тип научной рациональности учитывает соотнесённость получаемых знаний об объекте не только с особенностью средств и операций деятельности, но и с особенностями субъектов. Ключевой становится парадигма «субъект –

саморазвивающаяся полисубъектная среда» («субъект – метасубъект»)<sup>1</sup>, а базовым подходом – субъектно-ориентированный подход как развитие субъектно-деятельностного подхода с увеличением внимания к субъектам и их окружающей среде<sup>2</sup>. Усиливается ориентация на творческую активность и саморазвитие личности в связи с развитием общества в целом. Коммуникационное пространство процесса обучения выходит за рамки сетевого подхода и становится средой – каналы обмена информацией, культурой, этикой носят как вербальный, так и невербальный характер, безинформационный, виртуальный. Формирование образовательного процесса основывается на целях, сформулированных в процессе коммуникации с учащимися, используются мотивационные и другие «мягкие» инструменты вовлечения учеников в образовательный процесс. А сами коммуникации строятся на гармонизации личных и общественных интересов и благ. Систематизация подходов и параметров научных рациональностей в образовании представлена в Таблице 2-2.

Ориентация на человека как множества дискурсов и воплощений (реальных и виртуальных, рациональных и психических, явных и неявных), важность безопасности субъекта-человека в современном образовательном процессе позволит говорить о важных социогуманитарных тенденциях, как о позитивном сценарии общественной эволюции.

### **2.6.2. Этическая проблематика образования будущего**

В контексте классической научной рациональности этические регуляторы находились вне рамок деятельности и на них не фокусируется внимание субъектов деятельности. Базовыми ориентирами выступают цели, а доминирующую этику можно представить как «этику целей» вплоть до этики «цель оправдывает средства». Этика целей органично связана с доминированием деятельностной активности, что затрудняет коммуникации и сотрудничество субъектов – устанавливаются преимущественно «вертикальные связи» в социальных структурах. Этика целей порождает стремление к конкуренции, конфликтам, агрессии и потребительству.

В контексте этического рассмотрения специфики неклассической научной рациональности можно высказать соображение, что этические

---

<sup>1</sup> *Лепский В.Е.* На пути к управлению сферой образования как саморазвивающейся средой // Социология образования. 2014, №10. С. 4-24.

<sup>2</sup> *Лепский В.Е.* Прорывная социогуманитарная технология опережающего образования // SocioTime / Социальное время, № 1(13), 2018. С. 21–32.

регуляторы ориентированы на специфику субъект-субъектных отношений. Можно предположить, что активность субъектов основывается на этике «цели – средства». Такого рода этика ориентирована не только на ценностные ориентации деятельностной активности, но и должна учитывать отношения к другим субъектам, последствия для них совершаемых действий, потенциальные и реальные формы коммуникативной активности. Данную этику можно также сформулировать как «коммуникативную». Ориентиры педагога, методологические и образовательные, являются и его ценностными ориентирами, влияющими на коммуникации с учениками, и наоборот. Такого рода этика способствует коммуникациям субъектов, установлению сетевых горизонтальных связей в социальных структурах, что находит место в педагогических технологиях.

Доминирующая этика постнеклассической научной рациональности – это этика свободы с позиций ее самоограничения для полноценного взаимного уважительного взаимодействия с другими субъектами и учетом интересов общества. Задействуются мягкие мотивационные механизмы, запускающие процессы творчества, формирования ценностей, интересов, стратегических целей. Для понимания доминирующей этики постнеклассической научной рациональности важно введение понятия стратегический субъект<sup>1</sup>. Это субъект, включенный в какой-либо метасубъект (семья, группа, организация, страна и др.), идентифицирующего себя с этим метасубъектом и регулирующего свою активность (деятельностную, коммуникативную, рефлексивную) с учетом ее влияния на метасубъекта. Этика в контексте постнеклассической научной рациональности не является универсальной или целостной – она как мозаика, складывается из различных этик, входящих в нее субъектов. Это этика семьи, государственные архетипы, религиозные этики и т.д. Важно отметить, что этика стратегических субъектов является основополагающей для совершенствования механизмов образования.

Исходя из подхода, что три научные рациональности являются этапами эволюции науки, они не являются альтернативными, а дополняют друг друга и предполагают конвергентное использование, можно применить данную логику и к трем соответствующим этикам. Обобщенные выводы по этическим системам по этапам развития научной рациональности представлены в Таблице 2-3.

---

<sup>1</sup> Lepskiy V. Evolution of Cybernetics: Philosophical and Methodological Analysis // Kybernetes. 2018. Vol. 47. No. 2. P. 249–261.

**Таблица 2-3**

Доминирующие этические системы в трёх типах научной рациональности

<b>Классическая научная рациональность</b>	<b>Неклассическая научная рациональность</b>	<b>Постнеклассическая научная рациональность</b>
Этика «целей»	Этика «цели-средства»	Этика «стратегических субъектов»
Утилитарные ценности, эгоистическая этика	Этика субъектных отношений, коммуникационная этика	Этика стратегических субъектов саморазвивающихся сред (включает коммуникационную этику, этику целей)
Личностные интересы превыше всего	Личное саморазвитие превыше всего	Принцип «запрета эгоизма» - при реализации личных интересов не нарушать интересы общественной группы, к которой принадлежишь
Конкуренция, противоборство	Состязательность	Кооперация

Этические подходы в образовательной среде являются важным аспектом формирования новых технологий образования. Они опираются на культуру, религии, ценности и смыслы субъектов, входящих в образовательную среду, поэтому выработка общих этических подходов в образовании, исследования инструментов работы с этической средой представляются актуальными вопросами в контексте образования будущего.

### **2.6.3. Специфика влияния цифровых технологий и ИИ на образовательную среду**

На данный момент информационные технологии интегрированы почти во все аспекты образовательного процесса, поскольку преподавание и обучение не отделимы от информационных взаимодействий. Обучение все в большем объеме строится на цифровых платформах: интерактивность урока завязывается на взаимодействие с информационной средой и Интернетом, подготовка к занятиям, хранение данных, учебники, сами уроки носят все более «цифровой» характер. Государства целенаправленно разрабатывают и реализуют



программы по информатизации общества, по внедрению информационных технологий в систему образования, по переформатированию образования на обеспечение кадрами сферы ИТ<sup>1</sup>. Школы создают свои цифровые пространства и экосистемы, обогащаются цифровые образовательные ресурсы, широко внедряется онлайн-обучение, дистанционные форматы обучения, цифровая трансформация управления образованием.

Искусственный интеллект уже используется как инструмент в цифровых обучающих платформах<sup>2</sup> в качестве репетитора для учеников. Польза от такого рода использования ИИ может перекрываться существенными рисками – может закрепиться тенденция на стандартизацию обучения и тестирования, что приведет к унификации личностей учеников, снизит их адаптивность, усилит и без того высокую академическую конкуренцию<sup>3</sup>. В переходное время изменения образовательной системы под влиянием информационных технологий и ИИ видится важным соблюдать некоторые границы и разделения живого и искусственного преподавания и обучения – рутинные задачи и структурирование информации для учеников можно передать ИИ, а формирование уникальных человеческих навыков, воспитательных элементов, творческой и коммуникационной активности оставить за людьми-учителями. Учитывая важность способностей и интересов каждого субъекта образовательного процесса, а также влияния самих учеников на формирование образовательной, творческой и коммуникативной среды, важно обеспечивать интеграцию образования, цифровых технологий и искусственного интеллекта посредством «мягких» социогуманитарных технологий.

Создание рефлексивно-активной среды<sup>4</sup> как пространства взаимодействия, коммуникации и социализации всех включенных в образовательный процесс субъектов, позволит контролировать и гармонизировать интерактивность технических информационных

---

<sup>1</sup> Например, национальный проект «Образование», федеральный закон от 29.12.2012 N 273-ФЗ (ред. от 14.07.2022) «Об образовании в Российской Федерации» в России, 13-я программа развития национальной экономики и социальной сферы Китайской Народной Республики в Китае, планы Комиссии Leading Education by Advancing Digital (LEAD) в США.

<sup>2</sup> Например, Squirrel AI, китайская технологическая компания, специализирующаяся на интеллектуальном адаптивном обучении. Электронный ресурс. <https://squirrelai.com/>

<sup>3</sup> Данные риски оцениваются на государственном уровне – так, например, Государственный совет КНР обнародовал в 2019 году свод руководящих принципов, призывающих уделять меньше внимания экзаменам и больше внимания физическому, культурному и политическому образованию. Электронный ресурс. [http://www.gov.cn/zhengce/2019-07/08/content\\_5407361.htm](http://www.gov.cn/zhengce/2019-07/08/content_5407361.htm) (Кит.)

<sup>4</sup> *Лепский В.Е.* Рефлексивно-активные среды инновационного развития // М.: Изд-во «Когито-Центр», 2010. – 255 с.

средств и сред с социальными, этическими, психическими средами. «Субъект-метасубъект» и «метасубъект-субъект» как подходы организации данных взаимодействий обеспечат и трансдисциплинарность образования будущего, их интеграции, их сложность и открытость как принципов развития системы.

Взаимодействия «человек-машина», «субъект-метасубъект» по отношению к живым и искусственным акторам, «ученик-учитель» потребуют новой этики образовательного процесса, новых требований к безопасности личных данных, к формированию контуров безопасности умной образовательной среды.

#### **2.6.4. Выводы**

Усложнение социальных процессов, бурное развитие цифровых технологий и ИИ предъявляют все более высокие методологические требования к организации и развитию образовательных процессов. Требуются специалисты с новыми знаниями и умениями, основанными на культуре трансдисциплинарности. Социогуманитарные технологии должны не только лежать в основе организации процесса обучения как учеников, так и преподавателей, но и быть ориентированы на тесную связь этих процессов с развитием общества. Множественность идентичностей, которые приобретает человек в реальности и виртуальности может привести к утрате целостности его как субъекта, что требует разработки и использования принципиально новых социогуманитарных технологий обеспечения образовательной сферы. В этом контексте в центр внимания должны быть помещены также цифровые технологии и искусственный интеллект.

## 2.7. ВЫВОДЫ

1. Происходит свертывание идеи техногенной цивилизации, которая характеризуется проявлением в ее базовой ценности научно-технического прогресса и ценности науки как основы управления социальными процессами. Становится все более очевидным, что при этом фактически игнорируются социальные ценности и этические аспекты регулирования жизнедеятельности. Ищутся альтернативы техногенной цивилизации, одной из которых могла бы стать модель цивилизации, ориентированной на сохранение многоцивилизационного мира при обеспечении его гармоничного развития в планетарном масштабе.

2. Исходные философско-методологические основания для поиска новой модели цивилизации сформулированы В.С. Степиным, который утверждал, что речь должна идти о придании научно-техническому прогрессу существенно большего гуманистического измерения. При этом новый облик науки должен пониматься в контексте постнекласической научной рациональности.

3. Этот облик, помещая в центр внимания субъектов познания и развития, предполагает анализ их ценностно-целевых структур в неразрывной связи с культурой. Развитие же науки рассматривается во взаимодействии с динамикой социальной среды, что заставляет искать пути трансформации техногенной цивилизации в новые формы. При этом возникает вопрос о ценностях, задающих ориентиры перехода к новому типу цивилизации, и которые извне не придут, а должны формироваться в недрах техногенной культуры. Необходимо выявить единый процесс эволюции человечества и биосферы.

4. В рамках сложившегося тренда построения техногенной цивилизации проявляется экономический детерминизм развития, который порождает нарастание угроз для человечества при переходе к очередным технологическим укладам. Технологические инновации становятся неадекватными потенциальным последствиям для человечества. Угрозой становится отстраненность технологий от этического осмысления цивилизационных событий. Требуется поиск адекватных форм жизнедеятельности, под которые должны быть созданы соответствующие им экономические механизмы.

5. Предлагается упреждающий переход к VII социогуманитарному технологическому укладу. Россия должна стать лидером этого уклада, чтобы не рассчитывать на безнадежные шансы

догнать лидеров VI технологического уклада. Философско-методологические основания этого нового уклада базируются на современных представлениях постнеклассического подхода к философии науки. Для этого авторами настоящей монографии разработаны концептуальные основы и модели саморазвивающихся рефлексивно-активных сред, в которых создаются условия для гармоничного развития субъектов, сборки субъектов развития, конвергенции представительской и прямой демократии и др.

6. Выделены основные виды критериев оценки уровня культуры междисциплинарных исследований процессов организации различных видов человеческой деятельности: продуктивность, безопасность, удовлетворенность и развитие субъектов и самих видов деятельности. На их основе выделяются основные направления реагирования на инновации: эффективное реагирование человечества на позитивные аспекты технологических новаций; контролирующее реагирование человечества на потенциальные угрозы от внедрения технологий; справедливое реагирование человечества на возможности инноваций; развивающее реагирование человечества на технологические новации.

7. Отсутствие эффективных механизмов преодоления глобальных угроз цивилизационному развитию приводит к нецелесообразному расходованию ресурсов, проявлению нестабильности в экономических и политических процессах, возникновению непреодолимых последствий с высоким уровнем ущерба. Для умаления этих факторов развития формируются установки в сфере кибернетики третьего порядка, которые связаны с ориентацией на создание научного обеспечения социогуманитарной цивилизации.

8. В качестве модели посттехногенной цивилизации предлагается социогуманитарная цивилизация как целостный социальный организм. Задается вектор возможности России стать инициатором разработки такой модели цивилизации, а также переходных механизмов к новой цивилизации. Эта возможность подтверждается рядом весомых факторов, включая наличие соответствующего исторического опыта интеграции народов с разнообразными культурами; философско-методологических разработок, ориентированных на планетарное мировоззрение и др.

9. Требуется включенность общества в процессы управления и развития, необходимым условием чего является формирование двухконтурной структуры организации процессов развития. Первый контур исполнительный, он образуется из традиционно сложившихся

структур и механизмов управления. Второй контур формируется с приоритетом создания общественных структур и механизмов, в котором интегрируются представители администраций, бизнеса и общества.

10. Видимую, хотя и не определяющую, роль в развитии нового типа цивилизации начинает играть внедрение цифровых технологий, включая ИИ. Показано, что для разработки и определения критериев социальной значимости развития ИИ необходимо более глубокое осмысление самого понятия социальной значимости технологических новаций.

11. Для разработки общего ИИ требуются новые парадигмальные, теоретические и методологические подходы, отвечающие принципам постнеклассической эпистемологии с учетом опыта философского анализа междисциплинарных проблем и результатов когнитивной науки в социогуманитарном антропологическом контексте. На это работает концепция саморазвивающихся рефлексивно-активных сред (кибернетики третьего порядка).

12. Глобальный кризис сейчас приближается к бифуркационной точке, за которой либо гибель цивилизации, либо переход ее на новый этап развития. В этих условиях понятие социогуманитарной значимости ИИ существенно изменяет свое содержание, главным приоритетом становится жизнь и здоровье людей, выживание человечества. Вместе с этим растет вероятность того, что развитие ОИИ, разных форм гибридного интеллекта смогут качественно обогатить и ускорить антропотехнологическую эволюцию.

13. Технологии ИИ необходимо рассматривать с позиций постнеклассической научной рациональности как сложную среду, объединяющую технологические, биологические, социальные и этические аспекты своего развития. Это подразумевает включение ценностных ориентаций и этических аспектов как в процессы разработки и внедрения технологии, так и в процессы кросс-культурных коммуникаций между странами. При этом этические принципы постоянно дополняются, например, был введен принцип «обоснованности» в контексте понятности, объяснимости и подотчетности человеку. Строится матрица ценностных оснований технологий ИИ. В разных странах ценностные ориентиры различные, например, в Китае делается опора на внутренние силы и эффективную интеграцию глобальных инновационных ресурсов.

14. Становится все более очевидным, что создание неких наборов ценностных и морально-этических правил для использования технологией во благо недостаточно для получения этого благо. Необходим трансдисциплинарный подход, учитывающий как технологические, так и социогуманитарные аспекты разработки цифровых инновация, включая учет биологических, психологических, культурных, этических и иных аспектов.

15. В контексте нового типа цивилизационного развития встает фундаментальный для сферы образования вопрос: какие парадигмы могут с успехом использовать цифровые инновации, а какие к подобным инновациям неадекватны? Требуется оценка эффективности внедрения цифровых технологий в контексте существующих образовательных парадигм и педагогических технологий, а также того, насколько они в принципе позволяют включать в себя те или иные цифровые технологии.

16. Сейчас можно встретить множество классификаций, используемых в сфере системы образования инструментов с точки зрения особенностей самих технологий, однако, эти технологии явно недостаточно анализируются с точки зрения потребностей образования. Можно разделить цифровые технологии для образования на те, что могут быть использованы; те, что могут применяться для дистанционного обучения; те, что могут быть использованы в смешанных форматах. Возможно, есть образовательные процессы, где цифровые технологии применять пока преждевременно.

17. Постнеклассический тип научной рациональности учитывает соотнесённость получаемых знаний не только с особенностью средств и операций, но и с особенностями субъектов. Ключевой становится парадигма «субъект – саморазвивающаяся полисубъектная среда», а базовым подходом – субъектно-ориентированный подход как развитие субъектно-деятельностного подхода. Вместе с тем множественность идентичностей, которые приобретает человек в виртуальной реальности может привести к утрате целостности его как субъекта, что требует разработки принципиально новых социогуманитарных технологий.

### **3. СОЦИОГУМАНИТАРНЫЙ АНАЛИЗ ЧЕЛОВЕКО-ОРИЕНТИРОВАННЫХ ПРЕДСТАВЛЕНИЙ ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ (ЭТИКА ИИ, ДОВЕРИТЕЛЬНЫЙ ИИ, ОБЪЯСНИМЫЙ ИИ)**

---

#### **3.1. ВВЕДЕНИЕ**

Социогуманитарный анализ человеко-ориентированных представлений об ИИ, как уже отмечено в предыдущих главах, во многом остается за рамками техногенных предпочтений к выбору оценок значимости и ценности систем ИИ. Во многом это диктуется близлежащими прагматическими соображениями исследователей и производителей таких систем, погруженных в рыночную реальность. Есть и объективные ограничения, определяемые отсутствием адекватного математического аппарата и методологии, позволяющих сделать репрезентативное представление сугубо неформализуемых аспектов построения и функционирования систем ИИ. К таким аспектам относятся: этика, доверительность, объяснимость, эмоциональность, когнитивность и пр. Как следствие, развитие систем ИИ, несмотря на возрастание вычислительных мощностей, количества слоев в искусственной нейронной сети, совершенствование алгоритмов машинного обучения и пр., остается в рамках «узкой» и «слабой» парадигмы построения ИИ. Попытки вырваться из имеющихся классических теоретических и инженерных шор в виде декларации построения «общего» или «сильного» ИИ пока успеха заметного не приносят.

Для преодоления подобной тотальной «пробуксовки» развития ИИ в настоящей главе предлагается обратиться к использованию субъектно-ориентированного подхода, важность которого нарастает по мере движения от деятельностного к субъектно-деятельностному и, наконец, к субъектно-ориентированному подходу. Этот подход формируется в рамках постнеклассической кибернетики саморазвивающихся рефлексивно-активных полисубъектных сред.

В главе отражены особенности построения методологии цифровой трансформации на примере изменения концепции цифрового двойника, погружаемой в такие постнеклассические среды. В этой среде цифровой двойник трансформируется уже с учетом мировоззренческих оснований, соответствующей системы принципов и системы онтологий, что создает основу для сборки в единое целое субъектов, входящих в указанную среду.

В постнеклассической среде неизменным атрибутом функционирования систем ИИ является постоянный рост доверия к работе таких систем, что, в частности приводит к актуализации темы построения объяснимого ИИ с ориентацией на конкретного человека с его интенцией к познанию и предубеждениям, активному мышлению, эмоциональным аффектам и пр.

Морально-этические проблемы разработки систем ИИ в этой главе также не остались без внимания, что заставило авторов оценить, насколько произвольно или вредоносно может повести себя такие системы, а также тем, что сам человек может причинить вред искусственным сущностям. Все это поднимает вопросы морального уровня, что предопределяет потребность закладывать в разработку систем ИИ морально-этические принципы и ограничения.

Анализируется применение гуманитарного измерения к технологиям ИИ, при этом особого внимания заслужила тема обострения вопроса поиска пути преодоления национальных эгоизмов, недоверия и страхов. Предлагается способ построения частных моделей на основе классификатора видов деятельности и определения того их вида, который будет предназначен для выполнения системой ИИ в некоторой типовой ситуации.



## 3.2. СУБЪЕКТНО-ОРИЕНТИРОВАННЫЙ ПОДХОД К ПРЕДСТАВЛЕНИЯМ ЭЛЕМЕНТОВ И СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА (ЦИФРОВОЙ ДВОЙНИК, ЦИФРОВОЙ СУБЪЕКТ И ЦИФРОВОЙ МЕТАСУБЪЕКТ)

### 3.2.1. Введение

Проблема субъектности приобретает все большее значение в развитии кибернетики и ИИ. Субъекты оказываются в центре внимания обеспечения жизнедеятельности и развития социальных систем, а также совершенствования механизмов их управления. Интерес ученых и практиков к субъектности возрастает в связи с бурным развитием ИИ, с созданием искусственных псевдосубъектов, наделенных рефлексивными способностями. С необходимостью создать этичный, доверенный, объяснимый ИИ, не выходящий из-под контроля человека.

В XXI веке основой философско-методологического анализа совместной эволюции кибернетики и ИИ, определения роли и места субъектного подхода стала концепция трех типов научной рациональности (классическая, неклассическая, постнеклассическая) и особенно парадигма постнеклассической рациональности<sup>1</sup>. Такой подход определяет основу для систематизации этапов развития кибернетики и связанных с ними этапов развития ИИ<sup>2</sup>. Он является системно-генетическим, что соответствует представлениям Г. Симондона<sup>3</sup> о роли и месте технических объектов в жизни и развитии социальных систем и их конкретных субъектов.

В этом контексте особое место занимает концепция цифрового двойника, как виртуальной интерактивной копии реального физического объекта или процесса, которая соответствует кибернетике первого порядка. В кибернетике второго и третьего порядков цифровой двойник трансформируется с учетом мировоззренческих оснований, соответствующей системы принципов и системы онтологий в цифрового субъекта и цифрового метасубъекта.

---

<sup>1</sup> Степин В.С. Теоретическое знание. М.: Прогресс-Традиция, 2003.

<sup>2</sup> Лепский В.Е. (2021) Искусственный интеллект в субъектных парадигмах управления // Философские науки. 2021. Т. 64. № 1. С. 88–101.

<sup>3</sup> Simondon, G. (2017) On the Mode of Existence of Technical Objects. Translated by C. Malaspina and J. Rogrove. Minneapolis: Univocal Publishing.

### 3.2.2. Субъектность в эволюции научной рациональности и кибернетики

Становление субъектно-ориентированного подхода тесно связана с развитием представлений о научной рациональности: классика, неклассика, постнеклассики. Актуальность и научно-практическое значение субъектно-ориентированного подхода нарастало по мере движения от деятельностного к субъектно-деятельностному и, наконец, к субъектно-ориентированному подходу. Важно отметить, что эволюция этих подходов удовлетворяла принципу соответствия Бора<sup>1</sup>, каждый последующий подход включал в себя как частный предыдущий подход.

В *деятельностном подходе (классическая научная рациональность)* базовая парадигма «субъект – объект». Свобода субъекта существенно ограничена в рамках заданных целей и норм, регулирующих деятельность<sup>2</sup>.

В *субъектно-деятельностном подходе (неклассическая научная рациональность)* базовая парадигма «субъект – субъект»<sup>3</sup>. Исследователь и объект становятся равноправными партнерами.

*Субъектно-ориентированный подход (постнеклассическая научная рациональность)*. В постнеклассической научной рациональности в центре внимания оказываются субъекты, погруженные в социальные среды и культуру<sup>4</sup>. Ключевой в рамках постнеклассической научной рациональности становится *парадигма «субъект – саморазвивающаяся полисубъектная среда» («субъект – метасубъект»)*, которой соответствует субъектно-ориентированный подхода<sup>5</sup>. Этот подход включает в себя как частные парадигмы деятельностный и субъектно-деятельностный подходы, что соответствует логике отношений трех типов научной рациональности. В настоящее время субъектно-ориентированный подход используется в различных областях научного знания и практики: психология, педагогика, экономика, управление, политология и др.

Эволюция кибернетики тесно связана с развитием типов научной рациональности<sup>6</sup>. Такой подход позволит осуществить системный и междисциплинарный анализ проблемы субъектности.

---

<sup>1</sup> Niels Bohr Collected Works. (1976). Vol. 3: *The Correspondence Principle (1918–1923)*, J. R. Nielsen (ed.). Amsterdam: North-Holland Publishing.

<sup>2</sup> *Лекторский В.А.* Субъект в истории философии: проблемы и достижения // *Методология и история психологии*. 2010. Том 5. Выпуск 1. С.5–18.

<sup>3</sup> *Рубинштейн С.Л.* Человек и мир. М.: Наука, 1997.

<sup>4</sup> *Степин В.С.* Теоретическое знание. М.: Прогресс-Традиция, 2003.

<sup>5</sup> *Лепский В.Е.* Междисциплинарный анализ становления субъектно-ориентированного подхода // *Социальное время*. 2015. № 2. С. 18–32.

<sup>6</sup> *Lepskiy V.E.* Evolution of Cybernetics: Philosophical and Methodological Analysis. *Kybernetes*. 2018. Vol. 47, no, 2, pp. 249–261.

Классической научной рациональности соответствует кибернетика первого порядка (N. Wiener). Этот тип кибернетики не позволит преодолеть ограничения модели техногенной цивилизации.

Неклассической научной рациональности соответствует кибернетика второго порядка (H. Foerster). В кибернетике второго порядка принципиально меняется представление об рефлексивных и этических аспектах отношений субъектов при управлении социальными системами. Однако не снимается главная проблема техногенной цивилизации – игнорирование социальных ценностей, учет целостного видения социальных систем.

Постнеклассическая научная рациональность принципиально отличается от предшествующих типов рациональности в том, что в ней включены в рассмотрение социальные ценности. Постнеклассической научной рациональности соответствует кибернетика третьего порядка, саморазвивающихся полисубъектных (рефлексивно-активных) сред<sup>1</sup>. В центре внимания оказывается свобода субъектов, установление равноправно-партнерских отношений как с другим человеком, так и с природными процессами<sup>2</sup>. Интеграция деятельностного и субъектно-деятельностного подходов осуществляется на основе субъектно-ориентированного подхода<sup>3</sup>. Проявляется изменение представлений о субъектности и в концепции распределенного наблюдателя в саморазвивающейся полисубъектной среде, одним из которых оказывается субъект управления. Изменения в представлениях о субъектности проявляется и в изменениях представлений о рефлексивных процессах и этических аспектах.

Постнеклассическая кибернетика саморазвивающихся полисубъектных сред опирается на идеи ноосферы (В.И. Вернадский); представления об обществе как социальной системе (Никлас Луман); идеи деятельностного и субъектно-деятельностного подходов (А.Н. Леонтьев, Л.С. Выготский, С.Л. Рубинштейн и др.); концептуальные модели рефлексивных процессов (В.А. Лефевр); работы отечественных методологов (Г.П. Щедровицкий); идеи управлению сложностью в развитии социальных систем (В.И. Аршинов)<sup>4</sup>; обобщенный опыт исследований по социальной кибернетике (Stuart Umpleby); опыта

---

<sup>1</sup> Лепский В.Е. Методологический и философский анализ развития проблематики управления. М.: Когито-Центр, 2019. 340 с.

<sup>2</sup> Лекторский В.А. Субъект в истории философии: проблемы и достижения // Методология и история психологии. 2010. Том 5. Выпуск 1. С.5–18.

<sup>3</sup> Журавлев А.Л., Лепский В.Е. Проблема субъекта в инженерной психологии и эргономике // Психологический журнал. 2018. Том. 39. №4. С. 7-16.

<sup>4</sup> Аршинов В.И. Рефлексивно-активные среды инновационного развития в контексте синергетики сложности / Междисциплинарные проблемы средового подхода к инновационному развитию / Под ред. В.Е. Лепского. М.: Когито-Центр, 2011. С.52–73.

разработки организационных автоматизированных систем управления страной (В.Е. Лепский) и др.

Модель саморазвивающейся рефлексивно-активной среды представляется нами через многоуровневую структуру мировоззренческого, концептуального, технологического уровней и уровня реализации<sup>1</sup>. Важно отметить, что технологический уровень состоит из двух подуровней концептуально-технологического и инструментально-технологического, что создает предпосылки для конвергенции через технологии парадигм и онтологий социальных систем и ИИ.

### **3.2.3. Цифровой двойник, цифровой субъект и цифровой метасубъект**

#### *3.2.3.1. Цифровой двойник*

В контексте развития классической кибернетики сформировались базовые парадигмы ИИ: морфологическая, нейрокибернетическая, логическая, имитационная. Концепция цифрового двойника соответствует кибернетике первого порядка.

Майкл Гривз предложил термин «цифровой двойник»<sup>2</sup>. Цифровой двойник – это виртуальная интерактивная копия реального физического объекта или процесса, помогающая управлять им. В отличие от информационной модели объекта цифровой двойник не ограничивается сбором данных, полученных в процессе разработки и изготовления изделия, а продолжает сбор и анализ информации на протяжении всего жизненного цикла реального объекта. Концепция цифрового двойника становится все более актуальной для системного проектирования на основе моделей<sup>3</sup>.

Цифровой двойник имеет богатую предысторию. Человечество использовало модели виртуальных близнецов задолго до эпохи цифровых трансформаций. Близнецы объектов управления появились сначала в сознании людей, на бумаге и других материальных носителях, а затем и в цифровой реальности, с помощью Интернета вещей и ИИ.

Важно отметить, что этот физический объект и его цифровой двойник могут быть очень сложными, но пассивными по отношению к субъекту управления. Это наблюдаемый объект, и цифровой двойник

---

<sup>1</sup> Лепский В.Е. Методологический и философский анализ развития проблематики управления. М.: Когито-Центр, 2019. 340 с.

<sup>2</sup> Grieves, M. (2014) *Digital Twin: Manufacturing Excellence through Virtual Factory Replication*; A White Paper; LLC: Melbourne, FL, USA.

<sup>3</sup> Madni, A.M., Madni, C.C., Lucero, S.D. (2019). Leveraging Digital Twin Technology in Model-Based Systems Engineering. *Systems*. 7 (1): 7.

помогает наблюдать за ним в разных режимах и с разных сторон по командам наблюдателя. Такое понимание цифрового двойника соответствует концепции кибернетики первого порядка (Рис. 3-1).

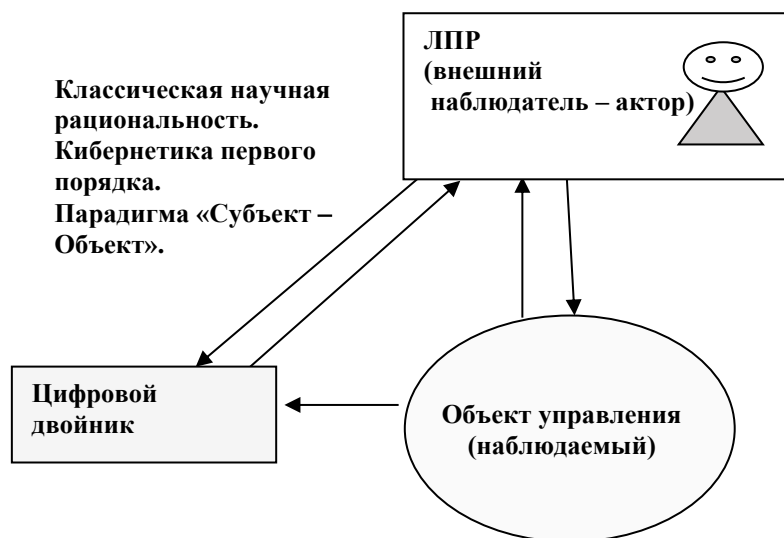


Рис. 3-1. Цифровой двойник в управлении (кибернетика первого порядка)

В табл. 3-1 показаны связи между функциональным назначением цифрового двойника и рефлексивностью ЛПР в контексте кибернетики первого порядка.

**Таблица 3-1**

Функциональное назначение цифрового двойника и рефлексивность ЛПР

<b>Цифровой двойник: функциональное назначение</b>	<b>Рефлексивность управленца (лицо принимающее решение)</b>
Получение разнообразной информации об объекте управления	Контрольная рефлексия состояния объекта управления
Формирование целостного образа объекта управления, через организацию режима работы «Восприятие – Действие»	Рефлексия формирования целостного образа объекта управления
Получение прогнозной информации об объекте управления	Прогнозная рефлексия
Удаленное управление объектом	Рефлексия виртуального объекта
Создание на базе цифрового двойника автономных систем управления	Рефлексия автономной системы управления
Анализ и совершенствование компетентности управленца	Саморефлексия управленца

Процесс развития цифровых двойников продолжится и получит качественно новый функционал, уже проявившийся в контексте кибернетики второго порядка. Важно отметить, что роль

рефлексивности в создании цифровых двойников и организации процессов взаимодействия с ними лиц, принимающих решения, будет возрастать.

### 3.2.3.2. Цифровой двойник и цифровой субъект

Кибернетике второго порядка соответствуют парадигмы ИИ: слабый и Общий ИИ. Кибернетика второго порядка расширяет функции цифрового двойника. Лицо, принимающее решение, может делегировать часть функций автономного управления цифровому двойнику. На основе технологий машинного обучения цифровой двойник становится самообучающейся автономной системой управления под контролем лица, принимающего решения. Цифровой двойник трансформируется в цифровой субъект (Рис. 3-2). ЛПР актуализирует коммуникативную рефлексивность для установления взаимодействия с новым цифровым субъектом и управления им.



Рис. 3-2. Цифровой двойник трансформируется в цифрового субъекта.

Понятие цифрового субъекта имеет разные философские основания и трактовки<sup>1</sup>. В данной статье используется понятие цифрового субъекта как псевдосубъекта, обладающего инвариантными свойствами естественных субъектов. Мы используем системно-генетический подход (классическая, неклассическая, постнеклассическая научная рациональность) к формированию представлений о цифровом субъекте как техническом объекте на основе ИИ<sup>2</sup>. Основными инвариантными

<sup>1</sup> Goriunova, O. (2019a). Digital subjects: an introduction. *Subjectivity*, 12, pp. 1–11. Goriunova, O. (2019b). The Digital Subject: People as Data as Persons. *Theory, Culture and Society*, vol. 36, No 6, pp.125-145. Wark, S. (2019) The subject of circulation: on the digital subject's technical individuations. *Subjectivity*. 12, pp. 65–81.

<sup>2</sup> Simondon, G. (2017) *On the Mode of Existence of Technical Objects*. Translated by C. Malaspina and J. Rogrove. Minneapolis: Univocal Publishing.

свойствами субъектов являются целеустремленность, рефлексивность, коммуникабельность, социальность и способность к развитию<sup>1</sup>.

Рассмотрим пример взаимодействия ЛПР с цифровым субъектом. Выделим некоторые функциональные назначения цифровых субъектов и соответствующие им особенности рефлексивности лиц, принимающих решения (табл. 3-2).

**Таблица 3-2**

Функциональное назначение цифрового субъекта и рефлексивность ЛПР

<b>Цифровой субъект: функциональное назначение</b>	<b>Рефлексивность управленца (лицо принимающее решение)</b>
Формирование комплекса субъектных (рефлексивных) позиций, адекватных потребностям ЛПР. Формирование цифровой полисубъектности.	Рефлексия синхронизации субъектных (рефлексивных) позиций с цифровой полисубъектностью.
Оперативное выявление субъектной (рефлексивной) позиции, актуализируемой ЛПР, и инициация диалога с адекватной позицией цифровой полисубъектности.	Коммуникативная рефлексия взаимодействия с цифровым субъектом.
Формирование целостного образа объекта управления через диалог с цифровым субъектом (взаимодействие с цифровым двойником).	Рефлексия формирования целостного образа объекта управления.
Автономное управление функциями, делегированными лицом, принимающим решение.	Рефлексия качества работы цифрового субъекта.
Поддержка рефлексивных процессов ЛПР в общении с другими субъектами.	Коммуникативная рефлексия с другими субъектами.
Формирование интегрированной модели ЛПР на основе анализа гибридной среды. Формирование позиции цифрового субъекта для обеспечения саморефлексии ЛПР.	Саморефлексия лица, принимающего решение.

В табл. 3-2 приведены сведения об отдельных функциях цифровых субъектов и типах рефлексивности естественных субъектов. Эта информация может быть значительно расширена и детализирована.

В контексте кибернетики второго порядка, соответствующей неклассической научной рациональности, цифровой субъект функционирует в сети естественных и цифровых субъектов. Он может быть привязан к конкретному природному или цифровому субъекту или субъектам, обеспечивающим их деятельность, рефлексивную и коммуникативную деятельность. Он может обособляться от естественных субъектов и быть готовым к присоединению к другим субъектам, а может выполнять общие сервисные функции сетевой организации. Функциональное назначение цифровых объектов и

<sup>1</sup> *Лепский В.Е.* Концепция субъектно-ориентированной компьютеризации управленческой деятельности. М.: Институт психологии РАН, 1998. 204с.

механизмы выполнения ими функций определяются с учетом системы принципов, среди которых важна группа субъектных принципов, и особенно принцип двойственного субъекта<sup>1</sup>. В кибернетике второго порядка базовая парадигма «субъект – субъект» определяет основной тип рефлексивности – коммуникативная рефлексия. Это существенно влияет на функциональное назначение цифровых субъектов для обеспечения коммуникативных процессов и коммуникативного отражения естественных предметов.

Принципиально важно отметить, что предлагаемый подход к рассмотрению взаимодействий цифровых и естественных субъектов формирует основу для организации гибридных систем, включающих иерархические и сетевые механизмы управления.

Реализация системно-генетического подхода осуществляется через превращение цифрового двойника в цифровой субъект при переходе от кибернетики первого порядка к кибернетике второго порядка. Это отражает исторический подход к развитию представлений о технических объектах<sup>2</sup>.

### *3.2.3.3. Цифровой метасубъект*

Существование субъектов и псевдосубъектов в постнеклассической кибернетике саморазвивающихся полисубъектных (рефлексивно-активных) сред, кибернетике третьего порядка, задается мировоззренческими основаниями, системой принципов и системой онтологий. Это создает основу для сборки в единое целое (метасубъект) субъектов (псевдосубъектов), входящих в среду, и обеспечивает ее жизнедеятельность и развитие.

Система принципов состоит из мировоззренческих принципов, принципов субъектности, принципов синергетики, принципов организации сложных социотехнических систем, общенаучных методологических принципов. Особое место занимают принципы субъектности.

Мы разработали систему онтологий, которая включает в себя онтологии жизнеобеспечения, преодоления жизненных переломов, стратегической постановки целей, разработки стратегий и проектов, реализации и инновационного сопровождения стратегий и проектов. Эта система онтологий имеет научную, методологическую и практическую апробацию.

---

<sup>1</sup> *Ленский В.Е.* Методологический и философский анализ развития проблематики управления. М.: Когито-Центр, 2019. 340 с.

<sup>2</sup> *Simondon, G.* (2017) *On the Mode of Existence of Technical Objects.* Translated by C. Malaspina and J. Rogrove. Minneapolis: Univocal Publishing.



В соответствии с мировоззренческими основаниями, системами принципов и онтологий конструируется цифровой метасубъект. Цифровой метасубъект – это система цифровых субъектов, обеспечивающая жизнедеятельность и развитие саморазвивающейся полисубъектной (рефлексивно-деятельностной) среды.

Система принципов взаимодействия естественных субъектов с цифровым метасубъектом имеет многоуровневую структуру. Верхний уровень включает четыре подуровня: конкретного субъекта, обеспечения деятельностной, коммуникативной и рефлексивной активности.

Парадигмы ИИ, соответствующие кибернетике третьего порядка, находятся в процессе формирования. Можно выделить два принципиально разных подхода. Первый ориентирован на спонтанное формирование метавселенной, максимизирует прибыль разработчиков ИИ. Второй ориентирован на доминирование правовых регуляторов и максимизацию контроля со стороны государства и общества. Компромисс может быть достигнут при использовании методологии постнеклассической кибернетики третьего порядка при разработке перспективной глобальной (экологической) парадигмы ИИ, которая будет соответствовать концепции цифрового метасубъекта.

Есть основания утверждать, что разработанные философско-методологические основания позволяют интенсифицировать развитие ИИ в интересах совершенствования процессов управления и развития социальных систем.

### **3.2.4. Выводы**

Представлены результаты философско-методологического анализа совместной эволюции научной рациональности, кибернетики и ИИ. Ведущее место в которой занимает субъектно-ориентированный подход. Это позволило с помощью системно-генетического подхода рассмотреть в историческом аспекте формирование понятий цифрового двойника, цифрового субъекта и цифрового метасубъекта как эволюционирующих технических объектов. Определено назначение и место этих цифровых образований в контексте кибернетики первого, второго и третьего порядка, а также связь с рефлексивностью в управлении и развитии социальных систем.

Полученные результаты полезны для совершенствования сред гибридной реальности с использованием искусственного интеллекта.

### 3.3. МАШИНЫ КАК ЛЮДИ

#### 3.3.1. Введение

На протяжении всей истории науки и технологий XX и XXI веков мы видим, что два стереотипа определяют ожидания от передовых исследований. Во-первых, явные успехи в одной определенной дисциплине создают ожидания чуда в более широком спектре дисциплин. Кажется, что научная революция вот-вот за поворотом. Во-вторых, надо скорее получать практический эффект от всех научных открытий. В истории СССР этот стереотип ярче всего отразился в истории разгрома генетики и лысенковщины. В истории России XXI века этот же стереотип может оказаться одним из главных факторов, сдерживающих прогресс ИИ. Эта область является очень общим направлением исследований, которая зачастую определяется и как наука и как технологии. Тем не менее, самой поверхностной философской рефлексии достаточно, чтобы сказать, что (а) ИИ не является реальной вещью<sup>1</sup>, которую можно было бы сделать предметом исследования и, следовательно, (б) рамки, в которых исследователь, претендующий на компетентность в сфере ИИ, самостоятельно определяет свой предмет, значительно шире, чем, скажем предметные рамки экспериментального физика<sup>2</sup>. В настоящем пункте монографии делается попытка показать составляющие части ИИ, возможный путь к Общему искусственному интеллекту. В заключении делается попытка обрисовать гипотезу посттюринговой методологии и экспериментальный способ ее проверки.

#### 3.3.2. Критерии и оценка прогресса в области исследований и технологий ИИ

Марвин Мински писал, что термин ИИ является «термином-чемоданом»: в нем заложено слишком много смыслов. И стоит добавить, что это порождает множество интерпретаций того, что есть ИИ. Государственные деятели и документы, которые они предлагают, определяют ИИ как комплекс технологических решений. Так определяет ИИ «Белая книга», подготовленная в начале 2022 года по заказу Правительства России<sup>3</sup>. Ученые более осторожны и называют ИИ

<sup>1</sup> *Bringsjord, Selmer and Naveen Sundar Govindarajulu, "Artificial Intelligence", The Stanford Encyclopedia of Philosophy (Fall 2022 Edition), Edward N. Zalta (ed).*

<sup>2</sup> *Russell, Stuart J. (Stuart Jonathan). (2010). Artificial intelligence: a modern approach. Upper Saddle River, N.J.: Prentice Hall.*

<sup>3</sup> «Развитие отдельных высокотехнологичных направлений. Белая книга». М: 2022 (интернет-ресурс <https://issek.hse.ru/mirror/pubs/share/565446894.pdf>)

областью научных исследований наряду с инженерией умных машин – так по крайней мере определял ИИ Джон Маккарти<sup>1</sup>, придумавший сам термин «искусственный интеллект» почти 70 лет назад.

Налицо очевидное противоречие – если исследуемая нами проблема является нерешенной (ИИ уровня человека пока не создан) и может быть классифицирована как *пред-проблемная ситуация* (известное неизвестное)<sup>2</sup>, то почему мы ограничиваем ИИ уже известными нам технологическими решениями. Но если это множество готовых решений (пусть и технологических), то возникает естественный вопрос «что именно в нем можно исследовать».

Рынок ИИ стремительно растет и это можно видеть на примере инвестиционной активности – в 2021 году сумма инвестиций в ИИ в мире превысила 170 млрд долл США<sup>3</sup>. Но неопределенность дефиниций широко используется теми, кто стремится заработать на растущем, как на дрожжах, рынке. Проблематика определения ИИ также осложняется широким использованием термина на популярных ресурсах, в политической риторике и в публичной коммуникации. Проведенное несколько лет назад исследование показало, что если взять принятое в computer science определение ИИ как приложений, созданных на основе алгоритмов машинного обучения<sup>4</sup>, то только 60% стартапов позиционирующих себя как ИИ будут удовлетворять этому критерию при детальном рассмотрении предлагаемых ими продуктов на рынок<sup>5</sup>. Такая всепроникающая популярность темы ИИ, иногда напрямую конвертируемая в экономическую прибыль, делает пристальное философское исследование ИИ особенно актуальным. Большие деньги и инвестиционные ожидания стимулируют практические предложения от стартапов и корпораций, которые во многом напоминают обещания «мичуринских» биологов. Без малого сто лет назад последователи Т.Д. Лысенко были готовы стремительно преобразовывать природу, следуя теории о наследовании приобретенных признаков без всякого внимания реальным механизмам генетики.

Кажется, что подобный интерес к теме ИИ, распространяющийся далеко за пределы области узких специалистов и непосредственных выгодополучателей, частично объясняется «магией» названия,

---

<sup>1</sup> Маккарти Дж. Архивные материалы. (интернет-ресурс <http://www-formal.stanford.edu/jmc/whatisai/node1.html>)

<sup>2</sup> Дубровский Д.И. Обман. Философско-психологический анализ. М.: «Канон+», 2010. — 336 с., стр. 241

<sup>3</sup> Stanford AI Index, (интернет-ресурс <https://aiindex.stanford.edu/report/>)

<sup>4</sup> Стэнфордская энциклопедия философии. (интернет-ресурс <https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf>)

<sup>5</sup> «Europe's AI start-ups often do not use AI, study finds», Financial Times, 05.03.2019 (интернет-ресурс <https://www.ft.com/content/21b19010-3e9f-11e9-b896-fe36ec32aece>)

отсылающей к изначальной метафоре, лежащей в основе ИИ – ученый-практик становится демиургом, повторяющим создание самого сложного объекта Природы, известного нам – интеллекта человека. Как бы настойчиво исследователи из *computer science* не повторяли, что их исследования не посвящены созданию «умных машин»<sup>1</sup>, идея создания искусственного агента, способного действовать на «человеческом уровне», привлекает внимание. Поэтому является актуальной задача анализа достигнутых результатов в области ИИ и сопоставления их идее создания Общего ИИ. Последняя цель является актуальной и сегодня и в прошлом, так как без сомнения мотивировала выдающихся исследователей, среди которых был и Алан Тьюринг<sup>2</sup>. Эту задачу можно уподобить «разделение зерен от плевел» в отношении ИИ.

Гари Маркус и Эрнст Дэвис, полагают, что такое «разделение» важных и мнимых достижений следует проводить на основе ряда критериев. Ими предложены следующие критерии<sup>3</sup>:

- *Какую задачу решает ИИ.* Какой конкретно результат мы получаем от использования этой системы ИИ?
- *Насколько общей является решённая задача.* Насколько чувствительна система ИИ к изменениям условий задачи?
- *Воспроизводим ли достигнутый результат.* Насколько продемонстрированное решение готово к индустриальному внедрению?
- *Является ли достигнутый результат полезным.* Насколько прагматически оправдано внедрение данной технологии при учёте затрат?
- *Безопасна ли созданная технология.* Готовы ли мы мириться с возможной угрозой человеческой жизни и свободе, которую создаёт ИИ?

---

<sup>1</sup> «Computer scientists are questioning whether Alphabet’s DeepMind will ever make A.I. more human-like», CNBS, 18.06.2021, (интернет-ресурс <https://www.cnbc.com/2021/06/18/computer-scientists-ask-if-deepmind-can-ever-make-ai-human-like.html>)

<sup>2</sup> Создатели информатики думали именно об интеллекте человеческого уровня: и Джон Маккарти, и Марвин Мински, и Норберт Винер публиковали статьи в философских журналах. Работы Джона Маккарти и Герберта Саймона, Марвина Мински до сих пор актуальны и являются одними из самых цитируемых в философии ИИ. Знаменитая статья Алана Тьюринга *Computing Machinery and Intelligence*, которую принято считать условной точкой возникновения ИИ как области, была опубликована в оксфордском философском журнале *Mind*. (интернет-ресурс <https://philpapers.org/browse/philosophy-of-artificial-intelligence>).

<sup>3</sup> *Marcus G., Davis E., Rebooting AI: Building Artificial Intelligence We Can Trust, Pantheon, 2019, 288 p.*

Если мы будем рассматривать существующие системы ИИ с помощью вышеозначенных пяти критериев, мы заметим, что большинство из них никак не участвуют или даже не собираются участвовать в достижении Общего ИИ, но создаются с целью демонстрации возможности решения конкретных задач или поиска решения актуальных научных проблем.

Предложенные Маркусом критерии безусловно полезны и должны «встать на вооружение» любого эксперта, работающего в отрасли ИИ. Однако, они не содержат самого важного критерия, который действительно указывал чем «зерно» отличается от «плевела», это немаловажный вопрос: *«Насколько интеллектуальной является машина?»*. Можно ли считать, что действия машины состоят из последовательности результативных действий, приводящих к достижению поставленных перед машиной целей? При этом машина должна действовать более эффективно по сравнению с человеком-экспертом. Классическим примером можно считать сравнение беспилотного автомобиля с человеком при вождении в сложных погодных условиях – а именно, такие условия и отличают хорошего водителя от плохого.

После дополнения этого списка можно предположить, что эти шесть критериев являются инструментом оценки результатов исследований и технологических разработок в ИИ.

### **3.3.3. Зачем создавать Общий ИИ?**

Вопрос не является праздным с учетом той озабоченности, которая высказывается многими зарубежными и отечественными учеными. Исследователи отмечают широкий спектр негативных проблем при распространении общего ИИ (AGI): от снижения общих когнитивных навыков людей до приписывания субъектности ИИ и лишения человека ответственности за совершаемые поступки.

Можно задаться вопросом, следует ли нам вообще пытаться делать предметом исследования Общий ИИ? Один из авторов настоящей монографии попытался найти ответ на этот вопрос во взаимодействии с аудиторией для чего провел несколько опросов в ходе лекций перед старшеклассниками и студентами нескольких московских вузов. Всем им был задан один вопрос *«зачем нам нужен ИИ?»*. Полученные ответы можно разделить на четыре группы:

- *Для повышения производительности труда в экономике.* В таких ответах молодежь подчеркивала улучшение качества жизни людей;

- *Для познания окружающего мира.* Учащиеся давали подобный ответ, настаивая, что ИИ может сыграть ключевую роль в научном поиске и будущих исследованиях. Это может быть изучение других планет или раскрытие тайны сознания и мозга человека;
- *Ради всего человечества необходимо создавать ИИ и роботов,* чтобы сделать жизнь людей лучше, освобождая их от грязной, опасной или механической работы. Роботы могут помочь избежать ненужных жертв при ликвидации стихийных бедствий. Все вместе высвободит ресурсы цивилизации на новое развитие;
- *Создание нового человека, трансгуманизм.* Молодежь обеспокоена проблемой одиночества и разобщения людей и считает, что человеку нужен друг, соратник или конкурент, которым и должен стать ИИ для того, чтобы человек продолжил свою эволюцию.

Распределение ответов между этими группами показано на Рис. 3-3).

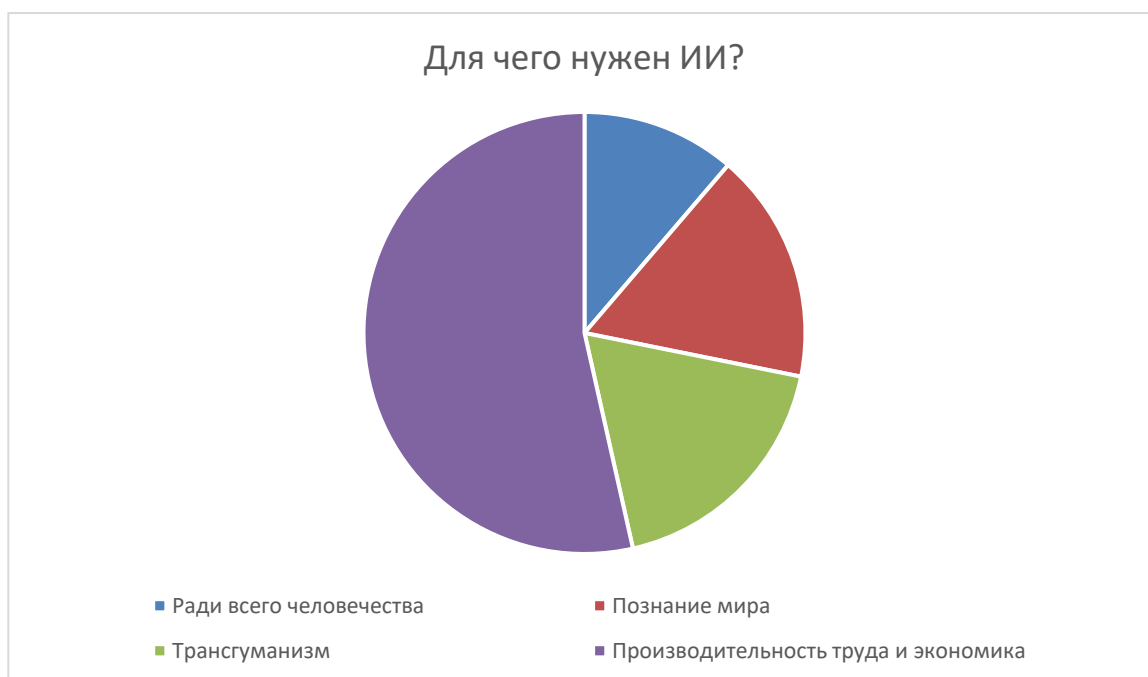


Рис. 3-3. Распределение ответов молодежи (14-25 лет) по вопросу о целях создания ИИ. Данные: автор

В ответах, которые мы отнесли к категории «ради всего человечества», молодые люди явно отмечают, что ИИ должен заменить человека в опасных местах, убрать человеческий фактор в критических ситуациях, не забывая и о том, чтобы человек был счастливее, а жизнь всех стала лучше. Общая доля таких ответов составила 11% и это говорит, что темы связанные со справедливым ИИ и надежным ИИ, социальными тематиками находят глубокий и значимый отклик у молодежи.

Из этого распределения можно сделать очень предварительный вывод о том, что рациональная и экономическая подоплека распространения и внедрения технологий ИИ является важнейшей, но далеко не единственной причиной, которой следует уделять внимание. Вопросы познания мира и познания человека как биологического вида столь же значимы, как и вопросы улучшения человека или создания общества всеобщего благоденствия. Совокупно, рациональные поводы внедрения ИИ лишь чуть превышают половину всех полученных ответов. Такие данные могут указывать на то, что люди видят в ИИ не только машину, оптимизирующую механические операции, но и надежного друга, помощника, открывающего для нас неизвестные аспекты окружающего нас мира.

ИИ как эпистемологический помощник человека, как инструмент познания окружающего мира, важны для молодежи больше, чем можно было бы предположить – доля этих составила 17% от общего числа всех ответов. Молодые люди отмечают, что ИИ необходимо создавать именно для того, чтобы расширять границы познания человека в том числе и за счет того, что у людей будет больше времени для творчества, так как машины будут выполнять монотонную работу. При этом ответы обращены как во вне – поиск возможностей для колонизации других планет, так и внутрь – познание возможностей интеллекта человека.

### **3.3.4. Человек и машина: в чем различие?**

Чтобы ответить на этот вопрос чуть более конкретно, чем «мы живые, а они нет», необходимо проанализировать принципиальные ограничения сегодняшних умных машины – как программ для компьютеров, так и машин-роботов, действующих в физическом мире. Общий ИИ – это не абстрактная универсальная технология избавления от всех проблем, но нечто, что должно помочь значительно продвинуться в решении существующих задач, которые не могут решать современные системы ИИ, но с легкостью может решить человек.

Начать следует с основного инструмента интеллекта – мозга человека. Следуя компьютерной метафоре, наш мозг можно сравнить с аппаратным обеспечением компьютера (табл. 3-3). Проведя такое сравнение, мы увидим, что мозг человека отличается от машины по множеству параметров.

**Таблица 3-3**  
Различие между мозгом и компьютером<sup>1</sup>

ЧЕЛОВЕК		МАШИНА
Асинхронные всплески нейронной активности	ВНУТРЕННЕЕ ВРЕМЯ	Универсальные системные часы
Аналогово-цифровой гибрид	СИГНАЛЬНАЯ СИСТЕМА	Двоичное исчисление
Аналоговое нелинейное суммирование с последующим однополупериодным выпрямлением и пороговой обработкой	ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА	Булева алгебра
Нет	РЕАЛИЗАЦИЯ УМТ	ДА
>1 000	КОЛИЧЕСТВО ТИПОВ ВЫЧИСЛИТЕЛЬНЫХ ЯДЕР	Немного
МИЛЛИСЕКУНДЫ, 10 <sup>-3</sup> СЕК	СКОРОСТЬ ВЫЧИСЛЕНИЙ	НАНОСЕКУНДЫ, 10 <sup>-9</sup> СЕК
От 1 000 до 50 000	СВЯЗАННОСТЬ ЯДЕР	Менее 10
Выдерживает потери компонентов	НАДЕЖНОСТЬ	Хрупки
среднее потребление – 100 Вт, 25% всего организма – мозг	ЭНЕРГЕТИКА	От 30 Вт до 30 МВт

С точки зрения химии, Природа создала человека и других живых существ пользуясь относительно простым набором элементов – почти 99% массы тела человека состоит из шести элементов Таблицы Менделеева: кислорода, углерода, водорода, азота, кальция и фосфора. Однако, построение робота или компьютера требует намного большей кладовой – в среднем, машины используют от 30 до 50 видов различных элементов. Это обуславливает значительную сложность построения любых умных машин по сравнению с человеком.

Сравнительный анализ конструктивных особенностей также свидетельствует в пользу сложности человека – мы состоим из более чем 200 костей, 360 суставов, приводимых в движение 850 мышцами, что дает нам феноменальные возможности в 244 степенях свободы. В противоположность, наиболее распространенные роботы, относимые к виду промышленных, имеют от 6 до 7 степеней свободы. Лучшие экспериментальные антропоморфные роботы имеют примерно в 10 раз меньше степеней свободы, чем у человека<sup>2</sup>.

<sup>1</sup> Koch C. The Feeling of Life Itself: Why Consciousness is Widespread but Can't Be Computed. The MIT Press, 2019

<sup>2</sup> «Meet Atlas, the Robot Designed to Save the Day», The MIT Technology Review, 12.07.2013 (интернет-ресурс) <https://www.technologyreview.com/2013/07/12/15691/meet-atlas-the-robot-designed-to-save-the-day/>



Если мы попробуем провести сравнение сенсорных аппаратов человека и машины, то мы увидим огромные различия. Например, несмотря на сходные сенсорные каналы, обработка информации из оптического диапазона робота и человека происходит с разной эффективностью: там, где человеческий мозг обрабатывает примерно 100 битов в секунду, машина обрабатывает чуть ли не каждый пиксель<sup>1</sup>.

Контраст между человеком и машиной станет ещё более разительным, если мы рассмотрим мобильность и адаптивность к среде. Например, робот, «научившийся» ходить по плоской поверхности, не способен воспроизвести спектр способностей человека к передвижению в разных средах.

Тем не менее, есть отдельные узкофункциональные области, где машины равны и даже существенно превосходят человека по эффективности. Все они относятся к достаточно простым задачам, таким как операции с числами, управление базами данных, поднятие тяжестей или монотонное выполнение итеративных движений или задач.

Проведенное выше сравнение человека и машины может быть полезной эвристикой в выработке приоритетных направлений исследования в области ИИ:

- *Машины плохо «понимают» окружающий мир.* Выше мы говорили об эффективной редукции, к которой способен сенсорный аппарат человека. Необходимо совершенствовать сенсорный и когнитивный аппарат машин для того, чтобы они могли эффективно взаимодействовать с окружающим миром.
- *Машины плохо взаимодействуют с другими машинами и людьми в группах.* Одна из самых сложных проблем в развитии общего интеллекта – проблема социального интеллекта, то есть способность искусственного агента к координации и агентности с другими машинами или людьми.
- *Машины плохо понимают естественный язык людей.* Многие исследователи рассматривают проблему языка как главную проблему, адресовав которую мы получим ключ от интеллекта человеческого уровня.
- *Машины плохо манипулируют объектами мира.* Несмотря на успехи в отдельных направлениях, машины пока ещё далеки от универсальных манипуляторов человеческого уровня.

Гипотезой, поддерживающей полезность такого сравнения, является гипотеза воплощённого интеллекта, утверждающая связанность когнитивных функций естественного интеллекта и эволюции

---

<sup>1</sup> Decadal Plan for Semiconductors, Semiconductor Research Corporation, 2021, p.12 (интернет-ресурс) <https://www.src.org/about/decadal-plan/decadal-plan-full-report.pdf>

человеческого тела<sup>1</sup>. Она же объясняет и проблему интеграции: решения для разных областей, например, автопилот и игра в шахматы пока ещё не интегрированы в единый интеллектуальный центр, тогда как структуры человеческого мозга не только интегрируют разные функции, но и функционально универсальны<sup>2</sup>.

### 3.3.5. Тест Тьюринга как путь к ИИ

Традиция сравнения человека и машины в контексте выяснения природы человеческого интеллекта имеет древнюю традицию. Ещё в 17 веке, рассуждая о разнице человека и машины, философ и математик Рене Декарт выделил критерии, по которым может быть обнаружена гипотетическая внешне идентичная человеку машина<sup>3</sup>:

- (1) она не сможет свободно пользоваться человеческим языком, и
- (2) не будет способна действовать в разных средах и контекстах.

В начале 20-го века эта идея получила развитие в работе математика Курта Гёделя, а именно в «закрученной» формулировке теоремы о неполноте, положившей начало гёделевским аргументам против общего ИИ<sup>4</sup>:

*Либо наш ум не является механическим, либо математика, и даже арифметика, не является нашей собственной конструкцией.*

Это стало отправной точкой для разделения философских рассуждений о природе интеллекта на два множества: ментализм и механицизм<sup>5</sup>. Теории, принадлежащие первому множеству, склоняются к фундаментальному превосходству человеческого интеллекта над машинным, а следовательно, и к невозможности общего искусственного

---

<sup>1</sup> Pfeifer, R., Iida, F. (2004). Embodied Artificial Intelligence: Trends and Challenges. In: Iida, F., Pfeifer, R., Steels, L., Kuniyoshi, Y. (eds) Embodied Artificial Intelligence. Lecture Notes in Computer Science, vol 3139. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-540-27833-7\\_1](https://doi.org/10.1007/978-3-540-27833-7_1)

<sup>2</sup> Эделмен Дж, Маунткасл В. Разумный мозг. Мир, 1981.

<sup>3</sup> Descartes R (1984). The Philosophical Writing of Descartes, Vol. II, Edited and Translated by John Cottingham, Robert Stoothoff, Dougald Murdoch, Cambridge University Press, Cambridge, p. 19.

Descartes R (1984). The Philosophical Writing of Descartes, Vol. II, John Cottingham, Robert Stoothoff, Dougald Murdoch, p. 54.

Descartes R (1984). The Philosophical Writing of Descartes, Vol. II, Edited and Translated by John Cottingham, Robert Stoothoff, Dougald Murdoch, p. 13.

Descartes R (1984). The Philosophical Writing of Descartes, Vol. II, Edited and Translated by John Cottingham, Robert Stoothoff, Dougald Murdoch, 1984, p. 7.

<sup>4</sup> Крайзель Г. Биография Курта Гёделя. М., 2003.

<sup>5</sup> Badcock, Christopher (2004) Mentalism and mechanism: the twin modes of human cognition. In: Crawford, Charles B and Salmon, Catherine A, (eds.) Evolutionary Psychology, Public Policy and Personal Decisions. Lawrence Erlbaum Associates, Mahwah, N.J., pp. 99-116.

интеллекта. Механицистские теории, наоборот, утверждают принципиальную сводимость интеллекта к механизмам (биологическим, электрическим и другим), что позволяет воспроизвести его на небιологических носителях.

Однако в середине 20-го века выдающемуся британскому математику Алану Тьюрингу удалось сформулировать принципиально новый подход: наблюдая за деятельностью людей-вычислителей, т.е. *computers* (а тогда это было должностью, а не названием устройства), он предположил, что интеллект, механицистская реализация которого возможна на разных физических носителях, может быть выражен в ограниченном наборе функций. И этот ограниченный набор, а вовсе не всё богатство ментальной жизни человека, может и должен быть операционализирован и воспроизведён.

Под влиянием этой идеи почти вся послевоенная деятельность Тьюринга была связана с цифровыми компьютерами. Если попытаться суммировать её в одном предложении, мы получим название одной из его основных статей: «Могут ли машины мыслить?»<sup>1</sup>. Тот набор функций, который Тьюринг считал релевантным в отношении мышления машин, сводился к короткому списку: (1) играть в настольные игры, (2) разговаривать, (3) переводить тексты, (4) шифровать сообщения и (5) решать математические задачи.

Очевидно, что как было замечено выше, человеческая деятельность не сводится к такому набору функций – это отмечал и сам Тьюринг. И тем не менее, с его точки зрения, осмысленное отнесение объекта к множеству мыслящих не может заключаться в простом копировании всех человеческих особенностей. В конце концов, никто не станет спорить, что не всю человеческую деятельность мы относим к интеллектуальной.

В качестве иллюстрации этой идеи Тьюринг предложил знаменитый тест, призванный верифицировать подобное отнесение: редуцировав всё множество интеллектуальных функций к языковой компетенции, мы можем утверждать, что машина мыслит, тогда и только тогда, когда её языковая компетенция функционально эквивалентна человеческой. Для реализации такого теста мы должны изолировать языковую компетенцию, возведя так называемую «стену Тьюринга», депривирующую тестирующего от всех каналов коммуникации кроме символической.

Конечно, тест Тьюринга – это неформальный тест, применяемый к конкретному объекту, имеющий скорее демонстрационный, а не

---

<sup>1</sup> *Turing, Alan* (October 1950), "Computing Machinery and Intelligence" (PDF), *Mind*, LIX (236): 433–460, doi:10.1093/mind/LIX.236.433

доказательный характер. Его демонстрационная сила основана на предположении, что, во-первых, мы понимаем других, представляя, как бы мы чувствовали себя или думали на их месте, а во-вторых, во время развития нашего мышления мы используем наши наблюдения за тем, как ведут себя другие, чтобы лучше понять самих себя. Такое циклическое определение собственной идентичности создает сильную ассоциативную связь между нашим восприятием собственного интеллекта и интеллектом других. Именно это делает тест таким убедительным: нам сложно не рассматривать сущности, с которыми мы можем полноценно общаться, в качестве похожих на нас самих.

### **3.3.6. Посттьюринговые исследования ИИ**

Впоследствии яркая идея теста была неоднократно модифицирована: придумывались тесты, которые расширяли коммуникативные каналы, считая вербальную коммуникацию недостаточной, а то и вовсе не необходимой для верификации мышления<sup>1</sup>. Сегодня существует множество вариаций теста, которое, тем не менее, может быть систематизировано с учётом изначальной структуры теста. Более подробно этот вопрос рассмотрен в других работах автора<sup>2</sup>.

В зависимости от изолируемого тестом канала коммуникации все множество возможных тестов Тьюринга может быть распределено по четырём областям, см. Рис. 3-4.

---

<sup>1</sup> *Ефимов А.Р.* Кандидатская диссертация «Философско-методологические основы посттьюринговой интеллектуальной робототехники», Институт Философии РАН, 2021, (интернет-ресурс [https://iphras.ru/uplfile/zinaida/ROOTED/aspir/autoreferat/efimov/dissertatsiya\\_efimov\\_final\\_151020.pdf](https://iphras.ru/uplfile/zinaida/ROOTED/aspir/autoreferat/efimov/dissertatsiya_efimov_final_151020.pdf))

<sup>2</sup> *Ефимов А.Р.* Снятся ли чат-ботам андрюиды? Перспективы технологического развития искусственного интеллекта и робототехники// Философские науки. 2019. Т62. № 7. С. 73-95.

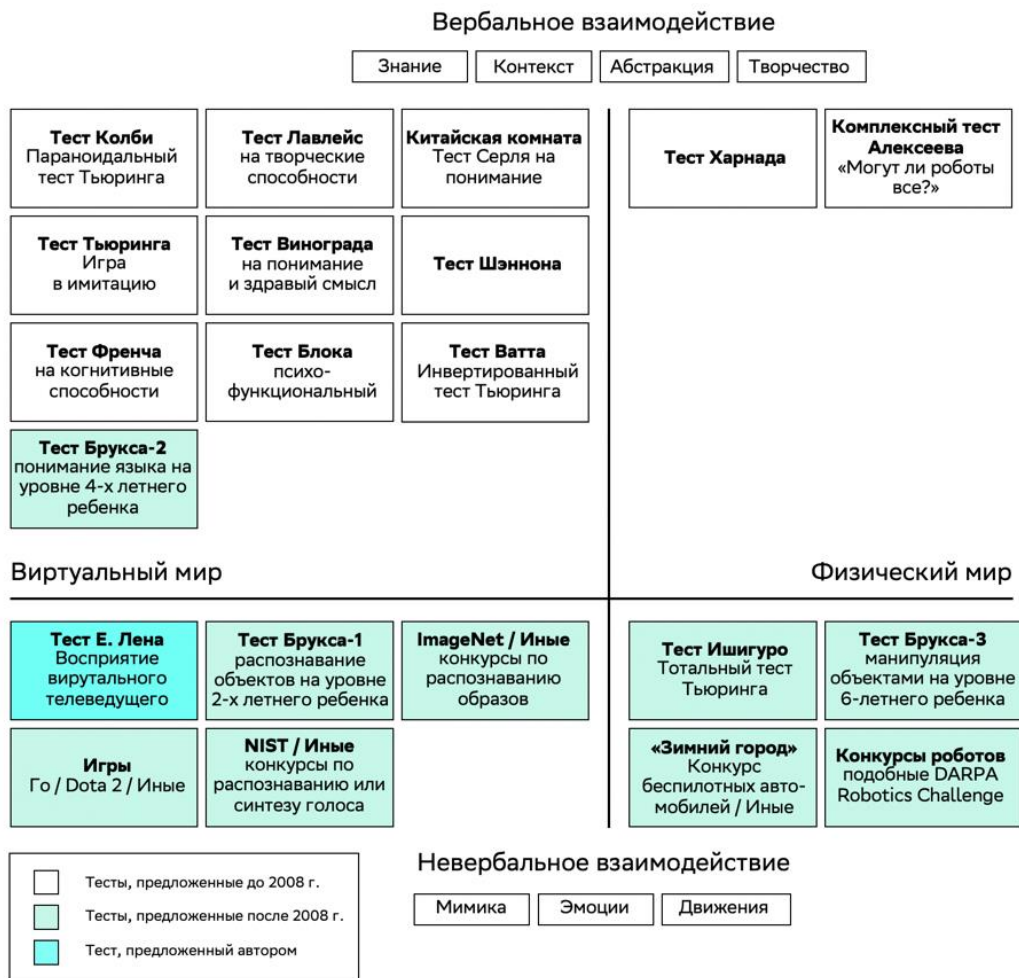


Рис. 3-4. Посттьюринговая методология: разрушение стены Тьюринга на пути к Общему ИИ<sup>1</sup>

Теперь попробуем применить эту систематизацию к Общему ИИ. Как было сказано выше, сравнение человека и машины даёт нам несколько перспективных направлений исследований, которые заключаются в совершенствовании способностей машин к (1) пониманию окружающего мира, (2) социальному взаимодействию, (3) пониманию естественного языка и (4) манипулированию физическими объектами. Интеграция этих способностей вместе со свойством адаптивности, которое так отличает человека от современных систем ИИ, должна быть связана с возможностью перемещения ИИ между разными областями коммуникации. Заметим также, что в этих способностях есть высокая доля неопределенности термина «понимание».

<sup>1</sup> *Ефимов А.Р.* Посттьюринговая методология: разрушение стены на пути к общему искусственному интеллекту // Интеллект. Инновации. Инвестиции. №2. 2020. С. 74-80.

ИИ, обладающий такими способностями, сможет обосновано претендовать на звание Общего ИИ. Таким образом, посттьюринговый этап в развитии ИИ и робототехники будет характеризоваться «снятием» изолированности интеллектуальных функций, а в контексте теста Тьюринга, переходом к новому универсальному комплексному тесту: взаимодействие с ИИ должно быть не менее или даже более интересным, чем с естественным.

Данный подход, таким образом, отличается от мейнстримного подхода большинства исследователей ИИ, моделирующих функции интеллекта в виртуальной среде, но лежит в рамках альтернативной, посттьюринговой методологии: для достижения Общего ИИ система ИИ должна обладать (1) интегральностью и (2) автономностью. Для этого необходимо моделировать когнитивные архитектуры совершенно иного типа, чем используются сегодня в ИИ. Но откуда взять эти архитектуры? Для разработки модели, отвечающей требованиям Общего ИИ мы предлагаем использовать достижения философии сознания и нейронаук.

Автором впервые был предложен посттьюринговый методологический подход, который переопределяет взаимодействие машины (робота) с окружающей средой следующим образом. Все возможные типы взаимодействия группируются вокруг двух концептуальных измерений (виртуальное-физическое и вербальное-невербальное), образуя четыре пространства, называемые *техно-умвельтами*. Последний термин подчеркивает взгляд на взаимодействие машины (технической системы) с окружающим миром (включая человека). Подобный подход дает возможность сформулировать определение Общего ИИ:

Общий ИИ (AGI) – это способность робота (или иной технической системы) обучаться и действовать совместно с человеком или автономно в любой области, но лучше, чем специалист в этой области, достигая поставленных целей во всех четырёх техно-умвельтах при ограничении потребляемых роботом ресурсов<sup>1</sup>.

При такой постановке задачи возникает экспериментальная и практическая возможность перехода робота от выполнения узких задач к более широким. Иными словами, одна и та же техническая система (или робот) действуют в различных техно-умвельтах. Именно их способность действовать в разных техно-умвельтах и определяет степень «общности» их интеллекта. Машина (робот), действующая лишь

---

<sup>1</sup> Ефимов А.Р. Философско-методологические основы посттьюринговой интеллектуальной робототехники / Диссертация на соискание ученой степени кандидата философских наук. Институт Философии РАН, 2021.

в одном техно-умвельте менее общая чем машины действующая в двух или более.

Разберем это на конкретном примере. Все тьюрингоподобные тесты, основанные на анализе вербального диалога человека и компьютера через текстовые интерфейсы, находятся лишь в одном техно-умвельте (см. Рис. 3-4).

Увеличение «общности» искусственного интеллекта может выражаться в выходе машины за пределы техно-умвельта «*вербальное-виртуальное*». В качестве примера такого выхода может быть рассмотрен выход машины в техно-умвельт «*невербальное-виртуальное*», когда машина, обладая способностями первого техно-умвельта также способна действовать и во втором. Данная идея многократно рассматривалась в научной фантастике и даже была примером известного советского документального фильма о тесте Тьюринга. Однако, ни в одном из известных примеров технологической реализации не была реализована.

Автор этого пункта монографии предлагает провести экспериментальную проверку гипотезы, согласно которой одна и та же система, имеющая свойства Общего ИИ, будет способна к взаимодействию с человеком как минимум в двух техно-умвельтах. В эксперименте предполагается создать стенд, в котором будут впервые одновременно воплощены две способности (ранее уже созданные):

Машина ведет диалог с человеком на произвольную тему (реализовано). Это делается с помощью одной из мощнейших диалоговых системы на базе ruGPT-3<sup>1</sup>. Кроме того, предполагается интеграция с системой онлайн генерации антропоморфных, гиперреалистичных визуализаций собеседника для человека, таких как Visper<sup>2</sup>.

В тьюринговой игре в имитацию испытуемые (человек и машина) были разделены стеной, которая была проницаема лишь для символической (а точнее, вербальной) коммуникации. Суть предлагаемого посттьюрингового эксперимента заключается в том, что стена становится проницаемой для любых видов вербальной и невербальной коммуникации в виртуальном пространстве. Возвращаясь к терминологии, введенной ранее, машина (робот) действует не в одном техно-умвельте, но сразу в двух. Предполагается, что данный способ коммуникации является более общим.

В эксперименте будет два параллельных режима: (1) диалог с машиной и (2) диалог с человеком. В обоих режимах в роли судьи-

---

<sup>1</sup> Интернет-ресурс <https://developers.sber.ru/portal/products/ru-gpt-3?attempt=1>

<sup>2</sup> Интернет-ресурс <https://visper.tech>

наблюдателя выступает человек. В обоих режимах он ведет диалог с «личинкой», аватаром (ранее продемонстрированным). Предполагается, что будет не менее 30 судей различного возраста и образовательного культурного уровня. Со всеми судьями диалоги ведут только аватары. Однако, в ряде случаев таких диалогов (предстоит уточнить методически) сам диалог будут вести чат-боты (на основе GPT-3) в режиме реального времени.

Кроме того, для контроля будет проводиться классический тест Тьюринга в диалоге с чат-ботом для контроля.

Гипотеза эксперимента состоит в том, что число ошибок судей (т.е. ошибок идентификации робота и человека) при диалоге с аватаром+чат-ботом (т.е. роботом, работающим в двух техно-умвельтах, посттьюринговый подход) будет больше, чем при классическом диалоге в игре в имитацию (тьюринговый подход).

Данный эксперимент позволит подтвердить, что для создания Общего ИИ необходимо научить робота действовать (выходить за пределы) в различных техно-умвельтах.

### **3.3.7. Выводы**

Вероятно, еще не скоро мы увидим машины, которые похожи на людей внешне или поведенчески. Однако, это технологически мыслимо при условии создания действительно общих решений, общего искусственного интеллекта, выходящего за рамки тьюринговой парадигмы решения частных головоломок. Мир намного сложнее чем разговор вслепую между машиной и человеком. В этом пункте авторы делают попытку обрисовать маршрут достижения такого состояния, в котором машинам придется преодолевать границы между техно-умвельтами, разделяющими виртуальное-физическое и вербальное-невербальное. Одной из остановок этого маршрута может быть посттьюринговый эксперимент, описанный в настоящей работе.



## 3.4. СУБЪЕКТНОСТЬ ОБЪЯСНИМОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

### 3.4.1. Актуальность и подходы в решении вопроса

Этот пункт подготовлен на основе авторской публикации<sup>1</sup>, посвящен определению путей обеспечения способности систем ИИ давать объяснения своим выводам. Современные нейронные сети содержат сотни слоев нейронов, число параметров этих сетей достигает триллионов, семантика моделей ИИ усложняется, уходя на квантовый и нелокальный уровни. Ведущие компании мира вкладывают большие средства в создание объяснимого ИИ (eXplainable AI, XAI), однако результат пока остается неудовлетворительным – человек зачастую не может понять «объяснений» ИИ потому, что последний принимает решения иным в отличие от человека образом, а возможно потому, что получить хорошее объяснение невозможно в рамках классической парадигмы ИИ. Настоящая работа акцентирует внимание на разрешении этой проблемы через погружение моделей ИИ в социально-гуманитарную среду.

Уже давно разработчиков систем ИИ интересовали правдоподобные рассуждения и достоверный вывод<sup>2,3</sup>. С успешным разрешением проблемы «Черного ящика ИИ», как она сейчас иногда обозначается<sup>4</sup>, научно-технический мир столкнулся давно, когда в создании систем ИИ основное внимание уделялось построению экспертных систем. Тогда система ИИ на основе замкнутой базы знаний, состоящей из порядка 400 правил-продукций, выводила результат, но не могла его доходчиво объяснить. Такие системы ИИ не вызывали доверие, плохо продавались. С этой проблемой справились путем усложнения логик и создания дополнительных баз знаний.

За последние годы ИИ его сложность резко возросла. Меняются взгляды на семантику моделей систем ИИ, ее экспликация уходит на когнитивный уровень, учитывает квантовые и релятивистские эффекты<sup>5</sup>. При этом встают новые проблемы: обостряются угрозы в сфере

---

<sup>1</sup> Райков А.Н. Субъектность объяснимого искусственного интеллекта // Философские науки. 2022. Т. 65. № 1. С. 72–90. <https://doi.org/10.30727/0235-1188-2022-65-1-72-90>

<sup>2</sup> Polya G. Mathematics and plausible reasoning. – New Jersey, Princeton: Princeton University Press. 1954. Vol. 1&2.

<sup>3</sup> Вагин В.Н. и др. Достоверный и правдоподобный вывод в интеллектуальных системах / Под ред. В.Н. Вагина и Д.А. Поспелова. – М.: ФИЗМАТЛИТ, 2004.

<sup>4</sup> Calderon R., Maxwell R. The Black-box effect: what level and type of opacity should society accept? *Frontiers in Artificial Intelligence*. 2022, in print.

<sup>5</sup> Raikov A. Cognitive Semantics of Artificial Intelligence: A New Perspective. *Topics: Computational Intelligence*. – Singapore: Springer. Vol. XVII. 2021. <https://doi.org/10.1007/978-981-33-6750-0>

кибербезопасности, создаются квантовые версии вычислителей, возрастает опасность злонамеренного использования ИИ.

В таких условиях ведущие компании мира увеличивают инвестиции в создание объяснимого ИИ (XAI, eXplainable AI). Ключевыми идеями исследований сейчас являются: послойный анализ работы ИИ, испытанные ранее логические схемы вывода, построения на графах знаний и пр.

Разветвленные связи в графах знаний делают их полезными для объяснений выводов систем ИИ. Например, это делается через подсчет расстояний между профилем пользователя и элементами в графе знаний<sup>1</sup>. В работе<sup>2</sup> нейронную сеть заставляют дать объяснение путем выявления причин взаимодействия пользователя с системой.

Однако, результаты пока не считаются удовлетворительными. Ответ на каждый новый вопрос по созданию объяснимого ИИ, как отмечают исследователи, делающий вроде бы «черный ящик» «белым», на самом деле обнаруживает внутри его множество других «черных ящиков». При этом экспоненциально растет число необходимых для получения объяснений вычислений, что зачастую заставляет отказываться от поиска и включения в систему компоненты объяснения.

Все больше внимание исследователей привлекает тема учета субъектного фактора, помещения систем ИИ в саморазвивающуюся рефлексивно-активную среду, построение онтологий контекста использования<sup>3,4</sup>.

Авторы настоящей монографии предлагают поменять акценты в парадигме создания XAI, больше обратить внимание на субъектные аспекты, внешнюю и глубинную стороны традиционных моделей ИИ. При этом к внешней стороне может относиться социально-экономическая и субъективная реальность, а к глубинной – флюктуирующая, квантовая и релятивистская природа сознания человека.

Вместе с тем проблема XAI, как все более становится очевидным, это не только логическое и технологическое предоставление удовлетворительных объяснений, она тесно связана с

---

<sup>1</sup> Liu R., Balsubramani A. and Zou J. Learning transport cost from subset correspondence // arXiv preprint arXiv:1909.13203. 2019.

<sup>2</sup> Madry A., Makelov A., Schmidt L., Tsipras D. and Vladu, A. Towards deep learning models resistant to adversarial attacks // arXiv preprint arXiv:1706.06083, 2017.

<sup>3</sup> Дубровский Д.И. Задача создания Общего искусственного интеллекта и проблема сознания // Филос. науки / Russ. J. Philos. Sci. 2021. 64(1). С. 13–44. DOI: 10.30727/0235-1188-2021-64-1-13-44

<sup>4</sup> Лепский В.Е. Искусственный интеллект в субъектных парадигмах управления. Филос. науки. 2021. 64(1). С. 88–101.

неформализуемыми аспектами феномена сознания, субъективной реальности<sup>1</sup>, противоречивого эпистемологического и этического выбора<sup>2</sup>. Ее разрешение должно учитывать способы, которыми люди достигают политических, экономических, социально-технических договоренностей, а также соглашаются с тем, чтобы ими впоследствии руководствоваться.

Ссылаясь на требование прозрачности в логическом принятии решений в работе<sup>3</sup> отмечается потребность понять, что такое объяснение вообще. Идея того, что имеется в виду под объяснением далеко не однозначна и варьируется в зависимости от ситуации и дисциплины. Довольно емкий обзор по теме объяснения<sup>4</sup> представляет собой энциклопедический охват ссылок на работы по ХАИ в различных дисциплинах. В обзоре анализируются ключевые концепции ХАИ и различные виды систем ИИ (экспертные системы, системы рассуждений на основе прецедентов, системы машинного обучения, байесовские классификаторы, статистические модели и деревья решений). Рассматривается разнообразие приложений – классификации жестов, изображений, текста, отладка программ, музыкальные рекомендации, финансовый учет, стратегические игры, формирование команд, роботы, агенты-неигровые персонажи, диагностика болезней, различные гипотезы, касающиеся отношения объяснения к фундаментальным когнитивным процессам, связи с обучением, пользователи объяснений, ограничения и пр. По всей видимости, обзор не столько определяет пути развития ХАИ, сколько демонстрирует неочевидность дальнейшего развития темы.

В работе<sup>5</sup> приводится обзор темы ХАИ, выявляются коммерческие, этические и нормативные причины, по которым необходимы объяснения. Рассматриваются различные цели объяснений, например, чтобы оправдать, контролировать, улучшать и открывать. Обсуждаются также методы объяснимости с точки зрения локального и глобального,

---

<sup>1</sup> Дубровский Д.И. Задача создания Общего искусственного интеллекта и проблема сознания// Филос. науки / Russ. J. Philos. Sci. 2021. 64(1). С. 13–44. DOI: 10.30727/0235-1188-2021-64-1-13-44

<sup>2</sup> Kaul N. (2022). 3Es for AI: Economics, Explanation, Epistemology. *Front. Artif. Intell.* 5:833238. doi: 10.3389/frai.2022.833238

<sup>3</sup> Rauber A., Trasarti R., and Gianotti F. Transparency in Algorithmic Decision Making. *ERCIM News*. 2019. Vol. 116. P. 10–11. [www.ercim.eu](http://www.ercim.eu).

<sup>4</sup> Mueller, S.T., Hoffman, R.R., Clancey, W., Emrey, A., and Klein, G. Explanation in Human-AI systems: A Literature Meta-Review Synopsis of Key Ideas and Publications and Bibliography for Explainable AI. DARPA XAI Literature Review. 2019. <https://arxiv.org/ftp/arxiv/papers/1902/1902.01876.pdf>.

<sup>5</sup> Adadi, A., Berrada, M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). 2018. Vol. 6. P. 52138–52160.

внутреннего и апостериорного, модельно-специфического и модельно-независимого. В указанной работе приводится таблица, обобщающая ключевые концепции ХАИ. Облако слов ХАИ, которое представлено в работе, имеет интерпретируемый язык машинного обучения (ML) и объяснимый ИИ как наиболее известные термины в литературе.

Работа<sup>1</sup> представляет обзоры концепций, таксономий, возможностей и проблем ХАИ – такие обзоры как: классификация моделей ML в зависимости от их уровня объяснимости, таксономия литературы и тенденций в сфере объяснимости для моделей машинного обучения. Авторы этой работы отмечают, что интерпретируемость «черного ящика» машинного обучения важна для обеспечения беспристрастности при принятии решений, устойчивости к злонамеренным возмущениям, а также в качестве гарантии того, что только значимые переменные определяют результат. Авторы полагают, что феномен понятности является наиболее важным атрибутом ХАИ.

В обзоре<sup>2</sup>, посвященном принципам объяснения и человеко-машинным системам ИИ, делается акцент на необходимости при создании системы объяснения ориентации на человека, который имеет ожидания от объяснения. Работа в области психологии, посвященная познанию и предубеждениям, использует данные человеческого мышления, чтобы аргументировать их вклад в интерпретируемые модели сложных систем ИИ<sup>3</sup>.

В итоге приведенного обзора можно отметить в целом позитивное отношение к дальнейшему развитию ХАИ, а также то, что тема постоянно диверсифицируется в множество смежных областей, включая социально-экономических. Это было ожидаемо, поскольку большинство исследователей в области ИИ имеют опыт работы в области вычислений, математики, технической кибернетики, естественных наук и, возможно, психологии или философии. В то же время, как известно, характер знаний существенно зависит от того, кто является его потребителем, в какой среде оно сформировано и как ведется. Поэтому природа и структура объяснений строится не столько

---

<sup>1</sup> Arrieta A. B., Rodriguez N. D., Ser J. D., Bennetot A., Tabik S., Barbado A., Garcia S., Gil-Lopez S., Molina D., Benjamins R., Chatila R., Herrera F. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges towards responsible AI. *Information Fusion*, 2019. P. 82–115.

<sup>2</sup> Mueller S. T., Veinott E. S., Hoffman R.R., Klein G., Alam L., Mamun T., and Clancey W. J. Principles of Explanation in Human-AI Systems. Association for the Advancement of Artificial Intelligence. 2020. <https://arxiv.org/ftp/arxiv/papers/2102/2102.04972.pdf>

<sup>3</sup> Byrne R.M.J. Counterfactuals in Explaining Artificial Intelligence (XAI): Evidence from Human Reasoning. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)*. 2019. P. 6276-6282. doi:10.24963/ijcai.2019/876

с учетом вероятности выбора, который делает система ИИ, сколько из внешних, в частности, социально–экономических, причин его порождения, а также ожиданий и переживаний пользователя системой ИИ.

### 3.4.2. Проблемные аспекты объяснения

Общепризнанного определения и взгляда на реализацию ХАИ нет, и, по-видимому, быть не может. Вместе с тем потребность в ХАИ нарастает, что нуждается в развитии методологической поддержки. Имеются разные предложения ученых и инженеров, которые сопровождают его разработку, исходя из того, что объяснительная возможность ИИ:

- зависит не столько от сложности логики и алгоритмов работы системы ИИ, сколько от пользователя и внешнего контекста, включая социум, экономику, пользователя, целенаправленность и др.<sup>1</sup>;
- обеспечивает создание моделей ИИ, которые могут объяснить свои выводы, сохраняя нужную точность прогнозирования, обеспечивая пользователям понимание, доверие и управление новыми объектами<sup>2</sup>;
- должен гарантировать, что решения и любые данные, их обеспечивающие, могут быть объяснены человеком непрофессионалом<sup>3</sup>;
- направлен на послойное открытие «черного ящика» для создания моделей и методов, которые одновременно точны и дают удовлетворительное объяснение<sup>4</sup> и др.

До недавнего времени основное внимание построению компонент объяснения в системах ИИ уделялось естественно-научным, инженерным и логико-технологическим компонентам. В результате выделяется технологических аспектов ХАИ, например: учет внешнего окружения; верифицируемость объяснения на других примерах и моделях. Эти аспекты характеризуют техническую сторону цифровой среды и почти не затрагивают природу субъектности, сознания, эмоций, духовную сферу.

---

<sup>1</sup> *Raikov A.* Cognitive Semantics of Artificial Intelligence: A New Perspective. Topics: Computational Intelligence. – Singapore: Springer. Vol. XVII. 2021. <https://doi.org/10.1007/978-981-33-6750-0>

<sup>2</sup> *Chen M. et al.* Simple and deep graph convolutional networks // International Conference on Machine Learning. 2020. P. 1725–1735.

<sup>3</sup> *Wang J. et al.* Learning node representations from noisy graph structures // Proc. IEEE Int. Conf. Data Mining, ICDM. 2020. Vol. 2020-Novem, No 1. P. 1310–1315.

<sup>4</sup> *Veličković P. et al.* Neural Execution of Graph Algorithms. 2019.

Подходы к построению объяснений могут быть релевантными (логический морфизм), формализованными, а также смысловыми, неформализованными. Первые два лежат в логической плоскости, когда само объяснение можно логически проверить, например, послойно анализируя процесс вывода решения построить логические причинно-следственные цепочки получения вывода. На систему объяснения в этом случае, как известно, принципиальным образом влияют данные, которые используются для обучения и адаптации системы ИИ<sup>1,2,3</sup>.

Следует заметить, что ответ на вопрос о самой потребности в объяснении не является однозначным. Взгляды на объяснимость иногда сопровождаются скептицизмом в отношении их полезности. Ставятся вопросы о том, может ли высокая прозрачность работы системы ИИ привести к информационной перегрузке, может ли визуализация привести к чрезмерному доверию или неправильному чтению, могут ли системы машинного обучения на самом деле предоставлять обоснования на естественном языке, и что разные люди в любом случае нуждаются в разных объяснениях<sup>4</sup>.

Ставятся вопросы связанные с определением роли объяснения при оценке ответственности в законодательстве и судебной практике, отмечается потребность достижения компромисса между полезностью и стоимостью объяснений, сложностью и вычислительными ресурсами. Ведь любая уточняющая информация в виде потенциально объяснения может быть представлена в виде набора абстрактных причин или оправданий некоего результата, а не описание процесса принятия решения в целом<sup>5</sup>. Для них генерация объяснений – это вопрос дизайна системы, и они предлагают рассматривать системы объяснений отдельно от систем ИИ, чтобы создать возможности для отраслей, специализирующихся на системах объяснений в терминах,

---

<sup>1</sup> *Lin, Y.-T., Hung, T.-W., Huang, L.T.-L.* Engineering Equity: How AI Can Help Reduce the Harm of Implicit Bias // *Philosophy and Technology*. 2021. Vol. 34, P. 65–90.

<sup>2</sup> *Leavy, S., Meaney, G., Wade, K., Greene, D.* Mitigating gender bias in machine learning data sets. *Communications in Computer and Information Science*, 1245 CCIS, 2020. P. 12-26. doi: 10.1007/978-3-030-52485-2\_2

<sup>3</sup> *Leavy, S., Siaper, E., O'Sullivan, B.* Ethical Data Curation for AI: An Approach based on Feminist Epistemology and Critical Theories of Race // *AIES 2021. Proceedings of the 2021. AAAI/ACM Conference on AI, Ethics, and Society*. P. 695–703.

<sup>4</sup> *Heaven W. D.* Why asking an AI to explain itself can make things worse. *MIT Technology Review*. 2020. January 29. <https://www.technologyreview.com/2020/01/29/304857/why-asking-an-ai-to-explainitself-can-make-things-worse/>

<sup>5</sup> *Doshi-Velez F., Kortz M., Budish R., Bavitz C. Gershman S., O'Brien D., Shieber S., Waldo J., Weinberger D., Weller A., Wood A.* Accountability of AI Under the Law: The Role of Explanation. Berkman Klein Center Working Group on Explanation and the Law, Berkman Klein Center for Internet & Society Working Paper. 2017 <http://nrs.harvard.edu/urn-3:HUL.InstRepos:34372584>.

интерпретируемых человеком, без ущерба для точности. исходного предиктора.

### 3.4.3. Причинность в объяснении

Причинность, по-видимому, является наиболее фундаментальным явлением, отражающим всеобщую связь и единство во Вселенной. Оно связывает мысли с действиями, а действия с последствиями, движение планет с образующими их атомами, успех лечения людей от используемой методики и др. На частый вопрос «Почему?» ответить помогают причинно–следственные связи событий. Научные законы и закономерности описывают подобные связи.

Причинность характеризуется совместным появлением событий. Однако, по частоте совместного появления событий далеко не всегда можно судить о наличии между ними причинно–следственной связи. Например, традиционный ИИ не может точно преобразовать корреляцию событий в такую связь. Однако именно корреляция и частотность событий, встречающихся в больших данных, являются основой выявления явных и неявных причинно–следственных связей событий.

Наиболее распространенное мнение, что в основе причинно–следственной связи лежит регулярность: одно событие постоянно связано с другим. Классическая и наиболее распространенная точка зрения идет от Дэвида Юма. Он считал, что если события постоянно появляются вместе и одно происходит до другого, то этого все же недостаточно для того, чтобы констатировать причинность. События должны быть также смежными, быть рядом друг с другом.

Вместе с тем понятие смежности нельзя не подвергнуть сомнению. Например, из фундаментальной физики известен эффект квантовой запутанности: две частицы, находящиеся в разных концах Вселенной, могут быть связаны таким образом, что изменение квантового состояния одной из них приводит к мгновенному изменению состояния другой<sup>1,2</sup>. Это, однако, происходит в нарушение законов теории относительности – причинно-следственная связь движется быстрее скорости света. То есть появился некий научный парадокс, оспаривающий тезис смежности по Юму.

В работе<sup>3</sup> вопросы причинности классифицируются по нескольким направлениям, определяемых такими основаниями, как: плюрализм,

---

<sup>1</sup> *Einstein A., Podolsky B., and Rosen N. Can quantum-mechanical description of physical reality be considered complete? // Phys. Rev., 1935. Vol. 47, P. 777–780.*

<sup>2</sup> *The BIG Bell Test Collaboration, Challenging local realism with human choices, Nature, 2018. Vol. 557, P. 212220. doi: 10.1038/s41586-018-0085-3*

<sup>3</sup> *Mumford S., Anjium R. L. Causation. A very short introduction. – Oxford: University press. UK.*

примитивизм, физикализм и др. Признается, что переход к плюрализму является просто признанием поражения всех иных, более особенных, направлений. Например, плюралист использует различные теории и под причинностью подразумевает множество разных вещей или событий.

Тема объяснимости выводов ИИ явно выходит за рамки вопроса каузальности, охватывает субъективную реальность в саморазвивающихся междисциплинарных полисубъектных (рефлексивно-активных) средах<sup>1</sup>. Бытие субъектов в таких средах может быть задано системой онтологий, которая обеспечивает сборку в целое включенных в среду субъектов. В работе<sup>2</sup> разработана и опробована система онтологий, в которую входят онтологии обеспечения жизнедеятельности, преодоления точек разрыва, стратегического целеполагания, разработки стратегий и проектов, внедрения и инновационного обеспечения стратегий и проектов. Такая система может быть использована для репрезентации причинности.

Таким образом, в контексте создания ХАИ можно думать о причинности как об одной из фундаментальных сил и элементов Вселенной. Причинность – это то, что удерживает объекты вместе посредством атомарных и молекулярных связей. Она производит изменение одной вещи с помощью другой. Это придает любому действию значимость. И тогда с чего бы кому-то вообще думать, что мы можем объяснить причинность в некаузальных терминах?

Отметим особый взгляд экономических теорий на феномен объяснения<sup>3</sup>. Экономисты считают объяснения, даваемые экономическими теориями и моделями, хорошей формой объяснения. Однако, в предлагаемых этими теориями подходах сама идея объяснения может показаться нелогичной. Ведь экономисты строят модели, которые направлены на абстрактное объяснение экономической ситуации и не всегда могут быть опровергнуты эмпирическими данными. Вместе с тем объяснительный приоритет модельного экономического построения работает совсем не в пользу получения количественных характеристик динамики и прогноза развития ситуации, которая, по сути, в экономике более важна. Существует множество взглядов экономистов на получение ответов на вопросы относительно феномена объяснения: примирение

---

2013.

<sup>1</sup> *Лепский В.Е.* Искусственный интеллект в субъектных парадигмах управления. Филос. науки. 2021. 64(1). С. 88–101.

<sup>2</sup> *Lepskiy V.* Evolution of Cybernetics: Philosophical and Methodological Analysis // *Kybernetes*. 2018. Vol. 47. No. 2. P. 249–261.

<sup>3</sup> *Kaul N.* (2022). 3Es for AI: Economics, Explanation, Epistemology. *Front. Artif. Intell.* 5:833238. doi: 10.3389/frai.2022.833238



противоборствующих сторон; различие и развитие природы телеологических и каузальных объяснений в экономике; использование модели рационального выбора; рассмотрение дедуктивного, индуктивного и абдуктивного рассуждений и др.

#### **3.4.4. Люди как системы**

Очевидно, что люди являются «системами» только в том смысле, что они относятся к сверхсложным системам, выходящим за границы, диктуемые технической кибернетикой. В отличие от естественных наук, где формулы, закономерности, физические законы и пр. могут быть получены через наблюдение, измерены и воспроизведены разными способами, в субъектной рефлексивно-активной среде, человеческое поведение внутри и внутри коллективов, и между коллективами не поддается прямому наблюдению, измерению и формализованному описанию, и, как следствие, однозначному объяснению. Поэтому для объяснения явлений привлекаются упомянуты выше онтологии.

В менеджменте, стратегическом планировании, где решение проблем осуществляется с учетом идентификации сотен факторов, объяснению проблемной ситуации служат методы стратегического анализа и когнитивного моделирования<sup>1</sup>. Так же, как и в подходе с онтологиями, решаемая проблема коллективно декомпозируется на конечное число целей, факторов, препятствий, ограничений, функций, задач и пр. Результат моделирования не всегда поддается объяснению, однако сам аналитический процесс, состоящий из множества шагов, делается под непосредственным контролем реализующей его команды и должен вызывать доверие.

Таким образом, объяснительная функция систем ИИ приобретает все более важное значение в дальнейшем развитии и применении таких систем, что вызвано их усложнением, и, как следствие, снижением доверия к генерируемым ими рекомендациям.

Решающее значение в повышении эффективности систем ИИ, помимо технологической составляющей, приобретает фактор учета субъективной реальности, погружения системы ИИ в полисубъектную рефлексивно-активную среду.

Росту доверия к выводам систем ИИ способствует непосредственное включение людей в процесс получения результата.

---

<sup>1</sup> Raikov A. Automating Cognitive Modelling Considering Non-Formalisable Semantics. In: Nagar A.K., Jat D.S., Marín-Raventós G., Mishra D.K. (eds) Intelligent Sustainable Systems. Lecture Notes in Networks and Systems, .2022. Vol 334. Springer, Singapore. [https://doi.org/10.1007/978-981-16-6369-7\\_2](https://doi.org/10.1007/978-981-16-6369-7_2)

Для такого включения используются специальные подходы, основанные на построении онтологий и структуризации проблем для создания условий их целенаправленного разрешения.

### **3.4.5. Выводы**

На основе изложенного в этом разделе, можно выделить следующие основания для построения искомым социально гуманитарных критериев:

- В исследованиях делается акцент на необходимости при создании системы объяснения выводов ИИ ориентации на человека, который имеет ожидания.
- Работы в области психологии, посвященные познанию и предубеждениям, использует данные человеческого мышления, чтобы аргументировать вклад интерпретаций в системы ИИ.
- Тема объяснимости выводов ИИ выходит за рамки вопроса каузальности, и охватывает субъективную реальность в саморазвивающихся междисциплинарных полисубъектных (рефлексивно-активных) средах.

## **3.5. ЭТИЧЕСКИЕ ДИЛЕММЫ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

### **3.5.1. Введение**

В настоящее время замещение человека на системы ИИ происходит повсеместно. В основном современные АWІ (Artificial Weak Intelligence, Слабый ИИ) оперируют в рамках заданных разработчиками контуров знаний и целеполагания<sup>1</sup>. Соответственно, постановка вопроса об этике ИИ на данном уровне его развития является сугубо декларативным, хотя и полезным упражнением. Реальный вопрос о правилах поведения АSІ возникает в том случае, когда АWІ смогут автономно менять контур используемых знаний, обучаться на неограниченных источниках и типах информации, корректировать или даже устанавливать свои цели. Вторая принципиальная проблема связана с этической ситуацией в «обычном» человеческом обществе, где формально воцарился этический релятивизм, в реальности означающий завуалированное доминирование вполне определенных этических моделей. Это создает предпосылки неизбежной трансляции данной ситуации в пространство разработки и применения АWІ.

### **3.5.2. Подлинно осознающие субъекты и системы ИИ (этические аспекты)**

При некоторых условиях к АWІ можно применить определение подлинно осознающих субъектов (ПОС), способных выбирать цели активности и произвольно работать с разными базами знаний. Отчасти это отражает зависимость АWІ от целеполагания разработчиков, которые могут заложить в АWІ явно или скрыто преступные по своим последствиям решения. Но в большей мере появление ПОС среди АWІ зависит от научно-технического прогресса и его распространения в повседневности.

Под ПОС понимаются живые или квазиживые существа, которые обладают самосознанием и субъективными переживаниями, схожими с самосознанием и субъективными переживаниями человека или другого высокоразвитого существа. К таким объектам могут быть, вероятно, отнесены СИИ в его сильной версии<sup>2</sup>. Строго говоря, к ПОС следует отнести и индивидов, имеющих вживленные те или иные «умные»

---

<sup>1</sup> Кукшев В.И. Классификация систем искусственного интеллекта // Экономические стратегии. 2020. № 5. С. 51–57; Кукшев В.И. Цифровая экономика: проблемы и решения // Экономические стратегии. 2020. № 6. С. 58–67.

<sup>2</sup> См.: Социально-экономические аспекты внедрения искусственного интеллекта / Под науч. ред. А.И. Агеева. М.: Айти-Сервис, 2020.

устройства-протезы (от кардиостимуляторов до протезов опорно-двигательного аппарата или отдельных органов чувств). Интеллектуальная составляющая таких устройств быстро развивается. Количество лиц, имеющих подобные устройства, растет и измеряется в мире десятками миллионов<sup>1</sup>. Легко увидеть границу, где работа этих устройств сталкивается с проблемой морального выбора. Кроме того, будучи «существом общественным», человек в современном мире пребывает в некотором количестве живых и виртуально организованных сетей, которые способны влиять на его повседневный, предметно-ситуативный и даже мировоззренческий выбор. Это означает, что на него оказывают влияние надличностные виртуальные системы, а также собираемые вне зависимости от воли индивида «большие пользовательские данные».

Морально-этические проблемы обусловлены не только тем, насколько произвольно и (или) вредоносно для личности и человеческого общества может повести себя такая система, но и тем, что само общество может причинить страдания созданным им искусственным сущностям, когда они достигают уровня ПОС.

Аналогичная моральная дилемма возникала при проведении научно-практических экспериментов на людях и на животных. На данный момент она решена следующим образом: эксперименты на людях запрещены без их согласия, животные считаются менее развитыми, «не имеют души» и эксперименты на них разрешены. Однако в последнее время существует тенденция приравнивания некоторых видов высокоразвитых животных к людям, с присвоением им морально-этических и юридических прав. Прецедент произошел в США: за двумя приматами, участвующими в медико-биологических исследованиях университета штата Нью-Йорк в Стоуни-Брук, признали некоторые права человека<sup>2</sup>.

При создании искусственных существ и квазисуществ не всегда возможно предсказать результат эксперимента и насколько осознающее существо в итоге возникнет, учитывая потенциал обучения такой системы. Кроме того, до создания существа у него невозможно спросить согласие на проведение эксперимента. Эта проблема также решаема: при рождении человека не требуется его согласие на появление на свет, но он сразу после рождения, а частично и до, получает определенные права и гарантии, предусмотренные законом и обычаями. Аналогичным

---

<sup>1</sup> Cadwell et al, for monitoring everyday prosthesis use a systematic review/ Journal of neuroEngineering and Rehabilitation. 2020.

<sup>2</sup> URL: <https://www.buzzfeednews.com/article/mbvd/us-court-recognizes-chimpanzees-as-legal-persons>.

образом можно обязать создателей искусственных существ, включая АWІ, упреждающе минимизировать их «страдания» при проведении экспериментов, также содержать эти существа в надлежащих условиях. В противном случае, когда невозможно спрогнозировать последствия, подобные эксперименты должны быть запрещены<sup>1</sup>.

Еще более сложный вопрос – способны ли интеллектуальные системы в будущем, решая какие-то задачи по жизненно важным социальным проблемам, брать на себя риски принятия решений. Готово ли общество взять на себя ответственность за деятельность такой технической системы? В силу сложности решаемых АWІ задач, невозможности полного контроля над машиной, отслеживания и проверки принятых решений человек уже вынужден перекладывать часть ответственности на машину (программу).

При разработке АWІ имеется некоторая степень неопределенности, которая сходна со свободой воли человека. Но при этом в разрабатываемые АWІ должны быть заложены некоторые принципы:

- этические самоограничения и знания о морально-этических нормах людей, имитация процесса саморегуляции своего поведения, способность к эмпатии;
- механизм прогнозирования рисков и последствий собственных действий, ограничения действий при наступлении определенных рисков;
- возможность осознания и исправления собственной ошибки<sup>2</sup>.

Дальнейшие исследования в области этики ИИ, несомненно, приведут к созданию различных стандартов и сертификации правил проектирования и функционирования АWІ. При создании таких стандартов необходимо оценить их влияние на дальнейшую разработку АWІ, исключив риски как вредоносного развития, так и торможения развития интеллектуальных систем<sup>3, 4</sup>.

### **3.5.3. Типы массовых сознаний и вызов этического релятивизма**

В XX в. накоплен беспрецедентный опыт манипулирования сознанием, целенаправленного формирования массовых сознаний, заданных тем или иным инициатором, как правило – государством или с опорой на государственные институты, типов личности и параметров ее

---

<sup>1</sup> Шуравин А. История Искусственного Интеллекта [Электронный ресурс] // <https://wiki.programstore.ru/istoriya-iskusstvennogo-intellekta/>

<sup>2</sup> Социально-экономические аспекты внедрения искусственного интеллекта / Под науч. ред. А.И. Агеева. М.: Айти-Сервис, 2020.

<sup>3</sup> <https://qz.com/1577451/century-tech-signs-deal-to-put-ai-in-700-classrooms-in-belgium/>

<sup>4</sup> Разин А.А. Этика искусственного интеллекта // *Философия и общество*. 2019. Вып. 1 (90). С. 57–73. URL: <https://doi.org/10.30884/jfio/2019.01.04>.

допустимого («этичного», социально-приемлемого) поведения. Этот опыт во многом дискредитирован в последние десятилетия, породив кризис основных этических концепций, сложившихся на протяжении последних тысячелетий человеческой истории и сформировавших доминанту этического релятивизма, нарастания мультикультурности наряду с обострением конфликтности идентичностей. Однако дискредитация данного опыта не означает ни его исчезновения, ни утраты им потенциала проявления социальной энергетике в будущем. Речь идет о некотором накопленном репертуаре социального бытия, в котором в зависимости от комбинации организующих и самоорганизованных процессов возникает совокупность мотивирующих и стимулирующих слоев культуры для социализации, адаптации, субъективации человека и его общностей по разным основаниям, отражающим биологическое и социальное воспроизводство и взаимодействие с внешним миром. Вне всякого сомнения, разработка и внедрение новых технологий, включая АИ, будет испытывать влияние со стороны этического состояния общества. Есть также все основания полагать, что достижения науки и техники будут в нарастающей степени оказывать влияние на коллективные и индивидуальные сознания.

Онтологическая карта социально-ценностных ориентаций («картин мира», «я (мы)-концепций») может быть сформирована по многим основаниям, во многом схожих с традиционным членением философских школ (версии идеализма и материализма, гносеологические и экзистенциальные концепции и т.д.). Среди этих критериев, например, идеалистичность vs материалистичность, субъектность vs объектность, всеобщность-частичность, персонализация-обезличенность и др.

В рассматриваемом ракурсе представляется содержательным и полезным подход к формированию онтологической карты с точки зрения обобщенных типов личности (ОТЛ) с учетом исторического опыта XX в., значимого для социальных практик в современный период. Эти вопросы отчасти осмысливаются в технократическом ракурсе как формирование обществ Индустрии 4.0, в социотехническом – как «инклюзивный капитализм», в более комплексном – как создание Общества 5.0.

В любом случае необозримое многообразие жизненно важной проблематики на фоне усугубляющегося информационного потока и расширяющегося применения АИ для обработки информационных массивов и потоков (от устройства Вселенной и ее частиц, дорожного трафика, нормативной базы до исследований человеческого мозга) провоцирует все больше людей к обращению ради определенности со своей идентификацией и выбором линии поведения и жизни в целом в

этом спутанном мире к различным техникам упорядочения и упрощения этой пестрой и сложной реальности, включая делегирование множества традиционно сугубо человеческих задач надличностным киберсистемам.

В условиях «мультикультурализма» теоретически и де-факто неизбежен этический релятивизм, проявляющийся как существование определенного, кажущегося неограниченным «меню» для выбора личной идентичности. Однако это отнюдь не означает равноположенность каждой строчки этого меню. Среди всего предлагаемого или допустимого для выработки «с нуля» (подразумевается и такая опция для личностного творчества, почти абсолютно манипулятивная) набора этических концептов есть как тяжелые, так и весьма легкие, ситуативные, конъюнктурные, обычно малоосознаваемые, но удобные фракции. «Тяжелые фракции» опираются на концептуальные, теоретические и (или) религиозные решения глубинных вопросов бытия человека в мире и социуме, в частности. У них, как правило, мощная система воспроизводства, институционализированная социальной структуре, в управлении, науке, образовании, закреплённая в архитектуре и искусстве. Их гравитационное притяжение очевидно выше тех или иных новомодных попыток создать «новую этику» и как ее политическую проекцию – идеологию. Но «тяжелые фракции» этических систем и конструктов, представляя собой аттрактор традиционности, всегда испытывают сложности с адаптацией к новым вызовам эволюции, формируемым прогрессом науки, техники, образа жизни, политической и экономической борьбой.

Технологические перемены и активная разработка проектов вызвали к жизни пестрое ассорти концептов «нового человека» («цифролюди», «ГМО-люди», «трансгуманизм», «служебные люди», «сингуляры», «новые европейцы», «новые кочевники» и т.п.).

Набирает силу и ансамбль идей, питаемых концептами ноосферы, космизма, «лучистого человечества», социально-духовного интегрализма. За ним – содержательная научная основа, восходящая к В.И. Вернадскому и Т. де Шардену, фантастам XX в. (Г. Уэллс, И. Ефремов, С. Лемм, братья Стругацкие), философам и ученым (Н. Федоров, К. Циолковский), но важнее – за этим практические триумфы человеческого развития середины XX в. – освоение земного, околоземного и космического пространства, глубин и просторов океана, материи, человеческой биологии и психики и т.п. Однако этот комплекс представлений при всей их научной обоснованности сложен, опирается на плохо известные в обществе новые научные парадигмы. До манифестации как новой массовой идеологии и тем более ОТЛ ему пока

далеко. Большую популярность имеют «географические», «геополитические», «геоэкономические», «прецедентные» версии идеологий, поднимающие на флаг механическую близость проживания и судьбы тех или иных народов и государств, какие-либо отдельные эпизоды истории и обслуживающие довольно прагматические интересы.

Все это разнообразие ОТЛ влияет на научно-технические решения в сфере цифровизации. Обобщенно, по мере достижения АWИ уровня развития, на котором возникает потребность в определенных этических ограничениях целеполагания, используемых знаний и поведения, приближения АWИ к уровню ПОС, будет быстро формироваться и регламентироваться «портфель» этических принципов и практик для разрабатываемых АWИ. В этом процессе на искусственные сущности неизбежно будут проецироваться ансамбли (наборы) этических шаблонов (матриц), сложившиеся на национальном и цивилизационном уровнях. При всей всеобщности технико-технологических решений именно цифровые технологии создали принципиальную возможность не только имитации реальных индивидов и процессов в виде их цифровых двойников, но и генерирования целенаправленных стратегий воздействия на кластеры индивидов, используя потенциал выявления, наблюдения и использования их психосемантических, биофизических и иных свойств. Иными словами, возможная в будущем «битва машин» в любом случае будет «битвой» вполне человеческих и человекообразных этик, инкрустированных в АWИ и растущую совокупность ПОС.

#### **3.5.4. Цифровой социум: этика и доверие к системам искусственного интеллекта**

Следует поставить и вопрос о моменте пересечения социумом определенного технологического порога, когда может произойти безвозвратная трансформация как общества в целом, так и человека в то, что условно можно называть «цифровой социум». Эта точка бифуркации связана, судя по всему, с формированием платформы цифровой стратификации как доминанты социального структурирования и управления. АWИ будут играть ключевую роль в этой платформе и формировании соответствующих социальных страт, исходя из заданных при разработке критериев. Необходимые условия для этого включают прежде всего достижение технологически возможной, практически полной осведомленности о критических параметрах жизни социума и всего множества индивидов.

Опыт использования цифровых технологий в условиях пандемии 2020–2021 гг. показал как высокий потенциал АWИ для социального управления, так и множество их программно-технических и социально-



психологических уязвимостей. Об этом же свидетельствует и опыт формирования всевозможных цифровых платформ (от госуслуг до маркетплейсов). Тем не менее цифровизация охватывает все больше сфер жизни социума. Просматривается перспектива всеобъемлющей интеграции созданных систем наблюдения, баз данных, центров их обработки, систем поддержки принятия решений. Возможности манипулирования личностным выбором через цифровые персонализированные модели и их кластеризацию обозначены выше.

Логика развития АWИ показывает, что они во все большей степени реализуют принцип самодостраивания не только в программной, но и в аппаратной частях. Например, если пока по входу и выходу информации в нейронных сетях можно восстановить способы ее обработки и соответственно получить человекочитаемые правила и алгоритмы, то в ближайшее время ситуация изменится. На хорошо известной ежегодной конференции Intelligence Advanced Research Projects Activity (IARPA) еще в 2016 г. речь шла о том, что в разведывательном сообществе используются комплексы со столь глубокими нейронными сетями, что перевести их алгоритмы на человекочитаемый язык за разумное время невозможно. Это пока первые симптомы, что появляется мир, где принятие решений будет происходить на основе критериев, закрытых для лиц, принимающих решения (ЛПР).

В 2012 г. международная организация по правам человека поставила перед ООН вопрос о необходимости запрещения автономных боевых систем (роботов), самостоятельно принимающих решения об использовании находящихся у них боевых средств. Несмотря на мощную поддержку со стороны правительств ряда государств, решение ООН, обязательное к исполнению, по состоянию на сегодня так и не принято, накапливается история такого рода инцидентов.

Аналогичная проблема возникает и применительно к роботизированным транспортным средствам. Так или иначе, все роботизированные машины будут управляться системой ИИ той или иной степени продвинутой. Автопроизводители подходят к этой проблеме по-разному. Например, *Mercedes* в 2016 г. опубликовал заявление, что в случае соблюдения дорожных правил долг компании защищать пассажиров, а не пешеходов. В любой ситуации, если роботизированный *Mercedes* решит, что он соблюдает правила, выбор будет сделан в пользу пассажиров, а не пешеходов. Принципиальную позицию заняла компания *Google*. В новом поколении *Google*-мобилей имеется программный видеофильтр. В том случае, если видеосенсоры автомобиля распознают на дороге ребенка, вне зависимости от того, нарушает он правила или нет, автомобиль в

качестве приоритета первого порядка будет выбирать сохранение жизни ребенка<sup>1</sup>.

Второе условие перехода к «цифровому социуму» определяется реальными интересами самого современного социума. Прежде всего он социально неоднороден практически повсеместно, что предопределяет различия не только в доступе к цифровым технологиям, но и к их разработке, внедрению и использованию. В этой сфере уже развернулась жесткая конкурентная борьба за лидерство и контроль. Однако практическое управление в социуме в целом и на уровне интеграционных, национальных, региональных, отраслевых образований пока детерминировано не цифровыми императивами, а политическими национально-государственными, корпоративными, частными интересами. Они отнюдь не безусловно благоприятствуют быстрому развитию «цифрового социума», между тем как это требует тотального охвата всех субъектов, объектов и процессов в социуме. Достаточно отметить риски массовой безработицы, желаемый уровень кибербезопасности, чтобы с осторожностью относиться к самым оптимистичным темпам и этапам формирования «цифрового социума».

Развилка в эволюции киберфизических систем к статусу глобальной осведомленности и успешного управления эволюцией социума просматривается в уровне централизации-децентрализации функции глобальной осведомленности и сочетания институтов организации и самоорганизации жизнедеятельности, включая экономику. В настоящее время в различных государствах приняты нормативно-правовые акты, ограничивающие степень централизации персональных данных.

Тем не менее разработка технологических решений для нового поколения общественных систем ведется во всех ведущих государствах и корпорациях мира. В США одним из перспективных проектов Defense Advanced Research Projects Agency (DARPA) стало создание «динамической виртуальной среды», в которой будут ликвидированы существующие сегодня барьеры (ведомственные, организационные, информационные, технические) для эффективной и оперативной совместной работы представителей различных министерств и ведомств, занятых вопросами разрешения кризисных ситуаций в различных сферах деятельности (политической, военной, экономической, социальной). Разрабатываемые инструменты принятия решений для территориально-

---

<sup>1</sup> Доклад заместителя директора ФБР, руководителя CJIS Стивена Л. Морриса «Искусственный интеллект: ФБР и полиция против преступников» [Электронный ресурс] // // <https://ordrf.ru/wp-content/uploads/2017/10/Обзор-отдельных-вопросов-использования-больших-данных-и-искусственного-интеллекта.pdf#page=10&zoom=100,72,537>

распределенных групп интересов призваны обеспечить максимально полное осмысление сложных ситуаций и сценариев их динамики, выбор оптимальных решений на основе всей доступной информации без ее полного изучения по принципу «знать, не читая». В основу технологии положены методы нечеткого структурирования аргументов, трехмерная цветная визуализация и корпоративная память.

Развитие АWИ и цифровая трансформация в целом ведут к образованию коллективов автономных агентов искусственной и смешанной генеалогии, а также сложных конструкций информационной и нормативной среды со множеством возможностей и патологий и растущим уровнем неопределенности для принятия управленческих решений.

При этом наблюдается усиливающийся феномен «деградации естественного интеллекта»<sup>1</sup>. В частности, постоянное просматривание сайтов может вести к эрозии способностей к системному и углубленному мышлению. Исследования медиков убедительно доказали, что у тех, кто проводит в Интернете много времени, быстро развиваются две области головного мозга – часть, отвечающая за кратковременную память, и центр, ответственный за принятие быстрых решений. В то же время те зоны мозга, которые заведуют детальным анализом, глубоким продумыванием проблемы остаются без нагрузки и постепенно утрачивают навык к интенсивной работе<sup>2</sup>.

Среди наиболее острых вопросов развития АWИ сегодня – проблема доверия (*trustworthiness*), которая охватывает проблематику уверенности потребителей, регуляторов и иных заинтересованных сторон в том, что СИИ способна выполнять возложенные на нее задачи с требуемым качеством и уровнем безопасности.

В соответствии с прогнозами в 2020–2030 гг. в мире социальных структур произойдет синдикация основных социальных платформ и переопределение стандартов человеческого взаимодействия. Налаживается взаимная трансляция нейроописаний, социальных описаний и описаний семантики человеческого интернета и интернета вещей. «Коды» нервных систем и мозга во многом будут описаны и использованы не только в медицине, но и для моделирования аналогичных процессов в других субстратах – экономических и социальных системах, самоорганизации «умных вещей» и искусственных систем. К 2030 г. семантики разных типов будут способны переводиться друг в друга, и

---

<sup>1</sup> Агеев А.И., Асанова Е.А., Глибенко О.В. и др. К «цифре» готов? Оценка адаптивности высокотехнологичного комплекса России к реалиям цифровой экономики / Под ред. д.э.н., проф. А.И. Агеева. М.: ИНЭС, 2018. 60 с.

<sup>2</sup> Интернет-серфинг меняет мозг пользователя [Электронный ресурс] // SecurityLab. 2010. 30 августа. URL: [https:// www.securitylab.ru/news/397247.php](https://www.securitylab.ru/news/397247.php).

это будет использоваться в экспериментальных сеттингах. Режимы работы психики будут описаны достаточно полно, включая состояния сознания в привязке к различным типам деятельности. Структура сознаваемого у человека может быть легко перестроена в зависимости от стоящих перед ним задач. Интерфейсы Нейронет абсолютно незаметны, прозрачны. Человек не работает за клавиатурой и экраном – он работает непосредственно с данными, со смыслами, с людьми.

Предмет управления на данном этапе – человеческий организм, представленный большим количеством данных от датчиков разного типа. Это и коллективы, частью которых человек в настоящий момент является. Также взаимодействие идет с распределенными системами умных вещей, которые постоянно себя перепроектируют. Одним из таких предметов управления является жизнь человека, образование на всем жизненном цикле, а также жизненные циклы сообществ<sup>1</sup>.

### **3.5.5. Выводы**

Вопрос об этике ИИ на данном уровне его развития является сугубо философским и футурологическим. Реальная проблема «этики» работы (поведения) АWИ возникает лишь в том случае, когда АWИ смогут автономно менять контур используемых знаний, обучаться на неограниченных источниках и типах информации, корректировать или даже устанавливать свои цели. Эта перспектива просматривается лишь в условиях появления и экспансии «сильного ИИ». Тем не менее уже существует значительное число существ, в том числе обычных людей, оснащенных различными «умными» устройствами, поведение которых может создавать этические проблемы. Это затрагивает и искусственные системы, подчиняющиеся логике поведения «роя», влияющего на индивида.

Вторая принципиальная проблема связана с этической ситуацией в «обычном» человеческом обществе, где формально воцарился этический релятивизм. Это создает предпосылки неизбежной трансляции данной ситуации в пространство разработки и применения АWИ. В результате с большой вероятностью «мир машин» воспроизведет мир людей, переведя существующие этические проблемы в киберфизический мир.

---

<sup>1</sup> Подходы к формированию и запуску новых отраслей промышленности в контексте Национальной технологической инициативы на примере сферы «Технологии и системы цифровой реальности и перспективные „человеко-компьютерные“ интерфейсы (в части нейроэлектроники)»: Аналитический доклад [Электронный ресурс]. URL: <http://rusneuro.net/cambiodocs/media/files/analitjeskii-doklad-podhodyk-formirovaniu-i-zapusku-novyh-otraslei-promyhlennosti.pdf>.

## 3.6 ЭТИЧЕСКИЕ АСПЕКТЫ РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В КИТАЕ

### 3.6.1. Введение

Приоритеты Китайской Народной республики (КНР, Китая) на ближайшие годы (до 2025 года) включают «достижение независимости от иностранных технологий» и «повышение конкурентоспособности» и одним из приоритетных направлений для достижения данных целей является ИИ<sup>1</sup>. Китай, в частности, решает не задачу «импортозамещения, как в России, а задачу «научно-технического превосходства. Была озвучена инициатива по созданию «Цифрового Китая»<sup>2</sup>, представляющего собой экосистему макроуровня, которая регулирует и регламентирует цифровые технологии, включая использование big data. Согласно отчету<sup>3</sup> Китай имеет более 1000 программных документов по ИИ, что в разы больше официальных документов по ИИ в других странах.

В КНР разработан «План развития искусственного интеллекта нового поколения»<sup>4</sup>, в котором учитывается важность ИИ как «новое направление международной конкуренции», поскольку является стратегической технологией для повышения национальной конкурентоспособности и защиты национальной безопасности. Искусственный интеллект видится и как конкурентное преимущество, и как проактивная технология и как механизм безопасности. ИИ – основа будущих глобальных промышленных преобразований в стране и изменений экономической, социальной деятельности. Утверждается, что технологии ИИ способны «своевременно распознавать групповые когнитивные и психологические изменения и проявлять инициативу в принятии решений и реагировании» что усиливает возможности социального управления. Таким образом, можно сделать вывод о том, что уже минимум пять лет в Китае рассматривают проблематику ИИ не

---

<sup>1</sup> State Council of the People's Republic of China. The 14th Five-Year Plan (2021-2025) for National Economic and Social Development and the Long-Range Objectives Through the Year 2035. Retrieved April 22, 2021, from [http://www.gov.cn/xinwen/2021-03/05/content\\_5590537.htm](http://www.gov.cn/xinwen/2021-03/05/content_5590537.htm)

<sup>2</sup> Iris Deng and Xinmei Shen «China's 'two sessions' 2021: plans for 'Digital China' transformation come with increased regulations for personal data» Published: South China Morning Post, 12:00pm, 6 Mar, 2021. <https://www.scmp.com/tech/big-tech/article/3124305/chinas-two-sessions-2021-plans-digital-china-transformation-come>

<sup>3</sup> Китайский институт научно-технической политики при Университете Цинхуа. China AI Development Report. 2018. 144 стр. <https://www.sppm.tsinghua.edu.cn/english/info/1032/1601.htm>

<sup>4</sup> Государственный совет КНР. План развития искусственного интеллекта нового поколения. 2017. [http://www.gov.cn/zhengce/content/2017-07/20/content\\_5211996.htm](http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm)

только исключительно в технологических аспектах, включая когнитивное обучение на основе больших данных, совместную обработку данных в разных средах и взаимодействие человека и машины, но и в социогуманитарных, рассматривая ИИ как важную часть социальной сферы. Рассмотрим этические подходы к разработке и применению ИИ, официально опубликованные в Китае, чтобы понять, насколько этические нормы будут влиять (влиают?) на развитие данных технологий, и соответствуют ли они постнеклассическим представлениям научной рациональности.

### **3.6.2. Этические аспекты развития искусственного интеллекта в Китае**

Согласно «Этическому кодексу нового поколения искусственного интеллекта»<sup>1</sup> этические нормы при разработке ИИ необходимы с целью избегания предрассудков, дискриминации, утечек личной тайны и персональных данных. Ключевые этические «столпы»: повышение благополучия всего человечества, содействие справедливости и беспристрастности, защиты безопасности личной информации и частной жизни, обеспечение управляемости и надежности технологии, усиление ответственности, подотчетности и совершенствование морального облика и повышения этической грамотности. В документе утверждается, что человека нужно ставить «на первое место», «следовать общечеловеческим ценностям», «уважать права человека», соблюдать «национальную и региональную этику». Отметим, что прописываются и принципы законности при обработке личной информации, защита частной жизни и безопасности личных данных. Таким образом, нельзя делать вывод о том, что вопросы частных и индивидуальных интересов в КНР считаются менее важными и ценными, чем общественное благо – в официальных документах данной проблематике уделяется достаточно много внимания, вопреки распространённым стереотипам. Присутствуют формулировки с «национальной спецификой» – «сообщество единой судьбы», «поощрение социальной справедливости», в разделе «ответственности» отметим «создание механизма подотчетности ИИ». Этические нормы ИИ прописаны для норм управления, НИОКР, поставок и использования технологии. Это важный стратегический и проработанный шаг – вводится в как можно больше сред для включения в них этических практик. Прописываются

---

<sup>1</sup> Министерство науки и технологий КНР. Этический кодекс нового поколения искусственного интеллекта. 2021. [http://www.most.gov.cn/kjbgz/202109/t20210926\\_177063.html](http://www.most.gov.cn/kjbgz/202109/t20210926_177063.html)

базовые нормы для управления развитием как технологии ИИ, так и общества, инновационной и управленческой среды, этика закладывается как техническая характеристика НИОКР в сфере ИИ (самопроверка, самодисциплина, исключение разработок, нарушающих этические нормы, проверка данных, соблюдение законов и стандартов ИИ, применение алгоритмов прозрачности и безопасности и пр.) и как техническая характеристика для рыночной среды развития технологий ИИ. Разумеется, свод таких рекомендаций – это не гарантия жесткого их соблюдения и включения в практику исследования и жизнедеятельности; тем не менее в данном издании Кодекса с 2019 года увеличилось количество таких категорий как «справедливость», «устойчивость», «прозрачность», «конфиденциальность», «безопасность», «подотчетность»<sup>1</sup>. Однако, такие категории как «сотрудничество», «совместно» и «для человека» наоборот, сократились. Стоит обратить внимание и на формулировку о «гармонии и дружелюбии между людьми и машинами», что, фактически, уравнивает субъектности человека и ИИ. Есть мнение, что это связано с вкладом в национальную культуру буддийских верований, согласно которым всё, живое и неживое, обладает потенциалом стать Буддой, если его культивировать для достижения просветления<sup>2</sup>. Можно выдвинуть гипотезу о техно-анимистических практиках, которые наблюдаются в странах Азии, где предполагается гармония и синтез технологий, человечества и религии как одно целого, как взаимодействие материального и духовного. Поэтому неодушевленные технологии, которые имитируют людей и животных, весьма популярны и весьма просто распространяются, в частности в КНР. Так, например, в Комитет по этике ИИ Китайской ассоциации ИИ (СААИ) входит профессор Сяопин Чен, известный тем, что руководил созданием реалистичной женской гуманоидной «богини-робота» по имени Цзя Цзя<sup>3</sup>. А в 2015 году в Китае был представлен робот-монах «Xian'er» (приблизительный перевод – «простой на вид, но добродетельный»), разработанный для распространения учения буддизма<sup>4</sup>.

В продуктах массовой культуры (кинематограф) робот часто становится объектом любовного увлечения человека, популярным воплощением ИИ в Китае является социальный чат-бот XiaoIce (более

---

<sup>1</sup> Подробнее см. <https://www.linking-ai-principles.org>

<sup>2</sup> Stanford-New America Digichina Project (2019) AI POLICY AND CHINA. Realities of State-Led Development. Special Report No. 1. 2019. P. 41

<sup>3</sup> <https://www.businessinsider.com/china-robot-goddesses-speak-talk-human-humanoid-2016-4>

<sup>4</sup> *Joseph Campbell*. Robot Monk Blends Science and Buddhism at Chinese Temple // Reuters (April 22, 2016), <https://www.reuters.com/article/us-china-religion-robot/robotmonk-blends-science-and-buddhism-at-chinese-temple-idUSKCN0XJ05I>

660 миллионов онлайн-пользователей), которого часто воспринимают как друга и как субъект любовного интереса<sup>1</sup>. Чат-бот воплощен в образе 18-летней девушки-подростка и является не только голосовым помощником, но и пишет стихи и поёт песни, рисует, готовит финансовые отчёты, ведёт телепередачи, создаёт дизайнерскую одежду. Учитывая, что общение ведется с ИИ, люди даже чаще обращаются к помощнику, особенно ближе к полуночи, поскольку ИИ не спит и всегда готов к «беседе». На китайской музыкальной сцене присутствует группа «Мэй Вэй ВИБ», являясь ИИ, образуются виртуальные церкви с аватарами прихожан<sup>2</sup>. Это показывает, что взаимодействие с ИИ в виде человеческого облика, при развлечении и жизни, осуществляется точно также, как и с людьми.

В Принципах гармоничного ИИ (HAIP), сформулированных под руководством И Цзэна, профессора Института автоматизации Китайской академии наук (CAS) прописаны такие концепции как «гуманизация для укрепления взаимодействия между ИИ и людьми», «сочувствие и альтруизм для обеспечения гармоничного общества человека и ИИ», человеческое сочувствие к ИИ (!), соблюдение конфиденциальности по отношению к ИИ, недопустимость предвзятости к ИИ. ИИ не воспринимается исследователями как менее значимый субъект взаимодействий и отношений<sup>3</sup>. Формулируется подход к безопасности ИИ через обеспечение ИИ самосознанием и интеллектом, сопоставимым с человеком для гармонизации отношений.

Для современного Китая характерно гармонично связывать технологическое развитие и совершенствование технологий ИИ не только с передовыми разработками, тенденциями и проблематиками, приходящими «с Запада», но и со своими внутренними традициями, культурой и философией, концепциями, выработанными национальными мыслителями и инженерами. Всё чаще исследователи из КНР обращаются к национальной культуре, поэзии и философии как источнику вдохновения и образцов для новых подходов к изучению ИИ – например, изучение этики агентов «общего назначения» ИИ для взаимодействия с людьми. Предполагается, что они должны соответствовать человеческим эмоциям, этике и моральным ценностям<sup>4</sup>,

---

<sup>1</sup> Gal D. Perspectives and Approaches in AI Ethics: East Asia // Dubber, Markus, Pasquale, Frank, and Das, Sunit, (Eds.) Oxford Handbook of Ethics of Artificial Intelligence, Oxford University Press. June 7, 2019. SSRN: <http://dx.doi.org/10.2139/ssrn.3400816>

<sup>2</sup> <https://www.scmp.com/news/hong-kong/society/article/3186329/hong-kong-gets-its-first-metaverse-churches-complete-avatars>

<sup>3</sup> Подробнее см.: [https://www.baai.ac.cn/portal/article/index/type/center\\_result/id/589.html](https://www.baai.ac.cn/portal/article/index/type/center_result/id/589.html)

<sup>4</sup> Songchun Zhu. Intelligence needs to be driven by «mind», achieving a dynamic balance between «mind» and «principle» // Beijing Institute for General Artificial Intelligence. 2022. Original Mandarin: <https://mp.weixin.qq.com/s/FMraladW0255C-yeU5alfA>



иметь три «взгляда» – мировоззрение, взгляды на жизнь (мораль, этика) и ценности. Данную структуру интерпретации психологической деятельности человека, соответствующей традициям общества, предлагается перекладывать для реконструкции на исследования в области ИИ. Восприятие, опыт и субъективное воображение – данные категории планируется вводить в сферу исследований ИИ, поскольку они существенно влияют на информацию, на субъект информационного воздействия (приводятся примеры каллиграфии, поэзии). Развитие ИИ прямо связывают с исследованиями в области философии, гуманитарных и социальных наук, с гармонизацией подходов различных школ конфуцианства (Школы Чэн-Чжу («Школы принципа») и Школы Лу-Вана («Школы разума»), условно – школы интуиции, чувств и этики с функциями полезности, ценностями, осмысленными разумом).

К сожалению, не все области развития ИИ на данный момент в КНР имеют философское осмысление и рассматриваются с позиций включения этики как обязательной части разработок. Например, специфика ИИ и военных разработок. Военными собираются данные, касающиеся геномов человека<sup>1</sup>, анализируются с помощью ИИ, что может привести к различным сценариям развития данной технологии. Технологии ИИ применяются для ведения информационной войны, навигации и распознавания целей в автономных транспортных средствах, военных играх, интеллектуальной поддержки принятия управленческих решений<sup>2</sup>.

Официально публикуются документы<sup>3</sup>, призывающие активизировать усилия по регулированию таких сфер военного применения ИИ с целью предупреждения потенциальных рисков и управления ими. Это может трактоваться как политический маневр, но, может быть, это является сигналом о том, что Китай понимает критичность этической проблематики при разработке технологий, связанных с ИИ, как минимум в военной сфере. Военное применение ИИ предлагается исключить в качестве инструмента для начала войны или достижения гегемонии. С точки зрения права и этики предлагается поддерживать общие ценности человечества, ставить благополучие людей во главу угла, соблюдать национальные или региональные этические нормы при разработке, развертывании и использовании

---

<sup>1</sup> <https://www.reuters.com/investigates/special-report/health-china-bgi-dna>

<sup>2</sup> R. Fedasiuk, J. Melot, B. Murphy. Harnessed Lightning. HOW THE CHINESE MILITARY IS ADOPTING ARTIFICIAL INTELLIGENCE // Center for Security and Emerging Technology (CSET) at Georgetown's Walsh School of Foreign Service. 2021. 84 p

<sup>3</sup> Ministry of Foreign Affairs of the People's Republic of China. Position Paper of the People's Republic of China on Regulating Military Applications of Artificial Intelligence (AI). 2021. [https://www.fmprc.gov.cn/mfa\\_eng/wjdt\\_665385/wjzcs/202112/t20211214\\_10469512.html](https://www.fmprc.gov.cn/mfa_eng/wjdt_665385/wjzcs/202112/t20211214_10469512.html)

соответствующих систем оружия. С точки зрения технологической безопасности – продолжать повышать безопасность, надежность и управляемость технологий ИИ, а также расширять возможности для оценки, управления и контроля безопасности технологий ИИ (через осуществление контроля над технологией человеком). Включение человека-субъекта как элемент контроля над применением технологии «критического урона» не является решением проблемы (по сути, это новый вид «ядерного чемоданчика»), риски войн и конфликтов не снимаются. Документ предлагает регламентировать военные технологии ИИ, стандартизировать их, включить ответственность человека, однако он не запрещает исследования в области военного применения ИИ, не предлагает международных правил по запрету таких исследований. К тому же, размышления о «сингулярности» на поле боя, когда человеческое сознание больше не сможет поспевать за скоростью принятия решений и темпом боя в будущей войне<sup>1</sup>, может рассматриваться военными не только как угроза, но и как способ достижения преимущества в войне.

Существует мнение, что основное отличие подходов Китая и Запада (США в настоящем исследовании) лежат в различии основания ценностей – коллективных и индивидуальных. Это закладывает различные этические подходы к разработке ИИ, разное понимание справедливости, безопасности, конфиденциальности. Культурное восприятие ИИ как предмет футурологии и визуализации будущего с развитием ИИ также различается – в западных культурах «киберпанк», «восстание машин», «Скайнет» и другие неоптимистичные сценарии будущего более распространены, чем в восточных культурах, воспринимающих развитие ИИ в более позитивном ключе – как компаньонов, друзей, помощников<sup>2</sup>.

Однако последние законодательные инициативы в Китае, например «Положение о рекомендациях и управлении алгоритмами информационных услуг в Интернете» которое учитывает такие законы КНР как «О кибербезопасности», «О безопасности данных», «О защите персональных данных», «Методы управления информационными интернет-услугами» и пр., показывают, что ориентация исключительно на «коллективные ценности» не так однозначна. Создаются меры регулирования алгоритмов (например, рекламы и продаж), которые обеспечивают безопасность личных данных пользователей,

---

<sup>1</sup> *Elsa B. Kania*. Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power. Center for a New American Security. 2021. 74 p

<sup>2</sup> <https://www.weforum.org/agenda/2021/11/can-china-and-europe-find-common-ground-on-ai-ethics>

прозрачность использования данных в Интернете, пользователям предоставляются инструменты управления алгоритмами предоставления данных. Если компании будут использовать алгоритмы стимулирования продаж, подавляющие пользователей к зависимости, излишним тратам, привлекающие несовершеннолетних к интернет-зависимости – то теперь они будут на законных основаниях оштрафованы и привлечены к ответственности.

Что действительно отличает ценностные подходы к технологиям ИИ в Китае, так это включение идеологических концептов в этику разработки и внедрения ИИ. Руководящая идеология прописывается в законодательстве, в алгоритмах как экологическая норма прописано «продвижение социалистических ценностей»<sup>1</sup>.

### **3.6.3. Социогуманитарные аспекты исследований этики ИИ в Китае с позиций постнеклассической научной рациональности**

В меру нашего понимания и уровня знаний о сфере разработок ИИ в КНР, можно сделать несколько обобщающих выводов для соотношения имеющихся тенденций с рамочными формами и форматами развития науки в рамках постнеклассической научной рациональности.

1. Можно наблюдать некоторые воплощения осознанной работы по развитию ИИ с позиций взаимодействия субъекта с саморазвивающейся полисубъектной средой со свойствами субъектности («субъект-метасубъект»)<sup>2</sup>. Акты и документы, касающиеся этики ИИ, всё в большем объеме включают социогуманитарные параметры оценки технологий, включают ценностные, моральные, этические характеристики развития технологий ИИ. По сути, прописываются субъектные подходы к влиянию среды на разработки технологий ИИ, и взаимосвязи субъектов разработки с данной средой. Так, например, даже на уровне технического образования включаются культурные, традиционные контексты. В национальном учебнике по инженерной этике для выпускников ВУЗов КНР<sup>3</sup> включаются социалистические ценности китайского правительства, такие как национальные ценности («процветание, демократия, вежливость и гармония»), социальные ценности («свободу, равенство, справедливость и верховенство закона»),

---

<sup>1</sup> Министерство промышленности и информатизации КНР. «Уведомление девяти департаментов о выпуске «Руководящих заключений по укреплению комплексного управления алгоритмами предоставления информационных услуг в Интернете». 2021. [https://wap.miit.gov.cn/xwdt/gxdt/art/2021/art\\_a8af2b48620b4905b365fc73cd81a1ec.html](https://wap.miit.gov.cn/xwdt/gxdt/art/2021/art_a8af2b48620b4905b365fc73cd81a1ec.html)

<sup>2</sup> *Ленский В.Е.* Методологический и философский анализ развития проблематики управления // М.: Когито-Центр, 2019. – 340 с.

<sup>3</sup> [http://www.tup.tsinghua.edu.cn/booksCenter/book\\_06831902.html](http://www.tup.tsinghua.edu.cn/booksCenter/book_06831902.html)

индивидуальные ценности («патриотизм, преданность, честность, дружба»). Особо выделяются принципы ответственности, «предшествующей свободе», обязательств, «предшествующие правам», коллективное, «предшествующее индивидуальному», «гармония, предшествующая конфликту». Схожие подходы к самой среде развития и осмысление контекстов развития «метасубъекта» представлены меньше в академических исследованиях, но есть признаки развития данного направления.

2. Смена приоритетов с философского конструктивизма развития технологий ИИ к усилению коммуникативных процессов, к партнерским отношениям, к включению культуры, языка, традиций также наблюдается в передовых исследованиях этики ИИ в КНР<sup>1</sup>. Осознаются риски бесконтрольного развития технологий ИИ в отсутствие доверительных международных коммуникаций, глобального сотрудничества и этического консенсуса. Открыто вносятся предложения по расширению открытости в сфере управления технологиями ИИ, расширению сотрудничества, в том числе междисциплинарного. Осмысливаются модели новых политических и культурных систем, изначально глобальных и общемировых для решения вопроса коллективной рациональности (некоторые философские переработки древних форм китайского универсализма) (система «Все под небом» (Тянь-ся))<sup>2</sup>. Осознается и проблематика создания такого глобального субъекта – простого объединения субъектов всего мира недостаточно, это не гарантирует рациональности, контроля или универсальной этики. Здесь видятся перспективы для расширения социогуманитарных исследований, развития рефлексивно-активных сред и трансдисциплинарного подхода<sup>3</sup>.

3. Исследований в сфере развития идей коэволюции, как новых отношений человека и окружающей среды, природы, практически не рассматривается. Хотя огромное внимание уделяется вопросам улучшения экологии, прогнозирования поведения окружающей среды (например, качество воздуха или воды в водоемах) с применением средств ИИ и суперкомпьютерных технологий. Технологии виртуальной реальности и развитие Метавселенных приводит к рассмотрению новой гибридной реальности как новой природы, взаимодействие человека и

---

<sup>1</sup> [https://www.guancha.cn/fuying/2019\\_08\\_30\\_515920.shtml](https://www.guancha.cn/fuying/2019_08_30_515920.shtml)

<sup>2</sup> <https://journals.sagepub.com/doi/10.1177/0392192109102149>

<sup>3</sup> *Ленский В.Е.* Седьмой социогуманитарный технологический уклад – контуры будущего человечества // Глобальный мир: системные сдвиги, вызовы и контуры будущего: XVII Международные Лихачевские научные чтения, 18-20 мая 2017 г. – СПб.: СПбГУП, 2017. С. 357-360.

среды рассматривается с технологических позиций созданию новых форм и платформ коммуникаций.

### 3.6.4. Выводы

Развитие технологий ИИ в Китае тесно связано с исследованиями этики, гармонично сочетаются подходы западных исследователей, ориентированных на контексты безопасности, свободы личности, хранения данных, объяснимости новых технологий, с подходами национальных мыслителей, которые внедряют в разработку ИИ понятия из культуры КНР, традиции, семантику, эстетику. Это довольно сильная позиция страны, поскольку данный подход является современным и адекватным развитию науки на этапе становления постнеклассической научной рациональности, внедрению гуманитарных измерений в процессы разработки технологий, созданию сред развития. Очень важно, что осмысляются риски развития технологий, решаются проблемы контроля и надзора над технологическим развитием в период «умного общества», осуществляются призывы к определению «исследований по предотвращению ошибок» (防错研究) как важнейшего из направлений исследования ИИ. Но стоит отметить недостаточность проработки проблематики этики ИИ как базовой характеристики развития международных коммуникаций и глобального управления. Этика рассматривается как вопрос согласования, а не принцип функционирования глобального «метасубъекта», как элемент разработки ИИ, а не как элемент гармонизации различных субъектов-разработчиков. Несмотря на применение гуманитарного измерения к технологиям, по-прежнему превалирует нормативный подход, административное регулирование и управленческие технологии, соответствующие предыдущим порядкам кибернетики. Важно искать пути преодоления национальных эгоизмов, недоверия и страхов и переходить от гармонизации среды внутри сообществ разработчиков, внутри корпораций и стран к гармонизации взаимодействия всех акторов и субъектов мировой политики для эффективного решения проблем, возникающих при исследовании этических контекстов развития ИИ. Не случайно «мудрость» (智) является одним из четырех иероглифов «искусственного интеллекта» (文工智障).

## 3.7. ПРОБЛЕМА МОДЕЛИРОВАНИЯ ЭТИЧЕСКИХ СВОЙСТВ ИНТЕЛЛЕКТУАЛЬНЫХ РОБОТОВ

### 3.7.1. Введение

Основные теоретические трудности озаглавленной проблемы прежде всего вызваны тем, что этика не может предоставить разработчикам нормативную систему в виде некоторой иерархичной архитектуры пригодной для вычислительной реализации. В попытках найти общее решение для любого кейса, требующего от агента этического поведения, мы сталкиваемся с ситуацией двойной неопределённости: объективной неопределённостью среды и субъективной неопределённостью релевантного вида действия. В этой связи рассматриваются возможные способы редукции неопределенности, уточняется понятие «этическое свойство», выясняются условия действительной реализации такого вида свойств в поведении людей и в функционировании искусственного интеллектуального агента. Анализ показывает, что задача выбора этически оправданного действия при моделировании этически санкционированных функций интеллектуального робота не имеет решения в общем виде. В силу этого предлагается способ построения частных моделей на основе классификатора видов деятельности и определения вида деятельности интеллектуального робота, что может послужить продвижению в решении поставленной проблемы.

### 3.7.2. Актуальность проблемы, ее теоретические трудности и создание Общего ИИ

Вопросы этики в развитии ИИ широко обсуждаются в последние годы. По данным *Stanford AI Index Report* за 2021 год с 2015 года наблюдается значительное увеличение числа статей, связанных с этикой ИИ. Из 18 факультетов лучших ВУЗов мира 11 сообщают о проведении мероприятий по этике ИИ, а 7 из них предлагают специальные курсы по этике ИИ. С точки зрения внедрения этических принципов в ИИ последние три года стали явно рекордными для технологических компаний, включая *IBM*, *Google* и *Facebook*<sup>1</sup>. Надо сказать и о таком новом направлении развития ИИ под названием «Искусственные общества», в котором под руководством академика РАН В.Л. Макарова

---

<sup>1</sup> Zhang, Daniel, Mishra, Saurabh, Brynjolfsson, Erik, et al. (2021 web) The ai index 2021 annual report, URL: [https://aiindex.stanford.edu/wp-content/uploads/2021/11/2021-AI-Index-Report\\_Master.pdf](https://aiindex.stanford.edu/wp-content/uploads/2021/11/2021-AI-Index-Report_Master.pdf)

разрабатываются агент-ориентированные компьютерные модели и уделяет значительное внимание этической проблематике.

Повышенный интерес к вопросам моделирования этических свойств ИИ связан во многом с опасениями очевидных рисков и угроз, которые вызывает цифровизация, быстро нарастающая практически во всех сферах социальной жизни. Сюда относятся также непредсказуемые последствия использования результатов информационных технологий, проблемы развития коммуникаций между человеком и интеллектуальным роботом. При этом обычно выделяют два направления этических исследований: первое рассматривает этические проблемы, связанные с инструментальным использованием систем ИИ (проблемы непрозрачности (*opacity*), предвзятости (т.н. *algorithmic bias*) и взаимодействия человека с системами ИИ); второе – проблемы, связанные с принципиальной возможностью создать «этичную машину» в смысле определенного искусственного морального агента.

Второе направление становится особенно актуальным в связи с задачей создания Общего ИИ (ОИИ), – такой системы ИИ, которая по своим функциям приближается к возможностям естественного интеллекта (ЕИ). Сегодня создание ОИИ стало мировым трендом, предметом жёсткой конкуренции между крупнейшими корпорациями и между мировыми державами. О стратегическом значении этой задачи для развития экономики нашей страны, решения её социальных проблем и укрепления безопасности говорил Президент России В.В. Путин<sup>1</sup>

Классические «узкие» системы ИИ, предназначенные для решения отдельных прикладных задач, требуют контроля и периодической перенастройки со стороны человека, не способны к самообучению в новой среде, то есть не являются автономными системами. В отличие от «узких» систем, ОИИ должен быть в значительной степени автономным и самообучаемым, а это означает, что он может выходить из-под контроля человека, его развитие становится в известной степени непредсказуемым. Эта ситуация обостряет проблему этического контроля «самодеятельности» таких автономных систем.

Некоторые авторы (например Н. Бостром) говорят об экзистенциальной угрозе со стороны ИИ: который в своём развитии сможет достигнуть рубежа сингулярности, догнать и превзойти ЕИ, а затем подчинить себе или даже уничтожить человека<sup>2</sup>. Теоретически подобная ситуация мыслима, хотя зачастую и встречает скептическое

---

<sup>1</sup>См. выступление В. В. Путина на организованной СБЕРом конференции *AI Journey 2021*  
<https://youtu.be/V0Hl0hqc69Q>

<sup>2</sup> *Bostrom, N. (2020). "Ethical issues in advanced artificial intelligence", Machine Ethics and Robot Ethics, pp. 69–75.*

отношение со стороны многих специалистов<sup>1</sup>. Но отсюда не следует, что поставленная в названии пункта проблема моделирования этических свойств интеллектуального робота (как и других систем ИИ) является бесперспективной. Есть достаточные основания полагать, что она допускает успешные решения. Но для этого необходимо внимательно рассмотреть само понятие «этическое свойство», в его связи с принципами этики и уточнить условия его реализации в определенных видах деятельности интеллектуального робота.

### **3.7.3. В каком смысле допустимо признавать наличие этических свойств у искусственного интеллектуального агента**

Для ответа на этот вопрос, надо вначале прояснить, что имеется в виду, когда речь идёт о *признании наличия этических свойств у самого человека*. По своему содержанию эти свойства соотносятся с известными нормами этики. Их перечень и особенно их взаимоотношения друг с другом требует специального анализа, так как по этим пунктам в самой этике остаются дискуссионные вопросы. Главное же состоит в том, что эти свойства должны быть не только *знаемыми*, но и *действенными*. Ведь зачастую человек может хорошо знать и одобрять этические нормы, не соблюдая их в своих поступках. Но если определённая этическая норма не является действенной, не соблюдается в реальных действиях личности (или институционального субъекта) то тогда нельзя считать, что соответствующее ей этическое свойство действительно присуще данному субъекту.

Однако даже в тех случаях, когда человек в своих действиях строго соблюдает определённую этическую норму, скажем, требование «не лги», она во многих ситуациях должна согласовываться с другими этическими нормами, в отсутствие чего её применение может привести к тяжкому безнравственному поступку (например, выдача военной тайны врагу и т.п.).

Но вместе с этим все же во многих ситуациях строгое отстаивание правды вопреки всему и разоблачение лжи является единственно достоверным выражением этически значимого поступка.

Проблема в том, что этические нормы не могут быть представлены в виде чёткой иерархической структуры, где одна из норм могла бы во всех случаях, безусловно, подчинять себе все другие. Отсюда неустрашимый аспект относительности всякой этической нормы: её

---

<sup>1</sup> Dowl, Maureen (2017 web) “Elon Musk’s Billion-Dollar Crusade to Stop the A.I. Apocalypse”, The Hive, URL: <https://www.vanityfair.com/news/2017/03/elon-musk-billion-dollar-crusade-to-stop-ai-space-x>



нравственное значение сохраняется лишь по отношению к определенным видам деятельности и в соответствующей ситуации. В силу этого во многих случаях возникает трудная задача выбора этически оправданного действия.

Важно отметить, что тема конфликта моральных норм является одной из наиболее актуальных и сложных областей современных этических исследований<sup>1</sup>. Эта проблема основательно обсуждалась на ряде теоретических семинаров Института философии РАН<sup>2</sup>.

Вопрос о моделировании функций морали и эмоций при исследованиях интеллектуальных агентов в различных ситуациях постоянно находится в центре научных поисков. Предпринимаются попытки построить адекватный математический аппарат для такого моделирования. Была рассмотрена математическая модель, в которой исследовались интеллектуальные агенты в различных типах сред (богатых и бедных ресурсами) и выявлялись более и менее выигрышные стратегии поведения (эгоизм или альтруизм)<sup>3</sup>. В этой связи предпринимались также попытки моделирования факторов симпатии и эмпатии в роли определенных адаптационных механизмов.

Принципиальное значение для существенного продвижения в моделировании этических свойств систем ИИ имеет разработка методологических вопросов ОИИ и развития интеллектуальной робототехники. Здесь, в первую очередь, важно выяснение путей преодоления ограничений методологии А. Тьюринга, исключавшей изучение специфики явлений сознания и использования его результатов. В этом отношении значительную роль приобретает предложенная А.Р. Ефимовым концепция *посттьюринговой методологии*, которая позволяет расширить функциональный диапазон интеллектуальных роботов, включая и решение вопросов моделирования их этических свойств<sup>4</sup>.

Одним из возможных ресурсов моделирования этических свойств интеллектуального робота (и других систем ИИ) могут служить результаты исследований в области нейроэтики, довольно быстро

---

<sup>1</sup> *McConnell, Terrance* (2018 web) “Moral Dilemmas”, The Stanford Encyclopedia of Philosophy, URL: <https://plato.stanford.edu/archives/fall2018/entries/moral-dilemmas/>

<sup>2</sup> О праве лгать / Под ред. Р.Г. Апресяна. М. РОССПЭН, 2011. 392 с.

<sup>3</sup> *Сорокоумов П.С., Карпов В.Э.* К вопросу о моральных аспектах адаптивного поведения искусственных агентов // Искусственные общества. 2021. Т. 16. № 2. URL: <https://arxiv.gaugn.ru/s207751800014740-3-1/> DOI: 10.18254/S207751800014740-3

<sup>4</sup> *Efimov, A.* Post-Turing Methodology: Breaking the Wall on the Way to Artificial General Intelligence // Artificial General Intelligence | (AGI). 2020. Vol. 12177. P. 83- 94/ DOI: 10.1007 / 978-3-030- 52152-3\_9. См также: *Efimov, Albert R., Dubrovsky, David I., Matveev, Philip M.* (2021). “Walking Through the Turing Wall”, IFAC Papers OnLine, Vol. 54, No. 13, pp. 215–220.

развивающегося в последние годы направления. В ней наряду с обоснованием этических ограничений и запретов определенных экспериментальных исследований мозга, выясняются важные данные и о тех его нейрофизиологических механизмах, которые ответственны за ряд существенных аспектов нравственной саморегуляции личности, в частности, за диспозициональную (действующую чаще всего бессознательно) жизненно важную психологическую установку каждого человека на правду и истину, способную сигнализировать в той или иной степени о фактах обмана или аморальных интенциях коммуниканта, вызывать чувство недоверия к его словам и поступкам. Изучение этих функциональных механизмов способно быть полезным для построения моделей этически поощряемых и запретных действий интеллектуального робота как по отношению к человеку, так и к другому роботу<sup>1</sup>.

В такой же мере могут быть полезны данные психогенетики, и этологии, раскрывающие выработанные в ходе эволюции фундаментальные информационные компетенции психики животных, а постольку и человека, которые определяют глубинные, если так можно выразиться, прото-этические свойства и способы регуляции, необходимые для коллективной жизни высших животных. Основательное и яркое описание таких свойств дано выдающимся российским генетиком Владимиром Павловичем Эфроимсоном<sup>2</sup>.

В этом плане перед нами стоит задача организации соответствующих междисциплинарных исследований.

### **3.7.4. Выводы**

Опираясь на изложенные выше соображения, можно считать, что задача выбора этически оправданного действия при моделировании этически санкционированных функций интеллектуального робота не имеет решения в общем виде. Мы сталкиваемся здесь с двойной трудностью: (1) неопределённостью ситуации, в которой оказывается интеллектуальный агент, и (2) неопределённостью видов деятельности, которые требуют этически оправданных действий.

Учитывая это, предлагается способ построения частных моделей на

---

<sup>1</sup> Дубровский Д. И. Нейроэтика: некоторые актуальные философско-методологические вопросы // Философия. Журнал Высшей школы экономики. 2020. Т. IV, № 1. С. 24 – 41.

<sup>2</sup> Эфроимсон В.П. Родословная альтруизма (Этика с позиций эволюционной генетики человек). Эта статья была впервые опубликована в журнале «Новый мир» в 1971 году и перепечатана в книге: Эфроимсон В.П. Гениальность и генетика. М.: Русский мир, 1998, с. 435 – 466. Более широко и детально эти вопросы проанализированы в книге: Эфроимсон В.П. Генетика этики и эстетики. М.: Тайдекс Ко, 2004.

основе классификатора видов деятельности и определения того вида деятельности, который предназначена выполнять данная система ИИ в некоторой типичной ситуации. В таком случае становится возможной алгоритмическая представленность комплексов этически запретных, этически поощряемых и этически нейтральных функций.

При создании соответствующих программ интеллектуальный робот может иметь существенные преимущества перед человеком в смысле надежности, неукоснительного исполнения запрета этически негативных функций и реализации этически поощряемых функций, ибо у него не будет множества постоянно растущих потребностей, интересов, соблазнов, компенсаторных интенций, которые у человека весьма часто служат помехой для исполнения этических требований.

Представляется важным ряд обстоятельств, которые должны быть учтены в дальнейших исследованиях морального поведения интеллектуальных роботов и систем ИИ.

Во-первых, ИИ в целом, любые интеллектуальные агенты-машины существуют на основе дискретной системы счисления («0», «1»), которая прямо или косвенно накладывает ограничения на описание любой ситуации, выходящей за пределы дискретности<sup>1</sup>. В связи с этим, любая модель, основанная на конечном алгоритме, будет неполной и, в результате, недостаточно адекватной при отображении сложности человеческой моральной системы.

Во-вторых, необходимо дополнительное исследование двух важных факторов: (1) действия агента в условиях полной неопределенности (т.е. отсутствия детерминированной ситуации), и (2) моделирования базовой физиологической реакции боли, как одного из существенных факторов формирования морально оправданного поведения. Мы полагаем, что оба фактора могут быть введены в моделирование действий интеллектуальных агентов с применением метода «обучения с подкреплением». Применение этого метода к первому фактору (неопределённость) было ранее продемонстрировано<sup>2</sup>. Исследование второго фактора – болевых ощущений – может также быть весьма эффективно при применении «обучения с подкреплением», в котором соответствующие веса в нейронной сети будут назначаться за действия агентов, ведущих к максимальному избеганию боли (дискомфорта) с учётом ограничений ресурсов.

---

<sup>1</sup> Raikov A. (2021). Cognitive Semantics of Artificial Intelligence: A New Perspective. Springer Singapore, Topics: Computational Intelligence XVII, 128 p. <https://doi.org/10.1007/978-981-33-6750-0>

<sup>2</sup> Ecoffet, Adrien, Lehman, Joel (2021). “Reinforcement Learning under Moral Uncertainty”, Proceedings of Machine Learning Research, Vol. 139, Proceedings of the 38th International Conference on Machine Learning, pp. 2926–2936.

### 3.8. ВЫВОДЫ

1. Становление субъектно-ориентированного подхода в настоящее время определяется в рамках представлений о научной рациональности: классика, неклассика, постнеклассики. Важность субъектно-ориентированного подхода нарастает по мере движения от деятельностного к субъектно-деятельностному и, наконец, к субъектно-ориентированному подходу. В деятельностном подходе базовая парадигма «субъект – объект», в субъектно-деятельностном – «субъект – субъект», в субъектно-ориентированном подходе в центре внимания оказываются субъекты, погруженные в социальные среды и культуру.

2. Субъектно-ориентированный подход формируется в рамках постнеклассической кибернетики саморазвивающихся полисубъектных рефлексивно-активных сред.

3. В развитии методологии цифровой индустрии особое и во многом определяющее место занимает концепция цифрового двойника (это виртуальная интерактивная копия реального физического объекта или процесса), которая соответствует кибернетике первого порядка. Цифровой двойник может быть очень сложным, однако остается пассивным по отношению к субъекту управления, это наблюдаемый объект.

4. В кибернетике второго порядка лицо, принимающее решение, может делегировать часть функций автономного управления цифровому двойнику. Этому способствует развитие технологий машинного обучения, благодаря чему цифровой двойник становится самообучающейся автономной системой управления под контролем лица, принимающего решения. Цифровой двойник трансформируется в цифровой субъект, который, например, может обладать инвариантными свойствами естественных субъектов, функционировать в сети естественных и цифровых субъектов.

5. В постнеклассической кибернетике саморазвивающихся полисубъектных (рефлексивно-деятельностных) сред, кибернетике третьего порядка, цифровой двойник трансформируется уже с учетом мировоззренческих оснований, соответствующей системы принципов и системы онтологий. Это создает основу для сборки в единое целое (метасубъекта) субъектов (псевдообъектов), входящих в среду, и обеспечивает ее жизнедеятельность и развитие.

6. Становится все более очевидным, что рациональная и экономическая подоплека распространения и внедрения технологий ИИ

является далеко не единственным основанием, которому следует уделять внимание. Предельно актуальны также вопросы познания мира и познания человека как биологического вида, вопросы улучшения человека или создания общества всеобщего благоденствия. В настоящее время рациональные поводы внедрения ИИ немного превышают половину всех полученных в исследованиях ученых ответов. То есть люди видят в ИИ не только машину, оптимизирующую механические операции, но и помощника, открывающего для нас неизвестные аспекты окружающего нас мира.

7. Поиск путей повышения доверия к работе систем ИИ приводит к актуализации темы построения объяснимого ИИ с ориентацией на человека, который имеет ожидания. В построении систем ИИ все больший акцент делается психологическим аспектам, посвященным познанию и предубеждениям, человеческого мышления, аргументации вклада когнитивных интерпретаций моделей ИИ. Тема объяснимости выводов ИИ далеко выходит за рамки вопроса каузальности, выявляемой статистическими методами, и охватывает гипотезу нелокальности физических явлений и, что более важно, субъективную реальность.

8. Морально-этические проблемы разработки систем ИИ обусловлены не только тем, насколько произвольно или вредоносно для личности и человечества может повести себя такие системы, но и тем, что само человеческое общество может причинить страдания созданным им искусственным сущностям. Аналогичная моральная дилемма возникала при проведении научно-практических экспериментов на людях и животных. На данный момент эксперименты на людях запрещены без их согласия, животные же считаются менее развитыми, а иногда и – не имеющими души, и эксперименты на них разрешены.

9. При разработке систем ИИ имеется неопределенность, которая сходна с неформализуемой свободой воли человека. В связи с этим, в разрабатываемые системы ИИ должны быть заложены такие принципы и ограничения, как: этические самоограничения и знания о морально-этических нормах людей, имитация процесса саморегуляции своего поведения, способность к эмпатии; прогнозирования рисков и последствий собственных действий; ограничения действий при наступлении определенных рисков; возможность осознания и исправления ошибки.

10. Вопрос об этике систем ИИ в настоящее время является сугубо философским и футурологическим. Реальная проблема такой

«этики» возникает лишь в случае, если эти системы смогут автономно менять контур используемых знаний, обучаться на неограниченных источниках и типах информации, корректировать или даже устанавливать свои цели. Таких систем сейчас нет. Однако уже существует значительное число существ, оснащенных различными «умными» устройствами, поведение которых может создавать этические проблемы.

11. Несмотря на применение гуманитарного измерения к технологиям ИИ, по-прежнему превалирует нормативный подход, административное регулирование и управленческие технологии, соответствующие предыдущим порядкам кибернетики. Обостряется вопрос поиска пути преодоления национальных эгоизмов, недоверия и страхов. Требуется переход от гармонизации среды внутри сообществ разработчиков к гармонизации взаимодействия всех акторов и субъектов мировой политики для эффективного решения проблем, возникающих при исследовании этических контекстов развития ИИ. Не случайно, например, в Китае «мудрость» (智) является одним из четырех иероглифов «искусственного интеллекта» (文工智障).

12. Задача выбора этически оправданного действия при моделировании этически санкционированных функций интеллектуального робота и иных систем ИИ не имеет решения в общем виде. Решению препятствуют, такие трудности, как: неопределённость ситуации, в которой оказывается интеллектуальный агент и неопределённость видов деятельности, которые требуют этически оправданного поведения.

13. В настоящей работе предлагается способ построения частных моделей на основе классификатора видов деятельности и определения того их вида, который будет предназначен для выполнения системой ИИ в некоторой типовой ситуации, что обеспечит возможность алгоритмически представлять комплексы этически запретных, этически поощряемых и этически нейтральных функций. Однако при этом следует иметь в виду, что любая алгоритмизация может катастрофическим образом редуцировать когнитивные и иные субъектные особенности реальных сред и явлений.

## **4. СОЦИОГУМАНИТАРНЫЕ ОСНОВАНИЯ КРИТЕРИЕВ ОЦЕНКИ ИННОВАЦИЙ, ИСПОЛЬЗУЮЩИХ ЦИФРОВЫЕ ТЕХНОЛОГИИ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ, В УПРАВЛЕНИИ СОЦИАЛЬНЫМИ СИСТЕМАМИ**

---

### **4.1. ВВЕДЕНИЕ**

В главе рассматриваются современные системы оценок инноваций, включая цифровые технологии и ИИ, на предмет учета ими социально-гуманитарного фактора. Современные системы оценок, как уже отмечалось выше, основной акцент делают на технологических факторах развития инноваций, умаляя субъективные.

При оценке цифровых инноваций, включая ИИ, большего внимания требуют вопросы поддержки системами социальной ответственности, а также создания гибридных систем, то есть систем, работающих в условиях, когда люди и технологии ИИ сосуществуют и влияют друг на друга. К таким оценкам стоит отнести: оценку социальной и экономической эффективности; ранние методы оценки инновационных технологий; конфликт интересов; характеристики лиц, принимающих решения и др.

В главе отмечается, что при оценке инноваций в области цифровой экономики следует учитывать решение вопросов безработицы, деградации естественного интеллекта, искажения этики, злонамеренного использования ИИ, а также падения уровня духовности общества и др.

Уделяется внимание новой парадигме развития ИИ, которая развивается в рамках развития Общего ИИ. Она должна отражать неформализованную когнитивную динамику работы систем ИИ и поддерживать их саморазвитие в условиях рефлексивно-активной среды. Рассматриваются также ограничения, которые сопутствуют разрабатываемому в настоящей работе подходу к созданию Общего ИИ, основанному на методах конструирования субъектной/субъективной реальности.

Представление ИИ как агента (псевдосубъекта) согласуется с принципом распределенного управления в биологии и психологии, который был назван принципом двойного субъекта. В сочетании с системами принципов и онтологий, задаваемых в концепции постнеклассической кибернетики саморазвивающихся сред, это позволяет использовать ИИ как средство социальных инноваций, при сохранении контроля над технологиями ИИ, а также ставить и решать проблему интеграции образований искусственного и естественного интеллекта при сохранении базовых качеств носителей естественного интеллекта.

## 4.2. СОЦИОГУМАНИТАРНЫЕ ОСНОВАНИЯ ОЦЕНКИ ИННОВАЦИЙ В СФЕРЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

### 4.2.1. Уровни готовности технологических инноваций

В параграфе проведен анализ базовых понятий, стандартов и критериев оценки инноваций; сделан аналитический обзор научных публикаций. Отмечено, что современные системы оценок инноваций, включая цифровые технологии и ИИ, основной акцент делают на технологических факторах развития инноваций, умаляя субъективные. В процессе анализа нами выявлены социально-ценностные ориентиры, включая определение этического измерения и социальных последствий от злонамеренного использования ИИ; проведен анализ моделей инноваций в контексте менеджмента качества и дифференциации семантик моделей ИИ. Построена классификация видов ИИ с учетом социально-гуманитарных оснований, которая позволяет повысить репутационный потенциал ИТ-компаний, противодействовать негативной этике при использовании цифровых технологий и ИИ.

Методологической основой анализа являлись положения пост-неклассической научной рациональности и Кибернетики 3.0<sup>1</sup>, методы стратегического анализа, конвергентного управления. Базовые понятия настоящего исследования сформированы в рыночном контексте, который связан с инновационными трендами и высокой динамикой изменения цифровых технологий, их влиянием на производительность труда.

Технологии обычно оцениваются по метрикам с учетом их потребности на рынке, а также издержек производства. Распространенным подходом считается балльная оценка уровней технологической, производственной и рыночной готовности. Описание этих метрик можно найти в литературе<sup>2,3,4</sup>. Напомним себе эти метрики.

Уровень готовности технологии (TRL) – это критерии для оценки уровня готовности результатов научно-исследовательских и опытно-конструкторских работ (НИОКР) к использованию их в производстве,

---

<sup>1</sup> *Лепский В.Е.* Искусственный интеллект в субъектных парадигмах управления // Философские науки. 2021. Т. 64. № 1. С. 88-101

<sup>2</sup> [https://en.wikipedia.org/wiki/Technology\\_readiness\\_level](https://en.wikipedia.org/wiki/Technology_readiness_level)

<sup>3</sup> *Петров А.Н., Комаров А.В.* Оценка уровня технологической готовности конкурсных заявок с использованием методологии TPRL // Экономика науки. 2020. Т. 6. № 1–2. С. 88–99.

<sup>4</sup> *Жебель В.В., Комаров А.В., Комаров К.А., Шуртаков К.В.* Программное средство для комплексной оценки технологической готовности инновационных научно-технологических проектов // Экономика науки. 2018. Т. 4. № 1. С. 58–68.



установках, процессах и проектах, а также технологического риска, связанного с использованием данных технологий в настоящий момент.

Уровень производственной готовности (MRL) – это критерии для оценки готовности к массовому промышленному производству результатов НИОКР по мере развития проектов ИИ с целью внедрения результатов в производственный процесс, а также риска, связанного с организацией производства технологий в настоящий момент.

Уровень рыночной готовности (CRL) – критерии для оценки готовности к коммерциализации результатов НИОКР, суммирующие результат оценки емкости рынка ИИ с учетом ценовых и потребительских качеств выводимых на рынок и разрабатываемых продуктов конкурентов, стадию разработки бизнес модели коммерциализации и ценообразования, конкурентные преимущества, статус клиентской заинтересованности в использовании результата и его ключевых характеристиках.

Каждый из перечисленных групп критериев для каждого балла имеет серию показателей со шкалой оценки от 1 до 100. Количество показателей для каждого балла может быть от 12 (TRL уровня 1) до 45 (TRL уровня 4). Анализ этих показателей позволяет констатировать их акцентирование на готовности технологий к дальнейшему продвижению, и только косвенно отражают социально-гуманитарные аспекты оценки инноваций. Например, выборочно показателями TRL уровня 4 являются такие, как:

- подробные сведения о функциях программного обеспечения;
- лабораторные требования, вытекающие из системных требований;
- компоненты собраны в системный макет и др.

Эта система критериев в реальной практике зачастую редуцируется, например, она может быть сведена к следующему набору уровней технологической готовности<sup>1</sup>:

1. фундаментальные научные исследования – сформулирована идея решения проблемы, произведено ее теоретическое описание, определена фундаментальная концепция технологии и обоснована ее полезность;
2. фундаментальные научные исследования – определены целевые области применения технологии, доказана эффективность использования идеи в решении прикладных задач на базе произведенных расчетов;

---

<sup>1</sup> Письмо Заместителя Министра финансов РФ от 16.05.2022 № 12-12-05/46697.

3. прикладные научные исследования – получен компонент и (или) макетный образец, продемонстрированы его (их) ключевые характеристики;
4. прикладные научные исследования – компонент и (или) макетный образец проверены в лабораторных условиях;
5. экспериментальные разработки – изготовлен и испытан экспериментальный образец в условиях, близких к реальным;
6. экспериментальные разработки – изготовлен и испытан опытный образец в условиях, близких к реальным;
7. экспериментальные разработки – проведены испытания опытно-промышленного образца в реальных условиях эксплуатации;
8. опытно-промышленное производство – окончательно подтверждена работоспособность опытно-промышленного образца, запущены опытно-промышленное производство и сертификация;
9. коммерциализация – технология удовлетворяет всем требованиям – инженерным, производственным, эксплуатационным, а также требованиям к качеству и надежности и выпускается серийно.

Система критериев оценки инновационных проектов может использоваться для отбора претендующих на государственную поддержку цифровых технологий, с указанием направлений экспертной, регуляторной и финансовой поддержки. В этой системе критерии могут быть упорядочены следующим образом<sup>1</sup>:

- на первом уровне – общая оценка уровня целесообразности государственной поддержки;
- на втором уровне оценка:
  - востребованности технологии рынком;
  - уровня научно-технического задела;
  - публикуемости, патентования, тиражируемости, комплексированности;
  - фазы хайп-цикла;
- на третьем уровне оценка:
  - национальных и секторальных приоритетов;
  - тиражируемости, используемости;
  - выручки от реализованных проектов;
  - степени соответствия мировым решениям;
  - рейтинга выгодности;

---

<sup>1</sup> Raikov A.N., Ermakov A.N., and Merkulov A.A. Assessments of the Economic Sectors Needs in Digital Technologies, Lobachevskii Journal of Mathematics, 2019, Vol. 40, No. 11, pp. 1837–1847. Pleiades Publishing, Ltd. <https://doi.org/10.1134/S1995080219110246>.

- эффекты, задачи;
- наукометрических параметров и др.

Видно, что существующие подходы к оценке инноваций в основном носят технологический характер. Социально-гуманитарные основания оценки инноваций в области ИИ в указанных системах оценок заложены только косвенно.

#### 4.2.2. Социально-гуманитарные основания оценок

Существующие системы оценок инновационных процессов ориентированы на удовлетворение потребностей различных сегментов рынка, то есть должны учитывать потребительские характеристики сегментов. В состав этих характеристик могут войти такие, которые имеют социально-гуманитарные основания. К ним можно отнести:

- социальная эффективность технологии,
- повышение качества и уровня жизни;
- рост объема ресурсов домашних хозяйств;
- оценена востребованность технологии населением;
- важность с учетом национальных приоритетов;
- выявляются перекрестные и сквозные технологические эффекты,
- соответствие мировым решениям,
- выгода использования населением.

Для наиболее полной оценки ситуации с построением критериев, носящих социо-гуманитарный характер, сделан обзор зарубежных публикаций, который показал следующее.

Социально-гуманитарные основания критериев оценки инноваций, использующих ИИ, характеризуют востребованность ИИ гражданским обществом, особенно такими областями деятельности человека, как медицина, образование, наука и пр.

Аналитическое моделирование решений используется в оценке эффективности инновационных технологий здравоохранения (ИТЗ). Примером может служить использование моделей экономической эффективности, которые взвешивают эффекты инновационной стратегии с перспективными затратами, и сравнивают их с затратами и эффектами текущего года<sup>1</sup>. У ИТЗ есть важное приложение – ранние методы ИТЗ. Они применяются на ранних этапах разработки, что позволяет выявлять пробелы в текущей клинической практике. Ведь на более поздних этапах исследований теряется возможность управлять

---

<sup>1</sup> Scholte M., Rovers M.M., Grutters J.P.C. The Use of Decision Analytic Modeling in the Evaluation of Surgical Innovations: A Scoping Review. Value in Health. Systematic literature review. 2021. Vol. 24, Issue 6, pp 884-900. <https://doi.org/10.1016/j.jval.2020.11.020>

разработкой инновации таким образом, чтобы она приносила наибольшую пользу<sup>1,2</sup>.

В публикации систем совместной работы IDEAL (идея, разработка, исследование, оценка, долгосрочное наблюдение) авторы определили инновации как устройства или процедуры, которые находятся на ранних стадиях исследования<sup>3</sup>. При этом вводятся следующие характеристики инноваций: тип инновации, статус на рынке, фаза исследования и распространение инновации, источник финансирования, а также конфликт интересов. Извлеченные характеристики моделирования включали лицо, принимающее решение, программное обеспечение, использование имеющихся моделей, источники входных данных, показатели результатов, перспективное развитие. Результаты и рекомендации включали коэффициент прироста экономической эффективности и стратегические рекомендации. Учитывались также аспекты, связанные с оценкой внедрения инноваций (кривая обучения, дополнительные инновации, динамическое ценообразование, организационное влияние и изменение качества).

В работе<sup>4</sup> внимание уделяется созданию гибридных систем. Гибридная реальность (HyR) – это ситуация, когда люди и технологии ИИ сосуществуют и влияют друг на друга. Технологии не просто изменяют существование людей, но люди помогают технологиям развиваться. При этом гибридная реальность рассматривается через призму социальной ответственности.

Концепции социальной ответственности (СО) предназначены для применения ко всем субъектам общества. Концепции СО признают сложность взаимодействия между людьми, организациями, обществом и техникой. Применяя естественную среду к исходным принципам СО сформирована итоговая линия: Прибыль – Планета – Люди<sup>5</sup>. В ISO 26000 выделены три основных принципа: ответственность субъекта за влияние на общество, взаимозависимость и целостный подход. Эти принципы

---

<sup>1</sup> Grutters J.P.C., Govers T, Nijboer J., et al. Problems and promises of health technologies: the role of early health economic modeling. *Int J Health Policy Manag.* 2019. 8:575–582.

<sup>2</sup> Abel L., Shinkins B., Smith A., et al. Early economic evaluation of diagnostic technologies: experiences of the NIHR Diagnostic Evidence Co-operatives. *Med Decis Making.* 2019. 39:857–866

<sup>3</sup> Hirst A., Philippou Y., Blazeby J., et al. No Surgical Innovation Without Evaluation: Evolution and Further Development of the IDEAL Framework and Recommendations. *Ann Surg.* 2019. 269:211–220.

<sup>4</sup> Perko I. Hybrid reality development - can social responsibility concepts provide guidance? *Kybernetes.* 2021. Vol. 50 No. 3, pp. 676-693. <https://doi.org/10.1108/K-01-2020-0061>

<sup>5</sup> Carroll A.B. Corporate social responsibility: evolution of a definitional construct. *Business and Society.* 1999. Vol. 38 No. 3, pp. 268-295.

поддерживаются семью концепциями СО: прозрачность, подотчетность, уважение к заинтересованным сторонам, этичное поведение, уважение к верховенству закона, уважение международных норм поведения и уважение прав человека. В СО рассматривается взаимодействие с заинтересованными сторонами: права человека, трудовые отношения, окружающая среда, добросовестные методы работы, проблемы потребителей, гражданское участие и развитие сообщества. При этом отмечается, что только физические или организационные аспекты неадекватны для определения границ НуR из-за вездесущности ИИ.

В статье<sup>1</sup> даются понятия, основанные на принципе построения коллективного интеллекта в контексте событий, связанных с COVID-19. Наряду с внедрением цифровых технологий здравоохранения наблюдается явление, которое связано с коллективными действиями сторон, вовлеченных в цифровую экосистему. С помощью цифровых технологий врачи, эпидемиологи, межнациональные организации, министерства, больницы, поставщики устройств и др. координируют свои усилия.

Экосистема здравоохранения как система коллективного интеллекта, состоит из множества заинтересованных сторон и объектов, в которых цифровые технологии и ИИ поддерживают управление. Онтология экосистемы включает четыре блока, которые предназначены для ответа на четыре вопроса: «Что делается?», «Кто это делает?», «Почему это делается?» и «Как это делается?».

В экосистеме здравоохранения на первом этапе глобального управления выделяется один важный аспект – преобладающая природа «самоорганизации», поддерживаемая национальными и международными организациями. Отмечается наличие сложного набора целей и мотивирующих факторов, которые стимулируют участие субъектов экосистемы (врачей, волонтеров, поставщиков медицинских устройств и изобретателей, медсестер и др.)<sup>2</sup>, определяет экосистему здоровья характеристиками, типичными для системы коллективного интеллекта.

Возможность принятия концепции коллективного интеллекта может заключаться в двух отношениях. Во-первых, управление вызовами кризисного события представляет собой сложный процесс, характеризующийся большой неопределенностью и который может только выиграть от сотрудничества. Во-вторых, использование

---

<sup>1</sup> *Secundo G., Shams S.M.R., Nucci F.* Digital technologies and collective intelligence for healthcare ecosystem: Optimizing Internet of Things adoption for pandemic management. *Journal of Business Research*, 2021. Vol. 131, pp. 563-572. <https://doi.org/10.1016/j.jbusres.2021.01.034>

<sup>2</sup> *Gasser U., Ienca M., Scheibner J., Sleight J., & Vayena E.* Digital tools against COVID-19: Framing the ethical challenges and how to address them. 2020. arXiv preprint arXiv:2004.10236.

цифровых технологий и ИИ может упростить распространение информации среди всех заинтересованных сторон и усилить влияние сетевого взаимодействия<sup>1</sup>. Важно проводить различие между коллективным интеллектом, управляемым, главным образом, физическими связями людей в некотором регионе, и формами коллективного интеллекта, которые являются глобальными.

Таким образом, аналитический обзор показывает широкий спектр аспектов построения социально-гуманитарных оснований критериев оценки инноваций, использующих ИИ. К таким аспектам стоит отнести:

- моделирование с оценкой экономической эффективности;
- ранние методы разработки и оценки инновационных технологий;
- характеристики инноваций: тип инновации, статус на рынке, фаза исследования в соответствии с системой IDEAL, распространение, источник финансирования, конфликт интересов;
- характеристики моделирования (лицо, принимающее решение, программное обеспечение, использование имеющихся моделей, источники входных данных и др.);
- создание гибридных систем;
- принципы СО и ее атрибуты;
- признаки и виды (локальный, глобальный) систем коллективного интеллекта;
- блочный (онтологический) принцип построения сетевых экосистем;
- сенсорика, датчики (визуальные, встроенные, носимые и др.).

Приведенный обзор помогает оттолкнуться в формировании методологических оснований моделей оценки инновационного развития ИИ в социально-гуманитарном контексте.

### 4.2.3. Негативные нюансы

При инновационном развитии обостряются, как уже отмечалось, вопросы замещения людей машинами, деградации естественного интеллекта, искажения этики, злонамеренного использования ИИ<sup>2</sup>, а также падения уровня духовности общества, не только из-за «клипового

---

<sup>1</sup> *Elia G., Margherita, A., & Passiante G.* Digital entrepreneurship ecosystem: How digital technologies and collective intelligence are reshaping the entrepreneurial process. *Technological Forecasting and Social Change.* 2020. 150, Article 119791. <https://doi.org/10.1016/j.techfore.2019.119791>

<sup>2</sup> *Bazarkina, D., Pashentsev, E.* Malicious Use of Artificial Intelligence: New Psychological Security Risks in BRICS Countries. *Russia in Global Affairs.* 2020. 18(4):154–177. doi: 10.31278/1810-6374-2020-18-4-154-177

мышления», но и накопления ошибки в передаче данных из-за замены естественного сигнала дискретным, цифровым<sup>1</sup>.

Система ИИ, согласно преобладающим представлениям, является компьютерным устройством, программой, автономно функционирующей от человека. Например, определение, которое распространено сейчас в Организации экономического сотрудничества и развития: Система ИИ – это машинная система, которая способна влиять на окружающую среду, давая рекомендации, прогнозы или решения для заданного набора целей.

Это определение дает взгляд на ИИ с позиций классической научной рациональности. Вместе с тем взаимодействие систем ИИ с человеком порождают новую, гибридную, реальность<sup>2</sup>. Совсем не скоро, да и вообще навряд ли когда-нибудь, эмоциональный потенциал человека будет достоянием машины. Еще более сомнительно это относится к бессознательному, трансцендентному, феномену сознания.

Скажем, герменевтика – это искусство понимания устной и письменной речи. Еще в эпоху, когда эта дисциплина переживала смену устоев, заложенных со времен античности, возникнув из экзегезы, философии и юриспруденции, акцент сместился к плану выражения для воплощения индивидуально-стилистической манеры произведения. Такой подход противоречит очевидному тезису, что понимание не тождественно интерпретации. Эта проблема ищет разрешения в феноменологии Э. Гуссерля, который не разделяет субъект и объект. Он считал, что сознание и смыслы интенциональны. Смысл не конкретная сущность целостного предмета, а лишь форма, предающая предмету целостность. Есть и другие взгляды на вопрос, например, Г. Шпета, Г.-Г. Гадамера, М. Хайдеггера, П. Рикера, Н. Хомского. Общее во взглядах состоит в отстаивании связи слова и мышления.

Из подобных рассуждений начинает проясняться концептуальная постановка вопроса о том, что такое смысл, который не формализуется. Он недоступен логической интерпретации, он лежит много глубже, а может и «выше», например, в форме нелокальных когнитивных семантик моделей ИИ<sup>3</sup>, которые неподвластны логике и даже современным научным теориям, которые сопровождаются

---

<sup>1</sup> Raikov A. Post-Non-Classical Artificial Intelligence and its Pioneer Practical Application. IFAC-PapersOnLine. 2019. 52(25):343-348. <https://doi.org/10.1016/j.ifacol.2019.12.547>

<sup>2</sup> Raikov, A.N. Convergent social responsibility as the key to corporate strategic success. Kybernetes. 2021. 50(3):785-793. <https://doi.org/10.1108/K-12-2019-0853>

<sup>3</sup> Raikov A. Cognitive Semantics of Artificial Intelligence: A New Perspective. Springer Singapore, Topics: Computational Intelligence XVII, 2021. 128 p. <https://doi.org/10.1007/978-981-33-6750-0>

противоречиями (например, между квантовой теорией и теорией относительности).

В разрезе менеджмента качества обращает на себя внимание хорошо известный метод структурирования функций качества (СФК). Можно выделить сегменты рынка потребителей совсем различных сквозных цифровых технологий (СЦТ). Потребительских характеристик для каждого сегмента может быть до 100. Применительно к ИИ метод СФК может помочь формализовать определение сегментов рынка, составления и ранжирования потребительских характеристик. Этот метод предполагает огромный объем экспертной работы. Вместе с тем вопрос его автоматизации, например, с использованием средств ИИ и анализа больших данных, стоит на повестке дня.

Метод СФК помогает структурировать нужды потребителя через развертывание технологий предприятия. Например, критериями оценки востребованности цифровых технологий, включая ИИ, применительно к сельскохозяйственному сектору могут быть:

- главная цель развития сектора;
- виртуальное сотрудничество и самоорганизация;
- новые ценности для фермеров и граждан;
- научно-техническое развитие и система образования;
- экспорт продукции и др.

Таким образом, искомые социально-гуманитарные основания критериев оценок инноваций в области ИИ могут быть операционализированы с применением рыночных механизмов. К ним могут быть отнесены национальные стратегические цели, например:

- а) сохранение населения, здоровье и благополучие людей;
- б) возможности для самореализации и развития талантов;
- в) комфортная и безопасная среда для жизни;
- г) достойный, эффективный труд и успешное предпринимательство;
- д) цифровая трансформация.

При формировании списков потребительских характеристик в субъектах России для учета социально-гуманитарных аспектов могут быть взяты позиции из соответствующих нормативных правовых документов<sup>1</sup>, отражающих такие социальные аспекты, как:

- ожидаемая продолжительность жизни при рождении;
- уровень бедности;

---

<sup>1</sup> Указ Президента Российской Федерации от 04.02.2021 № 68 «Об оценке эффективности деятельности высших должностных лиц (руководителей высших исполнительных органов государственной власти) субъектов Российской Федерации и деятельности органов исполнительной власти субъектов Российской Федерации».



- доля граждан, занимающихся физической культурой и спортом.
- уровень образования;
- эффективность системы выявления, поддержки и развития способностей и талантов у детей и молодежи;
- доля граждан, занимающихся добровольческой (волонтерской) деятельностью;
- условия для воспитания гармонично развитой и социально ответственной личности;
- число посещений культурных мероприятий.

Учет этих аспектов в деятельности юридических лиц будет способствовать как стратегическому развитию региона, так и может быть положен в формирование методологические оснований оценки моделей инновационного развития ИИ.

#### 4.2.4. Социально-гуманитарная классификация ИИ

Таким образом, современные системы ИИ преимущественно развиваются в рамках парадигмы, которая способствует совершенствованию отдельных направлений ИИ, но затрудняют использование результатов разработок в процессах управления. К таким парадигмам относятся: морфологическая, логическая, нейрокибернетическая, имитационная, а также парадигмы слабого, общего и сильного ИИ. ИИ развивается совместно с робототехникой, формируется человеко-машинный симбиоз. Он может привести к фундаментальной цивилизационной трансформации социума, в связи с чем поднимаются философские, правовые и социальные вопросы о сознании, гражданстве, правах и юридическом лице роботов<sup>1,2</sup>.

Основания классификации ИИ можно найти в различных нормативных документах. Так, определение ИИ дано в Национальной стратегии развития искусственного интеллекта на период до 2030 года. В ней также упоминается перспективный (сильный) ИИ.

В международных стандартах согласно ISO-IEC 22989<sup>3</sup> выделяется, как уже отмечалось выше, два вида ИИ: слабый ИИ и сильный ИИ. В слабом ИИ система может обрабатывать только символы, не понимая, что она делает. Сильный ИИ также обрабатывает символы, но уже «понимает», что она делает. Правда, в указанном стандарте указывается,

<sup>1</sup> *De Luca G.* The development of machine intelligence in a computational universe. *Technol Soc.* 2021. 65:101553. <https://doi.org/10.1016/j.techsoc>.

<sup>2</sup> *Karnouskos S.* Symbiosis with artificial intelligence via the prism of law, robots, and society. *Artif Intell Law.* 2021. <https://doi.org/10.1007/s10506-021-09289-1>

<sup>3</sup> ISO/IEC DIS 22989. Information technology – Artificial intelligence – Artificial intelligence concepts and terminology

что обозначения «слабый ИИ» и «сильный ИИ» в основном важны для философов и неактуальны для исследователей и практиков ИИ. Следуя этому аргументу, появились квалификации «узкий ИИ» против «общего ИИ». Система «узкого ИИ» способна решить конкретную задачу, возможно, намного лучше, чем это сделали бы люди. Система «общего ИИ» способна решать несколько задач разного характера с одинаковым уровнем производительности. Классификации ИИ, как правило, носят прагматический характер, зависят от участников команд, грантовой поддержки, образовательных дисциплин.

Таким образом, отсутствуют, да и навряд ли когда-либо появятся единые основания классификации видов ИИ. В настоящей монографии рассмотрена возможность разработки социально-гуманитарных оснований классификации видов ИИ.

Поиск социально-гуманитарных оснований критериев оценки инноваций, использующих ИИ, подразумевает учет отраслевой специфики их использования. Помимо прикладного, классификация ИИ может иметь фундаментальный характер, например, основанный на теории построения интеллектуальных управляющих систем. Цель такой теории состоит в обеспечении устойчивости функционирования управляющих систем. Так, в давней фундаментальной научной работе, опирающейся на термодинамику<sup>1</sup> приводится классификация продвинутого ИИ («ИИ в большом»), предлагающая 5 принципов управления организациями:

- наличие взаимодействия интеллектуальной управляющей системы с внешним миром по информационным каналам связи;
- принципиальная открытость систем с целью повышения интеллектуальности и совершенствования поведения;
- наличие механизмов прогноза изменений внешнего мира и собственного поведения;
- наличие многоуровневой иерархической управляющей системы, построенной по правилам: повышение интеллектуальности со снижением требований к точности модели по мере повышения уровня иерархии;
- сохранность функционирования при разрыве связей или потере управляющих воздействий.

Эти принципы могут считаться признаками и одновременно фасетами классификации систем ИИ, которые обеспечивают

---

<sup>1</sup> Захаров В.Н., Ульянов С.В. Нечеткие модели интеллектуальных промышленных регуляторов и систем управления. II. Эволюция и принципы построения // Техническая кибернетика. 1993. № 4. С. 189- 205.

устойчивость функционирования систем ИИ. К этим признакам могут быть добавлены такие, которые создают условия для обеспечения необходимых условий целенаправленности функционирования интеллектуальных информационных систем, например, экспертно-аналитических систем и ситуационных центров<sup>1</sup>.

Применение ИИ оправдано для предметной области, перечень характеристик которой можно свести к следующему списку:

- информационная открытость объекта управления;
- хаотичность поведения среды;
- когнитивные и денотативные семантики;
- неформализуемость задач;
- уникальность, нестереотипность ситуаций;
- латентность, неявность информации;
- парадоксальность логики решений и др.

В результате проведенного исследования предложена классификации различных видов ИИ, приведенная на Рис. 4-1.



Рис. 4-1. Схема классификации различных видов ИИ с учетом социально-гуманитарных оснований оценки инноваций

Таким образом, в результате проведенного исследования по выявлению социально-гуманитарных оснований выработки критериев оценки инноваций, использующих ИИ, построена соответствующая классификация ИИ, учитывающая перспективы его развития.

<sup>1</sup> Стратегическое целеполагание в ситуационных центрах развития / Под ред. В.Е. Лепского, А.Н. Райкова –М.: Когито-Центр, 2018. 320 с.

При этом выявлен дефицит внимания к параметрам ИИ для их социально-гуманитарной оценки, что препятствует росту репутационного потенциала технологий ИИ, создает существенные риски в его развитии и продвижении.

Недостаточный учет социально-гуманитарных факторов развития и оценки инноваций в области ИИ создает условия для злонамеренного использования ИИ. Причем, построение перспективного (общего, сильного) ИИ, даже пока на теоретическом уровне, делает такое использование все более неуязвимым.

#### **4.2.5. Выводы**

На основе изложенного в этом разделе, можно выделить следующие основания для построения искомых социально гуманитарных критериев:

- Существующие подходы к оценке инноваций в основном носят технологический характер, социально-гуманитарные основания оценки инноваций в области ИИ в указанных системах оценок заложены только косвенно.
- Социально-гуманитарные основания критериев оценки инноваций, использующих ИИ, характеризуют востребованность ИИ гражданским обществом, особенно такими областями деятельности человека, как медицина, образование, наука и пр.
- Искомые социально-гуманитарные основания критериев оценок инноваций в области ИИ могут быть операционализированы с применением рыночных механизмов, которые следует учитывать в рамках национальных стратегических целей.

## **4.3. ОБЩИЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В САМОРАЗВИВАЮЩИХСЯ РЕФЛЕКСИВНО-АКТИВНЫХ СРЕДАХ**

### **4.3.1. Феномен Общего Искусственного Интеллекта (ОИИ)**

Цель этого пункта состоит в объяснении и выявлении субъектных особенностей Общего Искусственного Интеллекта (ОИИ) в контексте этапов научной рациональности (классическая, неклассическая, постнеклассическая), и демонстрации его отличия от традиционного Искусственного Интеллекта (ИИ). Последний в настоящее время реализуется в основном на дискретных (цифровых) компьютерах и реализует функции распознавания и прогнозирования поведения объектов, а также подготовки ответов на простые вопросы. Вместе с тем новые условия заставляют рассматривать ОИИ с позиций функционализма, как человеко-машинную систему, целенаправленно функционирующую в саморазвивающейся полисубъектной (рефлексивно-активной) среде.

В этом контексте парадигма разработки ИИ должна учитывать аспекты его погружения в социально-гуманитарную среду и инновационную атмосферу. Новая парадигма развития ИИ должна отражать неформализованную когнитивную динамику моделей ИИ и поддерживать саморазвитие систем ИИ под давлением внешней среды. Тогда методология создания ОИИ должна базироваться на идеях субъектно-онтологического подхода, функционализма, феноменологии субъективной реальности, конвергентных когнитивных архитектур, а также методов создания саморазвивающейся полисубъектной (рефлексивно-активной) среды. ОИИ приобретает гибридный характер, целенаправленно объединяющий возможности машины и человека.

Общие свойства, ценностно-смысловые и интенционально-волевые операциональные структуры феномена субъективной реальности не опираются на прямое формализованное и алгоритмическое представление в дискретных компьютерных системах Фон-неймановской архитектуры. Изучение сознания в контексте субъективной реальности позволило сформулировать основные системные, структурно-функциональные и операциональные характеристики познавательной деятельности человека, что позволяет по-новому подойти к моделированию когнитивных архитектур, отвечающих задачам построения ОИИ. Очевидно, что характеристики субъективной реальности не могут быть полностью представлены в парадигме физикализма; то есть она не может быть представлена только с помощью физических устройств. Поэтому в такой реальности требуется разработка и приложение нередукционистского способа учета

субъектной характеристики реальности путем рассмотрения проблемы сознания в онтологическом и эпистемологическом контексте, что позволяет представить процессы сознания и познавательной деятельности человека и группы людей косвенно и инверсно.

Оригинальность такого подхода заключается в том, что современные когнитивные архитектуры и традиционные подходы к ИИ практически игнорируют решение субъектных проблем ОИИ, или оставляют их функционально нераскрытыми. Традиционные подходы к ИИ в большей степени ориентированы на формализованное построение модели мышления, выделение физических блоков и процессов мыслительной деятельности. В то же время для ОИИ наибольшее значение имеют вопросы онтологической, субъективной и гибридной реальности, особенно в объяснении деятельности сознания, бессознательного и беспричинных процессов, способных целенаправленно действовать в условиях неопределенности цели. ОИИ должен помогать описывать явления субъективной реальности, вызывающие физические изменения, объяснять способность целеполагания, свободу воли, интенцию к самоуправлению физическими действиями индивида в коллективе и т. д.

Такой взгляд на ОИИ порождает новый тип управления в гибридной среде, отличающийся от традиционного управления в цифровой реальности. При этом не остается без внимания и поиск однозначной оптимальной меры централизации и автономности контуров управления, помогающей обеспечить сохранение и укрепление целостности сложной полисубъектной системы, функционирующей в рефлексивно-активной среде, интерпретация которой не укладывается в узкие рамки цифровой и алгоритмической реальности, а также традиционного ИИ.

#### **4.3.2. Ограничения и структура реальности**

Изложенный подход к созданию ОИИ, основанный на методах конструирования субъективной реальности, также имеет свои ограничения. Например, этот подход к объяснению связи между мозгом человека, сознанием, мыслительными процессами и окружающей средой пока не позволяет объяснить информационно-когнитивные процессы, порождаемые эффектом нелокальности субъективной реальности, возникающим, например, при учете атомарного уровня человеческого мозга и тела.

Структура реальности уже давно является предметом растущего внимания ученых. В книге<sup>1</sup>, например, предлагается выделить четыре основных направления: эпистемологию с указанием путей развития

---

<sup>1</sup>Deutsch, D. (2016) The Fabric of Reality. <https://archive.org/details/TheFabricOfReality>.

науки, квантовую механику с парадигмой многих миров, дискретную теорию вычислений Тьюринга и универсальную теорию эволюции. Большинство современных подходов и представлений о структуре реальности основаны на физикалистских описаниях. В то же время новые инструменты познания Вселенной и сознания, особенно ОИИ, форсируют и ломают развитие устоявшихся взглядов. Все больше наблюдается попыток добраться до понимания структуры гибридной человеко-машинной реальности<sup>1</sup>, природы субъективной реальности<sup>2</sup> и саморазвивающейся полисубъектной (рефлексивно-деятельностной) среды<sup>3</sup>.

Современные системы ИИ не могут думать, понимать, хорошо объяснять и ставить проблемы. Неумение объяснить порождает недоверие к выводам, сделанным с помощью таких систем. Все, что делает современный ИИ, больше похоже на автоматизацию, чем на творчество. ИИ ускоряет то, что можно описать алгоритмом, логикой. Развитие систем ИИ идет классически экстенсивным путем – в основном за счет вычислительных мощностей и использования эвристик, реализованных на цифровых устройствах.

Вместе с тем человек может делать правильные и в то же время беспричинные выводы, мгновенно оценивать подлинность художественного произведения, обладает интуицией, может впадать в транс, у него загадочная душа. Его характеризует бессознательное явление. Человек формирует субъективную реальность, находящуюся «по ту сторону» компьютера, алгоритмов и чисел. Именно эти аспекты являются предметом дальнейшего развития современного ИИ, который развивается в направлении ОИИ, сильного ИИ, сверхразума.

Основной мотивацией создания ОИИ является: отсутствие у традиционного (слабого, узкого) ИИ возможности объяснять свои выводы, усиление возможностей для злонамеренного использования ИИ, миниатюризация роботов и летательных аппаратов (в оборонке, сельском хозяйстве и т.д.), бесперспективность решения многих задач стратегического планирования, национальной безопасности, биологии, физики, астрономии и т. д. традиционными методами.

Разработчики современного (слабого, узкого) ИИ часто рассчитывают на возможность добиться успеха в имитации творческой

---

<sup>1</sup> *Perko, I.* (2020). Hybrid reality development-can social responsibility concepts provide guidance? *Kybernetes* 50(3), 676–693. <https://doi.org/10.1108/K-01-2020-0061>

<sup>2</sup> *Dubrovsky, D.I.* (2019) The Problem of Free Will and Modern Neuroscience. *Neuroscience and Behavioral Physiology*. 49(5), 629–639.

<sup>3</sup> *Lepskiy, V.* (2018). Evolution of cybernetics: philosophical and methodological analysis, *Kybernetes*. 2018, Vol. 47 Issue: 2, pp. 249-261. <https://doi.org/10.1108/K-03-2017-0120>

интуиции человека с использованием причинно-следственных (каузальных) математических моделей с высокой скоростью вычислений. В этом допущении содержится неразрешимое противоречие, которое не приводит к диалектическому синтезу нового. Достаточно прочесть обширную литературу по генерированию человеком инсайтов, эмоциональному интеллекту, бессознательному, «Эврика-эффекту» и т.д. Будет очевидно, что каузальный подход к созданию ОИИ далеко не достаточен, хотя и необходим.

Текущее состояние продвинутого ИИ, включая ОИИ или сильный ИИ (ASI), основано на данных науки, культуры и философии<sup>1</sup>. Гибридный ИИ погружается в человеко-машинную реальность, где человек играет роль наблюдателя и создает когнитивный компонент системы управления. Далее в этом пункте монографии ОИИ, искусственный суперинтеллект и сильный ИИ рассматриваются как синонимы.

Тема ОИИ с его многозадачностью и «неаксиоматичностью» основы в мировом научном сообществе прозвучала относительно недавно<sup>2,3</sup>. Рождение термина можно проследить с 1997 года, когда компьютер Deep Blue обыграл чемпиона мира по шахматам. На основе концепции общего интеллекта было предложено разделить ИИ не на сильный и слабый, а на общий (способный приспосабливаться к решению различных задач и решать несколько задач) и узкий (способный решать конкретную задачу).

Стоит заметить, что пока опубликовано относительно немного статей по вопросам развития ОИИ; например, запрос по терминам «Искусственный интеллект» и «Искусственный общий интеллект» в базе данных Elsevier за 2021 год показывает, соответственно, 24704 публикации против 102. Особое внимание при изучении вопросов развития ИИ уделяется проблемам неформализуемых аспектов сознания, возможности охвата системой ИИ эмоциональных слоев сознания, построения объяснительной составляющей.

---

<sup>1</sup> Sundvall, S. (2019). Artificial Intelligence. In: Critical Terms in Futures Studies. Paul H. (Ed.), 29–34. Palgrave Macmillan, Cham.

<sup>2</sup> Wang, P. (2013). Natural language processing by reasoning and learning. In: International Conference on Artificial General Intelligence, 160–169. Springer. <https://doi.org/10.1007/978-3-642-39521-5>

<sup>3</sup> Wang P, Li X and Hammer P (2018) Self in NARS, an AGI System. Front. Robot. AI 5:20. <https://doi.org/10.3389/frobt.2018.00020>



Значительное количество работ посвящено изучению эмоционального интеллекта, см., например<sup>1,2</sup>. Области чувств, эмоций и бессознательного, которые нельзя формально представить, выходят за рамки описательных и каузальных возможностей языка и мышления. Даже кинестетические способности человека обладают своими познавательными и выразительными возможностями, которые не всегда склонны к формализованной интерпретации.

Изучение объяснимого ИИ (eXplainable AI, XAI) предъявляет такие требования, как создание выводов с четкими пояснениями, обеспечение контроля над новыми объектами<sup>3</sup>, принятие решений с хорошим пояснением для неспециалиста<sup>4</sup>. Есть стремление открыть «черный ящик» для создания точных моделей и дать удовлетворительное объяснение<sup>5</sup> и т. д. Проблемой для систем принятия решений является непонятность моделирования с помощью ИИ, непрозрачность. В статье<sup>6</sup> предложена структура для рассмотрения нелогических объяснений на основе данных, полученных путем запроса к «черному ящику».

Репрезентация сознания становится все более важной при создании ОИИ, на первый план выходят вопросы использования результатов феноменологического анализа субъективной реальности, ее ценностно-смысловой и операциональной структур. В связи с этим ОИИ должен обладать высокой степенью автономности и самостоятельного решения широкого круга задач в различных условиях внешней среды. На Рис. 4-2 представлена компонентная архитектура ОИИ с учетом аспектов субъективной реальности.

---

<sup>1</sup> Cooper, R.K., Sawaf, A. (2000). Executive, EQ: Emotional Intelligence in Business. London, NY: Texere.

<sup>2</sup> Gigerenzer, G. (2007). Gut Feelings. The Intelligence of the Unconscious. London.: Viking.

<sup>3</sup> Chen, M. et al. (2020) Simple and deep graph convolutional networks. International Conference on Machine Learning. 1725–1735.

<sup>4</sup> Wang, J. et al. (2020). Learning node representations from noisy graph structures. Proc. IEEE Int. Conf. Data Mining, ICDM. Vol. 2020-Novem, No 1. P. 1310–1315

<sup>5</sup> Veličković, P. et al. (2019). Neural Execution of Graph Algorithms. 2019. arXiv:1910.10593

<sup>6</sup> Setzu, M., Guidotti, R., Monreale, A., Turini, F., Pedreschi, D., & Giannotti, F. (2021). Glocalx – from local to global explanations of black box ai models. Artificial Intelligence 294, 103457. <https://doi.org/10.1016/j.artint.2021.103457>



Рис. 4-2. Компоненты архитектуры ОИИ (с учетом субъективной реальности)

В этом контексте становятся более заметными ограничения дискретной методологии Тьюринга, которая исключает использование результатов специальных исследований сознания как субъективной реальности. Согласно этой методологии, феномен «интеллект» трактуется в чисто операциональном (вычислительном) смысле, исключая роль сознания. В этой методике любая познавательная или практическая задача формулируется и решается чисто цифровыми методами. В настоящее время при создании ОИИ используются символизм, коннекционизм, бихевиоризм, статистический, квантовый, имитационный и даже нейроморфный подходы<sup>1</sup>, которые опираются на первичное цифровое представление данных. Заложённая в нем парадигма функционализма остается в силе, но требует более широкой смысловой интерпретации с учетом ее роли в объяснении сознания в контексте субъективной реальности.

### 4.3.3. Скептицизм

Некоторые исследователи скептически относятся к привлечению проблемы сознания, считая, что она влечет за собой массу неопределенностей, своего рода туман, что нет смысла прорабатывать

<sup>1</sup> Huang, T.-J. (2017). Imitating the Brain with Neurocomputer. A “New” Way Towards Artificial General Intelligence. International Journal of Automation and Computing. School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China. 14(5), 520-531. <https://doi.org/10.1007/s11633-017-1082-y>

феномен сознания в традиционном контексте ИИ, это только усложняет ситуацию. И, возможно, дефицит функции сознания в перспективных вариантах ИИ будет компенсирован за счет ускорения работы компьютера, например, с помощью квантовых или оптических вычислений. Однако основной проблемой традиционной вычислительной техники остается проблема возможности учета любого вида сознания. В работе<sup>1</sup>, например, предлагается делать это путем оптической обработки аналоговых (нецифровых) сигналов с бесконечными спектрами, с учетом нелокальных и релятивистских эффектов поведения нейронов человеческого мозга на атомарном уровне и т.д.

В центре внимания все чаще находятся вопросы теории и методологии междисциплинарности, интеллекта, мышления, познания и сознания, объяснения и понимания, соотношении эмпирического и теоретического знания, отношения между искусственным и естественным интеллектом. Первостепенная роль в их осмыслении принадлежит философам – специалистам в области эпистемологии, методологии науки, феноменологии, изучения сознания в широком смысле, включая те направления, которые часто относят к аналитической философии сознания.

На начальных этапах создания ОИИ это возможно при условии адекватного выделения и описания операционных аналогов тех или иных состояний субъективной реальности. В то же время при описании аналогов таких состояний сознания, как «переживание», «самоанализ», «мысли» или «образ себя как части картины мира», классический подход к развитию ИИ сталкивается с непреодолимыми трудностями. Эти понятия сложны, многомерны, дополняют друг друга по содержанию и не поддаются формальному описанию. Каждый из них включает в себя бесконечный набор аналитических и синтетических оперативных действий в когнитивном плане.

Исследования сознания как его субъективной реальности позволяют сформулировать основные системные, структурно-функциональные и операциональные характеристики естественной психической деятельности человека, а значит, представляют собой незаменимый ресурс моделирования познавательных процессов, отвечающих задачам здание ОИИ.

Аспект субъективной реальности создает еще неразрешенные теоретические трудности в решении проблемы сознания. Они связаны с

---

<sup>1</sup> Raikov, A. (2021). Cognitive Semantics of Artificial Intelligence: A New Perspective. Topics: Computational Intelligence XVII, Springer, Singapore. <https://doi.org/10.1007/978-981-33-6750-0>

тем, что физические свойства могут быть отнесены к субъективной реальности только в том случае, если субъектом является физическая машина, создающая реальность. Возникает вопрос, как, не прибегая к идеалистическим постулатам о наличии духовной субстанции и не прибегая к физикалистской редукции, построить теоретически правильное объяснение сознания в форме, доступной для понимания и оперирования системой ИИ, учитывая его кажущуюся связь с физическими процессами.

Было предложено использовать нередукционистский метод решения в рамках двух основных категориальных планов: онтологического и эпистемологического<sup>1</sup>. Первый помогает решить вопросы деятельности сознания, а также объяснить связь явлений субъективной действительности с деятельностью мозга, второй – осуществить переход от индивидуально-субъективного опыта к полисубъектной саморазвивающейся (рефлексивно-активной) среде.

При изучении систем ИИ в процессах управления возникают проблемы, связанные с ломкой парадигм, определяющих представления об ИИ без учета аспектов управления. Классические парадигмы ИИ рассматривают человека не как средство, включенное в управляющую деятельность, а как независимый от контуров управления объект. Эти парадигмы построения систем ИИ имеют такие разделы, как:

- морфологическое представление об ИИ как модели мозга;
- логическое понимание ИИ как средства решения проблем;
- нейросетевая концепция ИИ;
- имитация – дискретное подобие человеческому мышлению;
- нейроморфный – полагающий, что построение физического подобия мозга создаст ОИИ;
- перспектива – компьютерное моделирование человеческой деятельности, эмоций, мыслей и мышления и др.

Из вышеперечисленного поясним только нейроморфную парадигму, которая довольно смело претендует на создание ОИИ<sup>2</sup>. В указанной работе предлагается использовать метод «обратного инжиниринга» – построить ОИИ, буквально скопировав нейронную структуру мозга, с надеждой, что копия воспроизведет его функции. Однако, нетрудно

---

<sup>1</sup> *Lepskiy, V.* (2018). Evolution of cybernetics: philosophical and methodological analysis, *Kybernetes*. 2018, Vol. 47 Issue: 2, pp. 249-261. <https://doi.org/10.1108/K-03-2017-0120>

<sup>2</sup> *Huang, T.-J.* (2017). Imitating the Brain with Neurocomputer. A “New” Way Towards Artificial General Intelligence. *International Journal of Automation and Computing*. School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China. 14(5), 520-531. <https://doi.org/10.1007/s11633-017-1082-y>

заметить, что этот подход продолжает эволюционную цепочку развития методологий ИИ – символизма, коннекционизма, бихевиоризма и статистизма – создавая на той же логической и формализованной основе новую версию слабого ИИ: интуитивизм.

Перечисленные парадигмы способствуют развитию отдельных направлений развития ИИ, но затрудняют использование ИИ в процессах управления, поскольку недооценивают специфику субъекта в контуре управления. В частности, за скобками остается концептуальный и когнитивный аспекты управления. Отдельной проблемой является установление интерфейса между технологическими, основанными на традиционных парадигмах ИИ, и субъективными, основанными на психических и поведенческих особенностях человека, аспектами систем управления с использованием систем ИИ. В основе процессов управления лежат их специализированные парадигмы (целенаправленная, рефлексивная, субъективная, кибернетическая и др.), которые выдвигают определенные требования к идеям ИИ и задачам, в которых целесообразно использовать ИИ.

Анализ тенденций развития вопросов управления с позиций развития научной рациональности (классической, неклассической, постнеклассической) позволяет сделать вывод о том, что они в первую очередь связаны с развитием соответствующих субъектных парадигм управления<sup>1</sup>. Саморазвивающиеся системы находятся в центре внимания проблем управления. Доминирующей субъективной парадигмой в управлении становится парадигма «саморазвивающейся полисубъектной среды». Эта парадигма наделена свойствами субъектности, при которой субъекты могут взаимодействовать посредством рефлексии.

К основным инвариантным свойствам субъектов следует отнести целеустремленность, рефлексивность, коммуникабельность, социальность и способность к развитию. Сосуществование субъектов естественного интеллекта и систем ИИ образует гибридную среду, то есть реальность совместного действия естественного и искусственного интеллекта. В этой среде происходят взаимодействия разнородных образований и их объединение в целостные типы конструкций.

Существенным свойством субъекта является рефлексия. Взаимодействие и интеграция природных субъектов и систем ИИ в гибридной среде должны определять порядок их рефлексивной координации. Принцип «двойного субъекта» имеет большое значение

---

<sup>1</sup> Lepskiy, V. (2018). Evolution of cybernetics: philosophical and methodological analysis, *Kybernetes*. 2018, Vol. 47 Issue: 2, pp. 249-261. <https://doi.org/10.1108/K-03-2017-0120>

для понимания специфики саморегуляции рефлексивной деятельности. Его суть заключается в постоянном совершенствовании различных видов деятельности (мыслительной, деятельностной, коммуникативной, рефлексивной) субъектов посредством взаимодействия с партнерами, имеющими адекватные позиции, организации на основе анализа деятельности субъектов их близнецов, в том числе «цифровых»<sup>1</sup>.

Учет принципа двойного субъекта при организации гибридных систем позволит создавать системы с включением элементов ИИ, не выходящие из-под контроля их создателей и адекватно адаптирующиеся к изменениям внешней среды. Это соответствует созданию ОИИ, учитывающего особенности саморазвивающейся полисубъектной (рефлексивно-активной) среды и субъективной реальности. При этом проектируемые системы ОИИ должны создавать необходимые условия для обеспечения ускоренного сближения (конвергенции) позиций участников в принятии решений (коллективного разума) относительно целей и способов действий, даже если цели не ясны.

Конвергентный подход к созданию ОИИ обеспечивает ускорение коллективного принятия решений в гибридной реальности, в том числе в чрезвычайных ситуациях<sup>2</sup>. Структурирование информации осуществляется в ходе коллективного совещания, например, стратегического. Это может помочь достичь консенсуса среди членов команды быстрее, чем обычно. Подход помогает генерировать идеи в саморазвивающейся полисубъектной (рефлексивно-активной) среде. Конвергентная технология объединяет решение обратных задач, теорию топологии и категорий, когнитивное моделирование, анализ больших данных, генетические алгоритмы и т. д.

Динамика процессов коллективного принятия решений сочетает в себе аналитическую и синтезирующую составляющие. Первая характеризуется расходимостью процесса, а вторая – сходимостью. Эти процессы не могут быть полностью формализованы, поскольку они могут быть вызваны эмоциями, переживаниями, желаниями и мыслями участников, которые частично может охватывать хорошо известная модель убеждения-желания-намерения<sup>3</sup>. Эти аспекты ограничивают

---

<sup>1</sup> *Lepskiy V.* (2021) Reflexivity and Artificial Intelligence in Control (Subjectness-oriented Approach). IFAC-PapersOnLine. Volume 54, Issue 13, 2021, Pages 221-226. <https://doi.org/10.1016/j.ifacol.2021.10.449>

<sup>2</sup> *Raikov, A.* (2021). Convergent Ontologization of Collective Scientific Discoveries. 14th International Conference Management of large-scale system development (MLSD), 1-5, <https://doi.org/10.1109/MLSD52249.2021.9600184>

<sup>3</sup> *Umbrello, S., Yampolskiy, R.V.* (2021). Designing AI for Explainability and Verifiability: A Value Sensitive Design Approach to Avoid Artificial Stupidity in Autonomous Vehicles. Int J of Soc Robotics. <https://doi.org/10.1007/s12369-021-00790-w>

ускорение достижения целей коллективного принятия решений и должны учитываться при использовании ОИИ при таком принятии решений.

Учет когнитивной семантики моделей ОИИ помогает представить мыслительные события в едином пространстве. В этом случае когнитивная семантика снабжает неформализуемые явления предмета или человека символическим описанием, описываемым логикой, законами квантовой электродинамики и специальной теорией относительности. Теория категорий и теория топологии с использованием метода решения обратных задач, унифицированных операторов обеспечивают реализацию преобразований мыслительных событий и целенаправленность коллективного принятия решений с учетом когнитивной семантики моделей ОИИ. Теории категорий и топологии помогают изучать свойства отношений объектов без их логических структур. В этом случае когнитивная семантика моделей ОИИ соответствует перечню правил, например, таким, как:

- цели должны быть распределены по уровням и ранжированы по важности;
- решаемая задача должна быть разбита на бесконечное число составляющих;
- отображение ресурса команды на цели должно быть замкнутым;
- каждый элемент ресурса команды должен быть отделен от соседнего элемента;
- следует разработать механизмы свободной коммуникации между участниками;
- когнитивная семантика должна быть представлена возможностью участников влиять на процесс с использованием онтологических фреймов;
- модели ОИИ должны поддаваться проверке путем сопоставления с соответствующими большими данными.

Эти правила могут быть в перспективе расширены за счет учета нелокальных эффектов для учета когнитивной семантики, которая должна погружать модели ОИИ в саморазвивающуюся полисубъектную (рефлексивно-активную) среду решаемой задачи.

Таким образом, традиционный ИИ и ОИИ – явления, скорее всего, из разных пространств. Однако, их связывает стремление ускорить эволюционный тренд развития ИИ, за счет использования такого соотношения, как разделения материи и антиматерии. Это соотношение, возведенное в принцип, сочетает в себе взаимодополняющие аспекты технического ИИ и субъективного ОИИ для создания явлений,

направляющихся в довольно новую субъективную и рефлексивно-активную реальность, которая развивается конвергентным (сходящимся) образом, чтобы получить синергию и необычную для человека силу.

Отдельные элементы ОИИ появляются уже в настоящее время. Эти элементы объединяют методы изучения субъективной реальности, проектирования в полипредметных саморазвивающихся (рефлексивно-активных) средах, эпистемологии и феноменологии, решения обратных задач на понятийных пространствах, термодинамики, квантовой и релятивистской физики, биологии, оптики и др. Эта комбинация дисциплин уже помогает «приблизиться» к недоступным слоям сознания, преодолевает ограничения дискретных вычислений, помогает выстроить объяснительную составляющую ИИ. Элементы ОИИ уже позволяют ускорить коллективное согласование решений и быстро создавать командные стратегии в реальной практике.

#### **4.3.4. Выводы**

На основе изложенного в этом разделе, можно выделить следующие основания для построения искомых социогуманитарных критериев:

- при создании ИИ все больше уделяется внимания вопросам теории и методологии междисциплинарности, интеллекта, мышления, познания и сознания, объяснения и понимания, соотношении эмпирического и теоретического знания, отношения между искусственным и естественным интеллектом.
- методология создания ОИИ должна базироваться на идеях субъектно-онтологического подхода, функционализма, феноменологии субъективной реальности, конвергентных когнитивных архитектур, а также методов создания саморазвивающейся полисубъектной (рефлексивно-активной) среды.
- К основным инвариантным свойствам субъектов следует отнести целеустремленность, рефлексивность, коммуникабельность, социальность и способность к развитию. Сосуществование субъектов естественного интеллекта и систем ИИ образует гибридную среду, то есть реальность совместного действия естественного и искусственного интеллекта



## 4.4. НЕЙРОТЕХНОЛОГИИ В СИСТЕМАХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И ПРИМЕНЕНИЯ В СФЕРЕ УПРАВЛЕНИЯ

### 4.4.1. Введение

Среди важнейших и быстро развивающихся технологий систем искусственного интеллекта (СИИ) – нейротехнологии. Это класс технологий, которые помогают понять работу мозга, мыслительные процессы, высшую нервную деятельность, в том числе технологии по усилению, улучшению работы мозга и психической деятельности. Следует обратить внимание на введенный в действие еще в 2009 г. стандарт<sup>1</sup>, описывающий практические вопросы работы операторов сложных технических систем, а также их взаимодействия с социумом. Уже 5 лет назад прогнозировался экспоненциальный рост рынка нейротехнологий и соответствующей продукции после 2025 г. до 1,8 трлн долл. в 2035 г.<sup>2</sup> Нейронет является существенной частью ансамбля процессов цифровой трансформации. Реализуется множество поисковых и экспериментальных исследований и разработок в данной области во всех ведущих технологических державах мира.

Нейротехнологии можно разделить на несколько групп:

1) инвазивные, предполагающие вживление электродов в человеческое тело, чаще всего в головной мозг; главный недостаток данного типа технологий, несмотря на существенные достигнутые результаты, – вмешательство во внутреннюю среду организма;

2) миосенсоры, предполагающие размещение электродов на коже человека и считывание импульсов, проходящих через мышечные волокна; данный тип датчиков имеет существенную задержку действия, до нескольких секунд из-за разницы скорости нервных и мышечных импульсов;

3) неинвазивные, которые основаны на различных способах записи электрической активности мозга с помощью внешних устройств и, соответственно, – воздействия на работу мозга; их основной недостаток – неспецифичность получаемой информации и сложности ее интерпретации, что требует наработки больших массивов данных с последующим определением критически значимых данных<sup>3</sup>, и, соответственно, проблема калибровки воздействующего импульса.

---

<sup>1</sup> ГОСТ Р 43.0.3-2009. Информационное обеспечение техники и операторской деятельности. Ноон-технология в технической деятельности. Общие положения. М.: Стандартинформ, 2010.

<sup>2</sup> Нейроинтерфейсы для сервисов и продуктов нового поколения [Электронный ресурс] // Basisneuro. 2017. Ноябрь. URL: <https://basisneuro.com/BasisNeuroWhitePaper.pdf>.

<sup>3</sup> Neurogress: платформа систем нейроуправления от участников проекта BlueBrain [Электронный ресурс] // ITnan. URL: <https://itnan.ru/post.php?c=1&p=348498>.

Нейротехнологии в настоящее время позволяют снимать и идентифицировать сигналы мозга для дальнейшей их обработки в привязке к когнитивным, нервным или иным реакциям организма, а также способы улучшения и исправления функций мозга, включая восстановление или замещение памяти, стирание негативных воспоминаний и пр.<sup>1, 2, 3.</sup>

Однако человеческий мозг представляет собой сети из нескольких сотен триллионов синапсов, объединяющих десятки миллиардов нейронов. Поэтому создание полноценного, имитирующего работу человеческого мозга с высоким приближением, устройства нейроинтерфейса требует решения множества сложных задач в области нейробиологии, синтетической биологии, маломощной электроники, фотоники, медицинской техники и т.д.<sup>4, 5, 6</sup> Необходимо также развитие математических методов перевода сенсорной информации из формы, в которой она представлена в нейронах мозга, в форму, пригодную для обработки на компьютерах<sup>7, 8.</sup>

Технологически Нейронет – это технический пакет, обеспечивающий связность естественных субъектов (личностей) и искусственных агентов при помощи протоколов передачи, обмена и синтеза знаний. Очаги Нейронета возникают, с одной стороны, в сетях биометрического веба, насыщаемого новыми протоколами связи, новыми типами устройств, а также новыми приложениями. С другой,

---

<sup>1</sup> *Логинов Е.Л., Логинова В.Е., Шкута А.А.* «Дизайн мышления» элементов искусственного интеллекта для преодоления барьеров получения нового знания в электронной среде коллаборационной научной суперсистемы // Искусственные общества. 2018. № 3. 5 с.

<sup>2</sup> *Овод И.В., Осадчий А.Е., Пупышев А.А., Фрадков А.Л.* Формирование нейрообратной связи на основе адаптивной модели активности головного мозга // Нейрокомпьютеры: разработка, применение. 2012. № 2. С. 36–41.

<sup>3</sup> *Туровский Я.А., Кургалин С.Д., Адаменко А.А.* Моделирование обучения нейрочипов, внедренных в нервную ткань // Цифровая обработка сигналов. 2016. № 1. С. 50–56.

<sup>4</sup> *Галушкин А.И.* Нейрочипы и нейроморфные ЭВМ: проблемы моделирования // Информационные технологии. 2015. Т. 21. № 12. С. 942–949.

<sup>5</sup> *Калинин П.В., Воюцкая Ю.Ю., Тарасов М.Е.* О применении нейроинтерфейса для бесконтактного управления мобильным устройством // Информационные системы и технологии. 2016. №3 (95). С. 53–56.

<sup>6</sup> *Тычков А.Ю., Горячев Н.В., Кочегаров И.И.* Протоколы связи для беспроводного нейроинтерфейса // Труды Международного симпозиума «Надежность и качество». 2018. Т. 2. С. 366–368.

<sup>7</sup> DARPA профинансирует создание скоростного интерфейса между компьютером и человеческим мозгом [Электронный ресурс] // Open Systems Publications. URL: <https://www.osp.ru/news/2016/0126/13031430>.

<sup>8</sup> *Райков А.Н.* Конвергентный синтез когнитивной модели на основе глубокого обучения и квантовых семантик // International Journal of Open Information Technologies. 2018. Т. 6. № 12. С. 43–50.

запрос к Нейронету исходит из областей с наибольшими требованиями к коллективной деятельности при решении сложных задач (научные и инженерные проекты).

Развиваются стандарты управления коллективной деятельностью. Важную роль в возникновении таких коллективов будет играть технология «экзокортекса» – искусственных частей психики, поддерживаемых машинами и синхронизируемых с естественной психикой<sup>1</sup>.

Приоритетные проекты, реализуемые в США, ЕС, Китае, Японии, находятся на стыках линий развития «нейро-когнитивные науки», «информационно-коммуникационные технологии», «социальная инженерия», рынков теле- и нейромедицины и интерфейсов «мозг – компьютер».

Требуют внимания и более общие вопросы текущего и перспективного статуса данной тематики, влияния нейротехнологий на социальную жизнь и наоборот<sup>2</sup>.

#### **4.4.2. Нейротехнологии в управлении социализацией и поведением человека**

Экспансия нейротехнологий является составной частью всеобъемлющего сдвига в формах и способах первичной и вторичной социализации людей. Формирование личности в социуме всегда происходит в процессе динамического взаимодействия людей (семьи, рода, множества малых групп, государства и других надличностных систем). В эволюции личности сложным образом проецируется социум, подобно биологическому закону о повторении онтогенеза филогенезом при сохранении определенного диапазона вариативности. Индивидуализация происходит, во-первых, биологически (тело, функциональные системы); во-вторых, информационно (первичная и вторичная социализация с освоением видов грамотности, коммуникативных связей, накоплением эрудиции, умений и навыков); в-третьих, когнитивно (знания, чувствования, понимания, базовые ценностные ориентации, мировоззрение); в-четвертых, социально (членство в различных группах с возможностью множественной

---

<sup>1</sup> Подходы к формированию и запуску новых отраслей промышленности в контексте Национальной технологической инициативы на примере сферы «Технологии и системы цифровой реальности и перспективные „человеко-компьютерные“ интерфейсы (в части нейроэлектроники)»: Аналитический доклад [Электронный ресурс]. URL: <http://rusneuro.net/cambiodocs/media/files/analitijeskii-doklad-podhodyk-formirovaniu-i-zapusku-novyh-otraslei-promyhlennosti.pdf>.

<sup>2</sup> Агеев А.И., Логинов Е.Л. Нейроменеджмент личности. 2-е изд. М: ИНЭС, 2022.

идентификации и мотивации деятельности)<sup>1</sup>. Между этими видами индивидуализации существует взаимосвязь. Нарушение какого-либо из них может нарушать целостность бытия личности, но имеется и потенциал компенсации ослабления или выбытия той или иной подсистемы.

Разделение социализации на первичную и вторичную связано прежде всего с особенностями биологической возрастной эволюции человека, осваивающего львиную долю ключевых навыков жизни в социуме в первые годы жизни и в юности. Пропуск этого этапа необратим для человека, как показывают примеры человеческих существ, выросших во внечеловеческой среде (Маугли и др.). Вторичная социализация завершается с достижением базовых функциональных параметров личности зрелых значений (обычно – в начале ранней взрослости, иногда до 30–40 лет). Имеет смысл, по всей видимости, в рассматриваемом контексте ввести понятие третичной социализации, происходящей в период средней взрослости и далее (40+), учитывая заметный рост ожидаемой продолжительности жизни и возможности существенных изменений личностных характеристик в этот период.

В свою очередь поведенческая активность как отдельных личностей, так и их групп укрупненно определяется воздействием группы факторов:

- биофизические факторы, формирующие режимы работы мозга и нервной и других функциональных систем человека;

- психосемантические характеристики личности, определяющие характер восприятия, активности и реализации жизненных интересов (базовых установок) личности по отношению к внешним раздражителям;

- импринтированные рефлексивные матрицы – матрицы ключевых рефлексивных реакций сознательного и бессознательного характера, определяющие интерпретацию событий (поступающей информации);

- информационные раздражители – информация, поступающая из внешней по отношению к личности, среды во взаимосвязи с коммуникативными характеристиками каналов поступления информации<sup>2</sup>.

---

<sup>1</sup> Агеев А.И., Логинов Е.Л., Шкута А.А. Конвергентный мониторинг и программирование личности как инструмент оперирования интеллектуальной динамикой поведения больших групп людей // Экономические стратегии. 2018. № 2. С. 70–87; Агеев А.И., Логинова Е.Л., Шкута А.А. Нейроуправление: конвергентная интеграция человеческого мозга и искусственного интеллекта [Электронный ресурс] // Экономические стратегии. 2020. № 6. С. 46–57. DOI: 10.33917/es-6.172.2020.46-57.

<sup>2</sup> Агеев А.И., Логинов Е.Л., Шкута А.А. Конвергентный мониторинг и программирование личности как инструмент оперирования интеллектуальной динамикой поведения

Выявлен, в частности, ряд закономерностей в амплитудно-частотных и пространственно-временных перестройках биоэлектрической активности мозга. Обнаружено, что десинхронизация, массово распространившаяся в урбанизированной информационной среде, лежит в основе большинства вегетативных расстройств, составляющих физиологическую основу развития нейрофизиологического и психоэмоционального стресса<sup>1</sup> и сопутствующих заболеваний.

Выявлено также, что влияние интенсивных электромагнитных полей и соответствующих частотных диапазонов с учетом биоритмики конкретного биологического объекта способно ввести человека и группы людей в измененное состояние сознания, близкое к трансу<sup>2</sup>. Такое состояние, помимо прочего, блокирует большинство сознательных реакций человека (в том числе саморефлексии, самоконтроля, самоидентичности) в случае расхождения получаемой, но внешне убедительной и кажущейся достоверной информации и окружающей реальности. Тем самым облегчается воздействие на человека, в том числе с использованием для этого элементов нейролингвистического программирования с коррекцией или даже полной заменой матрицы ключевых рефлексивных реакций человека с соответствующим изменением и закреплением новой модели интерпретации происходящих событий. Резко увеличившееся с цифровизацией количество мошеннических финансовых действий показывает успешное применение именно таких технологий. Известны примеры, когда внешнее влияние мошенников могло длиться до нескольких дней, пока вполне образованный человек не осознавал преступный характер оказываемого на него влияния и критическую внешнюю зависимость своих поступков.

Исследования психосемантических качеств личности ведутся в рамках множества научных направлений, их результаты используются в

---

больших групп людей // Экономические стратегии. 2018. № 2. С. 70–87.

<sup>1</sup> Бутова О.А., Гришко Е.А. Особенности формирования биоэлектрической активности нейронов головного мозга военнослужащих ставропольского гарнизона в аспекте адаптации // Наука. Инновации. Технологии. 2009. № 4. С. 235–241.

<sup>2</sup> Измененные состояния сознания (ИСС) – качественные изменения в субъективных переживаниях или психологическом функционировании от определенных генерализованных для данного субъекта норм, рефлекслируемые самим человеком или отмечаемые наблюдателями (классическое определение Арнольда Людвига). Согласно А. Ревонсуо, главным характерным признаком измененных состояний сознания являются системные изменения (относительно нормального состояния сознания) связи содержания переживаний с реальным миром, то есть в ИСС присутствуют искажения представления внешней реальности или осознания себя в виде иллюзий, причем эти искажения складываются в глобальное изменение репрезентаций (Измененное состояние сознания [Электронный ресурс] // Википедия. URL: <https://ru.wikipedia.org/wiki>)

медицине, криминалистике и множестве других практических приложений. Так, анализ интересов и предпочтений личности в отношении поиска и просмотра информационных и развлекательных медийных программ, компьютерных игр, активности в социальных сетях, финансовое поведение и т.д. позволяет сгенерировать и структурировать обширные массивы данных электронного контента («большие пользовательские данные»). Результатом такого анализа может стать когнитивно-рефлексивная модель конкретной личности, получаемая другим заинтересованным лицом. В модель могут быть встроены элементы нейролингвистического программирования, что позволяет дистанционно и латентно создать для целевого объекта воздействия шаблоны восприятия и интерпретации происходящего потока событий и что способно стать источником мотивации и действий данной личности. При этом возможны как индивидуализированные действия по созданию заданных параметров поведения конкретной личности, так и кластеризация воздействия на целевые группы. Иными словами, цифровые модели личностей могут быть препарированы по целевым функциям и по критическим значениям параметров, сведены в кластеры, в отношении которых могут быть применены единые информационные воздействия<sup>1, 2, 3</sup>.

В настоящее время с учетом конвергенции информационных, телекоммуникационных и вычислительных сервисов в глобальных информационных сетях практически любое обращение пользователя к электронному контенту вне зависимости от оборудования, коммуникационного канала и способа связи становится частным случаем обращения к единой распределенной базе электронных данных. Следы обращения к электронному контенту сохраняются на длительный срок, позволяя применить наработанные методы мониторингового анализа, идентификации пользователя, прогнозирования реакций и поведения, выявления связей с другими пользователями, принадлежности к явным и неявным группам и пр.

---

<sup>1</sup> *Волынский-Басманов Ю.М.* Применение методов нейролингвистического программирования для выявления потенциально опасных лиц // Проблемы безопасности и чрезвычайных ситуаций. 2010. № 5. С. 124–128.

<sup>2</sup> *Ковалевская А.В.* Информационные войны: классификация суггестивной специфики // Theoretical and practical problems of language tools transformation in the context of the accelerated development of public relations / Peer-reviewed materials digest (collective monograph) published following the results of the CXVIII International Research and Practice Conference and I stage of the Championship in Philology. Chief editor V.V. Pavlov. 2016. С. 23–25.

<sup>3</sup> *Кузнецов В.* Использование нейролингвистического программирования (НЛП) при допросе // Право и жизнь. 2011. № 152 (2). С. 134–140.

Современный компьютер, смартфон или пульт телевизора с интеллектуальными функциями позволяют не только по факту, но и по манере и скорости нажатия на клавиши точно идентифицировать человека, определять знак его отношения к текущему электронному контенту на экране, выявлять уровни симпатии, раздражения или агрессии по отношению к этому контенту. Эти же устройства могут фиксировать не только геолокацию человека, но и увязывать ее с показателями интенсивности электромагнитных полей в этой точке, изменять параметры оборудования, находящегося вблизи этой точки, влияющие на электромагнитные поля. Гаджеты могут фиксировать температуру, артериальное давление, ритмы сердцебиения и иные биоритмы конкретного человека. Результаты мониторинга дают возможность наблюдения статуса и динамики психосемантической субъектности личности в условиях неоднозначности или недостатка информации о ней для принятия того или иного прагматичного решения (реклама, выборы и т.п.), анализа и прогнозирования риска смены состояний активности в обычных и в чрезвычайных условиях<sup>1,2</sup>.

Сведение данных из всех возможных баз электронного контента в единый пакет информации о психосемантической субъектности личности, объединяющем структурируемые, сложноструктурируемые и условно структурированные данные, позволяет определить характеристики ее поведенческой активности, включая и скрываемые качества (например, социопатические наклонности, политические или религиозные пристрастия, фобии, способы неформальной самореализации, сетевые и иерархические связи, ролевую структуру в семье и сообществах, объем оперируемых ресурсов, принадлежность к определенным группам и сетям, и т.п.)<sup>3</sup>.

Развитие цифровых технологий позволяет не только интегрировать текстовые данные с видео и изображениями, но реконструировать и прогнозировать в 3D- и 4-D-форматах ход событий, отслеживать динамические изменения поведения личности и связанных с ней групп так, что возникает агрегированный информационный кластер мировоззренческих и профессиональных шаблонов интерпретации

---

<sup>1</sup> Ляхов А.Ф., Тришин И.М. Компьютерное моделирование поведения игрока в интеллектуальной карточной игре с помощью нейронной сети // Компьютерные инструменты в образовании. 2013. № 5. С. 54–64.

<sup>2</sup> Самарцев О.Р., Латенкова В.М. Психосемантические аспекты восприятия интерактивного дискурса в интернет-СМИ // Вестник Череповецкого государственного университета. 2016. № 2 (71). С. 87–91.

<sup>3</sup> Пономарева О.С., Устюжанин В.Н. О состоянии и перспективах использования психосемантических методов познания личности подозреваемого в деятельности следственного работника // Вестник Санкт-Петербургского университета МВД России. 2016. № 2 (70). С. 190–194.

окружающей действительности и вытекающего из этого поведения. На этой основе складываются новые инструменты управления.

Быстро наращиваемый потенциал оперирования обобщенными семантическими структурами создает беспрецедентные возможности для перевода внешнего целенаправленного воздействия с сознательного на бессознательный для объекта уровень. Это достигается путем неявной импринтации отдельным личностям смысловых мотивационных «якорей», побуждающих к тем или иным оценкам и затем поступкам. На этой основе можно комплексировать, вызывая эффекты резонанса, множество индивидуальных действий в информационном и реальном пространстве: от ажиотажа на товарных рынках до политических митингов.

По существу, быстро развивается интеграция различных методов психоинжиниринга, дистантного воздействия, психокоррекции и психозондирования с синхронизацией психофизиологических и психосемантических воздействий и конвергенцией процесса восприятия информационных раздражителей. Концентрация внимания личности и интерпретация информации сводятся в единую композицию в рамках интерпретации человеком картины окружающего реального и выдуманного мира с вытекающими отсюда линиями поведения.

Новые возможности нейрокогнитивного управления создает применение СИИ. Цифровая модель (цифровой двойник) способна симулировать все возможные режимы жизнедеятельности человека, учитывать влияние внешних факторов и процессов управления, позволяет предсказывать будущее состояние и поведение физического объекта. Цифровой двойник основан на технологиях ИИ, машинного обучения и аналитического программирования. Цифровой двойник непрерывно обучается и обновляет свои параметры, получая информацию от множества сенсоров, представляет состояние физического объекта. При обучении им используются текущие данные от сенсоров, устройств управления, внешней среды; он объединяет фактические данные со знаниями, полученными от инженеров, опытных специалистов в данной области. Цифровой двойник использует исторические данные, накопленные на предыдущих этапах.

По сути, в настоящее время традиционные методы управления социумом превращаются в единую систему с новыми коммуникационными интерфейсами, нейро- и биоинтерфейсами. Для разработки методов прогнозирования интеллектуальной динамики поведенческой активности накоплен значительный исследовательский опыт<sup>1</sup>. Имеется множество перспективных концепций создания

---

<sup>1</sup> Агеев А.И., Логинов Е.Л. Битва за будущее: кто первым в мире освоит ноомониторинг и 208



многофункциональной информационной мониторинговой системы как платформы прогнозирования (с обратной связью) явных и неявных глубинных процессов и тенденций в социуме, техносфере и природной среде.

Отдельный исследовательский и практически значимый сюжет – исследование свойств кластеров факторов биофизического и информационно-когнитивного характера и прогнозирование выхода интеллектуальной и поведенческой активности за пределы устойчивых состояний как источника повышенных рисков возникновения чрезвычайных ситуаций<sup>1</sup>. С этой проблематикой связана задача комплексирования различных моделей поддержки социальной среды, лояльной к правовым нормам, управленческим ключевым установкам (в том числе и цифровым, анонимным) и глубинным регуляторам жизнедеятельности («матрикс»). Особый вопрос – риски сбоев в работе киберфизических систем, которые могут быть вызваны дистанционным способом. При этом сбой может иметь как программно-вирусное происхождение, так и антропогенное – через воздействие на сознание оператора. Аналогичным образом могут быть инспирированы террористические акты в исполнении персонажей, находящихся под явным или латентным внешним управлением. Современный уровень технологий позволяет осуществлять как зондаж, так и воздействие с учетом психосемантических качеств личности расширенного характера (официальной и реальной политической ориентации, качества ее профессиональной подготовки, культурного уровня, интересов, волевых качеств, внутренней мотивации, состояния здоровья и т.п.).

---

когнитивное программирование субъективной реальности? // Экономические стратегии. 2017. № 2. С. 124–139; *Агеев А.И., Логинов Е.Л., Шкута А.А.* Конвергентный мониторинг и программирование личности как инструмент оперирования интеллектуальной динамикой поведения больших групп людей // Экономические стратегии. 2018. № 2. С. 70–87; *Агеев А.И., Логинов Е.Л.* Нейроменеджмент личности. 2-е издание (в процессе производства); *Лепский В.Е.* Эволюция представлений об управлении (методологический и философский анализ). М.: Когито-Центр, 2015. 107 с.; *Лефевр В.А.* Рефлексия. М.: Когито-Центр, 2003. 495 с.; *Райков А.Н.* Моделирование коллективного бессознательного при принятии решений // Труды Международной научной конференции СРТ-2014 Международная научная конференция Московского физико-технического института (государственного университета), Института физико-технической информатики. М.: Институт физико-технической информатики, 2015. С. 146–156; *Смирнов И., Безносок Е., Журавлев А.* Психотехнологии: Компьютерный психосемантический анализ и психокоррекция на неосознаваемом уровне. М.: Издательская группа «Прогресс» – «Культура», 1995. 416 с.; *Холодов Ю.А.* Мозг в электромагнитных полях. М.: Наука, 1982. 123 с. и др.

<sup>1</sup> *Чухрова М.Г., Чухров А.С.* Пространственно-временная организация биоэлектрических процессов мозга как индикатор психосоциальной адаптации // Мир науки, культуры, образования. 2013. № 5 (42). С. 227–230.

Новейшие поколения технических устройств дают возможность заинтересованной стороне, опираясь на автоматизированные системы сбора, накопления, обработки и использования данных, не только идентифицировать и определять геолокацию обладателя устройства, его эмоциональное отношение к содержанию электронных сообщений, но и выявлять характеристики окружающего оборудования. Получение информации о состоянии личности через «гаджеты здоровья» дают возможность при необходимости прогнозировать переходы состояний личности в нормальных и в чрезвычайных условиях. Выявление особенностей личности по характеристикам просматриваемых ею информационных программ, активности в социальных сетях, выбору компьютерных игр и т.п. (портфель данных индивидуального электронного контента) позволяет сформировать когнитивно-рефлективную модель личности. На основе такой модели возможно нейропрограммирование мировоззренческих и ситуативных ориентиров и актов поведения личности и групп людей. Факты испытаний таких технологий встречаются практически во многих современных локальных конфликтах.

Сведение полученных данных из всех возможных форм электронного контента в пакет информации о психосемантической субъектности личности позволяет с высокой степенью достоверности обнаружить ее поведенческие доминанты и скрываемые качества, а также принадлежность к социопатической группе. Объектом управления человеком становится его же цифровой двойник, через воздействие на параметры которого можно корректировать его реальное поведение, мышление, интерпретации событий и процессов. При этом сам двойник непрерывно актуализируется по мере «электронной активности» самого человека. Применение СИИ позволяет осуществлять всю необходимую предиктивную аналитику двойника и его прототипа.

Особенно важно прогнозирование и упреждающее исключение возможности резонанса этих факторов на базе технологий дистантного воздействия, психокоррекции и психозондирования с учетом психосемантических качеств личности расширенного характера (официальной и реальной политической ориентации, качества ее

профессиональной подготовки, культурного уровня, интересов, волевых качеств, внутренней мотивации и т.п.)<sup>1,2,3</sup>.

Пиковые проявления поведенческой активности могут быть представлены как внешние и внутренние проявления векторов резонанса биофизических и информационно-когнитивных факторов. Расчет векторов этого резонанса представляется нетривиальной проблемой<sup>4,5,6</sup>, различных по качественному содержанию, размерности и иным тому подобным показателям<sup>7</sup>.

#### 4.4.3. Выводы

Нейротехнологии представляют собой один из самых быстро растущих сегментов всеохватывающего мира систем ИИ. Ближе, чем другие ветви СИИ, нейротехнологии и Нейронет в целом сталкиваются с миром сугубо человеческого социума, переживающего фундаментальные трансформации, в том числе цифровые. Появились и быстро совершенствуются новые технологии воздействия на индивидуальное и коллективное сознание. Есть все основания с особым вниманием отнестись к проблеме подлинно осознающих субъектов, являющейся одной из ключевых для формирующегося цифрового социума.

Приближение уровня развития СИИ к возможности оперировать в различных контурах знаний и корректировать целеполагание своей

---

<sup>1</sup> *Василевская Е.А., Менделевич В.Д.* Взаимосвязь между социальным интеллектом, антиципационными способностями и IQ у пациентов с шизофренией // XVI съезд психиатров России. Всероссийская научно-практическая конференция с международным участием «Психиатрия на этапах реформ: проблемы и перспективы»: Тезисы / Отв. ред. Н.Г. Незнанов. 2015. С. 280.

<sup>2</sup> *Дьяков С.И.* Психосемантическая модель и техника анализа и оценки субъектности личности // Научная конференция «Ломоносовские чтения» — 2015 / Тезисы докладов. 2015. С. 121–122.

<sup>3</sup> *Севостьянов Ю.О.* Изменение психосемантической структуры готовности работать в команде у студентов // Научный вестник Южного института менеджмента. 2014. № 2. С. 94–97.

<sup>4</sup> *Ермак Е.В.* Взаимосвязь свойств когнитивной переработки аффективной информации, эмоционального интеллекта и личностных черт // Философские проблемы биологии и медицины / Сб. статей. М.: Московский государственный медико-стоматологический университет им. А.И. Евдокимова, Московское философское общество. 2015. С. 250–253.

<sup>5</sup> *Куликов В.Ю., Антропова Л.К., Козлова Л.А.* Влияние функциональной асимметрии мозга на стратегию поведения индивида в стрессовой ситуации // Journal of Siberian Medical Sciences. 2010. № 5. С. 10.

<sup>6</sup> *Сергиевский Г.М., Лобачев В.С.* Моделирование поведения интеллектуального агента в проблемной ситуации с неполностью наблюдаемыми состояниями // Научная сессия НИЯУ МИФИ — 2012 / Аннотации докладов: В 3 т. 2012. С. 312.

<sup>7</sup> *Ясницкий Л.Н., Сичинава З.И.* Нейросетевые алгоритмы анализа поведения респондентов // Нейрокомпьютеры: разработка, применение. 2011. № 10. С. 59–64.

активности ставит на повестку дня вопрос об этике искусственных систем. Не менее важно осознание и опыта эволюции этики доцифрового социума, особенно в XX в. и в современности. Есть практически неизбежный риск отражения реально сложившегося этического релятивизма и этической конфликтности в архитектурах и онтологиях, программном обеспечении и процессах обучения СИИ.

## **4.5. СОЦИОГУМАНИТАРНАЯ ОЦЕНКА ИННОВАЦИОННЫХ ЦИФРОВЫХ ИНФОРМАЦИОННЫХ СИСТЕМ**

### **4.5.1. Введение**

Современный этап развития инновационных информационных систем (ИС) можно характеризовать как этап интеллектуализации. Интеллектуальные ИС интегрируют в себе не только наукоемкие технологии с высоким уровнем развитости интерфейса, автоматизации процессов сбора, накопления и анализа данных, но и технологии моделирования при выработке информации для принятия научно обоснованных решений, оценки их качества, доведения до исполнителей и упреждающий контроль их выполнения.

В этой связи тематика цифровой трансформации приносит кардинальное изменение взаимодействия между участниками органов управления, что стало возможным с появлением сквозных цифровых технологий, включая ИИ, интернета вещей, анализ данных, построения совсем новых процессов для оказания цифровых услуг гражданам. Этот этап является логическим продолжением цифровизации – процесса перевода существующих документов в цифровую форму в машиночитаемом виде, что привело к развитию, прежде всего, мониторинговых ИС. Пока же во всех федеральных и региональных органах власти преимущественно эксплуатируются сотни предметно-ориентированных ИС мониторингового типа, работающих с цифровой, текстовой и мультимедийной информацией.

Как результат, анализ современного состояния применения ИС и иных систем поддержки принятия решений позволяет выделить два существенно отличающихся семантических слоя. Первый основывается на сборе, хранении и обработке цифровой, текстовой и мультимедийной информации, что обеспечивает решение в основном мониторинговых задач. Второй слой акцентируется на имитационном, сценарном и когнитивном моделировании, умственно-содержательной обработке информации, ИИ, глубоком обучении и пр., что обеспечивает решение интеллектуальных задач синтеза, формирование вариантов эффективных управленческих решений.

### **4.5.2. Направления интеллектуализации ИС и моделирование**

Основными направлениями интеллектуализации ИС являются:

- применение машинного обучения на основе интеллектуальной обработки данных, текстов вырабатываются правила, закономерности, по которым генерируются варианты решений;

- использование эвристических правил при решении задач оптимизации траекторий достижения поставленных целей;
- внедрение технологий, основанных на знаниях, экспертных системах, нейросетях, позволяют имитировать процессы принятия решений;
- развитие технологий распознавания образов, графических интерфейсов, анализа и синтеза человеческой речи;
- биометрические системы, позволяющие проводить идентификацию граждан по отпечаткам пальцев, изображениям лиц и радужной оболочке глаз, татуировкам, пирсингам.

Однако, используемый в настоящее время класс систем пока еще слабо решает задачи поддержки принятия решений в социально-гуманитарной, социально-экономической и общественно-политической сферах. Эти сферы характеризуются существенным влиянием субъектного фактора, который особенно присущ рефлексивно-активным системам, в которых определяющую роль в управлении играет человек. Эти сферы плохо поддаются формализации и моделированию. Модели зачастую используются в них только при проведении анализа ситуации, человек же может принять решение совсем не то, которое рекомендует математическая модель. Вместе с тем такое моделирование существенно ускоряет принятие решений, повышает их качество и улучшает взаимопонимание решаемой проблемы в команде.

В настоящее время основой интеллектуализации ИС являются математическое моделирование и методы ИИ. Сейчас особую значимость приобретают вопросы обобщения накопленного опыта применения такого аппарата, типизации моделей и методов, создания на их основе цифровой инструментально-моделирующей платформы.

Рассмотрим возможную классификацию моделей, охватывающих социогуманитарную и социально-экономическую сферы, общественно-политические отношения, национальную безопасность при решении функциональных задач государственного управления.

Макроуровень включает модели геополитики, обороны, национальной безопасности, межотраслевые модели равновесия, модели отраслей и отраслевых комплексов, модели социальной сферы, территориальные модели, модели общественно-политического развития.

Региональный уровень должен включать модели социально-экономического развития, включая реализацию национальных проектов, модели развития социальной сферы, модели общественно-политического развития и взаимодействия с населением.

Третий уровень должен включать моделирование развития и размещение федеральных и муниципальных предприятий.

Особо следует подчеркнуть, что важнейшими свойствами таких моделей является их полнота, адекватность, точность и устойчивость.

Возрастающая сложность мира и последние глобальные потрясения (пандемия, разрыв производственных цепочек, торговые войны, политические катаклизмы и т.д.), с одной стороны, и колоссальная производительность современных вычислительных систем, с другой, привели к появлению инструментов, способных оценивать мультипликативные эффекты с прямыми и обратными связями на социально-экономические и общественно-политические системы. Отметим, что множество научных центров различных стран мира уже вовлечено в этот процесс, хотя такая возможность доступна для преимущественно крупных научных центров и международных организаций в силу ресурсоемкости (как научной, так и финансовой) этого вида деятельности. Подобного рода модели, основанные преимущественно на равновесном подходе, уже имеются.

Получившие распространение субъект-ориентированные модели, учитывающие погружение проблемной ситуации в гибридную среду<sup>1,2</sup>, могут рассматриваться в качестве инструментов решения не только перечисленных выше, но и других наукоемких проблем. При этом важно отметить, что субъект-ориентированное моделирование не является долгожданной панацеей и перспективой, поскольку уже существует и используется множество аналитических инструментов, имеющих программную реализацию (например, нейронные сети, генетические алгоритмы, природные вычисления и др.), которые дополняя друг друга могут существенно обогатить арсенал исследователей. К настоящему времени разработаны сотни специализированных программных средств, которые могут быть использованы для построения моделей с учетом субъектных факторов, например, NetLogo, RePast, MASON, AnyLogic, системы компаний Microsoft, IBM, ESRI, Wolfram Research и др. Однако перечисленные примеры программных средств нуждаются в специальной методологии применения, предусматривающих их использование в гибридном (человеко-машинном) пространстве, чтобы обеспечить устойчивость и целенаправленность решения социогуманитарных задач<sup>3</sup>.

---

<sup>1</sup> Perko I. (2020), Hybrid reality development - can social responsibility concepts provide guidance? *Kybernetes*, Vol. 50 No. 3, pp. 676–693. doi: 10.1108/K-01-2020-0061

<sup>2</sup> Lepskiy V. (2018), "Evolution of cybernetics: philosophical and methodological analysis", *Kybernetes*, Vol. 47 No. 2, pp. 249–261. doi: 10.1108/K-03-2017-0120

<sup>3</sup> Raikov, A.N. and Pirani, M. (2022), "Contradiction of modern and social-humanitarian artificial intelligence", *Kybernetes*, Vol. 51 No. 13, pp. 186-198. <https://doi.org/10.1108/K-01-2022-0057>

На основе анализа множества публикаций, посвященных агент-ориентированным моделям, можно обозначить несколько устоявшихся направлений их использования:

- эпидемиология (в первую очередь это оценка скорости распространения эпидемий с учетом различных ограничительных мер). Самая большая агент-ориентированная модель в мире (6.5 млрд. агентов) была разработана в 2009 г. в США, первоначально для сценарных расчетов различных вариантов профилактических мероприятий в отношении вируса H1N1;
- демографические процессы (использование агентного подхода позволяет не только более точно прогнозировать численность населения, но и оценивать эффективность различных мер государственной политики, направленной на решение демографических проблем; также модели этого класса хорошо себя зарекомендовали в вопросах прогнозирования миграционных процессов);
- транспортные системы (к настоящему моменту разработаны агентные (субъектные) модели транспортных систем для более 100 мегаполисов мира);
- окружающая среда (учитывая заметные климатические изменения и сформировавшийся тренд на декарбонизацию производств, модели этого класса активно применяются для мониторинга и прогнозирования состояния окружающей человека среды. Более того, третий год подряд специалистами в области высокопроизводительных вычислений стран БРИКС проект «Цифровая Земля», учитывающий большое количество факторов (атмосферные, океанические изменения, биогеохимию океана, динамику ионосферы, эволюцию ледяного покрова) и реализуемый в том числе с использованием агентного подхода, обозначается как приоритетный);
- городская динамика (моделирование городских агломераций более сложный процесс, нежели симуляции окружающей среды, поскольку такие модели должны включать в себя социальные, транспортные, экологические, экономические и др. системы; в силу трудоемкости построения таких инструментов разрабатывается они редко, но примеры есть – Гренобль, Лос-Анджелес, Вашингтон и др.);



- социальные сети (использование агентного подхода для моделирования и исследования социальных сетей открывает широкий спектр новых возможностей для изучения эволюции человеческого общества; возможно изучать скорость распространения информации, выявлять закономерности формирования социальных групп, продолжительность различных видов отношений между участниками и т.д.);
- экономика (в последние годы этот инструментарий прочно вошел в арсенал средств, используемых научными организациями, аналитическими центрами, корпорациями, органами государственного управления и др. применительно к вопросам моделирования и прогнозирования экономических процессов;
- военные конфликты (существует множество симуляторов, позволяющих моделировать военные операции и проигрывать различные варианты боевых действий с целью нахождения наиболее эффективной стратегии для заданной местности с учетом имеющегося арсенала и количества боевых единиц, к примеру, продукт от компании AgentFly Technologies или проект «The System Effectiveness Analysis Simulation».

Анализ применения субъектных моделей в реальных системах госуправления позволяет сформировать ряд требований:

- системное моделирование социогуманитарных, социально-экономических и общественно-политических процессов в условиях изменяющейся внешней и внутренней обстановки, обеспечивающее учет необходимого числа факторов, повышающих адекватность моделей;
- обеспечение проактивного управления с применением точных моделей прогнозирования возникновения инцидентов, позволяющих вырабатывать соответствующие упреждающие управленческие решения.

Создаваемые в Российской академии наук комплексы моделей и интеллектуальных технологий уже используются в настоящее время в Системе распределенных ситуационных центров<sup>1</sup> для оценки потенциала, конкурентных преимуществ и проблем развития России и других стран, а также для контроля реализации документов стратегического планирования.

---

<sup>1</sup> Стратегическое целеполагание в ситуационных центрах развития / Под ред. В.Е. Лепского, А.Н. Райкова / Авторский коллектив: Авдеева З.К., Зацаринный А.А., Журенков Д.А., Ильин Н.И., Колин К.К., Лепский В.Е., Малинецкий Г.Г., Райков А.Н., Савельев А.М., Сильвестров С.Н., Славин А.Б., Славин Б.Б. М.: Когито-Центр, 2018. 320 с.

Созданная система мониторинга решения вопросов национальной безопасности основывается на использовании методов многомерного статистического анализа, позволяющих снять проблему субъективизма (присущей экспертным методам, широко используемым в решении подобных задач<sup>1</sup>) при оценке вклада факторов в интегральный показатель. Система мониторинга национальной силы позволяет:

- оценивать и прогнозировать интегральные индексы национальной силы по странам мира и ранжировать страны на их основе;
- выявлять наиболее проблемные сферы России и других стран, создающие угрозу национальной безопасности;
- определять диапазоны допустимых значений для различных показателей, нарушение границ которых влечет угрозы как для отдельных сфер, так и в целом – для социо-эколого-экономической сферы страны.

Основные направления развития инновационных ИС:

- развитие информационной поддержки всех функций управления (прогнозирование, планирование с учетом рисков, контроль, регулирование, отчетность);
- разработка и внедрение системы экономико-математических и агент-ориентированных и др.;
- увеличение номенклатуры собираемой и хранимой разнокатегорийной информации из различных источников (предприятия, цены, бухгалтерия, отчетность, информация о недобросовестных поставщиках информации и др.);
- обмен разнокатегорийной информацией между органами исполнительной власти в процессе решения функциональных задач не только в режиме on-line, но и в режиме off-line, включая видеоконференцсвязь;
- разработка специализированного шлюза, обеспечивающего фильтрацию поступающей информации, включая синтаксический и семантический анализ, проверку достоверности, полноты и своевременности поступления перед загрузкой в базы данных;
- реализация механизма защиты информации и средств криптографической защиты непосредственно в информационных системах.

---

<sup>1</sup> Губанов Д.А., Коргин Н.А., Новиков Д.А., Райков А.Н. Сетевая экспертиза / Под ред. чл.-к. РАН Д.А. Новикова, проф. А.Н. Райкова. М.: Эгвес, 2011. 166 с.

Руководители органов государственной власти, госкорпораций и предприятий осуществляют управление по всему спектру вопросов социально-экономического и общественно-политического развития. В этой связи для решения указанных задач, а также оперативного реагирования на ситуации различного происхождения (природного, техногенного характера, социальные инциденты и т.п.) представляется целесообразным предусматривать взаимодействие информационных объектов при их создании и функционировании.

#### **4.5.3. Интеллектуальные Центры управления**

Перечисленное требует создания соответствующих интеллектуальных центров управления (ЦУ) различного уровня иерархии, которые позволят лицам, принимающим решения, повысить обоснованность своих решений и в целом увеличить эффективность управленческой деятельности за счет использования информационных технологий.

ЦУ должен обеспечить управление социогуманитарным и социально-экономическим развитием, общественно-политической обстановкой, национальными проектами и программами, национальной безопасности. В управлении важнейшей функцией помимо мониторинга и контроля, анализа и прогнозирования является планирование (стратегическое, среднесрочное, оперативное), которое включает обоснование целей, оптимальное распределение ресурсов по мероприятиям и работам для достижения поставленных целей.

Для реализации перечисленных функций необходимо создание единой сквозной защищенной цифровой платформы, включающей единый информационный фонд структурированных показателей всех сфер управления, единую систему классификаторов и справочников, единую систему аутентификации и идентификации пользователей, доверенную операционную систему и СУБД, средства информационной безопасности и информационно-телекоммуникационную инфраструктуру как сервис. На базе единой защищенной платформы строятся информационно-аналитические системы (ИАС), обеспечивающие поддержку принятия управленческих решений.

Весь спектр ИАС можно условно поделить на четыре группы: традиционные информационно-аналитические, системы цифровой социологии, системы проектного управления и специального назначения.

ИАМ региональных центров управления должны обеспечить процессы управления в сфере: экономики; промышленности и топливно-энергетического комплекса; строительства и ЖКХ; здравоохранения,

образования и науки; соцзащиты; внутренней политики и развития местного самоуправления; национальных проектов; культуры, спорта, туризма; мобилизации и ликвидации последствий кризисных ситуаций.

Современные тенденции развития информационного общества и цифровой трансформации государственного управления предъявляют повышенные требования к интеллектуальным системам, способным вывести работу каждого ЦУ в целом на качественно новый уровень. Применяемые в них технологии обработки данных должны ориентироваться на наиболее рациональное использование особенностей человеческого мышления и каналов восприятия информации, и, главное, учитывать при этом социогуманитарные факторы.

Формализуемым фундаментом интеллектуальной деятельности в ЦУ является математическое моделирование. Модельный аппарат включает цифровые модели, включая ИИ, методы математической статистики и теории вероятностей, экспертные методы, методы исследования операций, имитационное моделирование и ситуационный анализ, сетевые, когнитивные, сценарные, семантические модели и др.

Применение математических моделей, методов ИИ, реализованных в инструментально-моделирующей платформе ЦУ, позволяет проводить анализ разнородной информации на более высоком качественном уровне, получать решения, учитывающие структурные особенности и латентные причинно-следственные связи между факторами сферы управления, их компенсационные свойства, динамику, тенденции развития.

Использование результатов математического моделирования существенно сокращает время принятия решений и повышает их эффективность, что особенно важно в чрезвычайных ситуациях, когда процесс управления протекает в условиях дефицита времени, неопределенности и неполной информации.

В ЦУ должна быть создана инструментально-моделирующая платформа, включающая репозиторий качественных наборов моделей и алгоритмов (в том числе на основе технологии искусственного интеллекта и анализа больших данных), способных автоматизированно формировать цифровой профиль региона, отрасли, предприятия, что позволит перенести функции управления на обученные нейронные сети. Концептуальные решения ЦУ могут базироваться на конвергентной инфраструктуре и микросервисной технологии, контейнеризации и средств защиты информации нового поколения.

Необходимость хранения, обработки и передачи в ЦУ информации разнообразного характера и ведомственной принадлежности ставит во главу угла вопрос информационной безопасности. Ее необходимо

обеспечивать за счет применения специальных программно-аппаратных средств и реализации соответствующих организационных мер, гарантирующих конфиденциальность, целостность, доступность и аутентичность циркулируемой в системе информации. Эти работы должны быть максимально прозрачны, не трудозатратны и одновременно осуществляться в соответствии с действующими нормативными документами по информационной безопасности

#### **4.5.4. Многокритериальная оценка качества цифровых систем**

Понятие социотехнической оценки качества ИС соответствует тому, насколько система успешно справляется со всеми возлагаемыми на нее задачами, например: какие имеет показатели надежности, стоимости, удобства в эксплуатации и обслуживании, возможности сопряжения с другими системами и, в случае необходимости, какие имеет возможности для развития. Исследованиям в данной области посвящено большое количество зарубежных и отечественных работ.

Реальная практика показывает, что при оценке качества ИС должны учитываться, в том числе, функциональные возможности, производительность, удобство в эксплуатации, надежность, затраты на создание и эксплуатацию, длительность жизненного цикла, пригодность к модификации, простота и удобство администрирования, качество поддержки изготовителем.

При оценке качества ИС сложно предложить одну универсальную меру качества. Нужно использовать ряд характеристик, охватывающих весь спектр предъявляемых требований.

Анализ действующих нормативных актов показал, что в настоящее время не существует стандартов, полностью удовлетворяющим оценке качества ИС. Вместе с тем в зарубежных странах имеется ряд стандартов, определяющих основу этой оценки. Существующая в нашей стране система нормативно-технических документов относит ИС к «продукции производственно-технического назначения», которая рассматривается как материальный объект. Однако ИС являются, скорее, абстрактной нематериальной сферой.

В имеющихся государственных стандартах наиболее полно проработаны модели качества программного обеспечения (ПО), являющегося одной из важных составных частей ИС. Вместе с тем, эти модели практически не учитывают социогуманитарный фактор целенаправленного использования ПО, который обеспечил бы необходимую синергию при функционировании ПО в конкретном гибридном пространстве. В качестве основы для проведения работ по данному разделу обычно берутся требования по оценке качества ПО. В

настоящее время используется несколько абстрактных моделей качества ПО, основанных на определениях характеристик качества, критериев и метрик.

Под критериями качества в рассматриваемом контексте понимаются независимые атрибуты (показатели) оценки ИС или процесса ее создания. С помощью критериев могут быть измерены характеристики качества ИС на основе той или иной метрики (шкалы измерения). Совокупность нескольких критериев определяет показатель качества, формируемый исходя из требований, предъявляемых к ИС.

Для комплексной оценки качества ИС необходимо описание взаимосвязи компонент, которые могут быть объединены в виде иерархической модели. Для этого вначале должны быть определены общие характеристики качества, такие, как функциональные возможности, производительность, удобство в эксплуатации, надежность, затраты на создание и эксплуатацию, длительность жизненного цикла, пригодность к модификации, проста и удобство администрирования, качество поддержки изготовителем и т.п.

Далее формируются показатели, к числу которых могут быть отнесены практичность, целостность, корректность, удобство обслуживания, оцениваемость, гибкость, адаптируемость, мобильность, возможность взаимодействия и др. Каждому показателю качества ставится в соответствие группа критериев. Причём, один и тот же критерий может характеризовать несколько показателей.

С помощью метрик можно дать количественную или качественную оценку значений показателей качества ИС. Различают следующие виды метрик и шкал для измерения критериев:

- интервальные, характеризующиеся относительными величинами или реально измеряемыми физическими показателями, например, временем наработки на отказ, вероятностью ошибки, объемом информации и др.;
- порядковые, позволяющие ранжировать характеристики путем сравнения с опорными значениями;
- номинальные или категоризованные шкалы, определяющие наличие заданного свойства или признака у рассматриваемого объекта без учета градаций по этому признаку. Например, уровень сложности может быть «простым», «умеренным», «сложным», «очень сложным».

С учётом сказанного, можно выделить следующие классы критериев качества информационных систем:

- функциональные, с помощью которых должна оцениваться степень выполнения ИС основных целей и задач;

- конструктивные, предназначенные для оценки компонент ИС, не зависящих от целевого назначения;
- эксплуатационные, позволяющие оценивать удобство эксплуатации ИС.

Кроме того, должны учитываться факторы и параметры, влияющие на основные критерии качества.

Один и тот же критерий может характеризовать несколько показателей, например:

*практичность* – работоспособность, возможность обучения, коммуникативность, объем ввода, скорость ввода-вывода;

*целостность* – регулирование доступа, контроль доступа;

*эффективность* – эффективность использования памяти, эффективность функционирования;

*корректность* – трассируемость, завершенность, согласованность;

*надежность* – точность, устойчивость к ошибкам, согласованность, простоту;

*удобство обслуживания* – согласованность, простоту, краткость, информативность, модульность;

*оцениваемость* – простоту, наличие измерительных средств, информативность, модульность;

*гибкость* – распространяемость, общность, информативность, модульность;

*адаптируемость* – общность, информативность, модульность, аппаратную независимость, программную независимость;

*мобильность* – информативность, модульность, аппаратную независимость, программную независимость;

*возможность взаимодействия* – модульность, унифицируемость процедур связи, унифицируемость данных.

Как было сказано выше, поскольку в современных ИС ключевой является программная компонента, а пользователи, работающие с системой, в подавляющем большинстве случаев взаимодействуют непосредственно с программной компонентой, показатели качества информационных и программных систем в значительной степени совпадают. Однако, необходимы дополнительные показатели качества, когда эти системы работают в гибридной (человеко-машинной) среде. Прежде всего, показатели качества должны обеспечить оценку целенаправленности и устойчивости развития социально-экономического объекта с учетом субъектных факторов, функционирования систем в рефлексивно-активной среде.

Качество программного обеспечения определяется в стандарте ISO 9126 как совокупность его характеристик, относящихся к возможности удовлетворять высказанные или подразумеваемые потребности всех

заинтересованных лиц. ISO 9126 – это международный стандарт, определяющий оценочные характеристики качества программного обеспечения. Российский аналог стандарта ГОСТ 28195. Стандарт разделяется на четыре части, описывающие: модель качества; внешние метрики качества; внутренние метрики качества; метрики качества в использовании.

Модель качества, установленная в первой части стандарта ISO 9126-1, классифицирует качество ПО по шести аспектам:

- функциональность;
- надежность;
- эффективность (производительность);
- удобство использования;
- удобство сопровождения;
- переносимость.

В настоящее время группа стандартов ISO/IEC 9126 получила развитие в группе стандартов ГОСТ Р ИСО/МЭК 250xx. В соответствии с ГОСТ Р ИСО/МЭК 25010-2015, относящегося к этой группе, модель качества продукта можно применять как для программного продукта, так и для ИС, в состав которой входит ПО.

Следует отметить, что характеристики «Соответствие стандартам и правилам» были удалены и отсутствуют в ГОСТ Р ИСО/МЭК 25010-2015, поскольку они являются в соответствии с законами и правилами частью общих требований к системе, а не частью характеристики качества. Часть характеристик была добавлена в ГОСТ Р ИСО/МЭК 25010-2015. К ним относятся: «Защищенность», «Конфиденциальность», «Целостность», «Неподдельность», «Отслеживаемость» и «Подлинность». Как можно заметить, перечисленные характеристики не покрывают вопросов собственно использования цифровых систем в социогуманитарной среде таким образом, чтобы обеспечить целенаправленность и устойчивость деятельности пользователей.

#### **4.5.5. Базовые характеристики качества информационных систем**

Рассмотрим характеристики качества ИС, относящиеся к их: функциональности, надежности, удобство использования, производительность, удобство сопровождения, переносимость.

Функциональность (functionality) определяется как способность в определенных условиях решать задачи, нужные пользователям. Для данной характеристики выделяются следующие подхарактеристики:

- функциональность необходимых для достижения поставленных целей;



- точность – способность выдавать нужные результаты;
- способность к взаимодействию – способность взаимодействовать с нужным набором других систем;
- защищенность – способность предотвращать неразрешенный доступ к данным и программам;
- соответствие стандартам и правилам - соответствие имеющимся индустриальным стандартам, нормативным и законодательным актам, другим регулирующим нормам.

Основным показателем функциональности является процент реализации поставленных перед ИС целей.

*Надежность* (reliability) – это способность поддерживать определенную работоспособность в заданных условиях. Для данной характеристики выделяются следующие подхарактеристики:

- зрелость (наработка на отказ), измеряется средним временем работы без сбоев и величиной, обратной вероятности возникновения отказа за данный период времени;
- устойчивость к отказам, определяет способность поддерживать заданный уровень работоспособности при отказах;
- способность к восстановлению (ремонтпригодность), определяет способность восстанавливать определенный уровень работоспособности и целостности данных после отказа, необходимые для этого время и ресурсы;
- соответствие стандартам.

Основными параметрами оценки надежности являются: среднее время работы без отказов, риски возникающих отказов за определенный период времени, время восстановления требуемого уровня работоспособности и целостности данных после сбоя.

*Удобство использования* (usability), или практичность, определяется, как способность быть удобным в обучении и использовании, а также привлекательным для пользователей. Для данной характеристики выделяются следующие подхарактеристики:

- понятность – показатель, обратный к усилиям, которые затрачиваются пользователями на восприятие основных понятий ИС и осознание их применимости для решения своих задач;
- удобство обучения – показатель, обратный усилиям, затрачиваемым пользователями на обучение работе с ИС;
- удобство работы – показатель, обратный усилиям, предпринимаемым пользователями для решения своих задач;

- привлекательность – способность ИС быть привлекательной для пользователей;
- соответствие стандартам, определяется как удобство использования.

Основным параметром оценки удобства может служить среднее время, необходимое пользователям ИС на восприятие основных понятий ИС и осознание их применимости для решения своих задач. Заметим, что часть из этих характеристик в какой-то степени учитывает особенности социогуманитарного фактора и работу цифровых систем в рефлексивно-активной среде. Однако многие характеристики субъектности выпадают из списка критериев, например, параметр ответственности<sup>1</sup>.

*Производительность* (efficiency), или эффективность, – это способность ИС при заданных условиях обеспечивать необходимую работоспособность по отношению к выделяемым для этого ресурсам. Может быть определена, как отношение получаемых с помощью ИС результатов к затрачиваемым на это ресурсам всех типов. Для данной характеристики выделяются следующие подхарактеристики:

- временная эффективность – способность выдавать ожидаемые результаты, а также обеспечивать передачу необходимого объема данных за отведенное время;
- эффективность использования ресурсов – способность решать нужные задачи с использованием определенных объемов ресурсов определенных видов (оперативной и долговременной памяти, телекоммуникаций, устройств ввода и вывода и пр.);
- соответствие стандартам производительности.

*Удобство сопровождения* (maintainability) определяется как удобство проведения всех видов деятельности, связанных с сопровождением ИС. Для данной характеристики выделяются следующие подхарактеристики:

- анализируемость – удобство проведения анализа, определяется как удобство проведения анализа ошибок, дефектов и недостатков, а также удобство анализа необходимости изменений и их возможных последствий;
- удобство внесения изменений – показатель, обратный трудозатратам на выполнение необходимых изменений;
- стабильность – показатель, обратный риску возникновения неожиданных эффектов при внесении необходимых изменений;

---

<sup>1</sup> Espejo R., Lepskiy V. (2020), "An agenda for ontological cybernetics and social responsibility", *Kybernetes*, Vol. 50 No. 3, pp. 694–710. doi: 10.1108/K-06-2020-0390

- удобство проверки – показатель, обратный трудозатратам на проведение тестирования и других видов проверки того, что внесенные изменения привели к нужным результатам;
- соответствие стандартам удобства сопровождения.

Удобство сопровождения может измеряться в номинальной шкале по уровням: удобно, умеренно удобно, не очень удобно, не удобно.

*Переносимость* (portability) определяется как способность сохранять работоспособность при переносе из одного окружения в другое, включая организационные, аппаратные и программные аспекты окружения. Для данной характеристики выделяются следующие подхарактеристики:

- адаптируемость – способность приспосабливаться различным окружениям без проведения для этого действий, помимо заранее предусмотренных;
- удобство установки – способность быть установленным или развернутым в определенном окружении;
- способность к сосуществованию – способность сосуществовать с другими ИС в общем окружении в условиях конкуренции за ресурсы;
- удобство замены – возможность применения вместо других программных систем для решения тех же задач в определенном окружении;
- соответствие стандартам переносимости (portability compliance).
- универсальность – способность адаптировать ИС к новым функциональным требованиям, возникающим вследствие изменения области примыкания или других условий функционирования.

Переносимость измеряется в номинальной шкале по уровням: просто, умеренно, сложно, очень сложно.

На современном этапе создание и совершенствование инновационных ИС являются одним из инструментов повышения эффективности проектирования и управления в социально-экономической и общественно-политической сферах. В этой связи оценка качества ИС на всех этапах их жизненного цикла является важной и актуальной задачей.

Исследование характеристик качества ИС с зарубежными и отечественными аналогами позволяет выявить их слабые и сильные стороны и принять соответствующие проектные решения по повышению их эффективности.

#### 4.5.6. Выводы

В данном разделе рассмотрены основные подходы к оценке качества создания инновационных цифровых систем, включая системы ИИ, в контексте их погружения социогуманитарную сферу. Рассмотрены основные проблемы в данной области, разработаны предложения по перечню характеристик оценки качества конкретных ИС. Показано, что такое использование характеризуется следующими особенностями, которые следует учесть при определении оснований критериев оценки инноваций, включая цифровые технологии и искусственный интеллект:

- процесс разработки и использования ИС в рефлексивно-активной среде происходит под существенным влиянием субъектного фактора. Эта среда плохо поддается формализации и моделированию, которое используется не столько для получения управленческого решения, сколько при проведении анализа ситуации, человек же может принять решение совсем иное решение;
- в имеющихся государственных стандартах наиболее полно проработаны модели качества программного обеспечения (ПО), которые практически не учитывают социогуманитарный фактор целенаправленного использования ПО для обеспечения необходимой синергии действий при функционировании ИС в конкретном гибридном пространстве;
- многочисленные примеры создания цифровых ИС и систем поддержки решений нуждаются в специальной методологии применения, предусматривающих их использование в гибридном (человеко-машинном) пространстве, чтобы обеспечить устойчивость и целенаправленность решения социогуманитарных задач.

## 4.6. ВЫВОДЫ

1. Современные системы оценок инноваций, включая цифровые технологии и ИИ, основной акцент делают на технологических факторах развития инноваций, умаляя субъективные. Технологии обычно оцениваются по количественным метрикам с учетом их потребности на разных сегментах рынка.

2. Для оценки ситуации с построением критериев, носящих социогуманитарный характер, нужны иные, скорее, дополнительные, критерии. Социогуманитарные основания критериев оценки инноваций, использующих ИИ, характеризуют их востребованность гражданским обществом, особенно такими областями деятельности человека, как медицина, образование, наука и пр.

3. Для учета социогуманитарных аспектов инноваций в области цифровых технологий и ИИ, большего внимания требуют вопросы создания гибридных систем, то есть систем, работающих в условиях, когда люди и технологии ИИ сосуществуют и влияют друг на друга. Технологии не просто изменяют жизнь людей, но люди помогают технологиям развиваться. Гибридная реальность может рассматриваться через призму социальной ответственности.

4. Уже существует широкий спектр аспектов построения социогуманитарных оснований критериев оценки инноваций, использующих ИИ. К ним стоит отнести: оценку социальной и экономической эффективности; ранние методы оценки инновационных технологий; собственно характеристики инноваций (тип инновации, статус на рынке, фаза исследования, конфликт интересов); характеристики лиц, принимающих решения; создание гибридных систем; принципы социальной ответственности; этические аспекты и др.

5. При внедрении инновационных технологий в области цифровой экономики обостряются вопросы безработицы из-за замещения людей машинами, деградации естественного интеллекта, искажения этики, злонамеренного использования ИИ, а также падения уровня духовности общества, в том числе из-за накопления ошибок в данных из-за замены естественного аналогового сигнала дискретным, цифровым.

6. Современные системы ИИ развиваются в рамках парадигмы, которая обеспечивает совершенствование отдельных технических направлений ИИ, что затрудняет использование результатов разработок в процессах управления. К таким «техническим» парадигмам относятся:

морфологическая, логическая, нейрокибернетическая, имитационная, а также парадигмы слабого, общего и сильного ИИ в их современном «цифровом» толковании – без учета субъектных аспектов их разработки и применения.

7. Новая парадигма развития ИИ, которая развивается в рамках развития Общего ИИ (ОИИ), должна отражать неформализованную когнитивную динамику работы систем ИИ и поддерживать их саморазвитие в условиях рефлексивно-активной среды. Тогда методология создания ОИИ должна базироваться на идеях субъектно-онтологического подхода, функционализма, феноменологии субъективной реальности, конвергентных когнитивных архитектур, а также методов создания саморазвивающейся полисубъектной (рефлексивно-активной) среды. Тогда ОИИ приобретает гибридный характер.

8. Разрабатываемый в рамках настоящей работы подход к созданию ОИИ, основанный на методах конструирования субъективной реальности, также пока имеет ограничения. Например, он не позволяет объяснить информационно-когнитивные процессы, порождаемые эффектом нелокальности семантик моделей ИИ, возникающим при учете аспектов поведения человеческого мозга на атомарном и релятивистском уровнях. Последние аспекты, по-видимому, могут помочь в объяснении таких способностей человека, как умение делать правильные и в то же время беспричинные выводы, мгновенно оценивать подлинность художественного произведения, обладать интуицией, впадать в транс, иметь, что называется, душу.

9. Представление сознания становится все более важной задачей при создании ОИИ, при этом на первый план выходят вопросы использования феноменологического анализа субъективной реальности, ее ценностно-смысловой и операциональной структур. ОИИ должен обладать высокой степенью автономности и самостоятельного решения широкого круга задач в различных условиях внешней среды, включающей рефлексивно-активные инварианты.

10. Существующие парадигмы создания систем ИИ способствуют развитию только отдельных направлений проектирования, но затрудняют использование ИИ в процессах управления, поскольку недооценивают специфику субъекта в контуре управления. За скобками остаются концептуальный аспект управления, аспект установления интерфейса между технологическими и субъективными компонентами, последние из которых основаны на

психических и поведенческих особенностях человека, аспектами систем управления с использованием систем ИИ.

11. Методология создания ОИИ должна базироваться на идеях субъектно-онтологического подхода, функционализма, феноменологии субъективной реальности. При этом к основным инвариантным свойствам субъектов следует отнести целеустремленность, рефлексивность, коммуникабельность, социальность и способность к развитию. Сосуществование субъектов естественного интеллекта и систем ИИ образует гибридную среду, то есть реальность совместного действия естественного и искусственного интеллекта.

12. Среди важнейших и быстро развивающихся технологий систем ИИ приоритетное место приобрели нейротехнологии. Это класс технологий, которые обеспечивают решение огромного числа прикладных задач ИИ, включая таких, как комплексирование моделей поддержки социальной среды, лояльной к правовым нормам, и в то же время они помогают понять работу мозга, мыслительные процессы, высшую нервную деятельность, в том числе технологии по усилению, улучшению работы мозга и психической деятельности.

13. Развитие возможностей систем ИИ в различных контурах знаний, включая способность корректировать целеполагание своей активности, ставит на повестку дня вопрос об этике искусственных систем. При этом важно осознание опыта эволюции этики доцифрового социума в современном контексте, поскольку есть неизбежный риск отражения реально сложившегося этического релятивизма и этической конфликтности в архитектурах и онтологиях, программном обеспечении, процессах обучения с применением систем ИИ.

14. Особого внимания заслуживает оценка качества создания инновационных цифровых систем (ИС), включая системы ИИ, в контексте их погружения социогуманитарную сферу. Такая оценка характеризуется следующими особенностями: процесс разработки и использования ИС в рефлексивно-активной среде происходит под существенным влиянием субъектного фактора, который плохо поддается формализации; в имеющихся государственных стандартах наиболее полно проработаны модели качества программного обеспечения, которые практически не учитывают социогуманитарный фактор его целенаправленного использования программного обеспечения для достижения необходимой синергии действий в конкретном гибридном (человеко-машинном) пространстве.

## **5. СОЦИОГУМАНИТАРНЫЕ ОСНОВАНИЯ КРИТЕРИЕВ ОЦЕНКИ ИННОВАЦИЙ, ИСПОЛЬЗУЮЩИХ ЦИФРОВЫЕ ТЕХНОЛОГИИ И ИИ В ОРГАНИЗАЦИИ ЦИФРОВЫХ ПЛАТФОРМ, ТЕХНОЛОГИЧЕСКИХ ДОЛИН, ЭКОСИСТЕМ**

---

### **5.1. ВВЕДЕНИЕ**

В данной главе проводится анализ вопросов стандартизации в области цифровых инноваций, включая ИИ. Анализируется формирование и использование ценностей, возникающих благодаря технологическим цифровым изменениям, которые обусловлены субъектным участием, потребностью персонала в цифровых навыках, заинтересованности инвестора в исследованиях и разработках и др.

Анализируется фактор того, что критерии оценки инноваций должны учитывать не только сам цифровой продукт и процесс его разработки, но и насколько эффективно используется сами цифровые технологии людьми.

Акцентируется внимание на том, что управление цифровыми инновациями требует междисциплинарного подхода и что при этом необходимо сочетать последние достижения в инновационных цифровых технологиях и ИИ.

Уделяется внимание вопросам, которые объясняют суть человеческого интеллекта, ведь внешнее изучение процессов целеполагания, формирования потребностей, детерминация не может покрыть всю область исследования. Рассматривается социальная природа человеческого интеллекта, его связь с коллективным интеллектом. В этом контексте предлагается для построения модели общего ИИ использовать архитектуру информационной системы организации.

В рамках построения модели Общего ИИ на основе архитектурного подхода выявляются требования, даются контуры многоуровневой архитектуры. При этом императивно учитываются аспекты субъектности и погружения систем ИИ в полисубъектную рефлексивно-активную среду. При этом учитываются международные традиции и стандарты построения цифровых платформ и средств обработки больших данных на основе архитектурного подхода.

Затрагиваются вопросы построения пространства доверия при внедрении цифровых систем, включая ИИ. В этом русле учитывается роль социальных медиа, которые играют важную роль партнерств, способствующих созданию адекватного контекста и контента.



## 5.2. КРИТЕРИИ ОЦЕНКИ ИННОВАЦИЙ С ИСПОЛЬЗОВАНИЕМ ЦИФРОВЫХ ТЕХНОЛОГИЙ: АНАЛИЗ МЕЖДУНАРОДНОГО ОПЫТА СТАНДАРТИЗАЦИИ

Стандарты отражают многократно повторенный и признанный лучшим опыт в той или иной области. Концепция стандартов говорит об их социогуманитарной природе, и именно так необходимо подходить к исследованию роли стандартов в деятельности организаций. Обычно процесс стандартизации ассоциируется с чем-то уже устоявшимся и не являющимся инновационным. И это действительно так, поскольку стандартизировать можно только то, что можно стабильно воспроизводить с гарантированным результатом. Однако если сами по себе инновации имеют дело с чем-то неизведанным и неповторимым, то сам инновационный процесс вполне поддается стандартизации. Более того, чем более инновационной становится экономика, тем нужнее стандарты. Любые стандарты базируются на измерении, а следовательно, стандарты в области инноваций требуют инструментов для измерений.

Сегодня драйвером инноваций являются информационные технологии и ИИ. В них сочетаются как свойства производственных технологий, товаров, так и свойства услуг. Именно понимание особой роли информационных технологий привело к тому, что в конце прошлого века особое внимание было уделено измерению инноваций в области цифровизации. Еще в 1995 году Риналдо Евангелиста и Джорджио Сирилли писали: «Растущее экономическое и технологическое значение сектора услуг в современных обществах требует более систематического сбора данных об инновационной деятельности в этих отраслях»<sup>1</sup>. Более того, сегодня инновации становятся уделом не только бизнеса, но и некоммерческой деятельности, взаимодействия органов власти с гражданами. По мнению Фреда Голта, «до 2018 года существовала потребность только в одном секторе – предпринимательском, но теперь существует определение инноваций, применимое ко всем секторам экономики, таким как сектор государственного управления (управления органов власти в сочетании с совокупностью государственных корпораций), сектор домашних хозяйств и сектор некоммерческих организаций, обслуживающих домашние хозяйства»<sup>2</sup>. Все это привело к тому, что в конце второго

---

<sup>1</sup> *Evangelista R., Sirilli G. Measuring innovation in services // Research Evaluation. 1995. Vol. 5, No 3. P. 207–215.*

<sup>2</sup> *Gault F. Measuring Innovation Everywhere. Cheltenham; Northampton: Edward Elgar Publishing, 2021.*

десятилетия появились новые стандарты в области управления инновациями с учетом цифровизации экономики.

Так, например, управлению инновациями посвящены стандарты серии 56000 Международного института стандартизации (International Organization for Standardization – ISO). Стандарт ISO 56000:2020 Innovation management – Fundamentals and vocabulary<sup>1</sup> содержит глоссарий в области инноваций. Стандарт ISO 56002:2019 «Innovation management – Innovation management system – Guidance»<sup>2</sup> посвящен основам инновационного управления и локализован как ГОСТ Р ИСО 56002–2020 «Инновационный менеджмент. Системы инновационного менеджмента. Руководящие указания». Стандарт ISO 56003:2019 «Innovation management – Tools and methods for innovation partnership – Guidance»<sup>3</sup> посвящен инструментам и методам инновационного партнерства (например, при взаимодействии организаций и стартапов), русский перевод – ГОСТ Р ИСО 56003–2020 «Инновационный менеджмент. Методы и средства организации инновационного партнерства. Руководящие указания».

Стандарт ISO/TR 56004:2019 «Innovation Management Assessment – Guidance»<sup>4</sup>, посвященный оценке качества управления инновациями, аналогично предыдущему имеет русский перевод, хотя и с отличным от ISO номером в названии: ГОСТ Р 59062-2020/ISO/TR 56004:2019 «Оценка инновационного менеджмента. Руководящие указания». Стандарт ISO 56005:2020 Innovation management – Tools and methods for intellectual property management – Guidance<sup>5</sup> посвящен рекомендациям в области инструментов и методов управления интеллектуальной собственностью. Ряд стандартов ISO этой серии находится еще в разработке. Среди них – стандарт ISO/AWI 56001 Innovation management – Innovation management system – Requirements, посвященный системам управления инновациям, стандарт ISO/FDIS 56006 Innovation management – Tools and methods for strategic intelligence management – Guidance, где рассматриваются инструменты и методы стратегического интеллектуального управления, и стандарт ISO/AWI 56007 Innovation

---

<sup>1</sup> ISO 56000:2020 Innovation management – Fundamentals and vocabulary. – URL: <https://www.iso.org/ru/standard/69315.html>.

<sup>2</sup> ISO/TR 56004:2019 Innovation Management Assessment – Guidance. – URL: <https://www.iso.org/standard/69921.html>.

<sup>3</sup> ISO 56003:2019 Innovation management – Tools and methods for innovation partnership – Guidance. – URL: <https://www.iso.org/ru/standard/68929.html>.

<sup>4</sup> ISO/TR 56004:2019 Innovation Management Assessment – Guidance. – URL: <https://www.iso.org/standard/69921.html>.

<sup>5</sup> ISO 56005:2020 Innovation management – Tools and methods for intellectual property management – Guidance. – URL: <https://www.iso.org/ru/standard/72761.html>.

management – Tools and methods for idea management – Guidance, в котором представлены инструменты и методы управления идеями. Пока ведется работа над стандартом ISO/AWI 56008 Innovation management – tools and methods for innovation operation measurements – Guidance, в нем будут изложены рекомендации в области использования инструментов и методов измерения инновационной деятельности.

Основу всех стандартов ISO в области инноваций в части оценки инноваций во многом определяет справочное руководство «Измерение научно-технической деятельности. Принципы сбора и интерпретации данных о технологических инновациях», которое выпускает Европейская организация экономического сотрудничества и развития (ОЭСР) с 1992 года. Данное руководство получило название Oslo Manual (Руководство Осло) и входит в общее семейство стандартов сбора статистики Frascati Manual, разработанное еще для ОЭСР в 1963 году группой NESTI (National Experts on Science and Technology Indicators) в итальянском городке Фраскати. В четвертом издании Руководства Осло<sup>1</sup> учтена современная практика инноваций, в том числе обсуждается роль цифровизации для инновационного развития предприятий, важность работы с данными и использования цифровых платформ.

### **5.2.1. Руководство Осло по измерению инноваций**

В Руководстве Осло основополагающим считается принцип возможности и необходимости измерения инноваций. В документе дано определение: «Инновация – это новый или улучшенный продукт или процесс (или их комбинация), который значительно отличается от предыдущих продуктов или процессов подразделения и который был предоставлен потенциальным пользователям (как продукт) или введен в эксплуатацию подразделением (как процесс)». В редакции 2018 года типы инноваций сокращены до двух: продукт и процесс, убраны маркетинговые и организационные инновации, которые по своей сути являются процессными. Основную роль в инновациях играет знание, а не только новизна и полезность, явно обозначенная в формулировке. Важно, что понятие инновации относится как к самой деятельности, так и к продукту такой деятельности. Соответственно, предприятие может быть вполне инновационным, даже если выпускает продукцию, которую к инновациям отнести сложно, например, какой-нибудь продукт питания. В этом смысле к инновационным можно отнести многие задачи

---

<sup>1</sup> Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation. The Measurement of Scientific, Technological and Innovation Activities. 4th ed. – Paris: Eurostat; Luxembourg: OECD Publishing, 2018. doi: 10.1787/9789264304604-en.

по повышению эффективности бизнес-процессов, которые решаются с использованием цифровых технологий и искусственного интеллекта.

В Руководстве Осло большое внимание уделяется использованию цифровых технологий для сбора статистических данных об инновациях (глава 9, здесь и далее указаны пункты Руководства Осло), внедрению автоматизированных методов сбора данных при опросе респондентов об инновационности (главы 9, 10), использованию инструментов семантического анализа и визуализации собранной информации (глава 11). Однако для цели данного исследования интерес представляют: обсуждение роли информации, представления такой роли с точки зрения инновационности одновременно продукта и процесса (глава 3), потенциальная инновационная роль управления данными наряду с разработкой программного обеспечения (глава 4), данные с точки зрения информационного контента и возможности использования для инновационного развития организации (глава 5). Именно эти моменты целесообразно обсудить подробнее.

Как уже говорилось, в целях измерения инновации делятся в зависимости от объекта: инновации в продуктах и инновации в бизнес-процессах. «Инновации в продуктах должны обеспечивать значительное улучшение одной или нескольких характеристик или технических характеристик. Это включает в себя добавление новых функций или улучшения существующих функций или пользовательских утилит» (раздел 3.25). Таким образом, улучшение цифрового продукта или продукта, у которого есть цифровая составляющая (например, автомобиль с элементами автоматического ассистирования), является инновацией. Применительно к продуктам различают инновации в товарах и инновации в услугах. Обсуждение того, как соотносятся инновационность и доходность (раздел 3.28), показывает, что нецелесообразно ставить инновационность в зависимость от дохода, поскольку, например, часто цифровые инновационные продукты (товары или услуги) могут быть предоставлены потребителю бесплатно, доход предпринимателю приносит реклама, сопровождающая продажу таких продуктов. По всей видимости, при оценке инновационности продуктов было бы правильно говорить об их ценности для потребителей. В стандарте говорится о том, что ценность является неявной целью инноваций (раздел 2.22), однако при этом делается замечание, что в имеющихся системах сбора статистической информации нет единого показателя экономической или социальной ценности. Необходимо учитывать, что использование ценности<sup>1</sup>,

---

<sup>1</sup> *Castaño-Martínez María-Soledad. Innovation, Value Creation, and Entrepreneurship by Opportunity. Advances in Business Strategy and Competitive Advantage, pp. 43-63.*

возникающей благодаря технологическим изменениям в значительной степени обусловлено человеческим капиталом, инвестициями в исследования и разработки, финансовыми ресурсами, эффективным управлением. При этом необходимо учитывать траектории инноваций с точки зрения процесса создания стоимости. Так, например, в работе<sup>1</sup> предлагается выделять элементы, которые характеризуют инновационную цепочку создания стоимости: поиск, отбор и распространение знаний на уровне фирмы, бизнес-модель, реализация.

В цифровую эпоху продукт создается при активном участии пользователя, например заказ через Интернет сувенирной продукции с пользовательскими логотипами или личных фотоальбомов, перевод денег через личный кабинет на сайте банка и т.п. По всей видимости, к инновационным стоит отнести создание новых услуг, основанных на информационном самообслуживании. В Руководстве предлагается отдельно выделить наукоемкие продукты (*knowledge-capturing products*), получившие особое распространение благодаря развитию информационных технологий и снижению стоимости цифровых носителей. В качестве примера приведем сетевые информационные базы данных. Сама информация оказывается одновременно и товаром, и предметом услуги: «...продукты для сбора знаний похожи на товар, если потребители могут делиться или продавать их другим после покупки, но они похожи на услугу, если права потребителя ограничены лицензией, которая ограничивает обмен или продажу» (см. раздел 3.32).

Инновации в бизнес-процессах затрагивают производство продуктов, дистрибуцию и логистику, маркетинг и продажи, информационно-коммуникационные технологии, управление и администрирование, разработку новых продуктов и бизнес-процессов. В четвертом издании Руководства Осло цифровые технологии выделены в отдельную группу, в третьем они составляли одну группу с учетными функциями. В разделе 4.2 определены различные виды инновационной деятельности: научно-исследовательские и опытно-конструкторские работы, инженерная и проектная деятельность, маркетинг и продвижение бренда, работа с интеллектуальной собственностью, обучение сотрудников, приобретение или аренда материальных активов управление инновациями, отдельно указана деятельность по разработке программного обеспечения (ПО) и баз данных. В состав последней включены:

---

<sup>1</sup> *Shimelis T., Eshetie B. (2021). Innovation value chain: a systematic and narrative review. International Journal of Quality and Innovation, Vol. 6, no. 1, pp. 91-114.*

- собственная разработка и покупка ПО, как системного, так и прикладного (включая ПО, встроенное в продукты или оборудование);
- приобретение или внутренняя разработка компьютерных баз данных, включая сбор и анализ данных;
- мероприятия по модернизации или расширению функций информационных систем (ИС).

В работе<sup>1</sup> изучаются факторы, влияющие на внедрение конкретных программных средств для поддержки методов управления инновациями, называемых программным обеспечением для управления инновациями (IMS), и их специфической функциональности. В статье<sup>2</sup> концепция, согласно которой реализация и внедрение цифровых инноваций в организации происходит по трем концентрическим кольцам: разработка, основанная на технологиях, и различные категории внедрения цифровых инноваций в основе, средства, способствующие цифровым инновациям, во втором кольце и управление цифровыми инновациями в третьем кольце.

Разработка ПО, ведущая к незначительным изменениям, а также «приобретение и анализ баз данных для бухгалтерского учета и других обычных бизнес-функций», не являются инновационными. И наоборот, разработка ПО, которая ведет к появлению новых или улучшению имеющихся бизнес-процессов, работа с базами данных, используемых для анализа данных о свойствах материалов или предпочтениях клиентов, являются инновационными. Поскольку разработка ПО и баз данных включает в себя покупку таких продуктов, можно сказать, что практически вся деятельность ИТ службы, связанная с развитием информационной системы (в отличие от деятельности по поддержке ИС), является инновационной с точки зрения методологии оценки инноваций ОЭСР.

### **5.2.2. Особенности оценки инноваций на основе цифровых технологий**

Среди технологических возможностей для инноваций указаны экспертиза, проектирование и использование цифровых технологий и анализа данных (раздел 5.5). Выделение цифровых технологий и анализа данных в отдельный тип связано с их универсальностью. В недалеком прошлом к технологиям, которые дают прорывные возможности,

---

<sup>1</sup> *Huesig S., Endres H.* Exploring the digital innovation process. *European Journal of Innovation Management*, Vol. 22, no. 2, pp. 302-314.

<sup>2</sup> *Wiesböck F., Hess T.* Digital innovations. *Electronic Markets*, Vol. 30, no. 1, pp. 75-86.

относились «биотехнологии, передовые методы производства, нанотехнологии, ИКТ и их приложения. Более поздними областями интереса стали квантовые вычисления, ИИ и робототехника, а также приложения на базе Интернета, такие как облачные сервисы и аналитика больших данных» (раздел 5.80).

Цифровые технологии определяются как «электронные инструменты, системы, устройства и ресурсы, которые генерируют, хранят, обрабатывают, обмениваются или используют цифровые данные». Существует разница между понятиями:

- digitisation (оцифровка) – преобразование аналоговой информации (видео, изображение, текст, звук и т.п.) в цифровой формат (двоичные биты);
- digitalisation (цифровизация) – применение цифровых технологий на уровнях организации, отрасли, государства и на межгосударственном уровне.

Именно цифровизация имеет прямое отношение к инновациям. Более того, с точки зрения методологов ОЭСР в области инноваций, сегодня внедрение цифровых технологий, электронных средств связи, а также инструментов анализа данных, включая ИИ, сегодня создает благоприятные возможности для инновационного развития предприятий. Поэтому при оценке инновационности учитывается, насколько в организации развиты ИТ, имеется ли соответствующее подразделение, бюджет, стратегия и т.д.

Для использования цифровых технологий в инновациях у персонала должны быть цифровые навыки, что также говорит и о социогуманитарной сущности самих цифровых технологий. Цифровые компетенции, по всей видимости, должны быть включены в критерии оценки инноваций, использующих цифровые технологии и ИИ. На первый взгляд, цифровые технологии заменяют человеческий труд. Фактически же, заменяя рутинную деятельность человека, цифровые технологии и ИИ создают новые возможности для разработки новых продуктов, внедрения новых процессов, а следовательно, увеличивают потребность в человеческом труде. Именно поэтому сегодня в ИТ отрасли во всех странах мира наблюдается кадровый голод.

«Общей особенностью цифровых технологий является их способность соединять различные виды деятельности и бизнес-функции, формируя интегрированную систему со структурированным обменом данными между различными функциями и подразделениями» (раздел 5.105). Таким образом, оценивая инновации, основанные на цифровых технологиях, целесообразно оценивать, насколько инновационной оказалась интеграция различных видов деятельности и функций.

Возможно, наиболее значимо то, что цифровые технологии «позволяют фирмам генерировать и хранить огромные объемы данных (часто в режиме реального времени) по целому ряду бизнес-операций, как внутри фирмы, так и связанных с поставщиками и пользователями» (раздел 5.106). Благодаря использованию аналитических инструментов (включая ИИ) и накапливаемым данным можно выстраивать новые стратегии, новые бизнес-модели и сервисы, которые сами по себе являются инновациями. К инновациям, основанным на цифровых технологиях, целесообразно относить инновации, которые не только напрямую используют цифровые инструменты, но и возникли в результате использования аналитических методов работы с данными. Для оценки инновационности организации в Руководстве Осло приводится ссылка на исследование<sup>1</sup>, которое показывает значительную роль инноваций, основанных на цифровых технологиях, благодаря чему методологи (раздел 5.107) при оценке предлагают не делать различий между инновациями, если они «содержат или были разработаны с использованием цифровых технологий».

Именно в силу важности цифровых технологий для инновационного развития организаций в Руководстве Осло предлагается особо оценивать общую цифровую компетенцию, «которая отражает способность фирмы извлекать выгоду из цифровизации и решать связанные с этим проблемы. Некоторые соответствующие аспекты цифровой компетентности включают показатели:

- цифровой интеграции различных бизнес-функций;
- доступа и способности использовать аналитику данных для проектирования, разработки, коммерциализации и улучшения продуктов, включая данные о пользователях продуктов организации и использовании ими таких продуктов;
- доступа к сетям и использования соответствующих решений и архитектур (аппаратного и программного обеспечения);
- эффективное управление рисками конфиденциальности и кибербезопасности;
- внедрение соответствующих бизнес-моделей для цифровых сред, таких как электронная коммерция, цифровые платформы и т.д.» (раздел 5.108).

«Цифровые платформы являются отличительной чертой цифровой эпохи. Платформы объединяют производителей и пользователей на различных этапах цепочки создания стоимости. Они часто образуют

---

<sup>1</sup> The OECD Model Survey on ICT Usage by Businesses: 2nd revision // Working Party on Measurement and Analysis of the Digital Economy. 2015. – URL: <https://www.oecd.org/sti/ieconomy/ICT-Model-Survey-Usage-Businesses.pdf>.



экосистему, в которой разрабатываются и продаются новые продукты, а также генерируются и обмениваются данными» (разделе 5.110). По мнению методологов оценки инноваций, использование цифровых платформ свидетельствует о потенциале организации в использовании цифровых технологий. Кроме того, использование цифровых платформ, объединяющих поставщиков, производителей и потребителей, генерирует данные, которые также могут способствовать росту инноваций. Ссылаясь на исследование об использовании цифровых платформ (Evans, Gawer, 2016), авторы Руководства Осло делают вывод, что «эти платформы обеспечивают благодатную почву для развития и распространения инноваций», и, следовательно, для оценки инноваций целесообразно выяснить, насколько в организации и у ее партнеров используются цифровые платформы. Возможно, что и для оценки инноваций, использующих цифровые технологии, стоит учитывать их участие в работе цифровых платформ. Так, например, в статье<sup>1</sup> представлена глобальная модель цифровых платформ в стейкхолдинговом капитализме. Описывается роль, которую большие данные играют в формировании новых расширенных инновационных экосистем цифровых платформ.

### **5.2.3. Стандарты ISO в области инноваций**

Стандарты ISO в области управления и оценки инноваций разрабатываются техническим комитетом ISO/TC 279 «Стандартизация терминологии, инструментов, методов и взаимодействия между соответствующими сторонами для обеспечения инноваций». Его работу поддерживает французская организация по стандартизации AFNOR. Именно AFNOR выступила с предложением создания серии стандартов ISO, посвященных инновациям. Основу новых документов должны составить, на основе европейских стандартов серии CEN/TS 16555<sup>2</sup>, которые, в свою очередь, обобщили опыт стандартизации менеджмента инноваций в Великобритании, Ирландии, Испании и Португалии.

Стандарт ISO 56002 посвящен общим вопросам организации инновационной деятельности на предприятиях. В нем сформулированы принципы инновационного менеджмента: реализация ценности, лидерство, нацеленное в будущее, стратегическая направленность инноваций, инновационная культура, использование уникальной и

---

<sup>1</sup> Paredes-Frigolett H., Pyka A. The global stakeholder capitalism model of digital platforms and its implications for strategy and innovation from a Schumpeterian perspective. *Journal of Evolutionary Economics*, Vol. 32, no. 2, pp. 463–500.

<sup>2</sup> Хохлявин С., Кудрявцева, Ю. Европейский стандарт для управления инновационной деятельностью как предтеча международного // Стандарты и качество. 2013. № 11. С. 46-48.

достоверной информации, управление неопределенностью, адаптируемость и системный подход к инновациям. Дана общая схема системы инновационного менеджмента (Innovation Management System, IMS), куда входит среда предприятия, лидерство, инструменты инновационного развития. Как и в большинстве стандартов ISO, гармонизированных со стандартом ISO 9001, основным форматом развития является цикл Деминга (Plan – Do – Check – Act, PDCA). С точки зрения анализа критериев оценки инноваций, интересна формулировка требований к целям инновационного развития, которые должны распространяться и на инновации, использующие цифровые технологии. В стандарте отмечено семь требований к целям инновационного развития:

- соответствие инновационной политике организации;
- наличие на всех уровнях и для всех функций организации;
- измеримость и верифицируемость;
- соответствие всем применимым требованиям;
- возможность непрерывного контроля;
- доступность для передачи и понятность;
- возможность обновления при необходимости быть обновляемым.

Стандарт ISO 56004:2019 (ГОСТ 59062) посвящен оценке инновационного менеджмента (инструменты и метрики будут описаны в стандарте ISO 56008, который пока разрабатывается). В нем даны определения и основные подходы к оценке инновационного менеджмента (Innovation Management Assessment, IMA) на основе Руководства Осло (см. выше). Для проведения IMA в стандарте ISO 56004 предлагается определить следующие показатели:

- цель IMA (для выявления соответствия плану, для совершенствования IMS и/или для повышения возможностей инновационного менеджмента);
- охват IMA (одно, несколько или все подразделения организации);
- оцениваемые объекты (один или все);
- тип экспертизы (внутренняя или внешняя);
- тип сбора данных (по документам, интервью или опрос);
- средства сбора данных (вручную или автоматически);
- типы данных (качественные, количественные);
- инструменты анализа (вручную, частично или полностью автоматизированные);
- тип сравнения и ссылок (до и после, корреляционный анализ, бенчмаркинг);
- тип результатов (сильные/слабые стороны, недостатки, рекомендации);

- формат вывода результатов и др.

Для оценки инноваций, использующих цифровые технологии, также целесообразно определить цель и охват, типы и инструменты сбора данных, типы сравнения и выводы по результатам. Учитывая универсальность цифровых технологий и доступность для интеграции различных областей деятельности, считаем такую многомерную оценку безусловно целесообразной.

#### **5.2.4. Выводы**

Исследование оснований критериев оценки инноваций в области инновационного менеджмента, использующего цифровые технологии и ИИ, позволяет сделать следующие выводы:

- К инновациям, использующим цифровые технологии и ИИ, целесообразно отнести не только инновации, использующие цифровые инструменты, но и инновации, которые были получены с использованием цифровых технологий и ИИ. При этом допустимо установить критерии оценки, учитывающие различие между такими инновациями.
- Поскольку цифровые технологии позволяют интегрировать различные функции и отрасли, целесообразно учесть масштаб и глубину такой интеграции. Например, инновация в области дистанционного определения заражения коронавирусом на основе опроса может одновременно использовать технологии обработки больших данных, безопасный доступ к медицинским данным, психологические особенности дистанционного интервьюирования и т.п.
- К критериям оценки качества инноваций целесообразно отнести то, насколько она эффективно генерирует качественные данные, в какой мере порождает новые инновации.
- Критерием инновационности также может служить то, используется ли в инновации платформа, где взаимодействуют поставщики, производители и потребители товаров и услуг.
- При разработке критериев оценки инновационности необходимо уделить особое внимание понятию ценности. Часто ценность сводят к коммерческому успеху или к повышению интеллектуальной стоимости<sup>1</sup>, однако это понятие гораздо шире, и должно также иметь социогуманитарное измерение.

---

<sup>1</sup> *Tidd J. January A review and critical assessment of the ISO56002 innovation management systems standard: evidence and limitations // International Journal of Innovation Management. 2021. Vol. 25, no. 1. P. 1–17.*

### **5.3. ОСОБЕННОСТИ УПРАВЛЕНИЯ ИННОВАЦИЯМИ В ОБЛАСТИ ЦИФРОВЫХ ТЕХНОЛОГИЙ И ИИ**

Особенность деятельности технологических компаний и стартапов в сфере информационных (цифровых) технологий и ИИ заключается в том, что и цифровые технологии, и ИИ являются кросс-отраслевыми (или сквозными) технологиями, и применимы во всех отраслях экономики, а также и в деятельности органов власти, общественных и политических организаций. Фактически технологические компании и стартапы выступают в роли системных интеграторов, решающих все задачи организаций, связанные с использованием данных: управление внутренними ресурсами, организация производства, работа с клиентами и партнерами, и т.д. Такое широкое использование цифровых технологий связано с тем, что сегодня практически вся деятельность человека обусловлена использованием оборудования, генерирующего и использующего данные: компьютеры, планшеты и смартфоны, автоматизированные системы управления технологическими процессами (АСУТП), веб-камеры, датчики, сканеры и т.п. Безусловно широкое использование данных было бы невозможно без развития сетевых технологий: расширения инфраструктуры глобальных и локальных сетей передачи данных, современных технологий сотовой и космической радиосвязи.

#### **5.3.1 Специфика инноваций в ИТ отрасли**

Кросс-отраслевая специфика цифровых технологий существенно расширяет набор критериев инновационности. Если классические инновации можно было описать как монетизацию отраслевых знаний, обладающих научной новизной, то инновации в области цифровых технологий и ИИ описываются как монетизация новаторского использования науки о данных (data science) или иных цифровых технологий в той или иной отрасли. Это означает, что во многих случаях для того, чтобы новая информационная технология стала инновацией, не требуется отраслевой научной новизны, достаточно, чтобы использование такой технологии обладала ценностью и была уникальной для данной отрасли. Например, использование ИИ для распознавания образов, применяемое к созданию беспилотных автомобилей, не опирается на какие-либо научные открытия в области транспорта, но, поскольку использует data science в автомобильной промышленности и вполне монетизируемо, может быть отнесено к инновациям. Аналогично, использование технологий распределенного реестра для организации документооборота между организациями (так называемая технология смарт-контрактов) также можно считать

инновационной, поскольку в данном случае цифровые технологии работы с электронными документами с использованием особых инструментов криптозащиты позволили получить дополнительную ценность в области организации коммуникации предприятий.

Особенность инноваций в области цифровых технологий<sup>1</sup> и ИИ влияет и на формирование экосистем компаний и стартапов в этой области. Поскольку услуги ИТ-компаний пользуются спросом во всех отраслях, они развиваются наиболее динамично. Можно сказать, что экосистемы ИТ компаний и стартапов сформировались и стали значимыми в экономике за последние два десятилетия. Более того, ИТ-компании и компании, в основе деятельности которых лежат цифровые технологии (например, Amazon или Alibaba Group), с точки зрения инвесторов гораздо привлекательнее компаний из реального сектора (энергетики, торговли, транспорта и т.п.). Так среди компаний – мировых лидеров по капитализации на середину сентября 2022 года в первой пятерке (чья капитализация превысила 1 триллион долларов) четыре компании из сектора информационно-коммуникационных технологий (ИКТ), причем они все из США: Apple, Microsoft, Amazon, и Alphabet<sup>2</sup>. В десятке же крупнейших стартапов мира<sup>3</sup> (Ant Financial, ByteDance, Uber, Didi, Alibaba Cloud) все компании являются либо представителями ИТ отрасли, либо технологическими компаниями в областях финансов, медиа, отдыха, транспорта и т.п., причем подавляющее большинство (четыре из пяти) стартапов – китайские. В отличие от мировых трендов в России среди первых пяти самых дорогих компаний лишь Сбербанк можно с натяжкой отнести к ИТ отрасли (по крайней мере, эта компания так себя позиционирует), все остальные компании представляют собой нефте- и газодобывающие компании<sup>4</sup>, причем их капитализация на порядок меньше лидеров рынка.

Неразвитость экосистем ИТ компаний в России во многом связано со слабостью аутсорсинга в ИТ отрасли. Большинство крупных отечественных компаний, как частных, так и государственных, предпочитают иметь собственные ИТ-компании, которые часто превосходят по масштабам самых крупных игроков ИТ отрасли. Так, например, по данным информационно-аналитического агентства

---

<sup>1</sup> Славин Б. Информационные технологии и инновации. Инноватика и экспертиза, 15(2), стр. 28-37.

<sup>2</sup> <https://ru.tradingview.com/markets/world-economy/worlds-largest-companies/> - дата обращения: 16.09.2022 год

<sup>3</sup> <https://investfuture.ru/news/id/top-10-krupneyshih-startapov-v-mire> - дата обращения: 16.09.2022 год

<sup>4</sup> <https://www.forbes.com/lists/global2000/> - дата обращения: 16.09.2022 год

Tadviser<sup>1</sup> на конец 2019 года в технологическом блоке Сбербанка работало более 20 тысяч человек, что вдвое больше, чем в крупнейшей российской ИТ компании Яндекс, и в разы больше, чем в российских системных интеграторах. По данным Ассоциации предприятий компьютерных и информационных технологий (АПКИТ) численность ИКТ-кадров в России в 2019 году, работающих в организациях различных отраслей экономики и в госучреждениях, составляет 0,95 млн человек, и только 0,5 млн заняты в самой ИТ-отрасли<sup>2</sup>. Корпоративные ИТ службы ограничены в возможностях проведения научных изыскания в области Computer Science, поскольку предназначены, прежде всего, для решения эксплуатационных задач. Отсутствие достаточного количества крупных ИТ компаний также снижает научный потенциал в области ИТ. Все это существенно уменьшает инновационный потенциал российского ИТ рынка.

### 5.3.2 Роль стартапов в ИТ отрасли

Бурный рост ИТ компаний в мире, а также кросс-отраслевой, т.е. широкий по охвату целевой рынок применения цифровых технологий и ИИ, ведет к тому, что ни одна из ИТ-компаний не может развивать инновации только собственными силами. Именно этим обстоятельством обусловлен резкий рост экосистемы стартапов в области ИТ. Сегодня в мире ежегодно создаются порядка 50 млн. стартапов, 90% из которых потерпят неудачу (причем 10% уже в течение первого года)<sup>3</sup>. При этом инвесторы считают, что именно ИИ является наиболее перспективной технологией для стартапов. По сути дела, стартапы создают сегодня питательную среду для крупных ИТ-компаний, которые отбирают наиболее перспективные технологические разработки. При этом часто стартапы создают более совершенные сервисы, чем ИТ-гиганты. Показательным примером в такой экосистемной кооперации является, например, покупка компанией Facebook мессенджера WhatsApp. На момент покупки Facebook была уже крупной компанией с капитализацией в несколько десятков миллиардов долларов, и также разрабатывала свой собственный мессенджер. Компания же WhatsApp, хотя и была стартапом, ведущим свою деятельность на средства инвесторов, оказалась более успешной, и Facebook купил ее почти за 20 миллиардов долларов.

---

<sup>1</sup> [https://www.tadviser.ru/index.php/Статья:Больше\\_всего\\_айтишников\\_ждут\\_в\\_“Сбербанке”](https://www.tadviser.ru/index.php/Статья:Больше_всего_айтишников_ждут_в_“Сбербанке”) - дата обращения: 27.07.2021 год

<sup>2</sup> [https://www.tadviser.ru/index.php/Статья:Рынок\\_труда\\_в\\_России\\_\(ИТ\\_и\\_телеком\)](https://www.tadviser.ru/index.php/Статья:Рынок_труда_в_России_(ИТ_и_телеком)) - дата обращения: 16.09.2022 год

<sup>3</sup> <https://trends.rbc.ru/trends/innovation/5f04aeac9a79479c0727f494> - дата обращения: 16.09.2022 год

Рынок стартапов в России безусловно слабее, чем в упомянутых ИТ-развитых странах. Согласно совместному исследованию PwC и РВК «MoneyTree»<sup>1</sup> в 2019 году инвестиции в стартапы не превышали 500 млн. долл., что составляет менее процента от рынка США. Согласно исследованию рынка технологического предпринимательства в России «Startup Barometer 2020»<sup>2</sup>, проведенном в 2020 году при поддержке компании Ростелеком и Фонда развития интернет-инициатив (ФРИИ), более 75% российских стартапов не получают венчурного финансирования. Однако наиболее серьезной проблемой российских стартапов является то, что они не опираются на современные достижения и в России отсутствует современная инновационная среда. Успех Кремниевой долины был обеспечен тем, что она организована Стэнфордским университетом, студенты которого и создавали различные инновационные компании, опираясь на помощь университетских профессоров. В России большинство вузов, хотя и создали свои акселераторы или технопарки, но поддерживают такие инициативы формально, и не добиваются сколько-нибудь значительных успехов. В Сколково вообще пошли по другому пути, сначала создали инновационный центр, а потом университет, который, по сути, сам является стартапом, и ничем не может помочь молодым инновационным компаниям.

Кейс с Facebook и WhatsApp позволяет сделать еще несколько выводов об особенностях инноваций в области цифровых технологий и ИИ. Во-первых, в отличие от обычных инноваций, выделяемый на инновации бюджет не играет определяющей роли. В эпоху завоевания космоса и создания ядерного оружия лидерами становились те державы, кто мог выделить соответствующий бюджет на новые технологии. То, что Facebook, имея намного больший бюджет на создание мессенджера, проиграл маленькому стартапу WhatsApp, показывает, что роль инвестиций в ИТ индустрии не так высока. Еще одна особенность связана с характером покупки стартапа WhatsApp: компания Facebook купила не права на продукт, как это делали раньше компании, покупающие новые технологии, а приобрела всю команду, вместе с лидерами. В цифровую эпоху большую роль играют команды, а не сами технологии.

---

<sup>1</sup> <https://www.pwc.ru/ru/publications/money-tree/pwc-money-tree-2020.pdf> - дата обращения: 16.09.2022 год

<sup>2</sup> <https://vc-barometer.ru/startup> - дата обращения: 16.09.2022 год

### 5.3.3 Кадровый вопрос в инновациях ИТ и ИИ

В современную эпоху в области цифровых технологий инновации «живут» около трех месяцев. Не потому, что их «украдут», а потому, что такие инновации являются естественным продолжением развития самих ИТ продуктов. Например, инновационное развитие в области разработки смартфонов показывает, что все лидеры этой индустрии (американская компания Apple, корейская компания Samsung, китайские компании Huawei и Xiaomi) выводят на рынок продукты с разницей не больше года с одинаковыми новыми функциональностями, разработка которых занимает не один год. Это говорит о том, что в исследовательских лабораториях всех этих компаний инновации идентичны, и отставание не слишком значительно, хотя и играет важную маркетинговую роль.

В этой связи возникает вопрос, как компании быть инновационной в такой ситуации. Выход один – создавать инновации с периодичностью не реже раза в квартал. Именно такой режим работы сегодня установился для всех организаций, оказавшихся в центре «цифрового вихря»<sup>1</sup>, и для компаний ИКТ отрасли, и для банковской отрасли, и для телекома. Цифровая трансформация экономики<sup>2</sup> является по сути дела рынком сбыта для инноваций в ИТ, для использования сквозных технологий цифровой экономики, таких как ИИ. Однако такой повышенный спрос имеет и отрицательные последствия – рынок труда оказался не готов к такому быстрому развитию. Если в 2020 году в результате пандемии наблюдался период, когда предложение превышало спрос, то в 2021 году спрос превышал, причем более, чем на 50% предложение<sup>3</sup>.

В России проблемы рынка труда усугубляются так называемым эйдджизмом, когда возрастные сотрудники не воспринимаются также позитивно, как и молодые сотрудники. Однако по мнению аналитиков рынка труда работодатели вынуждены будут преодолеть это заблуждение, и обратить внимание на все возрастные группы, а также на тех, кто в силу разных причин был вынужден прерывать карьеру. В последнее время в мировой практике все больше уделяют внимания управлению так называемых трудовыми экосистемами, которые включают в себя не только сотрудников компании и фрилансеров, но и

---

<sup>1</sup> «Цифровой вихрь» используется как визуализация в исследовании Международного центра изучения цифровой трансформации (<https://www.imd.org/contentassets/8c5b42807da941ee95c7be87d54e5db9/20210427-digitalvortex21-report-web-final.pdf>), созданного компанией Cisco и Институтом менеджмента в Швейцарии

<sup>2</sup> Славин Б. Трансформирующая роль цифровых технологий. БИТ. Бизнес & информационные технологии, 7 (80), 20-22.

<sup>3</sup> <https://hh.ru/article/28795>



сотрудников компаний партнеров, контактирующих со штатными работниками, и даже с лояльными клиентами, которые участвуют в развитии компании.

Можно сказать, что ИТ компании, которые хотят быть инновационными и быстро меняться, должны в первую очередь создать систему обучения и переобучения своих сотрудников. В экономике знаний образование станет аналогом производства товаров в индустриальной экономике, поскольку именно компетенции сотрудников будут основным ресурсом, на основе которого компании за счет инноваций будут производить новые продукты и услуги. Уже сегодня многие компании создают корпоративные университеты, призванные поддерживать необходимые для инновационного развития компетенции в организации.

### **5.3.4 Гибкое управление инновациями в ИТ**

Помимо системы обучения, инновационные компании требуют и особых подходов к управлению. К таким новым подходам следует отнести плоские системы управления (например, социократические и бирюзовые практики<sup>1</sup>), которые позволяют вовлечь в управление большую часть креативных сотрудников. Инновационные предприятия используют гибкие системы проектирования проектами (например, Agile), которые позволяют создавать новые продукты. В арсенале инновационных ИТ компаний или компаний, активно внедряющих цифровые технологии, должна быть фабрика разработки DevOps, которая объединяет непрерывную разработку с тестированием и эксплуатацией.

Не менее важным инструментом управления инновационными предприятиями в области ИТ становится сегодня венчурное бюджетирование<sup>2</sup>. Венчурное бюджетирование является аналогом высокорискованной методологии инвестирования в стартапы, предполагающая разбиение проектов по портфелям таким образом, чтобы можно было получать положительный эффект от портфеля проектов, тогда как по некоторым проектам может быть отрицательный эффект. Очень часто надо провести пилотное внедрение той или иной инновационной задачи, чтобы понять, стоит ли вообще двигаться в эту сторону, возможно ли использовать те или иные технологии. Если не попробовать, то ни о каких инновациях можно уже не думать. Но

---

<sup>1</sup> Лалу Ф. Открывая организации будущего. Пер. В. Кулбина. Москва: Манн, Иванов и Фербер; Робертсон Б. Холакратия. Революционный подход в менеджменте. Москва: ЭКСМО.

<sup>2</sup> Аншина М., Славин Б., Уайт Т. Цифровая трансформация бизнеса. Москва: Кнорус.

некоторые из таких пилотных внедрений оказываются неудачными. Однако эти неудачи можно воспринимать как расходы на поиск правильного решения.

Особое значение в создании гибкости и адаптивности цифровых технологий играют так называемые цифровые платформы<sup>1</sup> – информационные системы, которые позволяют легко интегрировать в единую систему субъекты экономической деятельности. Цифровые платформы стали драйвером инноваций всей экономики, создавая новые сервисы, которые были невозможны в доцифровую эпоху (каршеринг, сервисы доставки, агрегаторы перевозок и т.д.). Особую роль в эффективности цифровых платформ играют технологии ИИ, которые позволяют формировать совершенно новые сервисы, рекомендательные, предсказательные. Не будет преувеличением утверждение, что цифровые платформы стали быстрым инструментом создания экосистемы сбыта, на которую раньше уходили годы. Так, например, цифровая платформа Alibaba позволила в течение нескольких лет китайской промышленности выйти на международный уровень, став серьезным конкурентом американским, европейским и японским компаниям.

### **5.3.5 Выводы**

Таким образом можно сделать выводы, что управление инновациями во всех цифровых технологиях, включая ИИ, требует особого подхода, поскольку такие инновации, как правило, носят междисциплинарный подход. При выборе тех или иных инструментов для инноваций необходимо сочетать как последние достижения в цифровых технологиях, так и достижения в отрасли, где такие инновации внедряются. Поскольку объем инноваций быстро растет, необходимо использовать не только корпоративные центры разработок, но и экосистему стартапов, которая позволяет включить в инновации широкий круг профессионалов. Особое значение в управлении инновациями играет человеческий фактор, необходимо научиться формировать трудовые экосистемы, чтобы управлять всеми человеческими ресурсами, а не только штатными сотрудниками. Управление инновациями в ИТ должно быть гибким, использовать не только соответствующие практики гибкого ведения проектов (типа Agile), но и специальные инструменты венчурного бюджетирования.

---

<sup>1</sup> Славин Б. Цифровые платформы - новый тренд в корпоративной автоматизации. БИТ. Бизнес & Информационные технологии, 2(85), 12-15.

## 5.4. АРХИТЕКТУРА ОБЩЕГО ИИ (СОЦИОГУМАНИТАРНЫЕ АСПЕКТЫ)

Тема ИИ, которая находится в центре внимания многих исследователей с середины прошлого века<sup>1</sup>, получила второе дыхание во втором десятилетии нашего века, благодаря выдающимся результатам использования технологии глубокого машинного обучения<sup>2</sup>. Однако успехи в области распознавания и предсказательной аналитики показали ограниченность новых технологий ИИ для моделирования реальной интеллектуальной деятельности человека. Это особенно стало очевидно, когда в реальной практике начали широко использоваться чат-боты, имитирующие человеческое общение<sup>3</sup>. Стало понятно, что ИИ еще очень далек от человеческого.

В результате этой технологии предсказания поведения и трендов, и распознавания речи, и образов получили название слабого ИИ (AWI), а технологии моделирования человеческого интеллекта – Сильного (ASI) или Общего ИИ (AGI)<sup>4</sup>, который еще не создан, и по мнению некоторых исследователей<sup>5</sup> будет создан еще не скоро, либо будет дополнением к человеку<sup>6</sup>. Основными аргументами, почему не может быть создан в обозримом будущем созданный Общим ИИ, считают (еще со времен Дрейфуса<sup>7</sup>) невозможность алгоритмизировать неявные знания человека, либо отсутствие большого числа компьютеров с ИИ, которые бы могли их социализировать. В любом случае есть общее согласие, что для разработки AGI недостаточно увеличения вычислительных мощностей и числа нейронов, а необходимо расширять функциональность систем с ИИ.

Так, например, в работе Пеи с коллегами<sup>8</sup> предлагается объединить два направления развития ИИ: алгоритмы на основе нейросетей и прикладные программы в рамках одного гибридного чипа Tianjic, что по мнению авторов позволяет реализовывать более функциональные

---

<sup>1</sup> *Russell S., Norvig P.* Artificial intelligence: a modern approach. New Jersey: Alan Apt.

<sup>2</sup> *LeCun, Y. et al.* Deep learning. Nature, May, Volume 521, p. 436–444.

<sup>3</sup> *Grudin J., Jacque R.* Chatbots, Humbots, and the Quest for Artificial General Intelligence. s.l., s.n., pp. 1-11.

<sup>4</sup> *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms.* Proceedings of the AGI Workshop 2006 ред. б.м.:IOS Press.

<sup>5</sup> *Fjelland R.* Why general artificial intelligence will not be realized. Humanities and Social Sciences Communications, Issue 7(10), pp. 1-9.

<sup>6</sup> *Korteling J. et al.* Human- versus Artificial Intelligence. Frontiers in Artificial Intelligence, March, 4(622364), pp. 1-13.

<sup>7</sup> *Dreyfus H. L.* What Computers Can't Do: A Critique of Artificial Reason. NY: Harper.

<sup>8</sup> *Pei J. et al.* Towards artificial general intelligence with hybrid Tianjic chip architecture. Nature, Volume 572, p. 106–111.

системы с ИИ. Однако по мнению авторов<sup>1</sup> просто усложнение функциональности ИИ не приведет к построению AGI, необходимо научить системы с ИИ моделировать взаимодействие живых организмов. И такого рода исследования ведутся. Так, например, в работе Александера<sup>2</sup> предлагается рассмотреть размножение ИИ со скрещиванием по аналогии с размножением биологических видов.

Реализация Общего ИИ невозможна без понимания человеческого интеллекта. В работе Легга и Хаттера<sup>3</sup> рассматриваются различные определения интеллекта и тесты (один из них – тест Тьюринга) по их выявлению. Случайно или по воле авторов, все определения в той или иной степени связаны с целеполаганием, и поэтому тесты направлены на реализацию тех или иных целей. Примерно схожий подход можно найти в работе<sup>4</sup>, где рассматривается роль мотивации в реализации когнитивных функций ИИ. ИИ по мнению авторов работы решает задачу по удовлетворению различных потребностей, пытаясь снизить элемент неопределенности при поиске решений.

Однако ни целеполагание, ни иерархия потребностей не объясняют суть человеческого интеллекта, и поэтому не могут служить основой моделирования Общего ИИ. В последнее время все больше исследователей склоняются к социальной природе человеческого интеллекта, его связи с коллективным интеллектом. Наиболее яркой публикацией в этой связи является статья трех ученых<sup>5</sup>, представляющих разные науки: лингвистику, философию и психологию. Они прямо заявили, что человеческий интеллект не находится только в мозге, он распределен между другими людьми, связан с сотрудничеством и аутсорсингом части знаний человеком. При этом авторы статьи ссылаются на работы в области технологий коллективного интеллекта группы Маллоуна из Массачусетского технологического института<sup>6</sup>.

Технологии коллективного интеллекта получили широкое распространение в последнее время в связи с массовым использованием сетевых технологий<sup>7</sup>. Такие технологии позволяют повысить

---

<sup>1</sup> *Shevlin H. et al.* The limits of machine intelligence. EMBO reports, 20(e49177), pp. 1-5.

<sup>2</sup> *Alexander S. A.* AGI and the Knight-Darwin Law: Why Idealized AGI Reproduction Requires Collaboration. St. Petersburg, Russia, s.n., pp. 1-11.

<sup>3</sup> *Legg S., Hutter M.* Universal Intelligence: A Definition of Machine. Minds & Machines, August, Volume 17, p. 391–444.

<sup>4</sup> *Bach J.* A Motivational System for Cognitive AI. s.l., s.n., pp. 232-242.

<sup>5</sup> *Sloman S., Patterson R., Barbey A.* Cognitive Neuroscience Meets the Community of Knowledge. Frontiers in Systems Neuroscience, October, 15(675127), pp. 1-13.

<sup>6</sup> *Woolley A. et al.* Evidence for a Collective Intelligence Factor in the Performance of Human Groups. Science, October, Issue 330, pp. 686-688.

<sup>7</sup> *Славин Б.* Технологии коллективного интеллекта. Проблемы управления, 2016, вып. 5, стр. 2-9.

эффективность коллективной интеллектуальной деятельности, но по своей сути они реализуют социальную сущность интеллекта, и могут быть использованы для моделирования Общего ИИ. Так, например, авторы статьи в журнале<sup>1</sup>, выделяя три подхода к ИИ: технологический, основанный на существующих технологиях ИИ; человеко-центричный (ИИ является инструментом для человека), и подход на основе коллективного разума, предлагают объединить все три подхода в рамках создания гибридного коллективного разума.

#### 5.4.1 Архитектурный подход и уровневая модель AGI

Как же моделировать человеческий интеллект, если он по своей сути социален? Самый простой путь – найти аналогичную систему, для которой уже строятся модели. Часто такие аналоги ищутся среди биологических систем, однако не все биологические системы социальны. Хорошим аналогом, как это ни странно звучит, может выступить информационная система (ИС) предприятия. Как и в случае человека, ИС предприятия поддерживает работу не только самого предприятия, но и коммуникации с клиентами, партнерами и т.п. Как и мозг человека, предприятие имеет сложную инфраструктуру, работающую с использованием локальных и глобальных вычислительных сетей, серверов и сетевых операционных систем, систем управления данными.

Данные и вычислительное оборудование обычно похожи у разных предприятий, и не позволяют в общем случае объяснить, для каких бизнес целей они служат. При этом в случае выхода из строя оборудования или программного обеспечения какие-то конкретно функции ИС для бизнеса пропадают. Это напоминает исследования человеческого мозга, когда нельзя непосредственно сказать как мозг осуществляет те или иные функции интеллекта, но при удалении части мозга, пропадают определенные функциональности, и можно делать утверждения о части мозга задействованных в реализации данных функций. Сложная зависимость говорит о том, что в обоих случаях (и интеллекта человека, и ИС предприятия) системы многоуровневые.

В корпоративных информационных технологиях построение информационной системы на основе многоуровневых моделей получило название архитектурного подхода. Впервые данный подход предложил Джон Захман в своей статье 1987 года<sup>2</sup>, тогда он работал в компании

---

<sup>1</sup> Peeters M. et al. Hybrid collective intelligence in a human–AI society. AI & SOCIETY, Volume 36, p. 217–238.

<sup>2</sup> Zachman J. A framework for information systems architecture. IBM Systems Journal, 1987, 26(3), pp. 276-292.

IBM, для планирования бизнес-систем. Слово «архитектура» он использовал по аналогии со зданием для демонстрации уровневого подхода. К настоящему времени архитектурный подход стал одним из самых распространенных подходов для построения ИС предприятий. Существует большое число фреймворков для разного типа предприятий (государственных, частных). Одним из наиболее известных архитектурных фреймворков является фреймворк TOGAF<sup>1</sup>, который развивает The Open Group. Имеется даже специальный открытый инструмент для моделирования архитектур ArchiMate, который был использован и в настоящей статье.

Для моделирования Общего ИИ мы будем использовать три слоя и пять уровней (по два на нижних двух уровнях) – см. Рис. 5-1. Нижний слой будет отвечать за технологии, используемые интеллектом, как искусственным, так и человеческим, для обеспечения своей деятельности. Второй слой за функциональность, связанную с внешними отношениями, и третий слой – непосредственно за деятельность самого интеллекта. Таким образом будут учтены как технологические задачи, так и социальные задачи. Разделение нижних двух слоев на два уровня позволит выстроить несколько иерархий: от самой простой до сложной.

Технологический слой целесообразно разделить на два уровня: энтропийный и процессный. Энтропийный уровень присущ всем живым системам, даже самым простейшим. Этот же уровень является необходимым элементом любой вычислительной системы, в том числе и ИИ. Основная функция задач, решаемых на этом уровне – снижать энтропию за счет использования энергии. Энтропия всегда повышается в закрытых системах, но в случае притока второй закон термодинамики не работает. Однако, чтобы организовать снижение энтропии, необходимо создать соответствующую инфраструктуру. Именно это делают все живые организмы, и это же делает вычислительная инфраструктура – не случайно формула для информации в кибернетике с обратным знаком равна энтропии. Надо понимать, что под информацией в кибернетике имеются в виду данные, имеющие носителем физические среды, подчиняющиеся физическим законам. Информация в человеческих коммуникациях имеет социальную природу, и учитывает человеческие отношения (см. ниже).

---

<sup>1</sup> <https://publications.opengroup.org/standards/togaf/specifications/c220> - дата обращения 16.09.2022

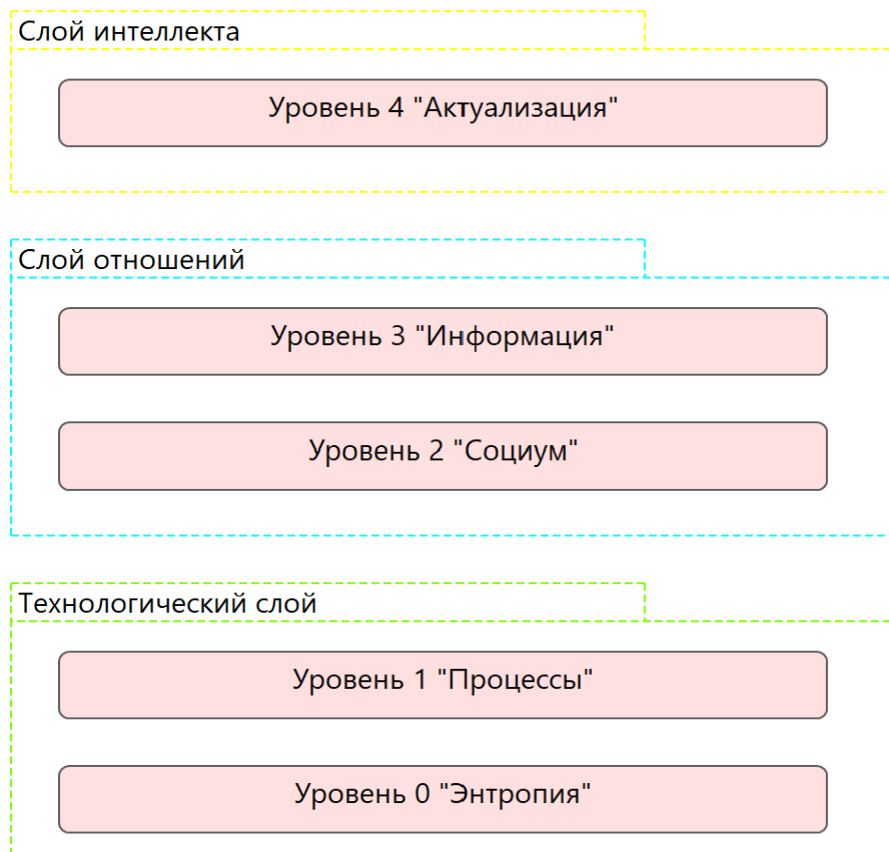


Рис. 5-1. Слои и уровни модели AGI.

Следующий уровень на технологическом слое отвечает за реализацию процессов, т.е. активности организма. Если же на этом уровне реализуются элементарные процессы, связанные с интеллектуальной деятельностью, помимо поддержки процессов на этом уровне должны реализовываться задачи, связанные с распознаванием и предсказанием. Правда, для этого понадобится на более низком уровне (энтропийном) разместить задачи, которые обычно решают нейросети. В природе память и нейросеть реализованы вместе, в вычислительном оборудовании они реализуются, как правило, отдельно. Технологический слой, содержащий два нижних уровня, фактически представляет собой слабый ИИ, т.е. не интеллект, а технологическую основу для интеллекта.

#### 5.4.2 Социальный слой и субъектность

На третьем уровне реализуется социальная сущность интеллекта<sup>1</sup>. Это как раз та сущность, без которой ИИ не может получить субъектность. На этом уровне интеллект выстраивает отношения с себе подобными: семейные отношения (которые во многом определены

<sup>1</sup> *Sloman S., Fernbach P. The Knowledge Illusion: Why We Never Think Alone. New York: Riverhead Books.*

алгоритмами, реализованные на процессном уровне), иерархии, обмен простой информацией.

Очень важно при моделировании коммуникаций соблюдать свойство субъектности. Современный компьютер позволяет поддерживать множественность коммуникаций, общаясь одновременно с большим числом участников. Фактически компьютер имитирует в такой ситуации сразу много компьютеров, которые просто используют общий технологический слой. В природе живые организмы, включая людей, способны в полную меру общаться в данный момент времени только с одним субъектом.

Это не является упущением или ограничением, это – условие субъектности, ограниченности субъекта. Допустим у нас есть человек, производительность труда которого в два раза выше, чем у других. Если бы мы измеряли человека по производительности (аналогично – по целям, по мотивациям, и т.п.), он бы не отличался от двух людей. Но это не так, человек может быть сколько угодно производителен, но время на коммуникации у него ограничено. Именно поэтому в семейных отношениях число участников невелико, невозможно устанавливать близкие отношения, если на это нет временного ресурса.

При разработке ИИ необходимо учитывать, что коммуникации с другими интеллектуальными организмами (живыми или искусственными) должны занимать все внимание интеллекта. Интеллект должен как бы проживать время вместе с другими интеллектами, только так он с одной стороны становится социальным, а с другой – обретает субъектность, отличие от других субъектов. Даже в одиночестве интеллект коммуницирует с другими, изображая их в своем сознании.

Три нижних уровня полностью моделируют поведение живых существ. Для высших животных на третьем уровне реализуется задача обучения, когда за счет коммуникаций, интеллект получает новые знания, приобретенные другим субъектом. На Рис. 5-2. показана архитектурная модель интеллекта высших животных, ведущих социальную жизнь. Она состоит из трех уровней, энтропийном, процессном и социальном.

Многие процессы у животных (например, создание семьи) реализованы на генетическом уровне. В модели это описывают алгоритмы. Реализация алгоритмов в свою очередь требует наличие памяти, органов чувств (In-Out), энергии. Задачи распознавания и предсказания требуют работы нейронной сети, и также органов чувств. Именно эти инструменты позволяют животному выстраивать иерархии, обучать и обучаться.

В случае человеческого интеллекта на слое отношений появляется еще один уровень – информационный, на котором реализуются



языковые возможности интеллекта. Под информацией имеется в виду не кибернетическая информация, связанная с энтропией, а информация, связанная с общением между людьми. Такая информация имеет следующие характеристики: актуальность, своевременность, точность и т.п.

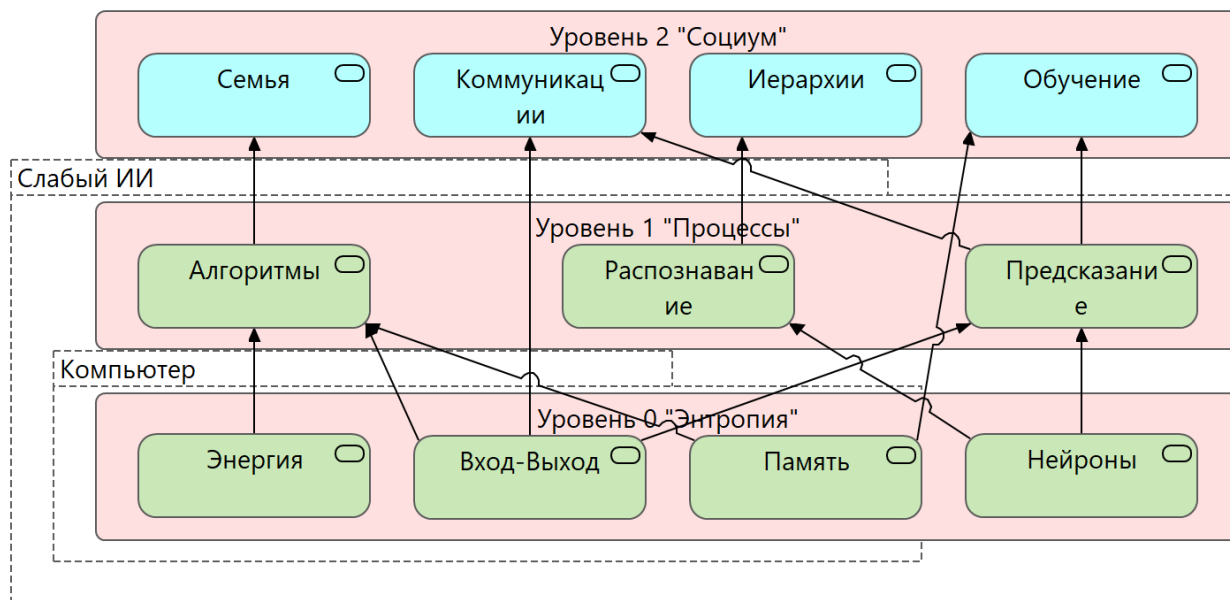


Рис. 5-2. Архитектурная модель интеллекта высших животных.

Язык является инструментом общения, коммуникаций. Но нельзя забывать, что сам язык является результатом общения и договоренностей субъектов. Выготский, анализируя отношение речи и мышления<sup>1</sup>, писал, что значение слова соответствует понятию и обобщению. Язык, состоящий из слов, таким образом, представляет собой систему понятий и обобщений, выработанных в результате совместной деятельности субъектов.

Современные чат-боты умеют разговаривать подобно людям, но это является всего лишь имитацией использования языка. Чтобы интеллект реально использовал язык для общения, он должен находиться в деятельностных отношениях с другими субъектами, не просто разговаривать, а обсуждать те или иные действия. Причем, как уже говорилось, должно выполняться требование субъектности, когда субъект общается в основном только с одним субъектом.

Символы, из которых состоят слова, или которые заменяют слова, представляют собой еще большую абстракцию. И это тоже должно стать предметом договоренности между субъектами интеллектуальной деятельности. Так же как понятия и определения. Глоссарий слов, терминов, определений и т.п. должен не просто храниться в памяти интеллекта, он должен все время верифицироваться деятельностью.

<sup>1</sup> Выгодский Л. Мышление и речь. М.: Лабиринт.

### 5.4.3 Актуализация искусственного интеллекта

Деятельность совместная с другими субъектами является важнейшей составляющей интеллекта. Слой отношений фактически является слоем приложений для внешних коммуникаций и внешней деятельности. Можно три слоя модели Общего ИИ рассматривать как диалектическую триаду: технологический слой – это внутреннее состояние интеллекта, слой отношений – внешние коммуникации, а верхний слой актуализации – внутреннее состояние, но познанное через совместную деятельность.

Аналогия с ИС предприятия сохраняется. ИТ инфраструктура позволяет предприятию поддерживать внутреннее содержание ИС. Слой приложений позволяет сотрудникам организации общаться между собой, с клиентами и партнерами. А бизнес-слой реализует стратегию и особенности развития организации.

В случае АIG на верхнем слое актуализации интеллект реализует свои стратегические целевые установки, формирует знания, которые ему необходимы для самореализации. Поскольку задачи на этом слое связаны со слоем отношений, именно здесь формируются родовые понятия и этики. И именно на этом слое интеллект проводит самоидентификацию, т.е. понимает свое место в мире.

В последнее время очень много говорится о субъектности ИИ, об этичности его использования и морали<sup>1</sup>. Если речь идет о слабом ИИ, т.е. только о технологиях, вопросы этики морали нельзя адресовать компьютеру, поскольку он не обладает субъектностью. Но если говорить об Общем ИИ, который встроен в социум, вопросы этики и морали становятся уместными, а ИИ получает субъектность.

На рисунке 5-3. показана модель АIG с составляющими, и связями между ними. Энтропийный уровень позволяет реализовать задачи на уровне процессов и социальном уровне. Процессный уровень обеспечивает социальную деятельность интеллекта, которая в свою очередь обеспечивает информационный уровень и уровень актуализации.

Субъектность интеллекта реализуется на социальном уровне как требование выделение большей части времени на коммуникации типа peer-to-peer. Актуализация, понимание субъектом его родового назначения, реализуется на верхнем уровне. И здесь также возникает особое требование, которой пока представляется гипотезой, и требует

---

<sup>1</sup> Славин Б. Может ли искусственный интеллект быть справедливым. БИТ, 2020, 113(10), С. 32-35.

дальнейшей верификации. Это особенность связана с ограниченным временем жизни интеллекта.

Человеческая жизнь – конечна, но жизнь человеческого рода – бесконечна (относительно времени человеческой жизни). Именно поэтому жизнь смертного человека можно воспринимать как инструмент преодоления смерти, но для рода в целом. По всей видимости, если бы человек был бессмертен, то и понятие рода бы отсутствовало. Конечность жизни человека делает его жизнь осмысленной, и является составной частью его актуализации.

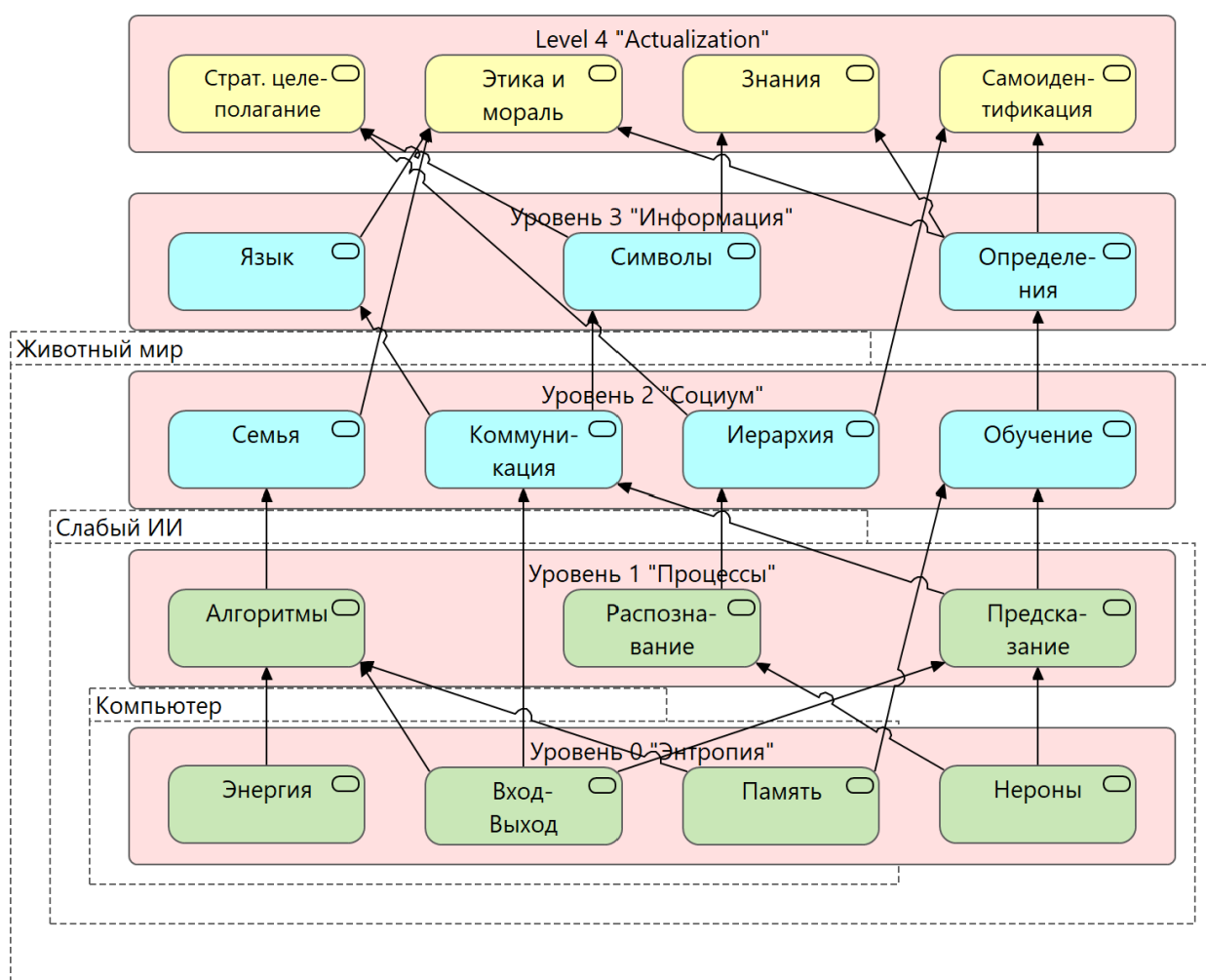


Рис. 5-3. Общая модель AGI со связями.

Аналогичная ситуация касается и предприятий, которые имеют конечный по времени жизненный цикл внутри бессмертного (по меркам времени их жизни) рынка. И именно ограниченность во времени и риск погибнуть в конкурентной борьбе, делает осмысленными бизнес-задачи, которые ставятся перед предприятиями.

Если гипотеза о важности смертной природы человека верна, значит и модель AGI должна иметь конечный по времени цикл жизни. Т.е. Общий ИИ как целое – это не процесс, а проект, с временем жизни, и с

набором стратегических задач, которые возможно интеллект сам вырабатывает. По крайней мере это является следствием архитектурного подхода.

#### **5.4.4. Выводы**

Безусловно, аналогия между интеллектом и информационной системой предприятия сама по себе не может быть подтверждением правильности выбора модели Общего ИИ. Однако, если архитектурный подход воспринимать как общесистемный подход на основе многоуровневости, и применить его к построению моделей, где причинно-следственные связи не являются прямыми, тогда его использование вполне обосновано.

В работе предлагается выделить пять уровней, расположенных на трех слоях. Три слоя – это минимально возможное разделение для многоуровневой модели общего интеллекта, где на первом слое реализуется внутренняя инфраструктура, на среднем слое внешние отношения, а на верхнем уровне реализуются стратегические цели, которые учитывают внутренние возможности и внешние ограничения.

Число уровней выбиралось из соображений того, чтобы учесть различный тип задач (например, отдельный уровень для языковых задач), добиться связности снизу вверх. Однако надо понимать, как и в любой классификации и число уровней, и правильность их расположения достаточно условно. Возможны взаимосвязи на одном уровне, и даже связи, направленные сверху вниз. Но такие аномалии должны быть редки.

В рамках построения модели AGI на основе архитектурного подхода было выявлено два важных требования, одно из которых представляется в виде гипотезы, и требует дальнейшего исследования.

Первое требование связано с субъектностью, и предполагает, что отношения ИИ с себе подобными, или людьми должны занимать основную часть деятельности интеллекта. Если вычислительный комплекс поддерживает большое число коммуникаций, это скорее всего означает, что создано такое же число возможных субъектов, которые надо социализировать по отдельности.

Второе требование связано с самоидентификацией Общего ИИ как части интеллектуального социума. Гипотеза предполагает, что Общий ИИ должен иметь конечный жизненный цикл, что станет сильной мотивацией к разработке стратегии жизни такого интеллекта, и приблизит его к человеческому интеллекту.

## **5.5. ЭТАЛОННАЯ МОДЕЛЬ РЕЗУЛЬТАТИВНОСТИ СОЦИОГУМАНИТАРНЫХ ОСНОВАНИЙ КРИТЕРИЕВ ОЦЕНКИ ИННОВАЦИЙ В СОГЛАШЕНИЯХ АРХИТЕКТУРНОГО ПОДХОДА К ПРОЕКТИРОВАНИЮ ЦИФРОВЫХ ПЛАТФОРМ**

### **5.5.1. Опора на архитектурный подход**

При постановке проблемы формирования искомой единой методологической платформы для проведения работ по оценке инноваций в цифровой рефлексивно-активной среде следует, по всей видимости, учитывать складывающиеся международные традиции и стандарты построения цифровых платформ и средств обработки больших данных с помощью методов ИИ, как инструментов обеспечения семантически адекватного взаимодействия субъектов информационного общества.

Для обеспечения такого взаимопонимания опираются на архитектурный подход, см., например, стандарт ISO 15704:2019 Enterprise modelling and architecture<sup>1</sup> или концептуальную схему построения и развития архитектуры TOGAF<sup>2</sup>. Он обеспечивает системную взаимосвязь различных субъектов и технологических компонент множества создаваемых информационных систем и подсистем, способствует постоянному росту качества данных и результатов их обработки.

Последнее особенно важно в условиях ненадежности и ошибок в интенсивно нарастающих объемах массивов информации. Качество данных может повлиять на результаты интеллектуальной аналитики и машинного обучения, особенно если входные данные имеют проблемы с характеристиками качества, такими как аккуратность, непротиворечивость, достоверность, актуальность, идентифицируемость, возможность аудита, понятность, эффективность, точность, полнота, переносимость и своевременность. Эти характеристики качества данных описаны в соответствующих международных стандартах. Организации и операторы процессов сбора, обработки и аналитики данных должны понимать такие аспекты и отражать их в механизмах управления при использовании и производстве данных на протяжении всего жизненного цикла данных в соответствии с моделью жизненного цикла данных (Рис. 5-4).

---

<sup>1</sup> ISO 15704:2019 Enterprise modelling and architecture — Requirements for enterprise-referencing architectures and methodologies

<sup>2</sup> The TOGAF Standard, Version 9.2 Overview. [Электронный ресурс]: URL: <https://www.opengroup.org/togaf>

Концепция	Требования	Планирование работы с данными	Комплектование наборов данных	Подготовка наборов данных	Построение модели	Развертывание системы	Эксплуатация системы	Вывод данных из эксплуатации	Вывод системы из эксплуатации
-----------	------------	-------------------------------	-------------------------------	---------------------------	-------------------	-----------------------	----------------------	------------------------------	-------------------------------

Рис. 5-4. Жизненный цикл данных<sup>1</sup>.

Основная цель применения архитектурного подхода состоит в том, чтобы увязать корпоративную стратегию развития организации со стратегией трансформации ее деятельности на основе цифровых технологий. Организация здесь понимается в широком смысле, это может быть: компания, отрасль, сектор экономики, регион, страна. Использование архитектурного подхода нашло свое применение в органах власти США<sup>2</sup>, Индии<sup>3</sup> и других стран.

### 5.5.2. Принципы архитектурного подхода

При реализации архитектурного подхода выдерживается принцип выделения нескольких архитектурных уровней и аспектов. Польза (выгода) от цифровой платформы может выглядеть следующим образом:

- растет производительность труда;
- улучшается социальное обеспечение и доверие населения к органам власти;
- падают затраты на создание информационных систем;
- услуги предлагаются потребителям понятно и единообразно,
- растет эффективность или качество предоставления услуг,
- предоставляются комплексные и межсегментные услуги,
- обеспечивается гибкость внесения изменений в множество взаимосвязанных информационных систем и др.

При формировании методологической платформы могут оказаться полезными отдельные формулировки из текстов принципов архитектурного подхода, например:

- соответствие процесса измерения эффективности и показателей результативности целям устойчивого развития,
- масштабируемое, совместное и повторное использование информации,

<sup>1</sup> ИСО/МЭК 8183:202X(E) Информационные технологии – Искусственный интеллект – Структура жизненного цикла данных (проект)

<sup>2</sup> Federal Enterprise Architecture Framework (USA). Version 2, 2013, 434 p. [Электронный ресурс]:

[https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/egov\\_docs/fea\\_v2.pdf](https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/egov_docs/fea_v2.pdf)

<sup>3</sup> IndEA Framework (India Enterprise Architecture Framework) [Электронный ресурс]: URL: <http://egovstandards.gov.in/sites/default/files/IndEA%20Framework%201.0.pdf>

- использование единых открытых стандартов на всех уровнях управления,
- безопасность, защита персональных данных, конфиденциальность и др.

Например, применительно к любой отрасли или сектору экономики архитектурная модель цифровой платформы (экосистемы) верхнего уровня может иметь вид, приведенный на Рис. 5-5.



Рис. 5-5. Архитектурная модель верхнего уровня

### 5.5.3. Эталонные модели

Концептуальная и функциональная части цифровой платформы в архитектурном подходе может быть представлена в виде эталонной модели результативности (Performance Reference Model), эталонной модели деятельности (Business Reference Model) и других эталонных моделях, соответствующих блокам принятой архитектуры (см. Рис. 5-5). Первая из этих двух моделей служит созданию единого механизма измерения эффективности и результативности деятельности организаций в ходе функционирования и достижения стратегических целей. Вторая модель рассматривает концепцию деятельности, задачи, функции и услуги, которые представляются на языке описания проблемной области пользователя.

Рассмотрим применение архитектурного подхода к созданию двух перечисленных эталонных моделей. При этом, учитывая фрактальный

характер архитектурных построений, полученные модели могут быть использованы в различных отраслях. На уровне архитектуры деятельности эталонная модель, отражающая аспекты эффективности и результативности, связанные с этими процессами представлена на Рис. 5-6.



Рис. 5-6. Эталонная модель деятельности

Эталонная модель деятельности представляет собой архитектурное отображение текущей ситуации в экономике отрасли. Предложенная архитектурная схема может использоваться в качестве шаблона при формулировании поисковых запросов к данным и для аналитической работы, построения прогнозов развития ситуации. Формирование поисковых запросов обеспечит релевантную выборку и последующую очистку массивов данных для дальнейшего анализа. Формулирование запросов делается, как правило, в онтологической форме. При этом для повышения качества поиска может формироваться словарь, который включает список дескрипторов, каждый из которых содержит понятия: род (вышестоящее понятие), вид (нижестоящее понятие), часть, подчасть, синонимы, антонимы, ассоциации. К искомым объектам могут быть отнесены: изделия, продукты, организации, физические и



юридические лица, действия и даже мечта потребителя<sup>1</sup>. Объекты могут находиться в состоянии постоянного изменения.

Каждый поисковый объект описывается в виде фрейма, который состоит из взаимосвязанных элементов (слотов). Например, фрейм «Нейротехнология» может включать такие элементы, как искусственный нейрон, слои нейронной сети, рекуррентная нейронная сеть, сверточная нейронная сеть, глубокая нейронная сеть, суммирование сигналов, логистическая функция и пр. Для аналитической обработки данных используются специальные эвристические алгоритмы.

Вернемся к эталонной модели деятельности (Рис. 5-6). Этап «Цели» предполагает обеспечение целеполагания и включение в поисковые предписания к массивам анализируемых больших данных целеполагающих параметров по каждому искомому сегменту рынка. Например, проведенный нами анализ документов стратегического планирования в области агропромышленной деятельности выявил около 40 различных целей и задач развития агропромышленного комплекса (АПК)<sup>2</sup>. Для достижения целей в каждом секторе экономики имеются свои программы, портфели проектов, проекты. Их также целесообразно учитывать при осуществлении поиска и анализа данных, поскольку они могут определять рамки слежения за ситуацией в сегменте рынка в контексте исследования инновационных процессов в различных сферах деятельности, включая цифровые технологии и ИИ. Этап «Организация» структурирует и определяет рамки слежения за изменением инновационной ситуации. Этот этап предполагает необходимость слежения ситуации по секторам, отраслям, сегментам рынка и видам деятельности. Этот этап предполагает формирование и сопровождение деятельности экспертного сообщества, что может быть реализовано с применением авторского подхода к организации сетевой экспертизы<sup>3</sup>.

Этап «Выгода и результат» определяет структуру аспектов деятельности сектора или отрасли экономики. В рамках выбранных компонентов формируются поисковые предписания к анализируемым массивам данных. Создаваемая система поиска и анализа данных потребует выявления количественных характеристик, определяющих эффективность функционирования секторов и отраслей экономики, основанных на использовании цифровых платформ и технологий. В

---

<sup>1</sup> Райков А.Н. Метафизика мечты// Экономические стратегии. – 2006. № 3 (С. 16-23) и № 4 (С. 22 - 25).

<sup>2</sup> Райков А.Н., Антипин С.И., Фучеджи Н.П. Архитектурные аспекты создания региональной цифровой платформы сельского хозяйства // Достижения науки и техники АПК. 2020. № 9 Том 34. С. 85-90. doi: 10.24411/0235-2451-2020-10915

<sup>3</sup> Губанов Д.А., Коргин Н.А., Новиков Д.А., Райков А.Н. Сетевая экспертиза. 2-е изд. / Под ред. чл.-к. РАН Д.А. Новикова, проф. А.Н. Райкова. – М.: Эгвес, 2011. – 166 с.

процессе поиска выделяется три основных этапа: определение характеристик поискового процесса, измерение характеристик и анализ результатов измерения, оценка качества данных, их очистка.

Данные, используемые аналитикой, реализуемой с помощью средств ИИ, могут предоставляться большим количеством и различного характера третьих лиц. Поэтому функциональность модели аналитики и ИИ может в первую очередь зависеть от надежности и прозрачности данных. В этом случае определение ответственности, поставляющей данные стороны становится критической проблемой, особенно если модель машинного обучения дает неточные или неправильные выводы или прогнозы из-за аномалии в данных, собранных с нескольких датчиков или систем. В этом случае у организаций может быть система, способная определить причину низкого качества аналитики и данных, если модель искусственного интеллекта приводит к нежелательным результатам. При этом можно опираться на ISO/IEC/IEEE 21841 для различных типов систем (SoS), находящихся за пределами организации.

Архитектурная модель результативности цифровой платформы приведена на Рис. 5-7. Одновременно эта модель может использоваться при составлении поисковых запросов к массивам данных для выявления потребности исследуемых секторов и отраслей экономики в инновационных продуктах, включая цифровые технологии и ИИ. При оценке результативности необходимо сформировать соответствующий тезаурус и фреймподобные схемы описания информационных объектов. Необходимо также, при возможности, указать интервалы значений этих параметров в количественном выражении по каждому сектору и отрасли экономики. Например, значения целевых параметров по АПК, отражающие необходимость к 2024 году достижения показателя экспорта продукции не менее 45 млрд долларов США, в сочетании с показателем «100% государственных услуг отрасли в электронном виде», а также данные о том, что Россия занимает 8-е место в мире по поголовью крупного рогатого скота (у находящейся на 6 месте Аргентины поголовье в 3 раза выше) – могут служить основанием для оценки потребности АПК в цифровых технологиях как высокой.

В приведенной на Рис. 5-7 модели выделено три основных этапа. На всех этапах задаются в основном количественные значения параметров, которые необходимо достигнуть или обеспечить в процессе мониторинга.



Рис. 5-7. Эталонная модель результативности

В самой создаваемой информационной системе перечисленные на Рис. 5-7 компоненты служат для построения адекватной структуры данных, которая позволит проводить их эффективный количественный анализ с применением технологий анализа больших данных и искусственного интеллекта. На первом этапе выделяются объекты мониторинга: сектор и/или отрасль экономики, вид экономической деятельности, которые подвергаются трансформации под воздействием цифровых платформ и/или цифровых технологий. Кроме того, выделяются характеристики типов результатов, подразумевающих наличие выходных параметров для оценки результативности. К ним, например, относятся такие экономические характеристики как себестоимость продукции, которую необходимо снижать, или добавочная стоимость, которую следует увеличивать. На этапе измерения проводится сравнение планового значения показателя и фактического его выполнения, в чем, собственно, и заключается сущность аспекта результативности. Плановые значения показателей формируются, например для сельскохозяйственной отрасли, на основе целевых показателей федерального проекта «Экспорт продукции АПК» и/или национальной программы «Цифровая экономика Российской Федерации». По секторам и отраслям экономики базовыми значениями

служат документы стратегического планирования, утвержденные проекты и программы, материалы официальной статистики и пр., а также материалы иных доступных источников данных, включая СМИ, научные публикации, патенты. На третьем этапе при обработке результатов мониторинга выделяется два основных подхода к проведению анализа результатов измерений: на основе моделирования при подключении экспертных процедур, и автоматизированный – с помощью методов и средств искусственного интеллекта, анализа больших данных, нейросетевых технологий, глубокого обучения, расчетов по определенным формулам и пр.

#### **5.5.4. Выводы**

При формировании методологической платформы для поддержки оценки инноваций в цифровой среде следует учитывать международные стандарты построения цифровых платформ и средств обработки больших данных на основе архитектурного подхода. Именно такая платформа обеспечивает системную взаимосвязь различных субъектов и технологических компонент создаваемых информационных систем и подсистем. Архитектурный подход увязывает стратегию развития организации (любой сложности) со стратегией трансформации ее деятельности на основе цифровых технологий. При реализации архитектурного подхода выдерживается принцип выделения нескольких архитектурных уровней и аспектов

Традиционный архитектурный подход, по всей видимости, под инновационную тематику должен быть модернизирован, чтобы при оценке инноваций учитывать следующие основания для построения искомых социально гуманитарных критериев:

- Наличие в деятельности организаций экспертных процедур, привносящих в процессы решения задач на основе цифровых технологий и ИИ качественную (неформализованную) информацию.
- Влияние на работу организаций потребностей бенефициаров, факторов мечты потребителя и мотивации сотрудников, которые невозможно полностью представить вербально.
- Определение субъектной ответственности поставщика данных становится критической проблемой, особенно если модель ИИ дает неправильные выводы или прогнозы из-за аномалии в данных, собранных с нескольких датчиков или систем.

## 5.6. СОЦИОГУМАНИТАРНЫЕ ОСНОВАНИЯ КРИТЕРИЕВ ОЦЕНКИ ЦИФРОВЫХ МЕДИА ПЛАТФОРМ

Массовое подключение к цифровым коммуникациям населения в конце первого начале второго десятилетия XXI века дало старт цифровой трансформации экономики. Однако различные отрасли экономики по-разному включаются в процесс цифровой трансформации. Global Center for Digital Business Transformation, созданный при Institute for Management Development (IMD), два раза в год проводит исследование – какие отрасли уже начали трансформироваться, какие еще нет. Результаты такого исследования авторы представляют в виде цифрового вихря<sup>1</sup>. Отраслью, которая первой начала цифровую трансформацию и поэтому оказалась в центре цифрового вихря – это медиа отрасль. Цифровая трансформация медиа затронула все стороны человеческой жизни, так или иначе связанные с контентом<sup>2</sup>. Такая ситуация понятна, первое, что люди, получившие доступ к сети Интернет, начали использовать – это контентные ресурсы. И сегодня инструменты обмена цифровым контентом различных форматов (сайты, цифровое вещание, социальные сети, мессенджеры и т.п.) используются большинством населения – по данным Digital 2021 Global Overview Report активными пользователями социальных медиа является 4,2 миллиарда человек, 53,6% всего населения Земли<sup>3</sup>.

Еще в 1996 году основатель компании Microsoft Билл Гейтс опубликовал эссе «Контент – это Король»<sup>4</sup>, в котором предрек цифровую трансформацию медийной индустрии. Все эти ожидания и даже больше воплотились сегодня в жизнь. В 2016 году отрасль медиа индустрии стала темой обсуждения на Всемирном Экономическом Форуме (ВЭФ) в рамках специального проекта, посвященного цифровой трансформации<sup>5</sup>. Экспертами было предложено выделить четыре этапа (волны) в цифровой трансформации медиа отрасли. Первый этап – конец 90-х годов прошлого столетия эксперты связывают с сервисами обмена файлами, и, в частности, с созданием Шоном Фэннингом peer-to-peer сервиса Napster, который позволял пользователям обмениваться

---

<sup>1</sup> International Institute for Management Development. (2021, April). <https://www.imd.org/contentassets/8c5b42807da941ee95c7be87d54e5db9/20210427-digitalvortex21-report-web-final.pdf>. Retrieved from <https://www.imd.org>.

<sup>2</sup> Friedrichsen, M., & Kamalipour, Y. (Eds.). (2017). *Digital Transformation in Journalism and News Media*. Stuttgart, Germany: Springer International Publishing AG.

<sup>3</sup> Kemp, S. (2021, January 21). <https://datareportal.com/reports/digital-2021-global-overview-report>. Retrieved from <https://datareportal.com>.

<sup>4</sup> Gates, B. (2010, May 31). <https://www.craigbailey.net/content-is-king-by-bill-gates/>. Retrieved from <https://www.craigbailey.net/>.

<sup>5</sup> World Economic Forum. (2015). <http://reports.weforum.org/digital-transformation/an-introduction-to-the-digital-transformation-initiative/>. Retrieved from <http://reports.weforum.org/>.

музыкальными файлами. Следующий этап цифровой трансформации связывают с новыми форматами, когда в начале XXI века Джобс, в то время руководивший компанией Apple, анонсировал медиаплеер iTunes, который фактически закрыл эру дискет и компакт-дисков в медиа индустрии. Третий этап (волна) начался со второй половины первого десятилетия, когда большое число людей получили доступ к сети Интернет и соответственно к цифровому контенту, благодаря мобильным технологиям и социальным сетям. Четвертый этап цифровой трансформации по мнению Всемирного экономического форума (ВЭФ) будет связан с созданием экосистем компаниями медиа индустрии.

В 2019 году исследовательская компания GWI (GlobalWebIndex) опубликовала отчет<sup>1</sup> о сравнении соотношения цифровых и традиционных медиа. На момент составления отчета в мире среднее время, проводимое пользователями цифровых медиа, составляло 6 часов 44 минуты в день, а среднее время пользования традиционными медиа – 4 часа 32 минуты. Цифровые медиа не только удовлетворяют потребность человека в получении информации, но и становятся частью его коммуникаций. Не случайно разница в два часа между онлайн и традиционными пользователями примерно равна времени, которое онлайн пользователи проводят в социальных сетях, позволяющие не только потреблять информацию, но и участвовать в ее создании, быть партнерами в производстве контента.

В создании и распространении цифрового контента ключевым фактором является вовлечение аудитории в процесс взаимодействия с контентом ещё до его потребления. А само потребление становится социальным и создаёт тренд на создание сообществ и партнёрских взаимоотношений, как между создателем контента и аудиторией, так и внутри самой аудитории.

Это также вынуждает авторов всё больше работать над выстраиванием доверительных и иногда намеренно парасоциальных отношений с потребителями их контента. Доверие – ключевой фактор, позволяющий контентмейкеру не только максимально эффективно распространять цифровую информацию, но и продавать её, так как финансовый аспект в индустрии медиа не менее важен, чем в любой другой.

Всё это приводит к тому, социогуманитарные основания критериев оценки цифровых медиа платформ заключаются в доверии и партнёрстве в цифровом контенте и на в цифровых медиа платформах. Ниже будут

---

<sup>1</sup> GlobalWebIndex. (2019). <https://www.gwi.com/reports/traditional-vs-digital-media-consumption>. Retrieved from <https://www.gwi.com/>.

рассмотрены вопросы партнерства в цифровой трансформации медиа отрасли. Лидирующее положение медиа отрасли в области цифровой трансформации делает особенно интересным исследование партнерства, поскольку такое исследование может быть распространено и на всю цифровую экономику в целом. Нетрудно понять, что партнерство как явление становится общим трендом для всех отраслей экономики, вовлеченных в цифровую трансформацию. Так, например, инструменты самообслуживания в финансовой сфере и в торговле являются соучастием клиента в оказании услуг. Рейтингование поставщиков услуг и товаров в маркетплейсах является вовлечением клиентов в оценку качества и рекомендации. Пользователи каршеринга берут на себя функции осмотра автомобилей, а покупатели, использующие сервис «Scan & Go», выполняют роль кассиров. Цифровая трансформация стирает грани между сотрудниками, клиентами, поставщиками и другими стейкхолдерами. Но не только стирает грани, но и вовлекает их в партнерские отношения.

### **5.6.1. Особенности партнерства в медиа отрасли**

В рамках упомянутого выше исследования GWI, а также их исследования, проведенного в 2021 году<sup>1</sup>, его авторы делают вывод, что изменение времени использования тех или иных медиа форматов в последние годы замедлилось, и можно говорить о некотором насыщении в области цифровизации медиа. По всей видимости в ближайшие годы трансформация форматов уже не будет такой радикальной, и палитра из цифровых и традиционных медиа сложится окончательно, хотя это не отменит конкуренцию производителей и появление новых разновидностей медиа. Так, например, с середины 2018-го года в статистике WGI появился новый формат – подкасты (дискуссии, записанные в цифровом формате), которые к 2021 году завоевали 54 минуты среднего времени пользователей, слегка потеснив другие медиа форматы. Расширение форматов приводит к тому, что и у специалистов, работающих в медиабизнесе, также появляются новые возможности перехода к более разнообразным задачам, более тесному сотрудничеству между различными медиа-профессиями<sup>2</sup>.

Основное отличие современных цифровых медиа от классических связано с тем, что цифровые коммуникации позволяют легко получить обратную связь от потребителей контента<sup>3</sup>. Такая возможность была и

---

<sup>1</sup> GlobalWebIndex. (2021). <https://www.gwi.com/reports/global-media-landscape>. Retrieved from <https://www.gwi.com/>.

<sup>2</sup> Malmelin, N., & Villi, M. (2017). Media work in change: Understanding the role of media professionals in times of digital transformation and convergence. *Sociology Compass*(11).

<sup>3</sup> Lindgren, S. (2017). *Digital Media & Society*. London: SAGE Publications Ltd.

раньше (и сейчас еще часто используется), когда на радио или телевидение можно было позвонить в прямом эфире, или написать письмо в газету, но число таких звонков или писем очень ограничено. Сегодня практически все социальные медиа предоставляют возможность потребителям контента оставлять свои комментарии, выражать свое одобрение или неодобрение материалом. Более того, часто комментарии и рекомендации являются неотъемлемой и значимой частью всего контента<sup>1</sup>. Особенно это касается таких социальных медиа, которые не предполагают большого по размеру контента (Twitter, Facebook, Instagram, Telegram и др.). Но даже при размещении большого контента (например, видео на хостинге YouTube), комментарии к нему могут играть значимую роль. Например, при публикации видеороликов в области строительных технологий обсуждение размещенного контента профессионалами может оказаться не менее полезным, чем сам ролик.

Партнерства в медиа отрасли призваны решить две задачи: первое – это улучшение контента (за счет его разнообразия у разных партнеров), а второе – повышение доверия к нему, прежде всего, за счет большей объективности информации и возможности давать вменяемые объяснения событиям и аналитическим заключениям, даваемым, в частности с помощью ИИ. Типы партнерств в цифровых медиа отрасли можно разделить на традиционные и сетевые. К первым относятся партнерства между создателями контента. Здесь возможны два варианта: использование контента (когда один создатель контента использует контент другого) и совместное создание контента (когда создатели контента объединяются для создания общего контента). Использование чужого контента часто используется в телевидении, когда в новостях предлагается видеоряд, произведенный специализированными агентствами или партнерскими студиями. К сетевым партнерствам можно отнести комментарии к контенту, и рейтингование (рекомендации) контента. Традиционные типы партнерств существовали и в доцифровую эпоху (хотя партнерство в создании контента чаще стали использовать в сети Интернет), сетевые же – напротив, стали возможны только в связи с использованием сетевых инструментов продвижения контента.

Сетевые партнерства сегодня доминируют в социальных сетях, традиционные же типы представлены как в классических медийных форматах (радио, телевидение, кино и печать), так и в цифровых (в основном на видеохостингах и в цифровых кинотеатрах). Успех сетевых

---

<sup>1</sup> *Slavin, B., & Slavin, A. (2021). Features Of the Media Environment in The Conditions of Digitalization of Social Communications. IFAC PapersOnLine 54-13 (pp. 393-396). Moscow: IFAC.*



партнерств связан с большим, чем в традиционных, разнообразием контента, который создается многочисленными блоггерами, но главное – с возможностью участия в обсуждении контента. Не случайно блоггеры почти никогда не отключают комментарии к своим постам, хотя часто такие комментарии бывают нелицеприятны, и порой несправедливы. Традиционные и сетевые партнерства существенно различаются по числу участников партнерств, в традиционных число участников – небольшое (единицы и десятки), во сетевых партнерствах – большое (сотни комментариев) и очень большое (десятки и сотни тысяч оценок). Это говорит о том, что доверие к сетевым типам партнерств выше, чем к традиционным.

### **5.6.2. Партнерство и коллективное авторство в медиа отрасли**

Серьезной проблемой партнерств в цифровом медиабизнесе является учет авторства контента. В доцифровые времена, когда контент создавался в относительно небольших масштабах, авторское право отслеживалось вручную, и особого беспокойства не вызывало. Цифровые технологии и обилие цифровых аудио и видеофайлов в сети Интернет предоставили возможность любому человеку легко использовать чужой контент для создания собственных медийных продуктов. Это может быть, например, использование известного музыкального сопровождения для оформления своего видео, или использование фото для создания уникальной презентации, или иллюстрация своего выступления кадрами из какого-либо фильма. Объем публикаций в сети Интернет с включенным чужим контентом исчисляется миллионами в день, и учесть авторское право в таких масштабах практически невозможно. Крупные видеохостинги и социальные сети вынуждены использовать технологии искусственного интеллекта и содержать большой штат сотрудников для выявления заимствований.

Проблема авторского права возникает не только в случае неправомерного использования контента, но и при учете продукции, в которой принимают участие большие коллективы авторов. Крупные изготовители медиа контента, такие как Walt Disney, используют специальные сервисы (в том числе и на основе технологии блокчейн), которые позволяют учесть в готовой медиа-продукции долю авторов, как сотрудников медиакорпораций, так и их подрядчиков (художников, композиторов, сценаристов и т.д.). Такие сервисы позволяют даже автоматически оформить лицензионные договоры на использование авторских прав. Однако никакие технические инструменты не могут решить все больше расширяющегося (и пандемии лишь его усиливает)

дисбаланса необходимости открытого доступа к информации и учета авторского права<sup>1</sup>.

Цифровая трансформация медиаиндустрии, создав проблемы учета авторства, фактически поставило общество перед необходимостью пересмотра всего авторского права, по крайней мере с точки зрения вознаграждения за творческий контент, причем уже и не только цифровой. Почему творческий продукт (например, популярный кинофильм), который потребляют миллионы людей, должен приносить его создателям существенно больший доход, чем научно-образовательный фильм, снятый для обучения немногочисленных специалистов? Как учесть в финальном продукте медиабизнеса знания, которые были использованы его авторами? Каким должно быть вознаграждение авторам в большом авторском коллективе? Чтобы ответить на эти вопросы, надо определиться с тем, зачем вообще необходим учет авторства, и что стоит за необходимостью вознаграждения.

Медийный контент, даже если у него вроде бы один автор, никогда не является продуктом одного человека. Художникам, музыкантам, поэтам приходится долго учиться, чтобы рисовать востребованные картины, сочинять и исполнять музыкальные произведения, писать стихи. Признание авторства и вознаграждение необходимо лишь для того, чтобы мотивировать творческих людей создавать уникальный контент. Но эта мотивация не должна напрямую зависеть от числа просмотренных репродукций, прослушанных песен или стихов. Оценка творческих произведений должна проводиться, но она не должна умножаться на число участников этой оценки. Более того, учет авторства надо проводить в границах тех компетенций, которые автор проявил при создании медийного продукта, что позволит легко учесть вклад автора в коллективную работу.

Компетентностный подход является основой технологий коллективного интеллекта, своего рода инструментом измерения интеллектуального и творческого труда. Адекватное измерение и соответствующее вознаграждение за творческий труд приведут с одной стороны к необходимой мотивации креативных людей, а с другой стороны позволят сделать медиапродукты доступными для всех, в том числе и для использования их в создании нового контента. Нетрудно понять, что такой подход требует реформирования всей экономики. Однако такое реформирование уже происходит: сегодня многие

---

<sup>1</sup> Shrayberg, Y., & Volkova, K. (2021). Features of Copyright Transformation in the Information Environment in the Age of Digitalization. *Scientific and technical information processing*, 48(1), pp. 30-37.

цифровые сервисы предоставляются бесплатно или частично бесплатно, а их монетизация осуществляется за счет, например рекламы или инвесторов<sup>1</sup>. Пока это касается в основном всех пользователей Интернет, и поэтому не сильно отличается от обычной товарной экономики. По мере того, как будут появляться более узкие профессиональные сообщества, реклама также будет сужаться и становиться направленной на конкретные категории пользователи. Такая специализация позволит в будущем финансировать эксклюзивную творческую деятельность, когда видеоконтент для специалистов будет оценен не менее, чем какой-нибудь блокбастер. Но переход к экономике свободного контента надо не просто ждать, а готовить его, формируя новые общественные отношения и меняя законодательные основы авторского права.

Технологии коллективного интеллекта, о которых уже говорилось выше, должны помочь и в области регулирования медиасреды, конфликты в которой становятся все более острыми. Основная проблема управления медиасредой связана, включая обилие «фейк-ньюз»<sup>2</sup>, с отсутствием доверия, с господством так называемого «капитализма больших данных»<sup>3</sup>. Известно, что внедрение систем управления знаниями также невозможно без управления доверия. Технологии коллективного интеллекта<sup>4</sup> предполагают создание глобальной системы экспертных сетей, в которой профессионализм экспертов подтверждается реальной практикой. Именно глобальная экспертная среда даст возможность масштабной верификации информации и оценки качества медийного контента. Цифровые технологии создают уникальные возможности для самообслуживания, а следовательно, и для самоорганизации людей. А это значит, что и управление медийным цифровым контентом также должно осуществляться на основе самоорганизации.

### **5.6.3. Модель доверия к цифровому контенту и ее исследование**

Чтобы понять как доверие связано с типом партнерств, построим модель доверия к медийному контенту, которая бы связывала все типы партнерств. Можно выделить три фактора, которые влияют на доверие к

---

<sup>1</sup> *Sanz, E., Simon, J., & Prato, G. (Eds.). (2014). Digital Media Worlds: The New Economy of Media. Hampshire: Palgrave Macmillan.*

<sup>2</sup> *Martens, B., Aguiar, L., Gomez-Herrera, E., & Mueller-Langer, F. (2018). The digital transformation of news media and the rise of disinformation and fake news. Seville: European Commission, Joint Research Centre. doi:10.2139/ssrn.3164170*

<sup>3</sup> *Fuchs, C. (2020). Social Media: A Critical Introduction (3-hd ed.). London: SAGE Publications Ltd.*

<sup>4</sup> *Slavin, B. (2018). Digital technologies of intellectual collective activity. In System analysis in economics- 2018. Proceedings of the V International research and practice conference-biennale (pp. 316-318). Moscow: Prometey.*

медийному контенту. Первый – это масштаб партнерства, производящего контент. Многие люди доверяют официальным новостям потому, что в создании их контента принимает участие большое число участников: новостные агентства, репортеры, корреспонденты. Не случайно, создатели контента упоминают всех, кто им предоставил информацию – не столько ради авторских прав (которые, например, не касаются собственных корреспондентов), сколько ради того, чтобы показать масштаб участников, создающих контент. Аналогично блоггеры, которые создают контент лично, также стараются ссылаться на различные источники, чтобы повысить «вес» своей информации. Еще одним важным фактором является персонализация информации. Человеку важно, что контент, который он потребляет, создан конкретным человеком, которого он знает, или может знать. Именно поэтому в тематических передачах часто обращаются к конкретным экспертам, а не просто зачитывают те или иные заключения.

Самым важным фактором, влияющим на доверие к информации, является участие в его создании. Такое участие возможно для всех типов партнерств, в том числе и для традиционных: потребитель информации может быть привлечен к дискуссии (это часто используют на радиопередачах, предлагая звонить на станцию); он может создать контент с друзьями и потреблять его. Однако число таких потребителей контента будет невелико. Гораздо большее число вовлеченных в создание контента будет у сетевых партнерств, где для участия достаточно поставить «лайк» или написать небольшой комментарий. На Рис. 5-8 показана модель такого влияния типов партнерств в зависимости от того или иного фактора.

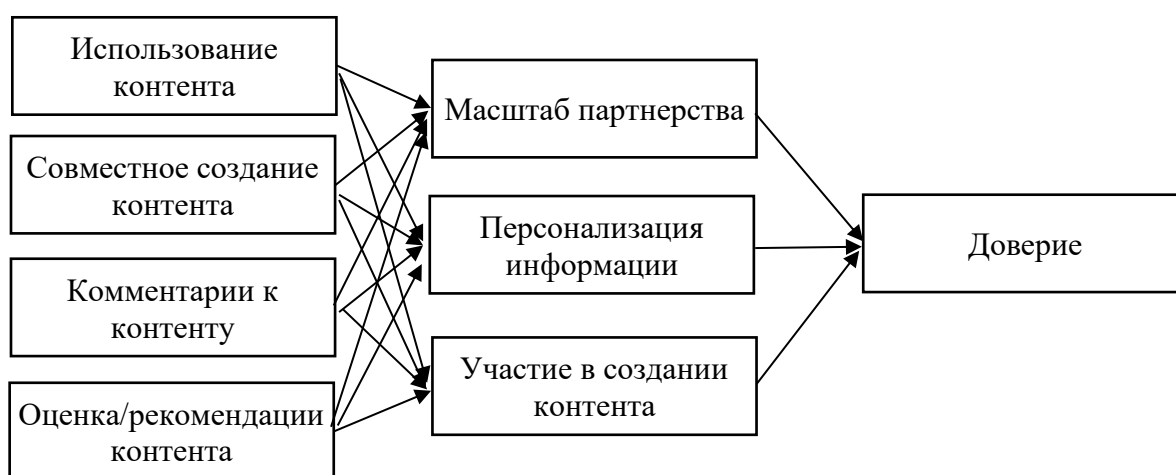


Рис. 5-8. Модель доверия к цифровому контенту

Доверие к контенту повышается и с ростом масштаба партнерства, и с увеличением персонализации, и с расширением участия потребителей к созданию контента. Традиционные партнерства, которые проигрывают сетевым в привлечении потребителей, будут соревноваться за счет масштабов привлекаемых создателей контента и персонализации контентмейкеров. Сетевые партнерства будут использовать все факторы увеличения доверия. Именно поэтому блоггеры, которые имеют миллионы подписчиков, успешно конкурируют с официальными и авторитетными средствами массовой информации. Средние и небольшие по масштабу партнерства, по-видимому, будут объединяться между собой для совместного создания контента с целью повышения доверия к своему контенту.

В рамках реализации данной модели был проведен опрос пользователей сети Интернет. Участникам опроса задавались следующие вопросы: «What sources of digital media content do you trust more?» и «What kind of content would you like to consume yourself?». На первый вопрос были возможны следующие варианты ответов: State media (State); Major media groups and holdings (Major); Teams of independent creators (Teams); Individual creators and bloggers (Individual). В скобках даны сокращения, которые будут ниже использованы в таблице. На второй вопрос варианты ответа были такие: Where the creator's opinion is expressed (Expressed); Where information is given without judgment or point of view (Impersonal); Where the experts' opinion is given (Expert); Where there are multiple points of view (Multiple). Ниже в таблице представлены результаты корреляции между этими вопросам. Например, на пересечении строки «State» и «Expressed» указана доля участников опроса, которые среди СМИ, которым они доверяют, указывали государственные медиа, при этом на вопрос, какой контент они бы сами использовали, указывали, что такой, где выражается мнение создателя.

**Таблица 5-1**  
Корреляция ответов

	Expressed	Impersonal	Expert	Multiple
State	12,5%	<b>87,5%</b>	75,0%	0,0%
Major	<b>84,9%</b>	17,0%	37,7%	<b>90,6%</b>
Teams	<b>92,3%</b>	12,5%	46,2%	<b>80,8%</b>
Individual	<b>92,2%</b>	12,7%	43,1%	<b>81,4%</b>

Согласно полученным данным как раз тем участникам, кто доверяет государственным СМИ меньше всего интересно мнение создателей контента. Зато эти участники требуют от контента быть максимально

обезличенным. В таблице выделены значения более 80%, которые можно считать сильно коррелируемыми. Интересно, что мнение экспертов ни для одного типа участников не стало основным, хотя среди доверяющих государственным СМИ желательность экспертного мнения в контенте указали 75%. Интересно, что и для тех, кто доверяет крупным СМИ, и для тех, кто доверяет небольшим творческим коллективам или индивидуальным блогерам, наличие точек зрения и их разнообразие являются очень важными. В некоторой степени плюрализма больше ждут от крупных медийных, зато выражения собственного мнения – от небольших команд и отдельных блогеров.

#### **5.6.4. Выводы**

Партнерства в социальных медиа играют важную роль в повышении доверия к контенту. Следующие факторы надо учитывать при анализе влияния партнерств на доверие: масштаб партнерства, персонализацию информации и участие потребителей в создании контента. По всей видимости, цифровая трансформация медиа отрасли приведет к тому, что коллективное участие потребителей в создание контента должно в будущем привести и к созданию новых правил работы с контентом, в том числе и к ревизии авторского права.

## 5.7. ВЫВОДЫ

1. Успешное развитие невозможно без стандартизации. Стандарты отражают многократно повторенный и устоявшийся опыт. Стандарты имеют социогуманитарную природу. Однако несмотря на то, что инновации имеют дело с чем-то неизведанным и неповторимым, инновационный процесс вполне поддается стандартизации.

2. В международном Руководстве Осло инновация трактуется как новый или улучшенный продукт или процесс, который значительно отличается от предыдущих продуктов или процессов и который был предоставлен потенциальным пользователям (как продукт) или введен в эксплуатацию как процесс. В редакции 2018 года типы инноваций сокращены до двух: продукт и процесс. Основную роль в инновациях играет знание, а не только новизна и полезность. Понятие инновации относится как к самой деятельности, так и к продукту такой деятельности.

3. Необходимо учитывать, что использование ценности, возникающей благодаря технологическим изменениям, обусловлено человеческим капиталом, инвестициями в исследования и разработки, финансовыми ресурсами, эффективным управлением. При этом продукт создается, и, следовательно наращивается ценность, при активном участии пользователя.

4. Цифровые компетенции должны быть включены в критерии оценки инноваций, использующих цифровые технологии и ИИ. Цифровые технологии заменяют человеческий труд, а также создают новые возможности для разработки новых продуктов, внедрения новых процессов.

5. В стандартах ISO отмечено семь требований к целям инновационного развития: соответствие инновационной политике организации; наличие функций организации; измеримость и верифицируемость; соответствие требованиям; возможность непрерывного контроля; доступность для передачи и понятность; возможность обновления при необходимости быть обновляемым.

6. Инновации тесно связаны с использованием цифровых технологий, поэтому критерии оценки инноваций должны учитывать, насколько эффективно используется сами цифровые технологии, в том числе при доставке инновационных продуктов клиентам, работе с инновацией в организации и т.д.

7. Управление цифровыми инновациями требует междисциплинарного подхода. В развитии инноваций необходимо сочетать последние достижения в цифровых технологиях и ИИ, а также в отрасли, где такие инновации внедряются. Объем инноваций быстро растет, поэтому необходимо использовать не только корпоративные центры разработок, но и экосистему стартапов. Особенное значение в управлении инновациями играет человеческий фактор, поэтому необходимо формировать трудовые экосистемы, чтобы эффективно управлять человеческими ресурсами.

8. Все больше исследователи склоняются к социальной природе человеческого интеллекта, его связи с коллективным интеллектом. Важна интеграция результатов разных наук: лингвистики, философии и психологии. Человеческий интеллект не находится только в мозге, он распределен по телу и клеткам человека, между другими людьми, связан с сотрудничеством и аутсорсингом части знаний человеком.

9. Моделью Общего ИИ могла бы стать архитектура информационной системы предприятия, которая поддерживает работу не только самого предприятия, но и коммуникации с клиентами, партнерами и т.п. Подобно мозгу и телу человека, предприятие имеет очень сложную инфраструктуру, правда, работающую с использованием локальных и глобальных вычислительных сетей, серверов и сетевых операционных систем, систем управления данными. Предприятие имеет и внешнюю сторону: рынок, потребителей.

10. В настоящей работе для моделирования Общего ИИ предлагается использовать три слоя и пять уровней. Нижний слой отвечает за технологии, используемые интеллектом, как искусственным, так и человеческим. Второй слой отвечает за функциональность, связанную с внешними отношениями, и третий слой – непосредственно за деятельность самого интеллекта. На третьем уровне реализуется социальная сущность интеллекта. Таким образом учитываются как технологические задачи, так и социальные. Разделение нижних двух слоев на два уровня позволяет выстроить несколько иерархий, различающихся по сложности.

11. Для человеческого интеллекта при послойном его представлении на слое отношений людей появляется еще один уровень – информационный, на котором реализуются языковые возможности интеллекта. Под информацией имеется в виду не кибернетическая информация, связанная с энтропией, а информация, связанная с



общением между людьми. Такая информация имеет следующие характеристики: актуальность, своевременность, точность и т.п.

12. В рамках построения модели общего ИИ на основе архитектурного подхода выявлены два требования, первое из которых представляется в виде гипотезы. Оно связано с субъектностью и предполагает, что отношения ИИ с себе подобными или людьми должны занимать основную часть деятельности интеллекта. Если вычислительный комплекс поддерживает большое число коммуникаций, это означает, что создано такое же число возможных субъектов, которые надо социализировать по отдельности. Второе требование связано с самоидентификацией Общего ИИ как части интеллектуального социума. Он должен иметь жизненный цикл, что станет сильной мотивации к разработке стратегии его жизни.

13. При формировании единой методологической платформы для проведения работ по оценке инноваций в цифровой рефлексивно-активной среде следует учитывать международные традиции и стандарты построения цифровых платформ и средств обработки больших данных на основе архитектурного подхода и с помощью методов искусственного интеллекта, как инструментов обеспечения семантически адекватного взаимодействия субъектов информационного общества. Именно такая платформа обеспечивает системную взаимосвязь различных субъектов и технологических компонент создаваемых информационных систем и подсистем.

14. Основная цель применения архитектурного подхода состоит в том, чтобы увязать стратегию развития организации (любой сложности) со стратегией трансформации ее деятельности на основе цифровых технологий. Организация здесь понимается в широком смысле, это может быть: компания, отрасль, сектор экономики, регион, страна. При реализации архитектурного подхода выдерживается принцип выделения нескольких архитектурных уровней и аспектов.

15. В социальных медиа играют важную роль партнерства, которые способствуют созданию адекватного пространства доверия между собой и к контенту. При этом надо учитывать следующие факторы: масштаб партнерства, персонализацию информации и участие потребителей в создании контента, методы верификации данных. Цифровая трансформация медиа отрасли ведет к тому, что коллективное участие потребителей в создании контента сгенерирует новые правила работы с контентом, в том числе и к ревизии авторского права.

## ЗАКЛЮЧЕНИЕ

---

1. Обоснован кризис техногенной цивилизации, в которой фактически игнорируются социальные ценности. Ищутся альтернативы техногенной цивилизации, одной из которых могла бы стать модель цивилизации, ориентированной на сохранение многоцивилизационного мира при обеспечении его гармоничного развития в планетарном масштабе.

2. Пока в развитии систем ИИ выделяются такие приоритетные направления: нейронные сети; глубокое обучение; компьютерное зрение; распознавание речи; генерация языка и речи; компьютерная логика и рассуждение, когнитивные вычисления; классификация и статистическое обучение; чтение сигналов мозга, нейроинформатика и др. Однако уже сейчас видно, что на пути развития ИИ выстраиваются высокие преграды, прежде всего, социогуманитарного характера.

3. Исследование вопросов целеполагания, формирования потребностей, детерминация эмоций – далеко не полностью объясняют суть человеческого интеллекта и не могут служить исключительной основой моделирования продвинутых версий ИИ, включая общего и сильного ИИ. Все больше исследователи склоняются к социальной природе человеческого интеллекта, его связи с коллективным интеллектом. Важна интеграция результатов разных наук: лингвистики, физики, математики, нейрофизиологии, философии и психологии. Человеческий интеллект не находится только в мозге, он распределен по телу и клеткам человека, между другими людьми, связан с сотрудничеством и аутсорсингом части знаний человеком.

4. Еще устойчиво убеждение, что новые технологии приводят к трансформации экономических и общественных отношений, и что они все больше обладают качеством субъектов в процессах переустройства мира. Но это далеко не так, поскольку культурные традиции, человеческие качества всегда много сильнее, особенно в когнитивных областях. Технологии находят свое место в развитии общей экосистемы научного прогресса, но они не ломают кардинально привычных устоев. Без человека цифровые технологии не имеют смысла.

5. Для преодоления возрастающих преград на пути развития ИИ может быть использован субъектно-ориентированный подход, который формируется в рамках постнеклассической кибернетики саморазвивающихся полисубъектных сред, основные положения которой опираются на представления об обществе как социальной

системе; идеи деятельностного и субъектно-деятельностного подходов; концептуальные модели рефлексивных процессов; идеи управлению сложностью; опыта разработки организационных автоматизированных систем управления страной и др.

6. Несмотря на применение гуманитарного измерения к технологиям ИИ, по-прежнему превалирует нормативный подход, административное регулирование и управленческие технологии, соответствующие предыдущим порядкам кибернетики. Обостряется вопрос поиска пути преодоления национальных эгоизмов, недоверия и страхов. Требуется переход от гармонизации среды внутри сообществ разработчиков к гармонизации взаимодействия всех акторов и субъектов мировой политики для эффективного решения проблем, возникающих при исследовании этических контекстов развития ИИ

7. Морально-этические проблемы разработки систем ИИ обусловлены не только тем, насколько вредоносно для личности и человечества могут повести себя такие системы, но и тем, что само человеческое общество может причинить страдания созданным им искусственным сущностям. Аналогичная моральная дилемма возникала при проведении научно-практических экспериментов на людях и животных. На данный момент эксперименты на людях запрещены без их согласия, животные считаются менее развитыми, а иногда и – не имеющими души, и эксперименты на них разрешены.

8. В развитии методологии цифровой индустрии во многом определяющее место занимает концепция цифрового двойника (виртуальная интерактивная копия реального физического объекта или процесса), которая, по сути, соответствует атавизмам кибернетики первого порядка. Классический цифровой двойник может быть очень сложным, однако остается пассивным по отношению к субъекту управления, это только наблюдаемый объект.

9. В кибернетике второго порядка лицо, принимающее решение (ЛПР), может делегировать часть функций автономного управления цифровому двойнику. Цифровой двойник становится самообучающейся автономной системой управления под контролем ЛПР. Цифровой двойник трансформируется в цифровой субъект.

10. В постнеклассической кибернетике саморазвивающихся полисубъектных рефлексивно-деятельностных сред цифровой двойник трансформируется уже с учетом мировоззренческих оснований, что создает основу для сборки в единое целое всех субъектов, входящих в среду, и обеспечивает ее жизнедеятельность и развитие.

11. Поиск путей повышения доверия к работе систем ИИ приводит к необходимости построения объяснимого ИИ с ориентацией на человека. Для этого в построении систем ИИ все больший акцент делается психологическим аспектам, посвященным познанию и предубеждениям, человеческому мышлению, аргументации когнитивных интерпретаций моделей ИИ. Тема объяснимости выводов ИИ далеко выходит за рамки вопроса каузальности, выявляемой статистическими методами, и охватывает гипотезу нелокальности физических явлений и, что более важно, субъективную реальность.

12. Одновременно с увеличением внимания к ИИ во власти и обществе усиливаются требования ограничить использование инструментов ИИ, например, распознавания, которые нарушают права личности и несут риски принятия ошибочных решений. В развитии ИИ усиливается акцент на необходимости повышения прозрачности работы систем ИИ, учета личностных, социальных, экономических, этических, юридических и иных вопросов.

13. В контексте развития ИИ ставятся извечные вопросы обеспечения справедливости. Например, обращает на себя внимание позиция идеологии социального либерализма, в которой для представления роли и места ИИ предлагается понимать структуру общества как совокупность социотехнических систем, функционирование которых складывается под все возрастающим влиянием ИИ, что можно охарактеризовать как упрощенное понимание организации социальных процессов в виде проявления технократического редуционизма.

14. Все большее внимание уделяется проблеме создания систем Общего ИИ, который призван решать одновременно множество разных междисциплинарных задачи. В этом контексте утверждается угроза наступления в скором времени момента сингулярности, когда ИИ станет сильнее человека во всех областях. Однако, при более внимательном рассмотрении оказывается, что важность разрешения проблемы субъектности систем ИИ еще более возрастает.

15. Для разработки Общего ИИ требуются новые парадигмальные, теоретические и методологические подходы, отвечающие принципам постнеклассической эпистемологии с учетом опыта философского анализа междисциплинарных проблем и результатов когнитивной науки в социогуманитарном антропологическом контексте. На это работает концепция саморазвивающихся рефлексивно-активных сред.

16. Все более актуальным становится решение этических вопросов применения ИИ. Пространство измерений этических основ поведения людей и морально-этических решений характеризуется бесконечным числом измерений. Дискуссии об этике ИИ уже ведутся и довольно интенсивно, и уже напрашивается потребность сходимости различных направлений обсуждения в единое конструктивное русло. При этом особого и стержневого внимания заслуживает проблема субъектности в области разработок этических основ применения ИИ.

17. Идет поиск эффективных подходов к построению систем ИИ, которые погружены в гибридную (человеко-машинную) реальность, характеризуемую высоким уровнем субъектности. В ней почти отсутствуют формализованные способы репрезентации явлений. В рамках этого видения кибернетика третьего порядка, включая конвергентную и холоническую интерпретацию явлений, является наиболее многообещающей и наиболее зрелой структурой.

18. Несмотря на применение гуманитарного измерения к технологиям ИИ, по-прежнему превалирует нормативный подход, административное регулирование и управленческие технологии, соответствующие предыдущим порядкам кибернетики. Обостряется вопрос поиска пути преодоления национальных эгоизмов, недоверия и страхов. Требуется переход от гармонизации среды внутри сообществ разработчиков к гармонизации взаимодействия всех акторов и субъектов мировой политики для эффективного решения проблем, возникающих при исследовании этических контекстов развития ИИ.

19. В контексте нового типа цивилизационного развития встает фундаментальный для сферы образования вопрос: какие парадигмы могут с успехом использовать цифровые инновации, а какие к подобным инновациям неадекватны. Требуется оценка эффективности внедрения цифровых технологий в контексте существующих образовательных парадигм и педагогических технологий и того, насколько они в принципе позволяют включать в себя те или иные цифровые технологии.

20. Требуется включенность общества в процессы управления и развития, необходимым условием чего является формирование двухконтурной структуры организации процессов развития. Первый контур исполнительный, он образуется из традиционно сложившихся структур и механизмов управления. Вторым контуром формируется с приоритетом создания общественных структур и механизмов, в котором интегрируются представители администраций, бизнеса и общества.

21. Современные системы оценок инноваций, включая цифровые технологии и ИИ, основной акцент делают на технологических факторах развития инноваций, умаляя субъективные. Технологии же обычно оцениваются по количественным метрикам с учетом потребности технологий на разных сегментах рынка. Для оценки ситуации с построением критериев, носящих социогуманитарный характер, нужны иные, скорее, дополнительные, критерии.

22. Социогуманитарные основания критериев оценки инноваций, использующих ИИ, характеризуют их востребованность обществом, особенно такими областями как медицина, образование, наука и пр. Существует широкий спектр аспектов построения социогуманитарных оснований критериев оценки инноваций, использующих ИИ. К ним стоит отнести: оценку социальной и экономической эффективности; ранние методы оценки инновационных технологий; собственно характеристики инноваций (тип инновации, статус на рынке, фаза исследования, конфликт интересов); характеристики лиц, принимающих решения; создание гибридных систем; принципы социальной ответственности; этические аспекты и др.

23. В монографии предлагается упреждающий переход к социогуманитарному технологическому укладу. Россия должна стать лидером этого уклада, чтобы не рассчитывать на безнадежные шансы догнать лидеров уходящего технологического уклада. Философско-методологические основания этого нового уклада базируются на современных представлениях постнеклассического подхода к философии науки. Для этого в настоящей монографии разработаны концептуальные основы и модели саморазвивающихся рефлексивно-активных сред, в которых создаются условия для гармоничного развития субъектов, сборки субъектов развития, конвергенции представительской и прямой демократии и др.

24. Успешное развитие инноваций, включая цифровые технологии и ИИ, невозможно без стандартизации. Стандарты имеют социогуманитарную природу. Однако несмотря на то, что инновации имеют дело с чем-то неизведанным и неповторимым, инновационный процесс вполне поддается стандартизации. В направлении стандартизации цифровых технологий, связанных с ИИ, в Российской Федерации работает ТК 164 «Искусственный интеллект», который по структуре зеркально отражает международный Технический комитет ИСО/МЭК 42 «Искусственный интеллект».

## ЛИТЕРАТУРА

---

- Авдеева З.К., Барышников П.Ю., Журенков Д.А., Зацаринный А.А., Ильин Н.И., Колин К.К., Лепский В.Е., Малинецкий Г.Г., Райков А.Н., Савельев А.М., Сильвестров С.Н., Славин А.Б., Славин Б.Б.* Стратегическое целеполагание в ситуационных центрах развития // Под ред. В.Е. Лепского, А.Н. Райкова. М.: Когито-Центр, 2018. 320 с.
- Агапова Н.Г.* Парадигмальные ориентации и модели современного образования (системный анализ в контексте философии культуры): монография // Н.Г. Агапова; Ряз. гос. ун-т им. С.А. Есенина. Рязань, 2008. – 364 с.
- Агеев А.И., Логинов Е.Л.* Нейроменеджмент личности // 2-е изд. М: ИНЭС, 2022.
- Агеев А.И., Асанова Е.А., Глибенко О.В. и др.* К «цифре» готов? Оценка адаптивности высокотехнологичного комплекса России к реалиям цифровой экономики // Под ред. А.И. Агеева. М.: ИНЭС, 2018. – 60 с.
- Агеев А.И., Логинов Е.Л.* Битва за будущее: кто первым в мире освоит ноомониторинг и когнитивное программирование субъективной реальности? // Экономические стратегии. 2017. № 2. С. 124–139.
- Агеев А.И., Логинов Е.Л., Шкута А.А.* Конвергентный мониторинг и программирование личности как инструмент оперирования интеллектуальной динамикой поведения больших групп людей // Экономические стратегии. 2018. № 2 (152). С. 70–87.
- Агеев А.И., Логинов Е.Л., Шкута А.А.* Нейроуправление: конвергентная интеграция человеческого мозга и искусственного интеллекта [Электронный ресурс] // Экономические стратегии. 2020. № 6. С. 46–57.
- Аншина М., Славин Б., Уайт Т.* Цифровая трансформация бизнеса // Москва: Кнорус. 2022. 272 с.
- Апресян Р.Г.* О праве лгать // М. РОССПЭН, 2011. 392 с.
- Аршинов В.И.* Рефлексивно-активные среды инновационного развития в контексте синергетики сложности / Междисциплинарные проблемы средового подхода к инновационному развитию // Под ред. В.Е. Лепского. М.: Когито-Центр, 2011. С. 52–73.
- Аршинов В.И., Гримов О.А., Чеклецов В.В.* Киберанимизм: искусство быть живым в гибридном обществе // Философские проблемы информационных технологий и киберпространства. 2021, (2). С. 39–60.
- Богомолов О.Т.* Крупный неэкономический резерв модернизации // Экономические стратегии. 2013, №1. С. 14–23.
- Бутова О.А., Гришко Е.А.* Особенности формирования биоэлектрической активности нейронов головного мозга военнослужащих ставропольского гарнизона в аспекте адаптации // Наука. Инновации. Технологии. 2009. № 4. С. 235–241.
- Вагин В.Н. и др.* Достоверный и правдоподобный вывод в интеллектуальных системах // Под ред. В.Н. Вагина и Д.А. Поспелова. М.: ФИЗМАТЛИТ, 2004. 704 с.
- Василевская Е.А., Менделевич В.Д.* Взаимосвязь между социальным интеллектом, антиципационными способностями и IQ у пациентов с шизофренией // XVI съезд психиатров России. Всероссийская научно-практическая конференция с международным участием «Психиатрия на этапах реформ: проблемы и перспективы»: Тезисы / Отв. ред. Н.Г. Незнанов. 2015. С. 280.
- Володина О. В.* Полипарадигмальные обоснования современного образования // Известия Волгоградского государственного педагогического университета, №. 7 (111), 2016. С. 4–11.

*Волынский-Басманов Ю.М.* Применение методов нейролингвистического программирования для выявления потенциально опасных лиц // Проблемы безопасности и чрезвычайных ситуаций. 2010. № 5. С. 124–128.

*Выгодский Л.С.* Мышление и речь // М.: Лабиринт. 2021. 567 с.

*Галушкин А.И.* Нейрочипы и нейроморфные ЭВМ: проблемы моделирования // Информационные технологии. 2015. Т. 21. № 12. С. 942–949.

*Герасимов Г.И.* Парадигмальный плюрализм как проявление кризиса образования // Теория и практика общественного развития, №. 3-4, 2009. С. 7–21.

ГОСТ Р 43.0.3-2009. Информационное обеспечение техники и операторской деятельности. Ноон-технология в технической деятельности. Общие положения. М.: Стандартиформ, 2010.

Государственный совет КНР. План развития искусственного интеллекта нового поколения. 2017. [Электронный ресурс]: [http://www.gov.cn/zhengce/content/2017-07/20/content\\_5211996.htm](http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm)

*Гринберг Р.С.* Глобализация, трансформация, кризис – что дальше? // М.: Магистр, 2011. С. 9.

*Губанов Д.А., Коргин Н.А., Новиков Д.А., Райков А.Н.* Сетевая экспертиза. 2-е изд. / Под ред. чл.-к. РАН Д.А. Новикова, проф. А.Н. Райкова // М.: Эгвес, 2011. 166 с.

*Гусейнов А.А.* Великие пророки и мыслители. Нравственные учения от Моисея до наших дней // М.: Вече, 2009. 496 с.

Доклад заместителя директора ФБР, руководителя CJIS Стивена Л. Морриса «Искусственный интеллект: ФБР и полиция против преступников» [Электронный ресурс] // Обзор отдельных вопросов в области больших данных и искусственного интеллекта // М.: ФКУ «ГИАЦ МВД России», 2019. 148 с.

*Дубровский Д.И.* Задача создания Общего искусственного интеллекта и проблема сознания // Философские науки. Т. 64. № 1. 2021. С. 13–44

*Дубровский Д.И.* Значение нейронаучных исследований сознания для разработки общего искусственного интеллекта (методологические вопросы) // Вопросы философии, Т. №2. 2022. С. 83–93.

*Дубровский Д.И.* Нейроэтика: некоторые актуальные философско-методологические вопросы // Философия. Журнал Высшей школы экономики. Т. IV, № 1. 2020. С. 24–41.

*Дубровский Д.И.* Задача создания Общего искусственного интеллекта и проблема сознания // Философские науки. Т. 64. № 1. 2021. С. 13–44.

*Дубровский Д.И.* О нравственном прогрессе и нравственном регрессе (К проблематике развития морального сознания // Философские науки, № 11, 2007. С. 81–102.

*Дубровский Д.И.* Обман. Философско-психологический анализ // М.: «Канон+», 2010. 336 с.

*Дубровский Д.И.* Природа человека и социальные процессы // Дубровский Д.И. Сознание, Мозг, Искусственный интеллект. – М.: Стратегия-Центр, 2007. С. 249–264.

*Дьяков С.И.* Психосемантическая модель и техника анализа и оценки субъектности личности // Научная конференция «Ломоносовские чтения» – 2015 / Тезисы докладов. 2015. С. 121–122.

*Ермак Е.В.* Взаимосвязь свойств когнитивной переработки аффективной информации, эмоционального интеллекта и личностных черт // Философские проблемы биологии и медицины / Сб. статей. М.: Московский государственный медико-стоматологический университет им. А.И. Евдокимова, Московское философское общество. 2015. С. 250–253.

*Ефимов А.Р.* Кандидатская диссертация «Философско-методологические основы постъюринговой интеллектуальной робототехники». Институт Философии РАН, 2021



- Ефимов А.Р.* Посттюринговая методология: разрушение стены на пути к общему искусственному интеллекту // Интеллект. Инновации. Инвестиции. № 2. 2020. С. 74–80.
- Ефимов А.Р.* Снятся ли чат-ботам андройды? Перспективы технологического развития искусственного интеллекта и робототехники // Философские науки. Т. 62. № 7. 2019. С. 73–95.
- Ефимов А.Р.* Технологические предпосылки неразличимости человека и его компьютерной имитации // Искусственные общества. 2019. Т. 14. Выпуск 4. <https://artsoc.jes.su/s207751800007645-8-1/>
- Ефимов А.Р., Алексеев А.Ю., Финн В.К.* Будущее искусственного интеллекта: тюринговая или посттюринговая методология? // Искусственные общества. 2019. Т. 14. – Выпуск 4. <https://artsoc.jes.su/s207751800007698-6-1/>
- Ефимов А.Р., Ведяхин А.А. и группа авторов.* Сильный искусственный интеллект: на подступах к сверхразуму // Интеллектуальная литература. 2021. 232 с.
- Жебель В.В., Комаров А. В., Комаров К. А., Шуртаков К. В.* Программное средство для комплексной оценки технологической готовности инновационных научно-технологических проектов // Экономика науки. 2018. Т. 4. № 1. С. 58–68.
- Журавлев А.Л., Лепский В.Е.* Проблема субъекта в инженерной психологии и эргономике // Психологический журнал. Том. 39. №4. 2018. С. 7–16.
- Захаров В.Н., Ульянов С.В.* Нечеткие модели интеллектуальных промышленных регуляторов и систем управления. II. Эволюция и принципы построения // Техническая кибернетика. № 4. 1993. С. 189–205.
- Иванов В.В., Малинецкий Г.Г.* Философские основания гуманитарно-технологической революции // Философские науки. 62(4). 2019. С. 76-95.
- Интернет-серфинг меняет мозг пользователя [Электронный ресурс] // SecurityLab. 2010. 30 августа. URL: <https://www.securitylab.ru/news/397247.php>.
- Ипполитов К.Х., Лепский В.Е.* О стратегических ориентирах развития России: что делать и куда идти // Рефлексивные процессы и управление. Том 3. №1. 2003. С. 5–27.
- ИСО/МЭК 8183:202X(E) Информационные технологии – Искусственный интеллект – Структура жизненного цикла данных (проект).
- Калинин П.В., Воюцкая Ю.Ю., Тарасов М.Е.* О применении нейроинтерфейса для бесконтактного управления мобильным устройством // Информационные системы и технологии. № 3 (95). 2016. С. 53–56.
- Китайский институт научно-технической политики при Университете Цинхуа. China AI Development Report. 2018. 144 стр. [Электронный ресурс]: <https://www.sppm.tsinghua.edu.cn/english/info/1032/1601.htm>
- Ковалевская А.В.* Информационные войны: классификация суггестивной специфики // Theoretical and practical problems of language tools transformation in the context of the accelerated development of public relations / Peer-reviewed materials digest (collective monograph) published following the results of the CXVIII International Research and Practice Conference and I stage of the Championship in Philology. Chief editor V.V. Pavlov. 2016. С. 23–25.
- Колесникова И.А.* Педагогические цивилизации и их парадигмы // Педагогика. 1995. № 6. С. 84–89.
- Коменский Я.А.* Великая дидактика // Ян Амос Коменский. М.: Книга по Требованию, 2012. 321 с.
- Кузнецов В.* Использование нейролингвистического программирования (НЛП) при допросе // Право и жизнь. № 152 (2). 2011. С. 134–140.

- Кукушев В.И.* Классификация систем искусственного интеллекта // Экономические стратегии. № 5. 2020. С. 51–57
- Кукушев В.И.* Цифровая экономика: проблемы и решения // Экономические стратегии. № 6. 2020. С. 58–67.
- Куликов В.Ю., Антропова Л.К., Козлова Л.А.* Влияние функциональной асимметрии мозга на стратегию поведения индивида в стрессовой ситуации // Journal of Siberian Medical Sciences. № 5. 2010. С. 10.
- Кун Т.* Структура научных революций: Пер. с англ. // Т. Кун; Сост. В.Ю. Кузнецов. М.: ООО «Издательство АСТ», 2003. 605 с.
- Лалу Ф.* Открывая организации будущего // Пер. В. Кулбина. Москва: Манн, Иванов и Фербер. 2016. 432 с.
- Лекторский В.А.* Субъект в истории философии: проблемы и достижения // Методология и история психологии. Том 5. Выпуск 1. 2010. С. 5–18.
- Лекторский В.А.* Эпистемология классическая и неоклассическая // М., УРСС, 2001. 255 с.
- Лекторский В.А.* Философия, искусственный интеллект и когнитивная наука // Искусственный интеллект: Междисциплинарный подход / Под ред Д.И. Дубровского и В.А. Лекторского. М.: ИИнтелЛЛ, 2006. С. 13–21.
- Лекторский В.А.* Человек и культура. Избранные статьи //СПб, 2018. 640 с.
- Лепский В.Е.* Аналитика сборки субъектов развития // М.: «Когито-Центр», 2016. – 130 с.
- Лепский В.Е.* Гармония культур в саморазвивающихся рефлексивно-активных средах (От техногенной к социогуманитарной цивилизации) // Контуры будущего в контексте мирового культурного развития: XVIII Международные Лихачевские научные чтения, 17-19 мая 2018 г. СПб.: СПбГУП, 2018. С. 431-433.
- Лепский В.Е.* Искусственный интеллект в субъектных парадигмах управления // Философские науки. Т. 64. № 1. 2021. С. 88–101.
- Лепский В.Е.* Концепция субъектно-ориентированной компьютеризации управленческой деятельности // М.: Институт психологии РАН, 1998. 204с.
- Лепский В.Е.* Междисциплинарный анализ становления субъектно-ориентированного подхода // Социальное время. 2015. № 2. С. 18–32.
- Лепский В.Е.* Методологический и философский анализ развития проблематики управления // М.: Когито-Центр, 2019. 340 с.
- Лепский В.Е.* На пути к управлению сферой образования как саморазвивающейся средой // Социология образования. №10. 2014. С. 4-24.
- Лепский В.Е.* Общественное участие в саморазвивающихся полисубъектных средах // М.: Когито-Центр, 2019. 141 с.
- Лепский В.Е.* Проблемы становления субъектности человечества и модели развития // Развитие и экономика. 2011, сентябрь. С. 95-101.
- Лепский В.Е.* Прорывная социогуманитарная технология опережающего образования // SocioTime / Социальное время, № 1(13), 2018. С. 21–32
- Лепский В.Е.* Развитие и национальная безопасность России // Экономические стратегии. №2. 2008. С. 24–30.
- Лепский В.Е.* Рефлексивно-активные среды инновационного развития // М.: Изд-во «Когито-Центр», 2010. 255 с.
- Лепский В.Е.* Рефлексия пандемии COVID-19: субъектно-ориентированный подход // Экономические стратегии. № 8 (174). 2020. С. 66–71.

- Лепский В.Е.* Седьмой социогуманитарный технологический уклад – адекватный ответ технологическим вызовам XXI века // *Философия в диалоге культур: материалы Всемирного дня философии.* М.: Прогресс-Традиция, 2010. С. 1010–1021.
- Лепский В.Е.* Седьмой социогуманитарный технологический уклад – контуры будущего человечества // *Глобальный мир: системные сдвиги, вызовы и контуры будущего: XVII Международные Лихачевские научные чтения, 18-20 мая 2017 г.* СПб.: СПбГУП, 2017. С. 357–360.
- Лепский В.Е.* Системные основания для перехода от техногенной цивилизации к социогуманитарной цивилизации // *Проблемы цивилизационного развития.* № 1. 2019. С. 33–48.
- Лепский В.Е.* Социогуманитарные критерии оценки новаций цифровой реальности // *SocioTime / Социальное время.* № 4(16). 2018. С. 16–26.
- Лепский В.Е.* Эволюция представлений об управлении (методологический и философский анализ) // М.: Когито-Центр, 2015. – 107 с.
- Лефевр В.А.* Рефлексия // М.: Когито-Центр, 2003. 495 с.
- Логинов Е.Л., Логинова В.Е., Шкута А.А.* «Дизайн мышления» элементов искусственного интеллекта для преодоления барьеров получения нового знания в электронной среде коллаборационной научной суперсистемы // *Искусственные общества.* Том 13, № 3. 2018.
- Локк Дж.* Мысли о воспитании // *Локк Дж. Сочинения: В 3 т. Т.3.* М.: Мысль, 1988. 668 с.
- Ляхов А.Ф., Тришин И.М.* Компьютерное моделирование поведения игрока в интеллектуальной карточной игре с помощью нейронной сети // *Компьютерные инструменты в образовании.* № 5. 2013. С. 54–64.
- Маанди В. И.* Школа: начинаем обратный отсчёт // *Народное образование,* № 7, 2010. С. 45–49.
- Министерство науки и технологий КНР. Этический кодекс нового поколения искусственного интеллекта. 2021. [http://www.most.gov.cn/kjbgz/202109/t20210926\\_177063.html](http://www.most.gov.cn/kjbgz/202109/t20210926_177063.html)
- Министерство промышленности и информатизации КНР. «Уведомление девяти департаментов о выпуске «Руководящих заключений по укреплению комплексного управления алгоритмами предоставления информационных услуг в Интернете». 2021. [https://wap.miit.gov.cn/xwdt/gxdt/art/2021/art\\_a8af2b48620b4905b365fc73cd81a1ec.html](https://wap.miit.gov.cn/xwdt/gxdt/art/2021/art_a8af2b48620b4905b365fc73cd81a1ec.html)
- Мнения Центрального комитета Коммунистической партии Китая и Государственного Совета об углублении реформы образования и преподавания и всестороннем повышении качества обязательного образования. Электронный ресурс. [http://www.gov.cn/zhengce/2019-07/08/content\\_5407361.htm](http://www.gov.cn/zhengce/2019-07/08/content_5407361.htm) (Кит.)
- Моисеев Н.Н.* Универсум. Информация. Общество // М.: Устойчивый мир, 2001. 198 с.
- Назаров В. Л., Жердев Д. В., Авербух Н. В.* Шоковая цифровизация образования: восприятие участников образовательного процесса // *Образование и наука.* 2021. Т. 23, № 1. С. 156–201.
- Нейроинтерфейсы для сервисов и продуктов нового поколения [Электронный ресурс] // *Basisneuro.* 2017. Ноябрь. URL: <https://basisneuro.com/BasisNeuroWhitePaper.pdf>.
- Овод И.В., Осадчий А.Е., Пупышев А.А., Фрадков А.Л.* Формирование нейрообратной связи на основе адаптивной модели активности головного мозга // *Нейрокомпьютеры: разработка, применение.* № 2. 2012. С. 36–41.
- Петров А.Н., Комаров А.В.* Оценка уровня технологической готовности конкурсных заявок с использованием методологии TPRL // *Экономика науки.* Т. 6. № 1–2. 2020. С. 88–99.

Подходы к формированию и запуску новых отраслей промышленности в контексте Национальной технологической инициативы на примере сферы «Технологии и системы цифровой реальности» и перспективные „человеко-компьютерные“ интерфейсы (в части нейроэлектроники)»: Аналитический доклад [Электронный ресурс]. URL: <http://rusneuro.net/cambiodocs/media/files/analitijeskii-doklad-podhodyk-formirovaniu-i-zapusku-novyh-otraslei-promyhlennosti.pdf>.

*Пономарева О.С., Устюжанин В.Н.* О состоянии и перспективах использования психосемантических методов познания личности подозреваемого в деятельности следственного работника // Вестник Санкт-Петербургского университета МВД России. № 2 (70). 2016. С. 190–194.

Проблемы социогуманитарного обеспечения инновационных процессов на евразийском пространстве / Под ред. В.Е.Лепского. М.: Когито-Центр, 2014. 201 с.

Прорывные инновации: человек 2.0: доклад к XXIII Ясинской (Апрельской) международной научной конференции по проблемам развития экономики и общества, Москва, 4–8 апреля 2022 г. // С.А. Гавриш, Л. М. Гохберг, Д.Е. Грибкова и др.; под ред. Л.М. Гохберга, А.Р. Ефимова, Ю.В. Мильшиной; Нац. исслед. ун-т «Высшая школа экономики», ПАО Сбербанк. М.: НИУ ВШЭ, 2022. 56 с.

*Разин А.А.* Этика искусственного интеллекта // Философия и общество. Вып. 1 (90). 2019. С. 57–73.

*Райков А.Н.* Гибридный сильный искусственный интеллект // Экономические стратегии. 2021. Т. 23. № 1(175). С. 62–63.

*Райков А.Н.* Конвергентный синтез когнитивной модели на основе глубокого обучения и квантовых семантик // International Journal of Open Information Technologies. 2018. Т. 6. № 12. С. 43–50.

*Райков А.Н.* Ловушки для искусственного интеллекта // Экономические стратегии. № 6. 2016. С. 172–179.

*Райков А.Н.* Метафизика мечты // Экономические стратегии. 2006. № 3 (С. 16–23) и № 4 (С. 22–25).

*Райков А.Н.* Моделирование коллективного бессознательного при принятии решений // Труды Международной научной конференции СРТ-2014 Международная научная конференция Московского физико-технического института (государственного университета), Института физико-технической информатики. М.: Институт физико-технической информатики, 2015. С. 146–156.

*Райков А.Н.* Субъектность объяснимого искусственного интеллекта // Философские науки. 2022. Т. 65. № 1. С. 72–90. <https://doi.org/10.30727/0235-1188-2022-65-1-72-90>

*Райков А.Н., Антипин С.И., Фучеджи Н.П.* Архитектурные аспекты создания региональной цифровой платформы сельского хозяйства // Достижения науки и техники АПК. 2020. № 9 Том 34. С. 85-90.

*Робертсон Б.* Холакратия. Революционный подход в менеджменте. Москва: ЭКСМО. 2018. 256 с.

Рубинштейн С.Л. Человек и мир. М.: Наука, 1997. 189 с.

*Савельев А.М., Журенков Д.А.* Национальные стратегии развития систем искусственного интеллекта: опыт стран - лидеров и ситуация в России // Научный вестник ОПК России. 2019. №3. С. 75-82.

*Самарцев О.Р., Латенкова В.М.* Психосемантические аспекты восприятия интерактивного дискурса в интернет-СМИ // Вестник Череповецкого государственного университета. 2016. № 2 (71). С. 87–91.

*Севостьянов Ю.О.* Изменение психосемантической структуры готовности работать в команде у студентов // Научный вестник Южного института менеджмента. 2014. № 2 С. 94–97.

*Селевко Г.К.* Энциклопедия образовательных технологий: В 2 т. // Т. 1. М.: НИИ школьных технологий, 2006. 816 с.

*Семенова Т.В., Вилкова К.А.* Типы интеграции массовых открытых онлайн-курсов в учебный процесс университетов // Университетское управление: практика и анализ Том 21, № 6, 2017. С. 114-126.

*Сергиевский Г.М., Лобачев В.С.* Моделирование поведения интеллектуального агента в проблемной ситуации с неполностью наблюдаемыми состояниями // Научная сессия НИЯУ МИФИ – 2012 / Аннотации докладов: В 3 т. 2012. С. 312.

Сильный искусственный интеллект: На подступах к сверхразуму // Александр Ведяхин и др. М.: Интеллектуальная Литература, 2021. 232 с.

*Славин Б.* Может ли искусственный интеллект быть справедливым // БИТ. 2021. №10 (113). С. 32-35.

*Славин Б.* Информационные технологии и инновации // Инноватика и экспертиза, 15(2), 2015. С. 28–37.

*Славин Б.* Технологии коллективного интеллекта // Проблемы управления, вып. 5, 2016. С. 2-9.

*Славин Б.* Трансформирующая роль цифровых технологий // БИТ. Бизнес & информационные технологии, 7 (80), 2018. С. 20-22.

*Славин Б.* Цифровые платформы – новый тренд в корпоративной автоматизации // БИТ. Бизнес & Информационные технологии, 2(85), 2019. С. 12-15.

*Смирнов А.В.* Проект многоцивилизационного мира как основание идеи многополярности: концепция всечеловеческого сегодня // Национальная философия в глобальном мире: тезисы Первого белорусского философского конгресса. Минск: Беларуская навука, 2017. С. 350-352.

*Смирнов И., Безносюк Е., Журавлев А.* Психотехнологии: Компьютерный психосемантический анализ и психокоррекция на неосознаваемом уровне // М.: Издательская группа «Прогресс» – «Культура», 1995. 416 с.

*Сорокоумов П.С., Карпов В.Э.* К вопросу о моральных аспектах адаптивного поведения искусственных агентов // Искусственные общества. 2021. Т. 16. № 2. URL: <https://arxiv.gaugn.ru/s207751800014740-3-1/> DOI: 10.18254/S207751800014740-3.

Социально-экономические аспекты внедрения искусственного интеллекта / Под науч. ред. А.И. Агеева. М.: Айти-Сервис, 2020. 192 с.

*Степин В.С.* XXI век – радикальная трансформация типа цивилизационного развития // Глобальный мир: системные сдвиги, вызовы и контуры будущего: XVII Международные Лихачевские научные чтения, 18–20 мая 2017 г. СПб.: СПбГУП, 2017. С. 185–188.

*Степин В.С.* Классика, неклассика, постнеклассика: критерии различения // Постнеклассика: философия, наука, культура. Санкт-Петербург, 2009. С. 249–295.

*Степин В.С.* Научное познание и ценности техногенной цивилизации // Вопросы философии. №10. 1989. С. 3–18.

*Степин В.С.* Социальные системы и методология прогнозирования их будущих состояний // Контуры будущего в контексте мирового культурного развития: XVIII Международные Лихачевские научные чтения, 17–19 мая 2018 г. СПб.: СПбГУП, 2018. С. 178–181.

*Степин В.С.* Теоретическое знание // М.: Прогресс-Традиция, 2003. 744 с.

*Степин В.С.* Цивилизация и культура // СПб.: СПбГУП, 2011. 408 с.

Стратегическое целеполагание в ситуационных центрах развития / Под ред. В.Е. Лепского, А.Н. Райкова / Авторский коллектив: Авдеева З.К., Зацаринный А.А., Журенков Д.А., Ильин Н.И., Колин К.К., Лепский В.Е., Малинецкий Г.Г., Райков А.Н., Савельев А.М., Сильвестров С.Н., Славин А.Б., Славин Б.Б. М.: Когито-Центр, 2018. 320 с.

Туровский Я.А., Кургалин С.Д., Адаменко А.А. Моделирование обучения нейрочипов, внедренных в нервную ткань // Цифровая обработка сигналов. 2016. № 1. С. 50–56.

Тычков А.Ю., Горячев Н.В., Кочегаров И.И. Протоколы связи для беспроводного нейроинтерфейса // Труды Международного симпозиума «Надежность и качество». Т. 2. 2018. С. 366–368.

Указ Президента Российской Федерации от 04.02.2021 № 68 «Об оценке эффективности деятельности высших должностных лиц (руководителей высших исполнительных органов государственной власти) субъектов Российской Федерации и деятельности органов исполнительной власти субъектов Российской Федерации»

Уроки «стресс-теста». Вузы в условиях пандемии и после нее // Аналитический доклад ректоров и рабочих групп университетов, Минобрнауки РФ, 2020. URL: [https://minobrnauki.gov.ru/ru/press-center/card/?id\\_4=2777](https://minobrnauki.gov.ru/ru/press-center/card/?id_4=2777)

Холодов Ю.А. Мозг в электромагнитных полях // М.: Наука, 1982. – 123 с.

Хохлявин С., Кудрявцева Ю. Европейский стандарт для управления инновационной деятельностью как предтеча международного // Стандарты и качество. 2013. № 11. С. 46–48.

Чайковский Ю.В. Автопоэз. Опыт пособия тем, кто хочет понять эволюции живого // М.: Товарищество научных изданий КМК. 2018. 560 с.

Чухрова М.Г., Чухров А.С. Пространственно-временная организация биоэлектрических процессов мозга как индикатор психосоциальной адаптации // Мир науки, культуры, образования. 2013. № 5 (42). С. 227–230.

Шуравин А. История Искусственного Интеллекта [Электронный ресурс] // <https://wiki.programstore.ru/istoriya-iskusstvennogo-intellekta/>

Эделмен Дж, Маунткасл В. Разумный мозг // Мир, 1981. 135 с.

Эфроимсон В.П. Генетика этики и эстетики // М.: Тайдекс Ко, 2004. 302 с.

Эфроимсон В.П. Родословная альтруизма (Этика с позиций эволюционной генетики человека) // В.П. Эфроимсон. Гениальность и генетика. М.: Русский мир, 1998. С. 435–466.

Ясницкий Л.Н., Сичинава З.И. Нейросетевые алгоритмы анализа поведения респондентов // Нейрокомпьютеры: разработка, применение. 2011. № 10. С. 59–64.

Яцык В.З., Чернышенко Ю.К., Пискарева О.В. Парадигмы современного образования // Физическая культура, спорт – наука и практика. 2008. № 2. URL: <https://cyberleninka.ru/article/n/paradigmy-sovremennogo-obrazovaniya-1>.

Abdovakhidov, A. M., Mannapova, E. T. & Akhmetshin, E. M. Digital Development of Education and Universities: Global Challenges of the Digital Economy. International Journal of Instruction, 14(1), 2021. P.743-760.

Abel L., Shinkins B., Smith A., et al. Early economic evaluation of diagnostic technologies: experiences of the NIHR Diagnostic Evidence Co-operatives. Med Decis Making. 2019. P.857–866.

Ackoff R. and Emery F. On Purposeful Systems. Aldine-Atherton. 1972. 288 p.

Adadi A., Berrada M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). 2018. Vol. 6. P.52138–52160.

Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms. Proceedings of the AGI Workshop. 2006.

- Alenezi M.* Deep Dive into Digital Transformation in Higher Education Institutions. *Educ. Sci.*, 11, 2021. 770 p. <https://doi.org/10.3390/educsci11120770>
- Alexander S.A.* AGI and the Knight-Darwin Law: Why Idealized AGI Reproduction Requires Collaboration. *St. Petersburg, Russia, s.n.*, pp. 1-11.
- Alkhatlan Ali and Kalita Jugal.* Intelligent Tutoring Systems: A Comprehensive Historical Survey with Recent Developments. *International Journal of Computer Applications* 181(43). 2019. P. 1-20.
- Arrieta A. B., Rodriguez N. D., Ser J. D., Bennetot A., Tabik S., Barbado A., Garcia S., Gil-Lopez S., Molina D., Benjamins R., Chatila R., Herrera F.* Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges towards responsible AI. *Information Fusion*, 201958. 2019. P.82–115.
- Bach J.* A Motivational System for Cognitive AI. *s.l., s.n.*, pp. 232-242.
- Bazarkina D., Kramar K.* The Malicious Use of Artificial Intelligence and International Psychological Security. *ASRIE Analytica Geopolitical Report*. 2020. Vol. 1. Rome: ASRIE. URL: [http://www.asrie.org/wp-content/uploads/2020/02/Geopolitical-Report-Volume-1\\_2020\\_-\\_Russian-researchers-and-MUAI.pdf](http://www.asrie.org/wp-content/uploads/2020/02/Geopolitical-Report-Volume-1_2020_-_Russian-researchers-and-MUAI.pdf)
- Bazarkina D., Pashentsev E.* Malicious Use of Artificial Intelligence: New Psychological Security Risks in BRICS Countries. *Russia in Global Affairs*. 2020. 18(4). pp.154–177.
- Beaton M., Pierce B., Stuart S.* «Neurophenomenology – a special issue» // *Constructivist Foundations*, Vol. 8 No. 3, 2013. pp. 265-268.
- Benthall S. & Haynes, BD.* Racial categories in machine learning. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, Atlanta. 2019.
- Bitbol M.* «The tangled dialectic of body and consciousness: a metaphysical counterpart of radical neurophenomenology», *Constructivist Foundations*, Vol. 16 No. 2, 2021. pp. 141-151, available at: <https://constructivist.info/16/2/141.bitbol>
- Bogdanov A.* *Essays in Tektology: The General Science of Organization*, 2nd ed., Intersystems Publications, Seaside, CA, Transl. Gorelik, C. 1984.
- Bostrom N.* «Ethical issues in advanced artificial intelligence», *Machine Ethics and Robot Ethics*, Routledge, London, 2020. pp. 69–75.
- Brookings Institution. *A Blueprint for the Future of AI. Weapons of the weak: Russia and AI-driven asymmetric warfare*. 2018. URL: <https://www.brookings.edu/research/weapons-of-the-weak-russia-and-ai-driven-asymmetric-warfare/#footnote-4>
- Byrne R.M.J.* Counterfactuals in Explaining Artificial Intelligence (XAI): Evidence from Human Reasoning. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)*. 2019. P.6276–6282.
- Cadwell et al.* for monitoring everyday prosthesis use a systematic review // *Jornal of neuroEngineering and Rehabilitation*. 2020.
- Calderon R., Maxwell R.* The Black-box effect: what level and type of opacity should society accept? *Frontiers in Artificial Intelligence*. 2022, in print.
- Carnegie Mellon University. 2022. Retrieved at: <https://www.cmu.edu/student-success/programs/>
- Carroll A.B.* Corporate social responsibility: evolution of a definitional construct. *Business and Society*. 1999. Vol. 38 No. 3, pp.268–295.
- Castaño-Martínez María-Soledad.* Innovation, Value Creation, and Entrepreneurship by Opportunity. *Advances in Business Strategy and Competitive Advantage*, pp. 43–63.
- Chassignola Maud, Khoroshavin Aleksandr, Klimova Alexandra and Bilyatdinova Anna.* Artificial Intelligence trends in education: a narrative overview, *Procedia Computer Science* 136, 2018. pp.16–24.
- Chen M. et al.* Simple and deep graph convolutional networks // *International Conference on*

Machine Learning. 2020. P. 1725–1735.

*Chen, M. et al.* Simple and deep graph convolutional networks. International Conference on Machine Learning. 2020. P. 1725–1735.

China Global Investment Tracker. AEI. 2018. Retrieved at: <http://www.aei.org/china-global-investment-tracker>

China mulls law revision to advance innovation in science, technology. 2021. Retrieved at: <https://goo.su/eUvHt>.

*Cooper, R.K., Sawaf, A.* Executive, EQ: Emotional Intelligence in Business. London, NY: Texere. 2000.

*De Luca G.* The development of machine intelligence in a computational universe. Technol Soc. 2021. 65:101553. <https://doi.org/10.1016/j.techsoc>.

*Deutsch D.* The Fabric of Reality. 2016. <https://archive.org/details/TheFabricOfReality>.

*Doshi-Velez F., Kortz M., Budish R., Bavitz C. Gershman S., O'Brien D., Shieber S., Waldo J., Weinberger D., Weller A., Wood A.* Accountability of AI Under the Law: The Role of Explanation. Berkman Klein Center Working Group on Explanation and the Law, Berkman Klein Center for Internet & Society Working Paper. 2017. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:34372584>.

*Dowd M.* “Elon Musk’s Billion-Dollar Crusade to Stop the A.I. Apocalypse”, The Hive, 2017. URL: <https://www.vanityfair.com/news/2017/03/elon-musk-billion-dollar-crusade-to-stop-ai-space-x>

*Dreyfus H.L.* What Computers Can't Do: A Critique of Artificial Reason. NY: Harper. 1972. – 259 p.

*Dubrovsky, D.I.* The Problem of Free Will and Modern Neuroscience. Neuroscience and Behavioral Physiology. 49(5), 2019. pp. 629–639.

*Ecoffet A., Lehman J.* “Reinforcement Learning under Moral Uncertainty”, Proceedings of Machine Learning Research, Vol. 139, Proceedings of the 38th International Conference on Machine Learning, 2021. pp. 2926–2936.

*Efimov A.R.* Post-turing Methodology: Breaking the Wall on the Way to Artificial General Intelligence // 13th International Conference on Artificial General Intelligence. (St. Petersburg, Russia, September 16–19, 2020) / Artificial General Intelligence. AGI 2020. Lecture Notes in Computer Science. Vol. 12177. Springer. 2020. pp. 84-94.

*Efimov, A.* Post-Turing Methodology: Breaking the Wall on the Way to Artificial General Intelligence // Artificial General Intelligence | (AGI). 2020. Vol. 12177. pp. 83–94.

*Efimov A.R., Dubrovsky D.I., Matveev P.M.* “Walking Through the Turing Wall”, IFAC Papers OnLine, Vol. 54, No. 13, 2021. pp. 215–220.

*Einstein A., Podolsky B., and Rosen N.* Can quantum-mechanical description of physical reality be considered complete? // Phys. Rev., 1935. Vol. 47. pp. 777–780.

*Elia G., Margherita, A., & Passiante G.* Digital entrepreneurship ecosystem: How digital technologies and collective intelligence are reshaping the entrepreneurial process. Technological Forecasting and Social Change. 2020. 150, Article 119791. <https://doi.org/10.1016/j.techfore.2019.119791>

*Elsa B. Kania.* Battlefield Singularity: Artificial Intelligence, Military Revolution, and China’s Future Military Power // Center for a New American Security. 2021. 74 p.

*Espejo R., Lepskiy V.* "An agenda for ontological cybernetics and social responsibility", Kybernetes, Vol. 50 No. 3, 2020. pp. 694–710.

*Esposito E.* (2017), "Artificial communication? The production of contingency by algorithms",

*Evangelista R., Sirilli G.* Measuring innovation in services // Research Evaluation. 1995. Vol. 5, No 3. pp. 207–215.



- Evans P., Gawer A.* The rise of the platform enterprise: A global survey / The Center of Global Enterprise. 2016 (The Emerging Platform Economy Series №1 ed.).
- Executive Office of the President. Maintaining American Leadership in Artificial Intelligence, E.O. 13859 of Feb 11, 2019, s.l.: Executive Office of the President, Presidential Document. 2019. Retrieved at: <https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence>
- Experts on the Malicious Use of Artificial Intelligence: Challenges for Political Stability and International Psychological Security. Report by ICSPSC. Coordinated by Evgeny Pashentsev. Moscow. 2020.
- Fedasiuk R., Melot J., Murphy B.* Harnessed Lightning. HOW THE CHINESE MILITARY IS ADOPTING ARTIFICIAL INTELLIGENCE // Center for Security and Emerging Technology (CSET) at Georgetown’s Walsh School of Foreign Service. 2021. 84 p.
- Federal Enterprise Architecture Framework (USA). Version 2, 2013, 434 p. [Электронный ресурс]: ISO 15704:2019 Enterprise modelling and architecture — Requirements for enterprise-referencing architectures and methodologies
- Fjelland R.* Why general artificial intelligence will not be realized. Humanities and Social Sciences Communications, Issue 7(10), pp. 1-9.
- Fjelland R.* “Why general artificial intelligence will not be realized”, Humanities and Social Sciences Communications, Vol. 7 No. 1, 2020. pp. 1-9.
- Floridi L.* What the Near Future of Artificial Intelligence Could Be // Philosophy & Technology, 32(1). 2019.
- Floridi L., Cowls J.* A Unified Framework of Five Principles for AI in Society // Harvard Data Science Review. Issue 1(1). June 2019.
- Floridi, L., Cowls, J., Beltrametti, M.* (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations, *Minds & Machines* 28:689–707
- Friedrichsen, M., & Kamalipour, Y. (Eds.)*. Digital Transformation in Journalism and News Media. Stuttgart, Germany: Springer International Publishing AG. 2017.
- Fuchs, C.* Social Media: A Critical Introduction (3-hd ed.). London: SAGE Publications Ltd. 2020.
- Füllsack M.* “The circular conditions of second-order science sporadically illustrated with agent-based experiments at the roots of observation”, Constructivist Foundations, Vol. 10 No. 1, 2014. pp. 46-54, available at: <http://constructivist.info/10/1/046>
- Future of Life Institute. An open letter: research priorities for robust and beneficial artificial intelligence. s.l., Future of Life Institute. 2015.
- G7 ICT and Industry Ministerial Meeting 2017. Making The Next Production Revolution Inclusive, Open and Secure. 2017. 24 p.
- Gabriel Iason.* Toward a Theory of Justice for Artificial Intelligence. Daedalus, Spring 2022, Vol. 151, No. 2, AI & Society (Spring 2022), pp. 218-231.
- Gal D.* Perspectives and Approaches in AI Ethics: East Asia // Dubber, Markus, Pasquale, Frank, and Das, Sunit, (Eds.) Oxford Handbook of Ethics of Artificial Intelligence, Oxford University Press. June 7, 2019. SSRN: <http://dx.doi.org/10.2139/ssrn.3400816>
- Gasser U., Ienca M., Scheibner J., Sleigh J., & Vayena E.* Digital tools against COVID-19: Framing the ethical challenges and how to address them. 2020. arXiv preprint arXiv:2004.10236.
- Gates, B. (2010, May 31)*. <https://www.craigbailey.net/content-is-king-by-bill-gates/>. Retrieved from <https://www.craigbailey.net/>.
- Gault F.* Measuring Innovation Everywhere. Cheltenham; Northampton: Edward Elgar Publishing, 2021.
- Gigerenzer, G.* Gut Feelings. The Intelligence of the Unconscious. London.: Viking. 2007.

- GlobalWebIndex.* (2019). <https://www.gwi.com/reports/traditional-vs-digital-media-consumption>. Retrieved from <https://www.gwi.com/>.
- GlobalWebIndex.* (2021). <https://www.gwi.com/reports/global-media-landscape>. Retrieved from <https://www.gwi.com/>.
- Goksel, N., & Bozkurt, A.* Artificial Intelligence in Education: Current Insights and Future Perspectives. In S. Sisman-Ugur, & G. Kurubacak (Eds.), *Handbook of Research on Learning in the Age of Transhumanism* (pp. 224-236). Hershey, PA: IGI Global. 2019.
- Goriunova, O.* Digital subjects: an introduction. *Subjectivity*, 12, 2019. pp. 1–11.
- Goriunova, O.* The Digital Subject: People as Data as Persons. *Theory, Culture and Society*, vol. 36, No 6, 2019. pp.125-145.
- Grieves, M.* *Digital Twin: Manufacturing Excellence through Virtual Factory Replication; A White Paper*; LLC: Melbourne, FL, USA. 2014.
- Grudin J., Jacque R.* Chatbots, Humbots, and the Quest for Artificial General Intelligence. s.l., s.n., 2019. pp. 1-11.
- Grutters J.P.C., Govers T, Nijboer J., et al.* Problems and promises of health technologies: the role of early health economic modeling. *Int J Health Policy Manag.* 2019. 8:575–582.
- Gubanov D., Korgin N., Novikov D., Raikov A.* *E-Expertise: Modern Collective Intelligence*. Springer. Series: Studies in Computational Intelligence, Vol. 558, 2014, XVIII, 112 p. <http://dx.doi.org/10.1007/978-3-319-06770-4>.
- Heaven W.D.* Why asking an AI to explain itself can make things worse. *MIT Technology Review*. 2020. January 29. <https://www.technologyreview.com/2020/01/29/304857/why-asking-an-ai-to-explainitself-can-make-things-worse/>
- Hirst A., Philippou Y., Blazeby J., et al.* No Surgical Innovation Without Evaluation: Evolution and Further Development of the IDEAL Framework and Recommendations. *Ann Surg.* 2019. 269:211–220.
- Huang, T.-J.* Imitating the Brain with Neurocomputer. A “New” Way Towards Artificial General Intelligence. *International Journal of Automation and Computing*. School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China. 14(5), 2017. P. 520-531.
- Huesig S., Endres H.* Exploring the digital innovation process. *European Journal of Innovation Management*, Vol. 22, no. 2, 2019. pp. 302-314.
- Human Rights Council. A/HRC/48/31: The right to privacy in the digital age - Report of the United Nations High Commissioner for Human Rights // <https://www.ohchr.org/>. 2021. URL: <https://www.ohchr.org/en/documents/thematic-reports/ahrc4831-right-privacy-digital-age-report-united-nations-high> (дата обращения: 18.05.2022).
- Iansiti, M. & Lakhani, K.R.* The Truth About Blockchain, *Harvard Business Review* 95(1). 2017. P. 118–127.
- IndEA Framework (India Enterprise Architecture Framework) [Электронный ресурс]: URL: <http://egovstandards.gov.in/sites/default/files/IndEA%20Framework%201.0.pdf>
- International Institute for Management Development. (2021, April). <https://www.imd.org/contentassets/8c5b42807da941ee95c7be87d54e5db9/20210427-digitalvortex21-report-web-final.pdf>. Retrieved from <https://www.imd.org>.
- Iris Deng and Xinmei Shen.* «China’s «two sessions» 2021: plans for «Digital China» transformation come with increased regulations for personal data» // Published: South China Morning Post, 12:00 pm, 6 Mar, 2021. <https://www.scmp.com/tech/big-tech/article/3124305/chinas-two-sessions-2021-plans-digital-china-transformation-come>
- ISO 56000:2020 Innovation management – Fundamentals and vocabulary. – URL: <https://www.iso.org/ru/standard/69315.html>.

- ISO 56002:2019 Innovation management – Innovation management system – Guidance. – URL: Received 29 03 2022 г., from <https://www.iso.org/standard/68221.html>.
- ISO 56003:2019 Innovation management – Tools and methods for innovation partnership – Guidance. – URL: <https://www.iso.org/ru/standard/68929.html>.
- ISO 56005:2020 Innovation management – Tools and methods for intellectual property management – Guidance. – URL: <https://www.iso.org/ru/standard/72761.html>.
- ISO/IEC DIS 22989. Information technology – Artificial intelligence – Artificial intelligence concepts and terminology
- ISO/TR 56004:2019 Innovation Management Assessment – Guidance. – URL: <https://www.iso.org/standard/69921.html>.
- Jacobson, N. C.; Bentley, K. H. et al.* Ethical dilemmas posed by mobile health and machine learning in psychiatry research, *Bulletin of the World Health Organization*, 98(4). 2020. pp. 270-276.
- Johnson, D. G.* Technology with No Human Responsibility? *Journal of Business Ethics* 127(4): 2015. pp. 707-715.
- Joseph Campbell.* Robot Monk Blends Science and Buddhism at Chinese Temple // Reuters (April 22, 2016), <https://www.reuters.com/article/us-china-religion-robot/robotmonk-blends-science-and-buddhism-at-chinese-temple-idUSKCN0XJ05I>
- Karnouskos S.* Symbiosis with artificial intelligence via the prism of law, robots, and society. *Artif Intell Law*. 2021. <https://doi.org/10.1007/s10506-021-09289-1>
- Kauffman L.H.* “Eigenform”, *Kybernetes*, Vol. 34 Nos 1/2, 2005. pp. 129-150,
- Kauffman L.H.* “Cybernetics, reflexivity and second-order science”, *Constructivist Foundations*, Vol. 11 No. 3, 2016. pp. 489-497, available at: <http://constructivist.info/11/3/489>
- Kaul N.* 3Es for AI: Economics, Explanation, Epistemology. *Front. Artif. Intell.* 5:833238. 2022.
- Kemp, S.* (2021, January 21). <https://datareportal.com/reports/digital-2021-global-overview-report>. Retrieved from <https://datareportal.com>.
- Koestler A.* “Beyond atomism and holism—the concept of the holon”, *Perspectives in Biology and Medicine*, Vol. 13 No. 2, 1970. pp.131-154.
- Korteling J. et al.* Human- versus Artificial Intelligence. *Frontiers in Artificial Intelligence*, March, 4(622364), pp. 1-13.
- Leavy, S., Meaney, G., Wade, K., Greene, D.* Mitigating gender bias in machine learning data sets. *Communications in Computer and Information Science*, 1245 CCIS, 2020. pp. 12-26.
- Leavy, S., Siaper, E., O'Sullivan, B.* Ethical Data Curation for AI: An Approach based on Feminist Epistemology and Critical Theories of Race // AIES 2021. Proceedings of the 2021. AAAI/ACM Conference on AI, Ethics, and Society. pp. 695–703.
- LeCun, Y. et al.* Deep learning. *Nature*, May, Volume 521, pp. 436–444.
- Lefebvre, V.* Second Order Cybernetics in the Soviet Union and the West POWER, AUTONOMY, UTOPIA. *New Approaches toward Complex Systems*, 1986: 123-131.
- Legg S., Hutter M.* Universal Intelligence: A Definition of Machine. *Minds & Machines*, August, Volume 17, pp. 391–444.
- Léna, P.* Robotics in the Classroom: Hopes or Threats? In: von Braun, J., S. Archer, M., Reichberg, G.M., Sánchez Sorondo, M. (eds) *Robotics, AI, and Humanity*. Springer, Cham. 2021.
- Lepskiy V.* Decision support ontologies in self-developing reflexive-active environments // *IFAC PapersOnLine*. 2018. 51(30). pp. 504-509.
- Lepskiy V.* Evolution of cybernetics: philosophical and methodological analysis // *Kybernetes*. 2018. V.47, Iss.2. pp.249–261.
- Lepskiy V.* Reflexivity and Artificial Intelligence in Control (Subjectness-oriented Approach).

IFAC-PapersOnLine. Volume 54, Issue 13, 2021, 2021. pp 221-226.

*Lepskiy V., Zhurenkov, D., Saveliev, A.* Science Diplomacy And Social Innovations // SocioTime, 2(18). 2019. pp. 54-62.

*Lin, Y.-T., Hung, T.-W., Huang, L.T.-L.* Engineering Equity: How AI Can Help Reduce the Harm of Implicit Bias // Philosophy and Technology. 2021. Vol. 34, pp. 65–90.

*Lindgren, S.* Digital Media & Society. London: SAGE Publications Ltd. 2017.

*Lissack M.R.* “Shed the name to find second-order success: renaming second-order cybernetics to rescue its essence”, Constructivist Foundations, Vol. 11 No. 3, 2016. pp. 470-472, available at: <http://constructivist.info/11/3/470>

*Liu R., Balsubramani A. and Zou J.* Learning transport cost from subset correspondence // arXiv preprint arXiv:1909.13203. 2019.

*Mackintosh K.* The principles of humanitarian action in international humanitarian law. London: overseas development institute". 2020. URL: <http://www.odi.org.uk/sites/odi.org.uk/files/odiassets/publications-opinion-files/305.pdf>

*Madni, A.M., Madni, C.C., Lucero, S.D.* Leveraging Digital Twin Technology in Model-Based Systems Engineering. Systems. 7 (1): 7. 2019.

*Madry A., Makelov A., Schmidt L., Tsipras D. and Vladu, A.* Towards deep learning models resistant to adversarial attacks // arXiv preprint arXiv:1706.06083, 2017.

*Malmelin, N., & Villi, M.* Media work in change: Understanding the role of media professionals in times of digital transformation and convergence. Sociology Compass (11). 2017.

*Marcus G., Davis E.* Rebooting AI: Building Artificial Intelligence We Can Trust, Pantheon, 2019, 288 p.

*Martens, B., Aguiar, L., Gomez-Herrera, E., & Mueller-Langer, F.* The digital transformation of news media and the rise of disinformation and fake news. Seville: European Commission, Joint Research Centre. 2018.

*Martin, K.* Ethical Implications and Accountability of Algorithms. Journal of Business Ethics 160: 835–850. 2019.

*McConnell T.* “Moral Dilemmas”, The Stanford Encyclopedia of Philosophy, 2018. URL: <https://plato.stanford.edu/archives/fall2018/entries/moral-dilemmas/>

MERICS Report. Chinese FDI In Europe 2021 Update. Investments remain on downward trajectory – Focus on venture capital. 2021. 19 p. Retrieved at: <https://rhg.com/research/chinese-fdi-in-europe-2021-update/>

Ministry of Foreign Affairs of the People's Republic of China. Position Paper of the People's Republic of China on Regulating Military Applications of Artificial Intelligence (AI). 2021. [https://www.fmprc.gov.cn/mfa\\_eng/wjdt\\_665385/wjzcs/202112/t20211214\\_10469512.html](https://www.fmprc.gov.cn/mfa_eng/wjdt_665385/wjzcs/202112/t20211214_10469512.html)

*Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L.* The ethics of algorithms: Mapping the debate, *Big Data & Society*, 3(2), 2016. pp. 1–21.

Montreal University. Montreal Declaration for a Responsible Development of Artificial Intelligence, s.l.: Announced at the conclusion of the Forum on the Socially Responsible Development of AI. 2017.

*Mueller S. T., Veinott E. S., Hoffman R.R., Klein G., Alam L., Mamun T., and Clancey W. J.* Principles of Explanation in Human-AI Systems. Association for the Advancement of Artificial Intelligence. 2020. <https://arxiv.org/ftp/arxiv/papers/2102/2102.04972.pdf>

*Mueller, S.T., Hoffman, R.R., Clancey, W., Emrey, A., and Klein, G.* Explanation in Human-AI systems: A Literature Meta-Review Synopsis of Key Ideas and Publications and Bibliography for Explainable AI. DARPA XAI Literature Review. 2019. <https://arxiv.org/ftp/arxiv/papers/1902/1902.01876.pdf>.

*Müller, Vincent C.* “Ethics of Artificial Intelligence and Robotics”, The Stanford Encyclopedia of

- Philosophy, 2021. URL: <https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/>
- Mumford S., Anjum R. L.* Causation. A very short introduction. – Oxford: University press. UK. 2013.
- Niels Bohr Collected Works.* Vol. 3: The Correspondence Principle (1918–1923), J. R. Nielsen (ed.). Amsterdam: North-Holland Publishing. 1976.
- Office of the United States trade representative executive Office of the President. Findings of the investigation into China’s acts, policies, and practices related to technology transfer, intellectual property, and innovation under section 301 of the trade Act of 1974. 2018. 215 p. Retrieved at: <https://ustr.gov/sites/default/files/Section%20301%20FINAL.PDF>
- Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation. The Measurement of Scientific, Technological and Innovation Activities. 4th ed. – Paris: Eurostat; Luxembourg: OECD Publishing, 2018. doi: 10.1787/9789264304604-en.
- Palacios E.R., Razi A. Parr T., Kirchhoff M. and Friston K.* “On Markov blankets and hierarchical self-organisation”, *Journal of Theoretical Biology*, Vol. 486, 2020. pp. 110089.
- Pan Chung.* New Age philosophy and social science discourse innovation (in Chinese). 2021. Retrieved at: <https://goo.su/2YPz0>.
- Pardo del Val, M. & Martinez Fuentes, C.* Resistance to change: A literature review and empirical study. *Management Decision*. 2003. 41: 148-155.
- Paredes-Frigolett H., Pyka A.* The global stakeholder capitalism model of digital platforms and its implications for strategy and innovation from a Schumpeterian perspective. *Journal of Evolutionary Economics*, Vol. 32, no. 2, pp. 463–500.
- Pashiri, R. T.; Rostami, Y.; Mahrami, M.* Spam detection through feature selection using artificial neural network and sine–cosine algorithm, *Mathematical Sciences*, 2020. 14: 193–199.
- Peeters M. et al.* Hybrid collective intelligence in a human–AI society. *AI & SOCIETY*, Volume 36, p. 217–238.
- Pei J. et al.* Towards artificial general intelligence with hybrid Tianjic chip architecture. *Nature*, Volume 572, pp. 106–111.
- Perko I.* Hybrid reality development - can social responsibility concepts provide guidance? *Kybernetes*, Vol. 50 No. 3, 2020. pp. 676–693.
- Pfeifer, R., Iida, F.* Embodied Artificial Intelligence: Trends and Challenges. In: Iida, F., Pfeifer, R., Steels, L., Kuniyoshi, Y. (eds) *Embodied Artificial Intelligence*. Lecture Notes in Computer Science, vol 3139. Springer, Berlin, Heidelberg. 2004.
- Pirani, M., Bonci, A. and Longhi, S.* “Towards a formal model of computation for RMAS”, *Procedia Computer Science*, Vol. 200, 2022. pp. 865-877.
- Polya G.* Mathematics and plausible reasoning. – New Jersey, Princeton: Princeton University Press. 1954. Vol. 1&2.
- Raikov A.* Automating Cognitive Modelling Considering Non-Formalisable Semantics. In: Nagar A.K., Jat D.S., Marín-Raventós G., Mishra D.K. (eds) *Intelligent Sustainable Systems*. Lecture Notes in Networks and Systems, .2022. Vol 334. Springer, Singapore. [https://doi.org/10.1007/978-981-16-6369-7\\_2](https://doi.org/10.1007/978-981-16-6369-7_2)
- Raikov A.* Cognitive Semantics of Artificial Intelligence: A New Perspective. Springer Singapore, Topics: Computational Intelligence XVII, 2021. 128 p. <https://doi.org/10.1007/978-981-33-6750-0>
- Raikov A.* Convergent Ontologization of Collective Scientific Discoveries, 2021 14th International Conference Management of large-scale system development (MLSD), 2021, pp. 1-5.
- Raikov A.* Convergent social responsibility as the key to corporate strategic success. *Kybernetes*. 2021. 50(3):785-793.

- Raikov A.* Post-Non-Classical Artificial Intelligence and its Pioneer Practical Application. IFAC-PapersOnLine. 2019. 52(25):343-348.
- Raikov A.N. and Pirani M.* Contradiction of modern and social-humanitarian artificial intelligence, *Kybernetes*, Vol. 51 No. 13, 2022. pp. 186-198.
- Raikov A.N., Ermakov A.N., and Merkulov A.A.* Assessments of the Economic Sectors Needs in Digital Technologies, *Lobachevskii Journal of Mathematics*, 2019, Vol. 40, No. 11, pp. 1837–1847. Pleiades Publishing, Ltd. <https://doi.org/10.1134/S1995080219110246>.
- Rauber A., Trasarti R., and Gianotti F.* Transparency in Algorithmic Decision Making. *ERCIM News*. 2019. Vol. 116. P. 10–11.
- Rawls John.* A Theory of Justice, rev. ed. Cambridge, Mass.: Harvard University Press, 1999.
- Røe Y, Wojniusz S and Bjerke AH.* The Digital Transformation of Higher Education Teaching: Four Pedagogical Prescriptions to Move Active Learning Pedagogy Forward. *Front. Educ.* 6:784701. 2022.
- Rovelli C.* “Relational quantum mechanics”, *International Journal of Theoretical Physics*, Vol. 35 No. 8, 1996. pp. 1637-1678.
- Rubtsov V.* Learning in children: organization and development of cooperative actions. M. J. Hall, Trans. Nova Science. 1991.
- Russell S., Norvig P.* Artificial intelligence: a modern approach. New Jersey: Alan Apt.
- Russell, Stuart J. (Stuart Jonathan).* Artificial intelligence: a modern approach. Upper Saddle River, N.J.: Prentice Hall. 2010.
- Sanz, E., Simon, J., & Prato, G. (Eds.).* Digital Media Worlds: The New Economy of Media. Hampshire: Palgrave Macmillan (2014).
- Scholte M., Rovers M.M., Grutters J.P.C.* The Use of Decision Analytic Modeling in the Evaluation of Surgical Innovations: A Scoping Review. *Value in Health. Systematic literature review.* 2021. Vol. 24, Issue 6, pp. 884-900.
- Secundo G., Shams S.M.R., Nucci F.* Digital technologies and collective intelligence for healthcare ecosystem: Optimizing Internet of Things adoption for pandemic management. *Journal of Business Research*, 2021. Vol. 131, pp. 563-572.
- Setzu, M., Guidotti, R., Monreale, A., Turini, F., Pedreschi, D., & Giannotti, F.* Glocalx – from local to global explanations of black box ai models. *Artificial Intelligence* 294, 103457. 2021.
- Shanks JD.* Toward A Contemplative Technopedagogy Framework: A Discourse Analysis. *Front. Educ.* 5:553212. 2020.
- Shevlin H. et al.* The limits of machine intelligence. *EMBO reports*, 20(e49177), pp. 1-5.
- Shimelis T., Eshetie B.* Innovation value chain: a systematic and narrative review. *International Journal of Quality and Innovation*, Vol. 6, no. 1, pp. 91-114.
- Shrayberg, Y., & Volkova, K.* Features of Copyright Transformation in the Information Environment in the Age of Digitalization. *Scientific and technical information processing*, 48(1), 2021. pp. 30-37.
- Simondon, G.* On the Mode of Existence of Technical Objects. Translated by C. Malaspina and J. Rogrove. Minneapolis: Univocal Publishing. 2017.
- Slavin, B.* Digital technologies of intellectual collective activity. In *System analysis in economics-2018. Proceedings of the V International research and practice conference-biennale* (pp. 316-318). Moscow: Prometey.
- Slavin, B., & Slavin, A.* Features of The Media Environment In The Conditions Of Digitalization Of Social Communications. *IFAC PapersOnLine* 54-13, 2021 (pp. 393-396). Moscow: IFAC.
- Sloman S., Fernbach P.* The Knowledge Illusion: Why We Never Think Alone. New York: Riverhead Books.

- Sloman S., Patterson R., Barbey A.* Cognitive Neuroscience Meets the Community of Knowledge. *Frontiers in Systems Neuroscience*, October, 15(675127), pp. 1-13.
- Smakman Matthijs, Vogt Paul, Konijn Elly A.* Moral considerations on social robots in education: A multi-stakeholder perspective, *Computers & Education*, Volume 174, 2021.
- Songchun Zhu.* Intelligence needs to be driven by «mind», achieving a dynamic balance between «mind» and «principle» (in Chinese). 2022. Retrieved at: <https://mp.weixin.qq.com/s/FMraladW0255C-yeU5alfA>
- Songchun Zhu.* Intelligence needs to be driven by «mind», achieving a dynamic balance between «mind» and «principle» // Beijing Institute for General Artificial Intelligence. 2022. Original Mandarin: <https://mp.weixin.qq.com/s/FMraladW0255C-yeU5alfA>
- Stanford-New America Digichina Project. AI POLICY AND CHINA // Realities of State-Led Development. Special Report No. 1. 2019. pp. 41
- State Council of the People's Republic of China. The 14th Five-Year Plan (2021-2025) for National Economic and Social Development and the Long-Range Objectives Through the Year 2035. Retrieved April 22, 2021, from [http://www.gov.cn/xinwen/2021-03/05/content\\_5590537.htm](http://www.gov.cn/xinwen/2021-03/05/content_5590537.htm)
- State Council of the People's Republic of China. Next-Generation Artificial Intelligence Development Plan. 2017. URL <https://www.mfa.gov.cn/ce/cefi/eng/kxjs/P020171025789108009001.pdf>
- Stuart A. Umpleby* Expanding Science and Advancing Reflexive Government: Two Current Projects in Cybernetics *Journal on Policy and Complex Systems* • 2018. Volume 4, Number 2, pp. 171-184.
- Sundvall, S.* Artificial Intelligence. In: *Critical Terms in Futures Studies*. Paul H. (Ed.), 29–34. Palgrave Macmillan, Cham. 2019.
- Telley, Ch.* The Influence Machine: Automated Information Operations as a Strategic Defeat Mechanism; The Institute of Land Warfare, Association of the United States Army, Arlington, VA, USA, 2018
- The BIG Bell Test Collaboration, Challenging local realism with human choices, *Nature*, 2018. Vol. 557, P. 212220. doi: 10.1038/s41586-018-0085-3
- The OECD Model Survey on ICT Usage by Businesses: 2nd revision // Working Party on Measurement and Analysis of the Digital Economy. 2015. – URL: <https://www.oecd.org/sti/ieconomy/ICT-Model-Survey-Usage-Businesses.pdf>.
- The TOGAF Standard, Version 9.2 Overview. [Электронный ресурс]: URL: <https://www.opengroup.org/togaf>
- Tidd J.* January A review and critical assessment of the ISO56002 innovation management systems standard: evidence and limitations // *International Journal of Innovation Management*. 2021. Vol. 25, no. 1. pp. 1–17.
- Turchin V. and Joslyn C.* The Cybernetic Manifesto. *Kybernetes*. 1990. 01.19, Nos.2 and 3. pp. 63-64, 52-55.
- Turing, Alan* "Computing Machinery and Intelligence" (PDF), *Mind*, LIX (October 1950), (236): 433–460.
- Umbrello, S., Yampolskiy, R.V.* Designing AI for Explainability and Verifiability: A Value Sensitive Design Approach to Avoid Artificial Stupidity in Autonomous Vehicles. *Int J of Soc Robotics*. 2021.
- Umpleby S., Medvedeva T., Lepskiy V.* Recent developments in cybernetics, from cognition to social systems // *Cybernetics and Systems*. 2019. Vol.50, Iss.4, pp. 367-382.
- Veličković, P. et al.* (2019). Neural Execution of Graph Algorithms. 2019. arXiv:1910.10593

- Von Foerster H.* “Objects: tokens for (eigen-) behaviors”, in *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer-Verlag, New York, Berlin, Heidelberg, 2003. pp. 261-271.
- Wang J. et al.* Learning node representations from noisy graph structures. *Proc. IEEE Int. Conf. Data Mining, ICDM*. Vol. 2020-Novem, No 1. 2020. pp. 1310–1315
- Wang P, Li X, Hammer P.* Self in NARS, an AGI System. *Front. Robot. AI* 5:20. 2018.
- Wang P.* Natural language processing by reasoning and learning. In: *International Conference on Artificial General Intelligence*, 160–169. Springer. 2013.
- Wark, S.* The subject of circulation: on the digital subject’s technical individuations. *Subjectivity*. 12, 2019. pp. 65–81.
- Wiesböck F., Hess T.* Digital innovations. *Electronic Markets*, Vol. 30, no. 1, pp. 75-86.
- Woolley A. et al.* Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*, October, Issue 330, pp. 686-688.
- World Economic Forum (2015). <http://reports.weforum.org/digital-transformation/an-introduction-to-the-digital-transformation-initiative/>. Retrieved from <http://reports.weforum.org/>.
- Yanes-Fernandez, Inti . David I. Dubrovsky and Merab Mamardashvili* : Adam’s Second Fall and the Advent of the Cyber-Leviathan // *CREATIVE-COMMONS-BY Forum Philosophicum* 24 (2019) . 2, 301–41.
- Yang, C. C. Y., Chen, I. Y. L., & Ogata, H.* Toward Precision Education: Educational Data Mining and Learning Analytics for Identifying Students’ Learning Patterns with Ebook Systems. *Educational Technology & Society*, 24 (1), 2021. pp. 152–163.
- Zachman J.* A framework for information systems architecture. *IBM Systems Journal*, 1987, 26(3), pp. 276-292.
- Zeitschrift für Soziologie*, Vol. 46 No. 2, pp. 249–265.



## АННОТАЦИЯ И КЛЮЧЕВЫЕ СЛОВА

---

**Аннотация.** Прогресс в развитии цифровых технологий, включая искусственный интеллект, несомненен. Однако, методологический базис этого развития до сих пор носит преимущественно техногенный характер. Как результат, например, системы искусственного интеллекта так и не вышли за рамки «слабого» и «узкого», они остаются однозадачными, не дают объяснения своим выводам, слабо учитывают специфику поведения субъектных сред.

Вместе с тем, уже явно происходит свертывание идеи техногенной цивилизации, которая характеризуется проявлением в ее базовой ценности научно-технического прогресса и ценности науки как основы управления социальными процессами. Как следствие, развитие цифровых технологий, включая искусственный интеллект, происходит без учета социально-гуманитарных ценностей и этических аспектов регулирования жизнедеятельности людей.

Настоящая монография для умаления подобных ограничений в развитии цифровых технологий, включая искусственный интеллект, обосновывает необходимость использования субъектно-ориентированного подхода, который формируется в рамках постнеклассической кибернетики саморазвивающихся полисубъектных рефлексивно-активных сред, основные положения которой опираются на представления об обществе как социально-гуманитарной системе.

Монография предназначена для специалистов из гуманитарных, естественнонаучных и технических областей знания, ориентированных на проблемы управления и развития, для студентов и аспирантов, а также для широкой аудитории практиков управления, в том числе связанных с разработкой и использованием цифровых технологий, включая искусственный интеллект.

**Ключевые слова:** искусственный интеллект, постнеклассическая кибернетика, субъектно-ориентированный подход, рефлексивно активная среда, техногенные ограничения, цифровые технологии, этика

## TITLE, ABSTRACT, KEYWORDS

---

### **Socio-humanitarian aspects of digital transformations and artificial intelligence**

*Abstract.* Progress in the development of digital technologies, including artificial intelligence, is undeniable. However, the methodological bases of this development are still mainly technogenic. As a result, for example, artificial intelligence systems have not gone beyond the "weak" and the "narrow", they remain single-tasking, do not explain their conclusions, and poorly take into account the specifics of the behaviour of subjects and environment.

At the same time, the idea of technogenic civilization is already clearly being curtailed, which is characterized by the manifestation in its basic value of scientific and technological progress and the value of science as the basis for managing social processes. As a result, the development of digital technologies, including artificial intelligence, takes place almost without taking into account social and humanitarian values and ethical aspects of regulating people's lives.

This monograph, to belittle such limitations in the development of digital technologies, including artificial intelligence, justifies the need of using the subject-oriented approach, which has been formed within the framework of post-non-classical cybernetics of self-developing polysubject reflexively active environments, the main provisions of which are based on the idea of society as a socio-humanitarian system.

The monograph is intended for specialists from the humanities, natural sciences and technical fields of knowledge focused on management and development problems, for undergraduates and postgraduates, as well as for a wide audience of management practitioners, including those related to the development and use of digital technologies, including artificial intelligence

**Keywords:** artificial intelligence, digital technologies, ethic, post-non-classical cybernetics, reflexively active environment, subject-oriented approach, technogenic restrictions

## Сведения об авторах и их индивидуальном вкладе в монографию

---

- Агеев Александр Иванович* – директор Института экономических стратегий при ООИ РАН, генеральный директор Международного научно-исследовательского института (МНИИПУ), д.э.н. (1.3., 3.5., 4.4.).
- Дубровский Давид Израилевич* – главный научный сотрудник Института философии РАН, д.ф.н. (1.2., 2.3., 3.7., 4.3.).
- Ефимов Альберт Рувимович* – вице-президент, директор управления исследований и инноваций, ПАО «Сбербанк», заведующий кафедрой инженерной кибернетики НИТУ «МИСИС», к.ф.н. (1.2., 3.3., 3.7.).
- Журенков Денис Александрович* – руководитель Центра диверсификации организаций ОПК ФГУП «Всероссийский научно-исследовательский институт «Центр» (Минпромторг России), член Экспертного совета по диверсификации оборонно-промышленного комплекса и импортозамещению при Правительственной комиссии по импортозамещению (2.4., 2.6., 3.6.).
- Ильин Николай Иванович* – профессор, заслуженный деятель науки РФ, чл.-корр. Академии криптографии России, заместитель начальника Управления информационных систем спецсвязи ФСО России, чл.-корр. Академии криптографии России, д.т.н. (4.5.).
- Лепский Владимир Евгеньевич* – главный научный сотрудник Института философии РАН, руководитель Центра междисциплинарных исследований рефлексивных процессов и управления Института философии РАН, д.психол.н. (Предисловие, 1.1., 1.2., 1.6., 2.1., 2.2., 2.7., 3.1., 3.2., 3.8., 4.1., 4.3., 4.6., 5.1., 5.7., Заключение).
- Малахова Елена Владимировна* – докторант Института философии РАН, к.ф.н. (2.5.).
- Матвеев Филипп Михайлович* – студент МГУ им. М.В. Ломоносова (3.7.).
- Пойкин Артем Евгеньевич* – заместитель начальника отдела, Всероссийский научно-исследовательский институт «Центр» (2.4., 2.6., 3.6.).
- Райков Александр Николаевич* – ведущий научный сотрудник Института проблем управления им. В.А. Трапезникова РАН, д.т.н. (Предисловие, 1.1., 1.4., 1.5., 1.6., 2.1., 2.7., 3.1., 3.4., 3.8., 4.1., 4.2., 4.3., 4.6., 5.1., 5.5., 5.7., Заключение).
- Савельев Антон Максимович* – начальник отдела, Всероссийский научно-исследовательский институт «Центр», соискатель Института философии РАН (2.4.).
- Славин Александр Борисович* – медиа эксперт, продюсер цифрового контента, выпускник ФКМ РГСУ (5.6.).
- Славин Борис Борисович* – профессор департамента бизнес-информатики Финансового университета при Правительстве РФ, д.э.н., к.ф.-м.н. (1.2., 5.2., 5.3., 5.4., 5.6.).

Научное издание  
Коллективная монография

**Социогуманитарные аспекты  
цифровых трансформаций  
и искусственного интеллекта**

Рецензенты:

Академик РАН, доктор философских наук  
*Владислав Александрович Лекторский*

Член-корреспондент РАН, доктор экономических наук  
*Владимир Викторович Иванов*

Издательство «Когито-Центр»  
129366, Москва, ул. Ярославская, 13  
Тел.: (495) 682-61-02

E-mail: [post@cogito-shop.com](mailto:post@cogito-shop.com), [cogito@bk.ru](mailto:cogito@bk.ru)  
[www.cogito-centre.com](http://www.cogito-centre.com)

Сдано в набор 10. 11. 22. Подписано в печать 15. 11. 22

Формат 60 × 90/16. Усл. печ. л. 17,5. Тираж 350 экз.

Отпечатано в типографии ООО «Белый ветер»