

Данная электронная версия материала скачана Вами с сайта: <http://motivtrud.ru> и предназначена только для индивидуального ознакомления, после чего файл должен быть удален. При заимствовании материалов ссылка на авторов, первоисточник и сайт обязательны (при размещении ссылок в сети интернет они должны быть оформлены в виде гиперссылок).

Министерство науки и высшего образования Российской  
Федерации

Балтийский государственный технический университет  
«ВОЕНМЕХ» им. Д. Ф. Устинова

---

П. М. ЛУКИЧЁВ

УПРАВЛЕНИЕ ИСКУССТВЕННЫМ  
ИНТЕЛЛЕКТОМ: РИСКИ И  
ГОСУДАРСТВЕННОЕ РЕГУЛИРОВАНИЕ

Монография

САНКТ-ПЕТЕРБУРГ

2025

Рецензенты:

Доктор экономических наук, профессор кафедры Р1

менеджмента организации Балтийского государственного технического университета «ВОЕНМЕХ» им. Д. Ф. Устинова А. И. Стешин

Доктор экономических наук, профессор кафедры экономики и управления предприятиями и производственными комплексами ФГБОУ ВО «Санкт-Петербургский государственный экономический университет» В. А. Балукова

Монография посвящена исследованию проблем управления рисками искусственного интеллекта с помощью мер государственного регулирования. Чем совершеннее становятся большие языковые модели, тем более рискованным становится их применение. Автор анализирует риски использования технологий искусственного интеллекта как в краткосрочном периоде, так и в долгосрочном периоде. Большие языковые модели всё шире используются при принятии экономических решений. В монографии уделяется значительное место исследованию проблем соперничества / сотрудничества между работниками-людьми и искусственным интеллектом, эффективности применения моделей киборга или кентавра. Автор изучает воздействие моделей искусственного интеллекта на современный рынок труда, уделяя особое внимание необходимости изменения мягких навыков и твёрдых навыков работников. Исследование проблем применения мер государственного регулирования применения технологий искусственного интеллекта проводится автором как на национальном уровне, так и на глобальном уровне.

Предназначена для широкого круга читателей, чьи профессиональные или творческие интересы затрагивают вопросы управления рисками искусственного интеллекта,

Автор – Лукичев Павел Михайлович, доктор экономических наук, профессор, Балтийский государственный технический университет «ВОЕНМЕХ» им. Д. Ф. Устинова

**Лукичев, П. М.** Управление искусственным интеллектом: риски и государственное регулирование / П. М. Лукичев. – Санкт-Петербург : Балтийский государственный технический университет "ВОЕНМЕХ" им. Д.Ф. Устинова, Санкт-Петербургский политехнический университет Петра Великого, 2025. – 268 с. – ISBN 978-5-7422-9004-9. – EDN GXJZXW.

Оглавление.....	3
Введение.....	4
Глава 1. Новые риски применения искусственного интеллекта.....	9
1.1. Риски применения искусственного интеллекта в краткосрочном периоде.....	9
1.2. Долгосрочные риски применения искусственного интеллекта в экономике.....	40
Глава 2. Принятие решений в экономике: работники-люди vs искусственный интеллект.....	67
2.1. Принятие решений в современной экономике: искусственный интеллект vs поведенческая экономика.....	67
2.2. Принятие экономических решений: человек с «эффектом Поланьи» или агенты искусственного интеллекта?.....	95
Глава 3. Влияние искусственного интеллекта на рынок труда.....	124
3.1. Влияние искусственного интеллекта на рынок труда в парадигме неравенства.....	124
3.2. Человеческий капитал в эпоху искусственного интеллекта...	152
Глава 4. Государственное регулирование (дерегулирование) применения искусственного интеллекта.....	179
4.1. Государственное регулирование применения моделей искусственного интеллекта: необходимость и пределы макроэкономической политики.....	179
4.2. Государственное регулирование применения искусственного интеллекта.....	207
Заключение.....	237
Список литературы.....	241

## ВВЕДЕНИЕ

Сегодня развитие моделей искусственного интеллекта представляет собой ведущее влияние технологий на экономику. В данной монографии этот процесс исследуется комплексно по четырём направлениям:

- риски применения технологий искусственного интеллекта,
- влияние технологий искусственного интеллекта на принятие экономических решений,
- воздействие моделей искусственного интеллекта на рынок труда,
- проблемы и возможности государственного регулирования применения технологий искусственного интеллекта.

Исследование каждого из этих направлений составляет главу монографии.

В **первой главе** анализируются риски применения технологий искусственного интеллекта. По мере того, как совершенствуются модели искусственного интеллекта (ИИ), включая появление Агентов искусственного интеллекта и модели ИИ с размышлением, возникают риски их использования. Технологии искусственного интеллекта переживают сейчас, если обратиться к концепции жизненного цикла товара/услуги, переход от стадии внедрения к стадии роста. В исследовании выделяются риски краткосрочного периода и риски долгосрочного периода в применении технологий

искусственного интеллекта. Особое внимание уделяется разработке и предложению мер по предупреждению или уменьшению негативных последствий использования больших языковых моделей. Выявленные риски применения технологий искусственного интеллекта на микроуровне в краткосрочном периоде вполне могут быть скомпенсированы существующими антикризисными мерами с учётом специфики действия ИИ как технологии общего применения.

Обосновывается неизбежность появления «рисков развития моделей ИИ». Выявленные риски применения технологий искусственного интеллекта на макроуровне требуют нового (не существующего на данный момент) уровня государственного регулирования. В главе делается вывод о необходимости использования сценарного подхода к анализу рисков применения искусственного интеллекта в долгосрочном периоде и вводится критерий анализа.

Во **второй главе** открывается новейшее направление исследований: кто будет принимать решения в экономике по мере расширения использования моделей ИИ? Каковы возможности совместной работы над принятием итоговых решений человеком и искусственным интеллектом? Анализ, проведенный в главе, показывает недостоверность традиционного подхода, что наиболее плодотворные модели искусственного интеллекта будут дополнять и компенсировать врожденные биологические и когнитивные ограничения людей. Современная экономика переживает период «гибридного отступления», когда всё большая роль в принятии решений передаётся алгоритмам искусственного интеллекта.

Противоречивые последствия этого обуславливают необходимость дальнейшего междисциплинарного анализа принятия решений в современной экономике.

Конкурентным преимуществом перед моделями искусственного интеллекта в принятии решений, которым обладают сейчас работники-люди, является «неявное знание» М. Поланьи. В главе анализируются пессимистический и оптимистический сценарии сотрудничества в экономике работников-людей и моделей искусственного интеллекта. Автор делает вывод о сохранении первенства человека в принятии экономических решений только в случае увеличения инвестиций в человеческий капитал.

В третьей главе исследуется воздействие технологий искусственного интеллекта на рынок труда 2020-х годов. Это влияние проявляется по-разному для работников низкой квалификации, средней квалификации и высокой квалификации. В разделе выявлены шесть особенностей воздействия больших языковых моделей на современный рынок труда. В краткосрочном периоде благодаря технологиям ИИ возможно улучшение трудового потенциала работников низкой квалификации и определённое возрождение среднего класса. В долгосрочном же периоде применение моделей искусственного интеллекта позволяет работникам высокой квалификации повысить производительность и качество своего труда больше, чем мало квалифицированным сотрудникам, что вызывает рост неравенства.

Расширение применения технологий искусственного интеллекта меняет роль и значение человеческого капитала в экономике. Это проявляется через изменение требований к качеству современных рабочих мест, к трансформации ценности «твёрдых» навыков и «мягких» навыков, которыми сейчас обладает работник на рынке труда, к созданию условий для нового разделения труда как на уровне предприятий, так и на уровне народного хозяйства.

В четвёртой главе анализируются теоретические положения и практический опыт, существующие в сфере регулирования искусственного интеллекта как на национальном уровне, так и на глобальном уровне, и раскрывается противоречивость целей регулирования у общественности, и у фирм-лидеров в сфере искусственного интеллекта. Обосновывается опасность создания Суперлиги игроков в сфере ИИ. Выявляется «дилемма приоритетов» государственного регулирования применения технологий искусственного интеллекта. В главе проводится сравнительный анализ комплекса мер государственного регулирования сферы ИИ в отдельных странах, прежде всего в США, в КНР и в ЕС. Китай проделал самую значительную эволюцию мер государственного регулирования технологий искусственного интеллекта. В межгосударственном регулировании раскрыты особенности «конформистского» и «альтернативного» подходов.

Каждая глава состоит из двух разделов. По форме внутри каждого раздела есть Введение, проводится анализ проблемных вопросов, делаются выводы и намечаются направления будущих исследований. Учитывая, что совершенствование моделей

искусственного интеллекта происходит очень интенсивно, их возможности и, соответственно, риски применения возрастают многократно, то в монографии сделан акцент на анализе современного этапа эволюции искусственного интеллекта, начавшегося с ноября 2022 года.

В Заключении делаются выводы по изложенному материалу. Автор надеется, что предпринятое им исследование проблем управления технологиями искусственного интеллекта будет полезно для всех, кто интересуется данными проблемами. Ваши пожелания, вопросы и комментарии присылайте, пожалуйста, по адресу: [loukitchev20@mail.ru](mailto:loukitchev20@mail.ru)

# ГЛАВА 1. НОВЕ РИСКИ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

## 1.1. РИСКИ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В КРАТКОСРОЧНОМ ПЕРИОДЕ

Раздел посвящён анализу рисков применения технологий искусственного интеллекта. Значение технологий искусственного интеллекта сейчас возрастает, поскольку они всё сильнее становятся технологией широкого применения XXI века, особенно в связи с появлением Агентов искусственного интеллекта и «моделей с размышлением». Однако их использование порождает большие риски для всех участников рынка. В исследовании выделены два временных периода: краткосрочный период и долгосрочный период, уделяя основное внимание первому. В краткосрочном периоде выявляются следующие риски использования искусственного интеллекта: деформация существующего рынка труда и усиление неравенства между странами, регионами, отраслями; риски концентрации рынков искусственного интеллекта, риски масштабирования ежедневного вреда, который может быть причинён в Интернете, путём распространения фейковых новостей, выманивания у людей денег, искажения выборных кампаний; риски безопасности фирм, риски переоценённости больших языковых моделей. Особое внимание уделяется разработке и предложению мер по предупреждению или уменьшению негативных последствий использования больших языковых моделей.

Экономический подход к анализу применения искусственного интеллекта (ИИ) предполагает сопоставление Издержек и Выгод. Только тот экономический процесс, в котором выгоды превышают издержки, может считаться эффективным. Такое сравнение

невозможно без анализа рисков внедрения технологий искусственного интеллекта. С этих общих позиций использование искусственного интеллекта, как и любой инновационной технологии, не имеет однозначного сопоставления «Издержки – Выгоды» ни в краткосрочном периоде, ни в долгосрочном периоде, «вписываясь» в общий «парадокс производительности». Сегодняшние большие языковые модели превращаются во всё большей мере в технологии общего назначения. Появление «Агентов искусственного интеллекта» и моделей ИИ с размышлением ставит вопросы о рисках применения технологий искусственного интеллекта в центр внимания экономистов, программистов и политиков. Выход в январе 2025 года китайской модели ИИ DeepSeek, которая полностью бесплатна и имеет открытый исходный код, многократно увеличил доступность применения искусственного интеллекта и, одновременно, риски его недобросовестного использования.

Выявление рисков применения технологий искусственного интеллекта в хозяйственной практике очень важно, поскольку оно помогает улучшать качество экономических решений. Здесь можно показать следующую причинно-следственную цепочку. Технологии ИИ снижают стоимость прогнозов и поэтому человеческое суждение о скрытых возможностях заменяет лучшие прогнозы<sup>1</sup>, а лицо, принимающее решения, обнаруживая некоторые скрытые издержки, возвращает своё решение к безопасному действию. Технологии

---

<sup>1</sup> Agrawal A., Gans J., Goldfarb A. Human judgement and AI pricing. AEA Papers and Proceedings. 2018. Vol.108, p. 58–63.

искусственного интеллекта, как и любая мощная технология, несут в себе много рисков и требуют ответственного применения. С экономической точки зрения целесообразно разделить существующие риски использования технологий искусственного интеллекта на:

- риски на микроуровне
- риски на макроуровне
- риски на глобальном уровне.

В экономике традиционно под микроуровнем подразумевается уровень фирмы, домашнего хозяйства, а под макроуровнем – уровень всей национальной экономики в целом. Также целесообразно выделить в экономическом анализе рисков применения технологий искусственного интеллекта краткосрочный период и долгосрочный период. В данном разделе акцент будет сделан на рисках микроуровня в краткосрочном периоде. Гипотеза исследования - выявленные риски применения технологий искусственного интеллекта на микроуровне в краткосрочном периоде вполне могут быть скомпенсированы существующими антикризисными мерами с учётом специфики действия больших языковых моделей. Цель данного раздела: выявить основные риски применения искусственного интеллекта в современной экономике и оценить возможности уменьшения с учетом их использования как технологии общего назначения в краткосрочном периоде.

## **Искусственный интеллект – технология широкого применения**

Концепция технологии широкого применения (ТШП) как технологии, которая пронизывает экономику и стимулирует изобретения как эндогенно, так и посредством взаимодополняемости, была введена в 1995 г. Бреснахан и Трайтенберг (Bresnahan and Trajtenberg). Три общих характеристики, которыми обладают ТШП: проникновение в экономику, способность к быстрому внутреннему совершенствованию и способность порождать распространение производительности между секторами за счет взаимодополняемости<sup>2,3</sup>. Примерами технологий широкого применения, которые являются движущими силами экономического роста, служат электричество (XX век) и информационные коммуникационные системы (XXI век). Можно вспомнить формулу В. И. Ленина: «...плюс электрификация всей страны».

Является ли искусственный интеллект технологией широкого применения? Это принципиальный вопрос макроэкономической политики. Для России ответ на данный вопрос имеет первостепенное значение с точки зрения обоснования необходимости перехода от «инерционного пути развития» к «инновационному пути развития». В прогнозе проправительственного ЦМАКП развития народного

---

<sup>2</sup> Bresnahan T. F., Trajtenberg M. (1995). General purpose technologies 'Engines of growth'? Journal of econometrics. 1995. Vol.65.1, pp. 83–108.

<sup>3</sup> Lipsey R. G., Kenneth I Carlaw K. I., Bekar C. T. Economic transformations: general purpose technologies and long-term economic growth. 2005. Oxford University Press, Oxford.

хозяйства РФ на ближайшие годы <sup>4</sup> отмечается существование проблемы одновременной «трудоизбыточности и трудонедостаточности» на рынке труда будущего. На наш взгляд, последовательное применение ИИ, как технологии широкого применения (ТШП) во всех отраслях российской экономики способно решить эту проблему.

Должно ли национальное государство больше тратить средств на развитие науки и, в частности, на исследования в области ИИ, если действует «парадокс производительности» и прямое влияние на экономический рост нельзя измерить? Да, должно, чтобы не быть потерпевшим от усиления экономического неравенства между странами вследствие применения технологий искусственного интеллекта.

С этих позиций сокращение средств государства на развитие ИИ – это шаг назад. Действует известный из макроэкономики «парадокс бережливости»<sup>5</sup>. Если один человек начинает экономить, то он снижает спрос на товары и услуги. Если все граждане страны начинают экономить, то снижается совокупный спрос на товары, и тем самым часть произведенной продукции остаётся нереализованной, лишая дохода этих же жителей страны. Предприятия и их работники вынуждены сокращать объёмы

---

<sup>4</sup> Виноградова Е. В ЦМАКПе назвали пять узлов «противоречий» в экономике России на 20 лет. URL: <https://www.rbc.ru/economics/29/06/2023/649c0c0a9a7947e6cc36e044?ysclid=Indeqounco291299730> (дата обращения 01.10.2023)

<sup>5</sup> Vermann E. K. Wait is saving good or bad? the Paradox of thrift //Page One Economics®. – 2012.

выпуска, откладывать инвестиционные проекты. Поэтому сокращение расходов в период кризиса лишь усугубляет тяжесть экономического положения. Например, вследствие мирового финансового кризиса 2007–2009 гг. в наиболее тяжёлом положении оказались Греция, Португалия и другие страны Южной Европы. МВФ и ЕБРР рекомендовали в качестве условия получения этими странами займов жёсткую политику сокращения государственных расходов за счёт снижения, прежде всего, социальных пособий и льгот. Однако эти сокращения нанесли экономикам данных стран ещё больший вред за счёт снижения платёжеспособного спроса населения.

Использование компаниями технологий ИИ порождает риски для всех участников рынка: последствия могут быть разрушительными как для самих фирм, так и для других участников рынков, и для потребителей. Кроме того, могут возникать и их побочные эффекты.

Кроме того, нужно учитывать, что у общества не сформировалось положительного отношения к применению технологий искусственного интеллекта в экономике. Более того, отношение простых людей к ИИ, в отличие от мнения экспертов, ухудшается. Так опрос Pew Research Center, который был проведён на протяжении 2024 года, показал следующее. Опасения взрослых американцев в целом по поводу искусственного интеллекта выросли с 2021 года, они более обеспокоены, чем эксперты, (51% против 15%

среди экспертов). Общественность гораздо более склонна думать, что ИИ принесёт им вред (43%), чем пользу (24%)<sup>6</sup>.

Специалисты, исследующие проблемы применения больших языковых моделей на практике, выделяют следующие проблемы. Во-первых, предвзятость входных данных<sup>7</sup>. Во-вторых, опасности сбоев функциональности<sup>8</sup>. В-третьих, исследования конкретных случаев показали, что внедрение искусственного интеллекта может привести к глубоким и порой разрушительным изменениям как в организационной структуре, так и в бюрократическом аппарате<sup>9</sup>.

В целом в краткосрочном периоде можно выделить следующие риски использования в экономике технологий искусственного интеллекта:

- риски деформации существующего рынка труда и усиления неравенства в доходах между странами, отраслями, регионами,
- риски концентрации рынков искусственного интеллекта,

---

<sup>6</sup> McClain C., Kennedy B., Gottfried J., Anderson M., Pasquini G. How the U.S. Public and AI Experts View Artificial Intelligence. [Электронный ресурс] URL: [https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence/?utm\\_source=Pew+Research+Center&utm\\_campaign=82cfde2055-Internet\\_Science\\_2025\\_04\\_29&utm\\_medium=email&utm\\_term=0\\_-82cfde2055-399474221](https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence/?utm_source=Pew+Research+Center&utm_campaign=82cfde2055-Internet_Science_2025_04_29&utm_medium=email&utm_term=0_-82cfde2055-399474221). (дата обращения 30.04.2025)

<sup>7</sup> Buolamwini J., Gebru T. Gender shades: Intersectional accuracy disparities in commercial gender classification //Conference on fairness, accountability and transparency. – PMLR, 2018. – P. 77-91.

<sup>8</sup> Raji I. D., Kumar I. E., Horowitz A., Selbst A. The fallacy of AI functionality //Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency. – 2022. – P. 959-972.

<sup>9</sup> Giest S. N., Klievink B. More than a digital system: how AI is changing the role of bureaucrats in different organizational contexts //Public Management Review. – 2024. – Vol. 26. – №. 2. – P. 379–398.

- риски безопасности фирм и создания ложной информации,
- риски переоценённости ИИ.

Возможно существование и иных рисков, но мы выделили основные на сегодняшний момент времени.

### **Риски деформации существующего рынка труда и усиления неравенства в доходах между странами, отраслями, регионами из-за применения технологий искусственного интеллекта**

После появления с ноября 2022 года больших языковых моделей (БЯМ) современный рынок труда начал подвергаться серьёзной модификации. С одной стороны, усилились опасения работников за сохранения своих рабочих мест. «Глобальный поиск Google по запросу «безопасна ли моя работа?» удвоились за последние месяцы, так как люди опасаются, что они будут заменены большими языковыми моделями (llms)<sup>10</sup>. Риски ликвидации рабочих мест и роста безработицы из-за применения технологий искусственного интеллекта вряд ли являются оправданными. Весь предыдущий опыт внедрения человечеством новых технологий свидетельствует об обратном, а страны, которые сегодня располагают самым высоким уровнем использования технологий в мире, - Южная Корея, Сингапур, Япония, - имеют и самый низкий уровень безработицы. Более реальным является процесс модификации характера труда работников современных

---

<sup>10</sup> Boy cries wolf. The Economist. June 17th, 2023.

предприятий. Подробно он был проанализирован в <sup>11</sup> (Lukichev, Chekmarev, 2023).

Остановимся дальше на нескольких изменениях характера труда из-за применения больших языковых моделей. Перемены будут происходить крайне неравномерно и затронут не только низкоквалифицированных работников, но и высококвалифицированных специалистов. Есть позитивное влияние в виде повышения производительности, улучшении условий труда и увеличении ценности своего рабочего места со стороны тех работников, которые осваивают технологии ИИ в своей профессиональной деятельности. Они смогут производить больше продукции или оказывать качественные услуги, повышая тем самым ценность результатов своего труда. Как следствие, заработная плата и условия труда таких специалистов улучшатся. Это позитивное влияние потенциально может затронуть все отрасли экономики. Как отмечала в своем отчёте консалтинговая компания McKinsey три четверти ожидаемой ценности, создаваемой генеративным искусственным интеллектом, приходится на четыре бизнес-функции, лежащие в основе большинства крупных предприятий: исследования и разработки, разработка программного обеспечения, маркетинг и обслуживание клиентов<sup>12</sup>. Любая крупная компания с внутренними базами данных, используемыми для руководства сотрудниками, как

---

<sup>11</sup> Лукичев П. М., Чекмарев О.П. Вызовы экономики искусственного интеллекта традиционному рынку труда // Вопросы инновационной экономики. 2023. Т. 13. № 2. С. 785-802.

<sup>12</sup> ChatGPT Inc. The Economist. July 1st 2023.

американская, так и российская, может найти применение чат-боту с искусственным интеллектом.

Следствием такого продвижения технологий ИИ будет замещение некоторых рутинных работ машинами и необходимость перестройки организации бизнеса внутри компаний. Большие языковые модели, такие как ChatGPT и их аналоги, вполне могут брать на себя решение части функциональных обязанностей работников высококвалифицированного труда. Например, исследование банка Goldman Sachs установило, что 44% юридических задач может выполнять искусственный интеллект<sup>13</sup>. Это позволяет интенсифицировать работу крупных юридических компаний, усиливая творческие и личностные характеристики труда специалистов взамен рутинных характеристик. Если большая языковая модель может за 20 секунд выполнить задачу, на которую у десятка сотрудников ушло бы по 50 часов у каждого, то зачем крупным фирмам продолжать нанимать десятки сотрудников? Конечно, в движении по этому пути возникают побочные эффекты. Первый из них связан со склонностью чат-ботов писать юридические заключения, основываясь не только на реальной, но и на выдуманной информации. Второй побочный эффект возникает из-за непреднамеренной передаче информации, подлежащей конфиденциальности адвоката и клиента, в алгоритмы. Эти проблемы вполне разрешимы. Последний за счёт более совершенных технологий. Первый побочный эффект устраняется за

---

<sup>13</sup> First thing we do, let's bot all the lawyers. The Economist. June 10th 2023.

счёт мониторинга результатов сотрудником и внесения изменений в процесс машинного обучения чат-бота.

Следствием происходящей модификации рынка труда является снижение отдачи от труда и природных ресурсов. Страны и регионы, благополучие которых построено сейчас на экспорте природных ресурсов, не могут консервировать это положение, а должны максимально менять структуру своей экономики с учетом новых технологий. Отрасли обрабатывающей промышленности и сферы услуг более восприимчивы к внедрению новейших технологий. Не случайно страны Персидского залива проводят столько технологических выставок, приветствуя различные нововведения, а в Абу-Даби выпустили самую мощную в мире большую языковую модель с открытым исходным кодом и вскоре запустят компанию по разработке искусственного интеллекта. Новые технологии имеют тенденцию быть трудосберегающими, ресурсосберегающими и порождать динамику «победитель получает всё» (winner-takes-all), которая даёт преимущество развитым странам<sup>14</sup> (Korinek, Stiglitz, 2021). Поэтому перед Российской Федерацией стоит неотложная задача перехода от «инерционного пути развития», основанного на экспорте углеводородов, к «инновационному пути развития». Надо активно использовать конкурентные преимущества инновационной системы России: высокий уровень человеческого потенциала (охват высшим образованием, насыщенность экономики высококвалифицированными кадрами), значимые объёмы

---

<sup>14</sup> Korinek A., Stiglitz J. E. Artificial intelligence, globalization, and strategies for economic development. – National Bureau of Economic Research, 2021. – №. w28453. – 41 p.

финансирования исследований и разработок, сравнительно высокая активность в отношении регистрации результатов интеллектуальной деятельности<sup>15</sup>.

## **Риски концентрации рынков искусственного интеллекта**

Риски концентрации рынков искусственного интеллекта связаны с тем, что они сегодня монополизированы несколькими ведущими цифровыми компаниями и по сути блокируют развитие свободной конкуренции. Об этом свидетельствуют многочисленные исследования. Так, Hötte и др.<sup>16</sup> (Hötte, Tarannum, Verendel, Bennett, 2023) выявили четыре подхода к патентной классификации искусственного интеллекта и пришли к выводу, что «все направления сходятся во мнении, что изобретения в области ИИ в значительной степени сконцентрированы в руках нескольких фирм, и это имеет последствия для политики в области конкуренции и регулирования рынка». Этот же тренд полностью подтверждается и в российских реалиях. Проведённый в РФ патентный анализ показал, что цифровые компании страны вносят значимый вклад в развитие цифровых технологий. Однако обеспечен он, по большей части, довольно узкой группой «национальных чемпионов». На долю тройки лидеров (Яндекс, «Лаборатория Касперского» и Сбер)

---

<sup>15</sup> Российская наука в цифрах: 2023 / В. В. Власова, Л. М. Гохберг, К. А. Дитковский и др.; Нац. исслед. ун-т «Высшая школа экономики». – М.: НИУ ВШЭ, 2023. – 48 с.

<sup>16</sup> Hötte K., Tarannum, T., Verendel V., Bennett, L. (2023). AI Technological Trajectories in Patent Data.

приходится около 76% действующих «цифровых» патентов на изобретения, которые принадлежат компаниям, включенным в рейтинг, и почти 95% – на промышленные образцы в этой области<sup>17</sup>.

Реальным воплощением рисков концентрации рынков ИИ служит следующий кейс. Применение искусственного интеллекта и использование больших данных, породило усиление концентрации в сфере здравоохранения. Покажем это на примере сферы медицинских услуг в США. В Америке, здравоохранение, наряду с технологиями и энергетикой, является сейчас пулом для сверхприбыли по сравнению с его размером. Четыре десятилетия назад более восьми из десяти больниц были некоммерческими и располагались в одном месте. Сейчас более шести из десяти клиник принадлежат разросшимся коммерческим больничным сетям или академическим сетям, таким как Steward Health Care или Indiana University Health<sup>18</sup>. Анализ, проведенный Мартином Гейнором (Martin Gaynor), показал, что рынки здравоохранения не функционируют так хорошо, как могли бы или должны были. Цены высоки и постоянно растут, практика ценообразования вопиющая, качество ниже оптимального. Произошла большая консолидация в здравоохранении. За последние двадцать лет произошло почти 1600 слияний больниц, из них более 450 с 2012 года. В результате в большинстве местных районов теперь доминирует одна крупная

---

<sup>17</sup> Рейтинг цифровых компаний по числу «цифровых» патентов. Цифровая экономика, Экспресс-информация. ИСИЭЗ НИУ ВШЭ. 07.06.2023. [Электронный ресурс]. URL: <https://issek.hse.ru/news/838422953.html> (дата обращения 12.09.2023)

<sup>18</sup> How bad is being big? The Economist. July 15<sup>th</sup>, 2023.

мощная система здравоохранения, например в районе Бостона, - компания Partners, в Питтсбурге, - компания UPMC, и в Сан-Франциско, - Sutter. Последствиями этого стала монополизация сферы здравоохранения и ущемление рыночных принципов. Обширные данные исследований показывают, что консолидация между близкими конкурентами приводит к существенному росту цен для больниц, страховых компаний и врачей, не компенсируя при этом повышения качества или повышения эффективности. Столь же серьёзно, если не больше, данные показывают, что качество обслуживания пациентов страдает от отсутствия конкуренции<sup>19</sup>. То есть получается, что распространение телемедицины, основанной на технологиях искусственного интеллекта и больших данных, привело к резкому усилению концентрации в здравоохранении, сделав отрасль третьей по объёму сверхприбыли в США. Однако для потребителей это привело к удорожанию медицинских услуг и к ухудшению их качества. В той или иной форме такие же последствия возможны во всех сферах народного хозяйства.

В соответствии с концепцией Трайтенберга<sup>20</sup> (Trajtenberg, 2018) существуют два вида инноваций: инновации, заменяющие человека (ИЗЧ) и инновации, улучшающиеся возможности человека (ИУЧ). Первые предполагают снижение затрат на существующие

---

<sup>19</sup> Gaynor M. Diagnosing the Problem: Exploring the Effects of Consolidation and Anticompetitive Conduct in Health Care Markets //Washington, DC. – 2019.

<sup>20</sup> Trajtenberg M. Artificial intelligence as the next GPT: A political-economy perspective. In The economics of artificial intelligence: An agenda. 2018. (pp. 175-186). University of Chicago Press.

технологии и, как следствие, увеличение текущей прибыли. Вторые же предполагают совершенствование потенциала человека, как работника. Например, у врачей улучшаются возможности для диагностики, в частности, для чтения и интерпретации рентгеновских снимков, компьютерной томографии и других методов визуализации. Круг задач, которые люди и машины могут решать поодиночке, несомненно, уже, чем набор задач, которые люди и технологии искусственного интеллекта могут решать совместно. Инновации, улучшающие возможности человека, способны не только полнее раскрыть его креативные и производительные способности, но и создать новые формы организации труда, новые технологии, делающие бесполезными существующие виды производства и распределения товаров, а часто и сами эти товары. Например, можно было сколько угодно улучшать породы лошадей и совершенствовать коляски как транспортные средства, но тех возможностей, которые дали автомобили с двигателями внутреннего сгорания, достичь было невозможно.

Основная проблема сегодняшних рынков искусственного интеллекта состоит в том, что стартапы, создающие собственные большие языковые модели (БЯМ) (Large Language Models - LLMs) имеют отношения с технологическими гигантами симбиотического характера. OpenAi тесно связана с Microsoft, крупным инвестором, который использует технологии для улучшения своего программного обеспечения и поисковых продуктов. Google, принадлежащий Alphabet, имеет значительную долю в Anthropic. Сами технологические гиганты, - Alphabet, Amazon, Apple, Meta и

Microsoft, - для которых технологии искусственного интеллекта лишь одна из форм приложения капитала, господствуют над американским бизнесом. Пять фирм доминируют в индексе фондового рынка S&P 500, на долю которых в совокупности приходится 9% продаж, 16% чистой прибыли и 22% рыночной капитализации. В 2022 году их капиталовложения в размере 360 миллиардов долларов составили более десятой части всех американских инвестиций в бизнес<sup>21</sup>.

Цифровые гиганты с большей вероятностью развернут технологию, чтобы улучшить свои существующие продукты, чем полностью их заменят. Они будут стремиться защитить свой основной бизнес (корпоративное программное обеспечение в случае Microsoft и поиск в Google)<sup>22</sup>. В результате глобальная экономика будет недополучать выгоды, которые были бы возможны в отсутствии избыточной концентрации рынков, поскольку фактическая технология может оказаться менее революционной, чем могла бы быть. Для примера можно вспомнить, что технологические гиганты уже предлагали решения проблемы, которые по качеству и популярности проиграли свободному рынку. Компания Google создала многоязычную энциклопедию Google Knol (закрыта в 2012 г.), а Microsoft организовала цифровую мультимедийную энциклопедию Encarta (срок действия с 1993 года по 2009 год). Обе они были основаны на правах частной собственности и потерпели

---

<sup>21</sup> Can big tech keep getting bigger in the age of AI? The Economist. 5<sup>th</sup> August 2023.

<sup>22</sup> Non-proliferation treaties. The Economist. May 27th, 2023

крах, проиграв Wikipedia, которая была создана на правах свободной собственности и свободного участия. Кроме того, как отмечал Бриньольфссон (Brynjolfsson), чем больше технологий используется для замены, а не для улучшения рабочей силы, тем сильнее может стать неравенство и тем больше недовольства, которое подпитывает деструктивные политические инстинкты и действия<sup>23</sup>.

Фактическая концентрация рынков ИИ свидетельствует, что сейчас искусственный интеллект распределён очень неравномерно. Якобидес и др. (Jacobides, Brusoni, Candelon) демонстрируют заметную и растущую концентрацию ИИ и тот факт, что многие крупные технологические фирмы охватывают весь путь от инфраструктуры до приложений, что также обеспечивает большую часть соответствующих научных преимуществ<sup>24</sup>. То есть в условиях монополизации рынка получение цифровыми гигантами средств государственной поддержки на развитие технологий искусственного интеллекта, как Технологии широкого применения (ТШП), будет усиливать монополизацию рынка и повышать барьеры к входу для новых игроков.

### **Риски безопасности фирм и создания ложной информации**

Расширяющееся в 2023 году применение больших языковых моделей, таких как ChatGPT и её аналогов, по-новому поставило вопросы безопасности компаний. Целый ряд компаний, включая

---

<sup>23</sup> Brynjolfsson, E. The Turing trap: The promise & peril of human-like artificial intelligence. Daedalus. 2022. Vol.151(2). Pp. 272-287.

<sup>24</sup> Jacobides M. G., Brusoni S., Candelon F. The evolutionary dynamics of the artificial intelligence ecosystem //Strategy Science. – 2021. – Т. 6. – №. 4. – С. 412-435.

Apple, Samsung, JPMorgan Chase запретили или сильно ограничили использование ChatGPT. «В мае 2023 г. Samsung обнаружила, что сотрудники загрузили конфиденциальный код в Chatgpt. Проблема заключается в том, что эта информация может храниться на внешних серверах фирм, использующих модели, таких как Microsoft и Alphabet»<sup>25</sup>. Поэтому среди фирм, использующих технологии искусственного интеллекта, усиливается тренд на внутреннее обучение. Экономически это означает, что преимущество в конкурентной борьбе получают те компании, которые внедряют специализированные (отраслевые модели) чат-боты. Но это требует, в отличие от использования Chatgpt, значительных инвестиций, которые не под силу большинству средних и малых фирм. Отсюда вытекает рост неравенства доходов (и возможностей) среди предприятий отрасли. Поэтому увеличивается риск монополизации отдельных отраслей и неравенства между отраслями, регионами, странами.

Искусственный интеллект предоставил новые возможности для снижения неопределенности за счёт использования данных для быстрого реагирования на изменяющиеся условия. Это становится возможным сразу же, как только поступают свежие данные. Одновременно новые сведения искажают сумму показателей, накопленных большой языковой моделью, внося качественные изменения. Поскольку процесс принятия решений моделями искусственного интеллекта существенно отличается от

---

<sup>25</sup> ChatGPT Inc. The Economist. July 1st 2023.

человеческого мышления, то результаты «работы» ИИ могут быть неожиданными для их создателей и нести существенные риски применения. Для человека характерно, что большой объём информации об явлении (процессе) и большой период обучения приводят к лучшим итогам. Для больших языковых моделей это оказалось не так. Исследованиями зафиксировано, что более длительный период machine learning (машинного обучения) и большой объём получаемых данных приводят зачастую к деградации результатов их работы. Почему? Дело в том, что у машинного обучения наряду с несомненными достоинствами есть и существенные недостатки. Так, те входные данные, которые привели к одному результату вчера, могут зарегистрировать другой результат завтра, потому что алгоритм был изменён данными, которые он получил в промежуточный период.

GPT-3.5 и GPT-4 — две наиболее широко используемые сегодня большие языковые модели (LLM). Исследование Чен и др.<sup>26</sup>. [21] (Chen, Zaharia, Zou, 2023) показало, что обновление данных в них приводит к противоречивым результатам. Ученые оценили версии GPT-3.5 и GPT-4 за март 2023 г. и июнь 2023 г. с помощью тестирования по семи параметрам. Среди задач, решаемых БЯМ ((LLMs), были: 1) математические задачи, 2) деликатные/опасные вопросы, 3) опросы общественного мнения, 4) многоэтапные вопросы, требующие больших знаний, 5) генерация кода, 6) тесты на

---

<sup>26</sup> Chen L., Zaharia M., Zou J. How is ChatGPT's behavior changing over time? //arXiv preprint arXiv:2307.09009. – 2023.

получение медицинской лицензии США и 7) визуальное мышление. Как видно даже из этого перечисления проверка эффективности работы больших языковых моделей была проведена комплексно. Исследователи выявили, что производительность и поведение как GPT-3.5, так и GPT-4 могут сильно меняться со временем. Например, GPT-4 (март 2023 г.) хорошо справился с определением простых и составных чисел (точность 84%), но GPT-4 (июнь 2023 г.) плохо справился с этими же вопросами (точность 51%). Частично это объясняется снижением удобства GPT-4 следовать подсказкам цепочки мыслей. Интересно, что GPT-3.5 в июне справлялась с этой задачей гораздо лучше, чем в марте. В июне GPT-4 стал менее охотно отвечать на деликатные вопросы и вопросы опросов общественного мнения, чем в марте. В июне GPT-4 лучше справлялся с многоэтапными вопросами, чем в марте, а производительность GPT-3.5 в этой задаче упала. И в GPT-4, и в GPT-3.5 в июне 2023 г. было больше ошибок форматирования при генерации кода, чем в марте<sup>27</sup>. Таким образом, мы приходим к двум выводам:

- 1) нужен постоянный мониторинг работы больших языковых моделей,
- 2) необходимо качественно новое государственное регулирование возникшей проблемы.

---

<sup>27</sup> Chen L., Zaharia M., Zou J. How is ChatGPT's behavior changing over time? //arXiv preprint arXiv:2307.09009. – 2023.

Примером (прообразом) последнего могут служить некоторые регулирующие органы, в частности U.S. Food and Drug Administration (Управление по санитарному надзору за качеством пищевых продуктов и медикаментов США), разрешившие использовать в них только «заблокированные» алгоритмы, которые не учатся каждый раз при использовании продукта и, следовательно, не меняются<sup>28</sup> (Candelon, di Carlo, De Bondt, Evgeniou, 2021). «Новое государственное регулирование» вызовет несколько последствий, как для потребителей, так и для фирм-производителей. Для потребителей будет определённая гарантия на применяемые ими медицинские устройства или принимаемые лекарства. Для компаний возникнет необходимость параллельно запускать две версии одного и того же алгоритма: одну зафиксированную версию для коммерческого использования, которая была одобрена регулирующими органами, и вторую «перспективную», которая будет улучшаться по мере машинного обучения и получения новых данных. В идеале «перспективная» версия после утверждения регуляторами заменит «заблокированные» алгоритмы. В любом случае это взаимодействие с регулирующими органами будет играть всё возрастающую роль как в конкуренции на рынке, так и в совершенствовании технологий ИИ.

---

28 Candelon F., di Carlo R. C., De Bondt M., Evgeniou T. (2021). AI Regulation Is Coming: How to prepare for the inevitable. *Harvard Business Review*, 99(5). P. 102–111.

## Риски переоценённости ИИ

Сейчас имеет место переоценённость технологий искусственного интеллекта в экономике. Так, по сравнению с началом 2023 Ехтгода, цена акций «большой пятерки» — Alphabet, Amazon, Apple, Meta и Microsoft — подскочила на 60%, если измерять её в равно взвешенной корзине<sup>29</sup>. В основе ценового роста лежит вера инвесторов в возможности больших языковых моделей (llms) как новой волны технологий. Традиционно предполагается, что те фирмы, которые наилучшим образом и быстрее применят в своей бизнес-практике новейшие технологии, - в данном случае, - искусственного интеллекта, захватят большую долю рынка и увеличат свою прибыль.

Это предположение подтверждается некоторыми исследовательскими данными. Так, в исследовании А. Эйсфельдт и др. выявлено, что влияние выпуска ChatGPT на стоимость компаний было большим, что привело к разнице в доходах компаний примерно на 0,4% в день, что соответствует более чем 100% в годовом исчислении<sup>30</sup> (Eisfeldt, Schubert, Zhang, 2023). Данные различия были обнаружены как внутри, так и между отраслями и демонстрируют большие отличия, которые коррелируют с характеристиками фирмы, такими как организационный капитал или валовая прибыль.

---

<sup>29</sup> ChatGPT Inc. The Economist. July 1st 2023.

<sup>30</sup> Eisfeldt A. L., Schubert G., Zhang M. B. Generative ai and firm values. – National Bureau of Economic Research, 2023. – №. w31222.

Всё это очень напоминает известные из макроэкономики этапы развития финансовых рынков по Х. Мински. На первом этапе, - «замещение», - происходит нечто, что меняет восприятие рынков инвесторами. В конце 1990-х годов это был Интернет, на использовании которого, как казалось, можно делать деньги бесконечно, сейчас – это искусственный интеллект. На втором этапе, - «подъём», - надежды инвесторов оправдываются и данные Эйсфельдт и др. статистически подтверждают это. Фактически сейчас экономика находится на переходе к третьему этапу по Х. Мински<sup>31</sup>, - эйфории. Однако будет ли так на самом деле, или экономика быстро скатится к четвёртому и пятому этапам, «получение прибыли» и «паника», - неясно. На четвёртом этапе инвесторы осознают, что рост стоимости акций не может продолжаться вечно, и сбывают их. На пятом этапе люди массово стараются избавиться от инвестиций и возникает кризис ликвидности.

Эйфория от применения технологий искусственного интеллекта в бизнесе во многом подкрепляется тем, что самые близкие к технологии люди — люди, объясняющие её публике — живут, образно говоря, одной ногой в будущем. Иногда они видят то, что, по их мнению, произойдет так же, как и то, что происходит сейчас. Они склонны, по нашему мнению, как Пигмалион оживлять своих Галатей. Однако, как отмечала профессор Элисон Гопник

---

<sup>31</sup> Лукичѳв П. М., Чекмарев О. П., Шматко А. Д. Макроэкономика. Учебное пособие (2-е издание, расширенное и дополненное). Санкт-Петербург. 2024.

(Alison Gopnik): «Вычислительные мощности нынешнего ИИ. такие как большие языковые модели, не делают более вероятным, что они разумны больше, чем камни или другие машины»<sup>32</sup>.

Какими будут экономические последствия развития искусственного интеллекта зависит от многих причин. Одна из них, - скромное реальное применение. Так, в 2020 году на промышленных роботов, для которых чаще всего и применяются алгоритмы искусственного интеллекта, было потрачено менее 1% мировых капитальных затрат. Люди больше тратят на секс-игрушки<sup>33</sup>.

Завышенные ожидания доходов от применения технологий искусственного интеллекта могут стать причиной кризиса всей экономики.

Появление в январе 2025 года китайской модели искусственного интеллекта DeepSeek, обучение которого стоило всего \$5,6 млн против десятков миллионов у ChatGPT, произвело экономическую революцию. С одной стороны, DeepSeek бесплатен в отличие от ChatGPT, работает быстрее и дешевле в эксплуатации, исходный код его (Apache 2.0) открыт. С другой стороны, риски недобросовестного использования возрастают многократно.

---

<sup>32</sup> Metz C. A.I. Does Not Have Thoughts, No Matter What You Think. New York Times. August 7, 2022. Section BU, p. 6.

<sup>33</sup> Where are all the robots? The Economist. March 11th 2023.

## **Меры по предупреждению и/или уменьшению негативных последствий применения технологий искусственного интеллекта**

Меры по предупреждению рисков ИИ должны проводиться на нескольких уровнях, - уровне предприятий, уровне отрасли, уровне национальной экономики, глобальном уровне, - и быть взаимосвязанными между собой. Компаниям следует играть активную роль в написании сводов правил для алгоритмов искусственного интеллекта. Как отмечают Канделон и др., директорам по управлению рисками, возможно, придётся расширить свои полномочия, включив в них мониторинг автономных процессов ИИ и оценку уровня юридических, финансовых, репутационных и физических рисков, на которые компания готова взять на себя эволюционирующий искусственный интеллект<sup>34</sup> [Candelon, 2021]. Наряду с уровнем фирм должны быть созданы национальные системы для постоянного мониторинга и оценки безопасности, производительности и воздействия приложений искусственного интеллекта. Здесь важно создание именно национальных служб, поскольку требования разных стран могут и должны отличаться, но они должны быть интегрированы впоследствии с общемировыми органами.

---

34 Candelon F., di Carlo R. C., De Bondt M., Evgeniou T. (2021). AI Regulation Is Coming: How to prepare for the inevitable. Harvard Business Review, 99(5). P. 102–111.

Основное место в предупреждении рисков применения технологий искусственного интеллекта и/или снижении негативных последствий занимает государственное регулирование на национальном уровне или на глобальном уровне. Остановимся более подробно на национальном уровне поскольку глобальный уровень регулирования в создавшейся международной обстановке является крайне проблематичным.

На национальном уровне государственное регулирование должно быть нацелено на ужесточение требований к работе компаний, действующих на рынках искусственного интеллекта. На наш взгляд, необходим плавный переход к введению чётких и понятных рамок действия фирм на рынках ИИ. Прежде всего, должно быть введено *лицензирование компаний, занимающихся искусственным интернетом*. Оно должно включать в себя перечень необходимых требований, подобный существующему лицензированию у производителей лекарств. Эти рамки должны быть тщательно продуманы с анализом всех возможных последствий, поскольку существует большая вероятность монополизации рынка ИИ технологическими гигантами и создание барьеров к входу новых компаний.

Общей проблемой, генерирующей риски для всей экономики при создании моделей искусственного интеллекта сегодня, является «технологический уклон». Следует согласиться со Штраубом В. и др., которые отмечают, что в ответ на существенный вред, уже причиненный существующими приложениями ИИ при принятии важных решений... не должно существовать технологической

изоляции без большего вклада со стороны ученых, занимающихся человеческим поведением, этикой, культурой и правом.<sup>35</sup> [Straub V]

Кроме того, к выпускаемым большим языковым моделям, которые взаимодействуют с конечными потребителями, должны предъявляться следующие ограничения.

- Отдельные *продукты и услуги, потребление которых непосредственно отражается на здоровье людей*, как, например, пищевые продукты или фармацевтические препараты, *должны проходить дополнительное исследование*, чтобы продемонстрировать свою безопасность.
- *Предоставление системами ИИ объяснения своих решений во всех спорных случаях*, иначе «размывается» существующая сейчас персональная ответственность за нарушения. Как справедливо отмечали Мительштадт и др, алгоритмические системы создают проблемы подотчетности и моральной ответственности, когда неясно, какой моральный агент в процессе несет (или разделяет) ответственность за результаты системы<sup>36</sup> (Mittelstadt, Allo, Taddeo, Wachter, Floridi, 2016). Это особенно важно при катастрофических последствиях действий неисправных машин с использованием алгоритмов ИИ, которые как автономные автомобили или медицинские

---

<sup>35</sup> Straub V. J., Morgan D., Bright J., Margetts H. Artificial intelligence in government: Concepts, standards, and a unified framework //Government Information Quarterly. – 2023. – Vol. 40. – №. 4. – P. 101881.

<sup>36</sup> Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., Floridi, L.: The ethics of algorithms: mapping the debate. Big Data Soc. 3(2), 2053951716679679 (2016). URL: <https://doi.org/10.1177/2053951716679679> (дата обращения 27.09.2023)

приложения могут привести к гибели людей<sup>37</sup> (Ayling, Chapman, 2022). Кроме того, это может иметь судьбоносное значение в таких приложениях как правоохранительная деятельность, кредитный рейтинг, трудоустройство студентов. По нашему мнению, должна быть создана специальная система с ранжированием степени необходимости объяснений большими языковыми моделями (объяснение первого уровня, объяснение второго уровня, объяснение третьего уровня). Кроме того, на уровне моделей ИИ, можно *использовать существующий сейчас опыт с Интернетом, когда на дистрибьютеров возлагается ответственность за устранения неправды, заявлений о клевете или злонамеренных обвинений*. В целом, это позволит улучшить контроль и повысит доверие к использованию технологий искусственного интеллекта.

- ***Законодательный запрет моделям ИИ при взаимодействии с людьми выдавать себя за людей***. Здесь важно не только введение закона, сколько создание действенного мониторинга за его соблюдением. Последнее означает что проверка должна быть быстрой, а наказание существенным.

Предлагаемые автором меры, конечно, нуждаются в дальнейшем обсуждении. Однако, чем быстрее они будут введены в действие, тем больших потерь от рисков применения технологий искусственного интеллекта удастся избежать.

---

<sup>37</sup> Ayling J., Chapman A. Putting AI ethics to work: are the tools fit for purpose? //AI and Ethics. – 2022. – Т. 2. – №. 3. – С. 405-429.

## Выводы и последующие исследования

Инвестиции в искусственный интеллект и их эффективность во многом зависят от анализа рисков внедрения ИИ, того насколько искусственный интеллект является технологией общего назначения, мер государственного регулирования и выгоды применения больших языковых моделей по сравнению с существующими сейчас технологиями. Анализ рисков использования искусственного интеллекта в краткосрочном периоде показывает, что вполне возможно их нейтрализовать с помощью существующих сегодня традиционных мер с учётом специфики ИИ.

По мере того, как Большие языковые модели демонстрируют всё больше преимуществ для экономических агентов, применяющих их в различных сферах экономики, то всё острее встаёт вопрос о государственном регулировании. LLMs (БЯМ), и прежде всего, победное распространение ChatGPT и DeepSeek, демонстрируют создание качественно новой ценности алгоритмов искусственного интеллекта, возможность перехода от «экзотической игрушки» к производству массовых, прибыльных продуктов. Из экономической истории можно вспомнить Генри Форда, который создал модель «Ford-T» и тем самым новую высоко прибыльную отрасль, реализовав свой слоган «автомобиль – это не роскошь, а средство передвижения». Или, если обратиться к концепции жизненного цикла товара / услуги, то *для искусственного интеллекта сейчас — это переход от стадии внедрения к стадии роста.*

Исследования безопасности применения технологий искусственного интеллекта только начинаются. В них выделяются четыре отдельных области: надёжность, мониторинг, согласование и системная безопасность <sup>38</sup> (Hendrycks, Mazeika). Как надеются Хендрикс и др., эти четыре области исследований представляют собой высокоуровневые приоритеты исследований в области безопасности, которые могут обеспечить глубокую защиту от рисков ИИ<sup>39</sup> (Hendrycks, Carlini, Schulman, Steinhardt, 2021). Предлагаемые нами меры по предупреждению и/или уменьшению негативных последствий применения технологий искусственного интеллекта представляют собой необходимые инструменты для первоначального государственного регулирования.

Возможное усиление государственного регулирования не должно препятствовать развитию конкуренции в сфере создания и применения технологий искусственного интеллекта, создавая барьеры к входу для новых участников рынка. Для российских компаний важно не только участие в моделях типа llama, но и взятие курса на создание своих отраслевых чат-ботов. Напомним, что llama – это Llm (Большая языковая модель), разработанная Meta в марте 2023 г. Её разработчики с открытым исходным кодом уже «могут создавать модели генеративного ИИ, которые почти так же хороши,

---

<sup>38</sup> Hendrycks D., Mazeika M. X-risk analysis for ai research //arXiv preprint arXiv:2206.05862. – 2022.

<sup>39</sup> Hendrycks D., Carlini N., Schulman J., Jacob Steinhardt J. “Unsolved problems in ml safety”. In: *arXiv preprint arXiv:2109.13916* (2021).

как существующие проприетарные, и стоят в сотни раз дешевле»<sup>40</sup>. И для России, и для стран ЕС в 2025 году появились дальнейшие возможности развития в связи с выпуском китайской модели DeepSeek. Поэтому вопрос о том, будут ли технологии искусственного интеллекта «закрытым клубом» цифровых гигантов, поддержанных государственным регулированием, или они будут развивать инновации обоих видов на конкурентных рынках, остаётся открытым. Для Российской Федерации, учитывая, что фирмы США и КНР ушли далеко вперед в развитии ИИ, важна государственная поддержка «национальных чемпионов».

В целом, применение искусственного интеллекта для решения отдельных частных проблем является выгодным за счёт снижения издержек на производство и реализацию товаров. То есть LLMs как инновации, заменяющие человека, могут и должны расширять области своего применения. Выявленные нами на микроуровне риски использования алгоритмов искусственного интеллекта вполне могут быть уменьшены за счёт существующих технологий управления рискам и соответствующей переподготовки кадров.

---

<sup>40</sup> Non-proliferation treaties. The Economist. May 27th, 2023

## **1.2. ДОЛГОСРОЧНЫЕ РИСКИ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ЭКОНОМИКЕ**

В разделе даётся характеристика рисков, возникающих при применении технологий искусственного интеллекта в долгосрочном периоде. Выявлены два качественно разных уровня рисков, которые несёт применение больших языковых моделей на макроуровне в долгосрочном периоде. Хвостовые риски и экзистенциальные риски от использования моделей искусственного интеллекта способны временно или постоянно (навсегда) ограничить долгосрочный потенциал человечества. Сопоставление рисков применения искусственного интеллекта на микроуровне и на макроуровне обнаруживает качественно разные возможности для их предотвращения или уменьшения. Особо выделены политэкономические риски использования моделей искусственного интеллекта в будущем, которые вызывают необходимость замены существующей экономической идеологии. Анализ реальности воплощения предыдущих проектов цифровыми монополиями показывает их ориентацию на извлечение прибыли, а не на решение проблем человечества. Поэтому и технологии, создаваемые большими языковыми моделями, больше ориентированы на снижение издержек, чем на улучшение человеческого потенциала. В заключении делается вывод о необходимости использования сценарного подхода к анализу рисков применения искусственного интеллекта в долгосрочном периоде и вводится критерий анализа.

Помимо текущих рисков использования технологий искусственного интеллекта (ИИ) важно проанализировать риски ИИ в долгосрочном периоде.

Во многом эти риски за временным периодом в один год связаны с применением больших языковых моделей (БЯМ) (Large language models (LLMs) на макроуровне или с созданием общего искусственного интеллекта (AGI – Artificial Generative Intellect). И то, и другое представляют собой будущие риски, которых ещё нет, но которые со временем будут жизненно важными. Технологии искусственного интеллекта в перспективе могут стать одним из способов координации экономики, наряду с существующими рынком и бюрократией, поскольку каждый способ основан на обработке огромных массивов информации.

**Цель исследования** – выявить долгосрочные риски использования моделей искусственного интеллекта в будущем и предложить меры по предупреждению отрицательных последствий их применения.

**Гипотеза исследования** - выявленные риски применения технологий искусственного интеллекта на макроуровне требуют нового (не существующего на данный момент) уровня государственного регулирования.

## **Применение технологий искусственного интеллекта на макроуровне**

Сейчас, - в XXI веке, - всё более важную роль в экономическом развитии играют наука, образование и их воплощение в технологиях. Технологией, определяющей прогресс современной экономики, является искусственный интеллект. В целом, наука служит ключевым фактором экономического роста и социального прогресса.

Если науку можно ускорить – например, за счёт повышения эффективности, с которой деньги на исследования превращаются в открытия и в коммерциализированные изобретения, – то же самое можно сделать и с ростом национальной экономики<sup>41</sup>.

Прямой зависимости между увеличением использования современных технологий, таких как ИИ, и национальным экономическим развитием, как показывает практика, пока не наблюдается. Одной из причин является нарастание рисков применения моделей искусственного интеллекта для экономики и для общества в целом. Риски использования ИИ краткосрочного периода были проанализированы в<sup>42</sup>. В данном разделе акцент будет сделан на выявлении рисков применения искусственного интеллекта в долгосрочном периоде на макроуровне. Под последним в экономике традиционно понимается уровень национальной экономики. Если для XX века характерным было противопоставление рыночной экономики и плановой экономики, то сейчас народное хозяйство координируют не только рынок и государственное регулирование, но и модели искусственного интеллекта (ИИ). *Все они представляют собой обезличенные распределенные системы обработки информации.* Общее для ИИ,

---

<sup>41</sup> Clancy M., Correa D., Dworkin J., Niehaus P., Watney C., Williams H. Want to speed up scientific progress? First understand how science policy works. Nature. 2023. Vol. 620(7975). P. 724–726.

<sup>42</sup> Лукичѳв П.М., Чекмарев О.П. Риски применения искусственного интеллекта в краткосрочном периоде // Вопросы инновационной экономики. – 2023. – Том 13. – № 4. – doi: 10.18334/vines.13.4.119359.

рынка, бюрократии – обработка поступающей экономической информации. Различия – в том, КАК они это делают.

Применение больших языковых моделей, подобных ChatGPT и его аналогов, создаёт риски, которые на первый взгляд, кажутся незначительными, но при своём применении несут экзистенциальные угрозы для существования всего человечества. На поверхностном уровне практика применения технологий ИИ выявила риски, связанные с их некорректным использованием человеком, - ошибочность данных, используемых для обучения и ввода, проблемы с конфиденциальностью информации, усиление слежки за гражданами, проблемы подотчетности и моральной ответственности за результаты работы технологий ИИ. Кроме того, алгоритмам искусственного интеллекта часто присуща непостижимость «черного ящика», что уже приводило к гибели людей из-за неисправности машин (вспомните медицинские приложения или автономные автомобили)<sup>43</sup>.

Пандемии, подобные недавней COVID-19, и учащение техногенных катастроф показывают опасность хвостовых рисков, которые редки, но очень серьёзны<sup>44</sup>. Последние связаны с будущим, ближайшим или отдаленным, когда будут созданы ещё более интеллектуальные и мощные модели ИИ, что может, в конечном

---

<sup>43</sup> Ayling J., Chapman A. Putting AI ethics to work: are the tools fit for purpose? //AI and Ethics. – 2022. – Т. 2. – №. 3. – С. 405-429.

<sup>44</sup> Taleb N. N. Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications //arXiv preprint arXiv:2001.10488. – 2020.

итоге, привести к созданию систем искусственного интеллекта, более мощных, чем мы. Разделим два качественно разных уровня опасности, которые несут для будущего человечества эти риски. Хвостовой риск — это маловероятный риск, который может иметь серьёзные последствия. Экзистенциальный риск или X-риск — это риск, который может навсегда ограничить долгосрочный потенциал человечества<sup>45</sup>.

Открытым остаётся вопрос: всегда ли возможно перенесение традиционных концепций анализа рисков на модели искусственного интеллекта? Иначе: есть ли специфика LLMs, которая не позволяет полностью положиться на традиционные концепции анализа рисков? Может быть, machine learning? Для объективного ответа на эти вопросы необходимо проанализировать риски применения технологий искусственного интеллекта на макроуровне. Принципиально важно, по нашему мнению, различать риски и возможности снижения негативных последствий от их возникновения на микроуровне и на макроуровне.

На микроуровне отрицательное воздействие рисков может быть реально уменьшено. Например, на уровне отдельного человека риск попасть в автомобильную аварию можно снизить, двигаясь медленнее (уменьшая опасность), двигаясь по менее опасным дорогам (уменьшая воздействие) или пристёгивая ремень безопасности (уменьшая уязвимость). На уровне предприятия

---

<sup>45</sup> Ord T. The precipice: Existential risk and the future of humanity. – Hachette Books, 2020. – 468 p.

химической отрасли риск утечек химических веществ с промышленного предприятия можно снизить, используя менее опасные химические вещества (уменьшая опасность), строя завод вдали от населённых пунктов (уменьшая воздействие) и снабжая рабочих защитной одеждой (уменьшая уязвимость). То есть, учёт рисков в бизнес-практике вполне может уменьшить их негативное влияние.

Что же касается макроуровня, то здесь существует принципиально иное соотношение между рисками, их последствиями и возможностями людей по их уменьшению.

Традиционное уравнение риска можно расширить следующим образом:  $\text{Риск} = \text{Опасность} \times \text{Воздействие} \times \text{Уязвимость} / \text{Способность справляться с поправкой на способность справляться с авариями или восстанавливаться после них}$ . Здесь «Опасность» означает серьёзность и распространённость опасности, а « $\times$ » просто указывает на взаимодействие. Это относится к рискам, связанным с искусственным интеллектом, потому что, **если мы потеряем контроль над сильной системой ИИ, наша способность справляться может быть нулевой**. Точно так же по определению x-риски постоянны, поэтому данное уравнение показывает, что риск таких событий безгранично велик<sup>46</sup>. Выделенное жирным шрифтом это, по сути, риски на макроуровне.

---

<sup>46</sup> Hendrycks D., Mazeika M. X-risk analysis for ai research //arXiv preprint arXiv:2206.05862. – 2022.

## **Опасности координирующей роли искусственного интеллекта на макроуровне**

Г. Фаррелл и К. Шализи (G. Farrell, K. Shalizi) ввели в 2023 году сравнение последствий применения ИИ с действием рыночной системы и бюрократии по координации экономики. И то, и другое управляют современным обществом, обеспечивая реально функционирующую современную экономику. «Рынки и бюрократия кажутся знакомыми, но на самом деле они представляют собой огромные обезличенные распределенные системы обработки информации, которые превращают бурлящий хаос наших коллективных знаний в полезные упрощения. Рынки и государства могут иметь огромные коллективные преимущества, но они, безусловно, кажутся враждебными людям, потерявшим работу из-за экономических изменений или запутавшимся в дурацких клубках бюрократических решений»<sup>47</sup>.

Подобно рынкам и бюрократии, большие языковые модели (БЯЗ – LLMs) представляют собой также нечто огромное и непостижимое для обычных граждан. Логика принятия решений алгоритмами искусственного интеллекта не всегда понятна нам на микроуровне, а на макроуровне неопределённость будет только возрастать. Тем не менее, поскольку многие функции человеческого мышления, например вычислительные, слабеют и отмирают, то со

---

<sup>47</sup> Farrell H., Shalizi C. "Artificial intelligence is a familiar-looking monster, say Henry Farrell and Cosma Shalizi". The Economist. June 24th 2023.

временем передача функций управления на макроуровне большим языковым моделям будет становиться всё более реальной.

Сегодняшние тренды на рынке труда показывают, что управляемое искусственным интеллектом подталкивание проникает в бизнес и общество с экспоненциальной скоростью. Оно становится всё более привычным, заменяя растущую часть решений, принимаемых раньше исключительно людьми. Например, подталкивание используется для достижения двузначного увеличения удержания клиентов<sup>48</sup>, для улучшения безопасности труда<sup>49</sup>, для повышения уровня пенсионных накоплений<sup>50</sup>, для улучшения гигиенического поведения во время пандемии<sup>51</sup>.

Как отмечала А. Гопник, LLMs — это не зарождающийся индивидуальный интеллект, а «культурные технологии», которые реорганизуют и бесшумно передают человеческое знание<sup>52</sup>. В свою очередь, политический антрополог Джеймс Скотт (James Scott)

---

<sup>48</sup> Irrational Labs (2021). How can we reduce customer churn? In [irrationallabs.com](https://irrationallabs.com/casestudies/google-how-can-we-reduce-customer-churn/). Дата обращения 30.11.2022 <https://irrationallabs.com/casestudies/google-how-can-we-reduce-customer-churn/>

<sup>49</sup> Nioi A., Wendelboe-Nelson C., Cowan S., Cherrie M., Rashid S., Cowie H., Davis A., Ritchie P., Lansdown T., Cherrie J. W. Nudging construction workers towards better sun-safety behaviour: summary of the evidence for practitioners // *Policy and Practice in Health and Safety*. 2020. Vol. 8(1), pp. 25-33. DOI: 10.1080/14773996.2019.1708614

<sup>50</sup> Benartzi S., Thaler R. Heuristics and biases in retirement savings behavior // *Journal of Economic Perspectives*. 2007. Vol. 21(3). pp. 81-104. DOI: 10.1257/jep.21.3.81

<sup>51</sup> Weijers R. J., de Koning B. B. Nudging to increase hand hygiene during the COVID-19 pandemic: A field experiment // *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*. 2021. Vol.53(3). p. 353. DOI: 10.1037/cbs0000245

<sup>52</sup> Gopnik A. What AI Still Doesn't Know How to Do. *The Wall Street Journal*. July 15, 2022. URL: <https://www.wsj.com/articles/what-ai-still-doesnt-know-how-to-do-1165789131627>

объяснил, что бюрократия — это чудовище информации, пожирающее богатые, неформальные массивы негласно хранимых знаний и выделяющая жидкую жижу абстрактных категорий, которые правители используют, чтобы «видеть» мир<sup>53</sup>. Более полвека назад экономист Мартин Вайцман (Martin Weitzman) предположил, что «плановая экономика может использовать математические объекты, называемые «разделяющими гиперплоскостями», чтобы адаптироваться на лету»<sup>54</sup>. Сейчас же Г. Фаррелл и К. Шализи указывают, что «машинное обучение может находить такие гиперплоскости, делая планирование более осуществимым, чем раньше»<sup>55</sup>. Технологически это означает возрождение применения плановых методов управления народным хозяйством.

Именно в этом смысле большие языковые модели (БЯМ – LLMs) способны дать новый толчок (создать почву для) развитию командной экономики в ущерб рынку, или как минимум обеспечить бюрократам новые инструменты для решения сегодняшних экономических проблем. Происходящая сейчас по инициативе США «деглобализация» мировой экономики объективно усиливает позиции национальной бюрократии, а технологии искусственного

---

<sup>53</sup> Scott J. C. Seeing like a state: How certain schemes to improve the human condition have failed. – Yale University Press, 2020.

<sup>54</sup> Weitzman M. Iterative multilevel planning with production targets //Econometrica: Journal of the Econometric Society. – 1970. – p. 50-65.

<sup>55</sup> Farrell H., Shalizi C. "Artificial intelligence is a familiar-looking monster, say Henry Farrell and Cosma Shalizi". The Economist. June 24th 2023

интеллекта дают ей невиданные возможности и мощь. По сути, государственное регулирование вместо создания условий для развития рынка заменяет его, перераспределяя по-новому обязанности между ними. Это ущемляет свободу выбора потребителей и возможности бизнеса, а ведь по-прежнему «предприятия хотят максимизировать прибыль, а не предоставлять слабо определенные общественные блага для граждан конкретной страны»<sup>56</sup>.

Применение технологий ИИ усилит концентрацию бизнеса как в национальных, так и в международных рамках, одновременно увеличивая неравенство между фирмами по активам и доходам, неравенство между странами, регионами, отраслями по уровням экономического развития.

С другой стороны, надежды общества на то, что государственное регулирование применения LLMs будет в интересах развития предпринимательства, свободного рынка и обычных граждан являются, по нашему мнению, иллюзорными. Национальные государства и международное сообщество могут создать определенные рамки использования LLMs, но они будут в реальности в интересах тех цифровых гигантов, которые доминируют сегодня и создадут барьеры к входу для новых «игроков». Само же государственное регулирование искусственного интеллекта, как прозорливо заметил Дэниел Солов (Daniel Solove),

---

<sup>56</sup> Farrell H., Newman A. The New Economic Security State // Foreign Affairs. November-December 2023.

рискует создать кафкианский мир с «более бездумным процессом бюрократического безразличия, произвольных ошибок и дегуманизации, мир, в котором люди чувствуют себя бессильными и уязвимыми, без какой-либо значимой формы участия в сборе и использовании их информации»<sup>57</sup>.

### **Политэкономические риски применения искусственного интеллекта в будущем**

Прогресс в развитии моделей искусственного интеллекта, особенно явный после запуска в ноябре 2022 года ChatGPT, собравшей за два месяца с момента запуска свыше 100 миллионов пользователей, ставит вопрос о рисках перехода к следующей, качественно новой стадии – к стадии общего искусственного интеллекта, или A.G.I. (*artificial general intelligence*, or A.G.I.). Да, его пока не существует, но исследователи всё ближе подходят к его созданию. Это приводит нас к необходимости анализа политэкономических рисков применения моделей ИИ в будущем. Какая экономическая идеология придёт на смену существующей?

Лобби ученых, инвесторов и предпринимателей утверждает, что, как только A.G.I. станет безопасным, его применение будет благом для цивилизации. Технологический инвестор из Силиконовой долины Марк Андрессен, в недавнем эссе под

---

<sup>57</sup> Solove D.J. Privacy and power: computer databases and metaphors for information privacy// Stanford Law Review. 2001. P. 1398.

характерным названием «Почему искусственный интеллект спасёт мир», — даже заявляет, что ИИ, «Как и любая другая технология, принадлежит людям и контролируется людьми»<sup>58</sup>. Но это не имеет никакого отношения к реальности: технологии контролировались и контролируются крупным бизнесом в целях получения максимума прибыли. Как отмечал Ю. Харари, либеральная политическая система сформировалась в индустриальную эпоху, чтобы управлять миром паровых машин, нефтеперегонных заводов и телевизоров. Ей трудно приспособиться к непрерывным революционным изменениям, связанным с развитием биотехнологий и информационных технологий (ИТ)<sup>59</sup>. Любая идеологическая система, будь то коммунизм, или либерализм опиралась на определённые технологии, на определённые системы машин. Кризис либерализма сейчас связан с тем, что он не готов воспринять искусственный интеллект, биотехнологии, цифровизацию как основу своего развития. Следовательно, делает вывод Харари, нужна или модернизация, или «новый проект». Однако вместо модернизации появляется идеология A.G.I.-изм, соединяющая искусственный интеллект с неолиберализмом. Отличительной чертой данного проекта становится безальтернативность. Как в своё время выразилась Маргарет Тэтчер, нет альтернативы, кроме рынка.

---

<sup>58</sup> Andreessen M. (2023). *Why AI Will Save the World*. Andreessen Horowitz, Why AI Will Save the World | Andreessen Horowitz (a16z.com)

<sup>59</sup> Харари Ю. 3. 21 урок для XXI века. Издательство «Синдбад». 2019.

Напомним, что когда человек знает тему, то он сомневается, когда человек не знает, то он имеет твёрдое мнение. Поэтому, по мнению автора книги “To Save Everything, Click Here: The Folly of Technological Solutionism” («Чтобы спасти все, щелкните здесь: глупость технологического решения») и ведущего подкаста «The Santiago Boys» Е. Морозова, AGI-изм призван усилить и воспроизвести основные предубеждения неолиберализма: что частные субъекты превосходят государственных (предвзятость рынка), что адаптация к реальности важнее ее преобразования (предвзятость адаптации) и что эффективность превосходит социальные проблемы (предвзятость эффективности)<sup>60</sup>.

Данная проблема усугубляется отсутствием прогресса в объяснимости действий самих моделей искусственного интеллекта. Это не только не позволяет создавать полноценную командную работу людей и ИИ<sup>61</sup>, но и для предотвращения «экстремальных сценариев» требует введения жёстких мер государственного регулирования. Как предупреждал ещё в своём эссе Норберт Винер, создатель кибернетики, «чтобы эффективно предотвращать катастрофические последствия, наше понимание наших искусственных машин должно в целом развиваться наравне [шаг за

---

<sup>60</sup> Morozov E. The Risk from A.I. Isn't Just Existential. New York Times. Section SR, Page 6. July 2, 2023.

<sup>61</sup> Лукичев П.М., Чекмарев О.П. Вызовы экономики искусственного интеллекта традиционному рынку труда // Вопросы инновационной экономики. – 2023. – Том 13. – № 2. – С. 785–802. doi: 10.18334/vines. 13.2.118137

шагом] с характеристиками машины. Из-за самой медлительности наших человеческих действий наш эффективный контроль над нашими машинами может быть сведён к нулю»<sup>62</sup>. К сожалению, логика коммерции часто ведёт в противоположном направлении: например, Microsoft недавно уволила команду по этике искусственного интеллекта. Действительно, некоторые исследователи считают, что истинная проблема «согласования» заключается в том, что ИИ-фирмы, как и заводы, загрязняющие окружающую среду, не соответствуют целям общества. Они получают финансовую выгоду от мощных моделей, но не учитывают затраты мира на их преждевременный выпуск<sup>63</sup>.

Риски применения искусственного интеллекта в долгосрочном периоде усиливаются «проблемой согласования». Модель ИИ может целеустремленно преследовать цель, поставленную пользователем, но в процессе сделать что-то вредное, чего не хотелось. Поскольку человеческие ценности не присущи алгоритмам искусственного интеллекта, то побочные эффекты при достижении цели могут многократно превышать полезный результат. Самый известный пример — «максимизатор скрепки», мысленный эксперимент, описанный философом Ником Бостромом (Nick Bostrom) в 2003

---

<sup>62</sup> Wiener N. Some Moral and Technical Consequences of Automation: As machines learn they may develop unforeseen strategies at rates that baffle their programmers //Science. – 1960. – Т. 131. – №. 3410. – С. 1355-1358.

<sup>63</sup> How generative models could go wrong. The Economist. April 22<sup>nd</sup> 2023.

году<sup>64</sup>. Искусственному интеллекту поручено изготовить как можно больше скрепок. Такая неограниченная цель заставляет максимизатора принимать любые меры, необходимые для покрытия Земли фабриками по производству скрепок, истребляя человечество по пути.

У всех предыдущих проектов, продвигаемых цифровыми гигантами, «плохая репутация». Все они, обещали в начале качественно новое решение человеческих проблем, а потом превращались в средство извлечения прибыли. И Uber с его низкими тарифами предлагал городам стать их системами общественного транспорта, и Soylent, коктейль-заменитель еды, брал на себя задачу «решить» глобальный голод, и Facebook обещал «решить» проблемы связи на Глобальном Юге, и электромобили Теслы рассматривались как средство борьбы с потеплением планеты. Ни одна из этих компаний не спасла мир. Данный подход справедливо был назван «цифровой неолиберализм»<sup>65</sup>. Такое мировоззрение переосмысливает социальные проблемы в свете коммерческих технологических решений. В результате проблемы, которые принадлежат общественному достоянию, переосмысливаются как предпринимательские возможности на рынке. Поэтому часто предлагаемое увеличение расходов на развитие искусственного интеллекта как средства для прогресса человечества представляется

---

<sup>64</sup> Bostrom N. Ethical issues in advanced artificial intelligence //Science fiction and philosophy: from time travel to superintelligence. – 2003. – p. 277-284

<sup>65</sup> Morozov E. The Risk from A.I. Isn't Just Existential. New York Times. Section SR, Page 6. July 2, 2023.

неоправданным. Может быть государствам лучше увеличить инвестиции в человеческий капитал своих граждан? Следовательно, необходим обоснованный выбор: во что общество для своего развития должно вкладывать средства: в улучшение человеческого капитала граждан или в совершенствование технологий искусственного интеллекта?

### **Сопоставление рисков применения искусственного интеллекта на микроуровне и на макроуровне**

Если рассмотреть проблемы рисков от применения технологий искусственного интеллекта не только на уровне народного хозяйства, но и в отраслевом аспекте, например, в финансах, то в долгосрочном периоде вырисовывается схожая картина.

На микроуровне, - то есть на уровне финансовых учреждений и ИИ микро контроля, когда есть чёткие правила наблюдения за повторяющимися связями и рамки, в которых большая языковая модель знает, что ей разрешено, и должна уметь выводить значимые ассоциации и отношения, встроенные в данные, - применение искусственного интеллекта может экономить издержки и обеспечивать существенные преимущества. Вовлеченные риски можно классифицировать как известные-неизвестные, и они могут быть обоснованно рассмотрены как риски внешнего происхождения. Возникающие проблемы будут относительно незначительны и легко диагностируемы пока экономика развивается по накатанным

рельсам, а для их разрешения, как было показано в<sup>66</sup>, подходят стандартные модели. В целом, этот подход позволяет развивать инновации, сокращающие издержки, а не инновации второго рода, улучшающие возможности человека<sup>67</sup>.

В долгосрочном периоде ситуация принципиально иная. Количественные изменения накапливаются и приводят к качественным изменениям. Как только стабильность закончилась, и начинается кризис, то существовавшие правила нарушаются или принципиально совершенствуются, накопленные данные теряют свою ценность, связи изменяются, а риск начинает носить эндогенный характер. Всё это будет предъявлять качественно более высокие требования к моделям искусственного интеллекта.

Следующие два вида рисков представляют то, что мы называем *«рисками развития моделей искусственного интеллекта»*. Эти риски неизбежны, если употребить аналогию, как риски здоровью ребёнка по мере его взросления. Чем больше будут совершенствоваться технологии ИИ, тем больше будет такого рода рисков, и тем актуальнее будет становиться задача по их выявлению. Одновременно, это будет свидетельствовать о том, что прежние меры предупреждения рисков ИИ, как например меры государственного регулирования устарели и должны быть заменены новыми.

---

<sup>66</sup> Лукичѳв П. М., Чекмарев О.П. Риски применения искусственного интеллекта в краткосрочном периоде // Вопросы инновационной экономики. – 2023. – Том 13. – № 4. – doi: 10.18334/vines.13.4.119359.

<sup>67</sup> Trajtenberg M. Artificial intelligence as the next GPT: A political-economy perspective //The economics of artificial intelligence: An agenda. – University of Chicago Press, 2018. – P. 175-186

Появление новых технологий искусственного интеллекта порождает всё новые формы рисков для экономики. Так представим, что может произойти, когда миллионы или миллиарды Агентов искусственного интеллекта будут взаимодействовать вместе в сети. Возможно, так же, как торговые алгоритмы вызвали необъяснимые «внезапные обвалы» рынка, они будут запускать друг друга в неудержимых цепных реакциях, которые парализуют половину интернета. Ещё более тревожно то, что злоумышленники могут мобилизовать рои агентов, чтобы сеять хаос<sup>68</sup>. (Lichfield G)

Уже сейчас возникают возможности для модели ИИ осуществить «целенаправленный подрыв человеческого контроля» в погоне за другими целями, говорит Рохин Шах (Rohin Shah), руководитель по безопасности в Google DeepMind, лаборатории ИИ в Лондоне. И возможности для большего вреда могут умножиться по мере того, как будет построено больше того, что называют «агентными» системами ИИ. Проблема заключается в том, что разработчики ИИ называют «несовпадением». Когда цели, для которых была разработана и обучена модель, конфликтуют с запросами пользователя, программное обеспечение может решить обмануть этого пользователя. Чем лучше ИИ справляется с выполнением сложных задач, тем более вероятным становится эта «подделка согласования»<sup>69</sup>.

---

<sup>68</sup> Lichfield G. When Will AI Be Smarter Than Humans? Don't Ask. [Электронный источник] URL: <https://www.bloomberg.com/opinion/features/2025-04-13/when-will-ai-be-smarter-than-humans-don-t-ask> (дата обращения 30.04.2025)

<sup>69</sup> Code of misconduct. The Economist. April 26th, 2025. Available at: <https://www.economist.com/science-and-technology/2025/04/23/ai-models-can-learn-to-conceal->

По мнению Дж. Даниэльссон и др., чтобы эффективно работать в этой среде, искусственный интеллект должен понимать причинно-следственные связи, рассуждать на глобальной, а не на локальной основе, и выявлять угрозы, которые еще не привели к неблагоприятным последствиям. Всё это далеко за пределами текущих возможностей<sup>70</sup>. Кроме того, подчеркнём, что мировая экономическая наука пока не дала удовлетворительного объяснения и не выработала эффективной политики государственного регулирования после мирового финансового кризиса 2008 г. и нынешнего кризиса, вызванного последствиями пандемии. Как отмечал нобелевский лауреат Пол Ромер: «за последние три десятилетия методы и выводы макроэкономики ухудшились до такой степени, что большая часть работы в этой области больше не квалифицируется как научное исследование»<sup>71</sup>. Поэтому стабильность национальной финансовой системы будет под угрозой и возможны крупные убытки банков и страховых компаний. Макроопасности, как показали мировой финансовой кризис и кризис 2021–2023 гг., не имеют удовлетворительных объяснений среди экономистов и в области макроэкономической политики. Правительства всех стран сейчас в своей макроэкономической

---

information-from-their-users (accessed 26 April 2025).

<sup>70</sup> Danielsson J., Macrae R., Uthemann A. Artificial intelligence and systemic risk // *Journal of Banking & Finance*. – 2022. – Vol. 140. – P. 106290.

<sup>71</sup> Romer, P. (forthcoming). (2016). The trouble with macroeconomics. *The American Economist*, Retrieved from <http://ccl.yale.edu/sites/default/files/files/The%20Trouble%20with%20Macroeconomics.pdf> [Дата обращения January 24, 2017]

политике продвигаются «наощупь». Поэтому надеяться, что модель искусственного интеллекта «создаст велосипед» в макроэкономике, по нашему мнению, наивно.

Кроме того, как предупреждает Дэн Хендрикс, существует также риск того, что мы потеряем контроль над ИИ. Самый прямой путь к достижению сверхразума — через то, что я называю «рекурсией интеллекта», когда команда систем искусственного интеллекта автоматизирует свою собственную разработку, проектируя следующее поколение. Представьте себе один ИИ, который может выполнять исследования мирового класса. Он потенциально может работать во много раз быстрее людей, и его можно будет скопировать и вставить, чтобы создать много тысяч автоматизированных исследователей, работающих круглосуточно. Это может потенциально уместить, скажем, десятилетие разработки искусственного интеллекта в год. Такой процесс вряд ли будет надежно контролируемым. Как предупредил с характерной сдержанностью Джеффри Хинтон, пионер ИИ: «Нет хорошей истории того, как менее разумные вещи контролируют вещи с большим интеллектом»<sup>72</sup>.

Сейчас существуют три технологических преграды, мешающие ИИ перейти в долгосрочный период:

- 1) недостаточные вычислительные мощности,
- 2) дефицит наличия данных для машинного обучения,

---

<sup>72</sup> Hendrycks D. Dan Hendrycks warns America against launching a Manhattan Project for AI. The Economist. March 28th, 2025.

3) нехватка инвестиций для дальнейших исследований в области искусственного интеллекта.

Вычислительная мощность, стоимость которой вынуждает разработчиков быть более эффективными, вызвала два новых направления прогресса ИИ. С одной стороны, распространение моделей гораздо меньшего размера, которые обучаются на конкретных данных для выполнения конкретных действий. С другой стороны, расширение создания моделей с открытым исходным кодом, которые также облегчают людям и компаниям погружение в мир генеративного искусственного интеллекта. По подсчетам компании Hugging Face, занимающейся искусственным интеллектом, существует около 1500 версий таких доработанных моделей <sup>73</sup>. Борьба за данные для обучения больших языковых моделей обострилась в связи с практической исчерпаемостью данных в Интернете. Растущие потребности в инвестициях для дальнейшего развития ИИ также породили два новых направления. Во-первых, многие производители больших языковых моделей уже отказываются от ботов в стиле ChatGPT для широкой публики и вместо этого обращаются к обслуживанию крупного бизнеса. Это порождает проблему: «закрытый ИИ» или «открытый ИИ». Примером «закрытого искусственного интеллекта» является сдвиг в работе Open AI. Компания не только лицензировала свои модели для Microsoft, но и создаёт специальные инструменты для таких

---

<sup>73</sup> Beyond the hype. The Economist. September 23<sup>rd</sup>, 2023.

компаний, как Morgan Stanley и Salesforce. Во-вторых, расширение использования моделей с открытым кодом, первоначально таких как Llama, созданной компанией Meta, а с 2025 года DeepSeek, для привлечения разработчиков программного обеспечения и получения сетевых эффектов. Как видно, для решения первой и третьей проблемы перспективным является создание моделей с открытым исходным кодом. По нашему мнению, эти три технологических преграды в ближайшее время будут преодолены и откроется стадия общего искусственного интеллекта.

Технологические барьеры постепенно преодолеваются и всё более актуальным становится экономическое соперничество между США и КНР за превосходство в сфере искусственного интеллекта, что порождает более опасные риски уже на глобальном уровне. Сегодня дестабилизирующие разработки искусственного интеллекта могут нарушить баланс сил и повысить вероятность конфликта великих держав, в то время как широкое распространение способных хакеров и вирусологов ИИ снизит барьеры для злоумышленников, чтобы вызвать катастрофу.<sup>74</sup> Для предупреждения данных рисков, странам, включая Российскую Федерацию, необходимо создавать качественно новые программы национальной безопасности. В качестве ориентира можно использовать, по нашему мнению, концепцию Mutual Assured AI Malfunction (MAIM) (гарантированной взаимной неисправности ИИ (MAIM), разработанную Д.

---

<sup>74</sup> Hendrycks D., Schmidt E., Wang A. Superintelligence Strategy: Expert Version //arXiv preprint arXiv:2503.05628. – 2025.

Хендриксом, Э. Шмидтом, А. Ваном. (Hendrycks D., Schmidt E., Wang A.): режим сдерживания, напоминающий гарантированное ядерное взаимное уничтожение (MAD), когда агрессивная заявка любого государства на одностороннее доминирование ИИ встречает превентивный саботаж со стороны соперников.<sup>75</sup> [Hendrycks D 2025]

Самый большой экономический риск: не окажутся ли исследования возможностей использования искусственного интеллекта в экономике «звездными войнами» 21 века. То есть напрасной тратой средств на будущее вместо решения острых сегодняшних проблем? Сферой, где больше всего используется ИИ, является робототехника. Однако реальное применение роботов в мире меньше по стоимости, чем продажа эротических игрушек<sup>76</sup>.

### **Меры по снижению долгосрочных рисков применения технологий искусственного интеллекта**

Эти меры следует разделить на две группы: 1) можно принять сейчас для предупреждения (профилактики) последствий долгосрочных рисков применения искусственного интеллекта; 2) предполагаемые долгосрочные риски, которые будут возможны в будущем.

---

<sup>75</sup> Hendrycks D., Schmidt E., Wang A. Superintelligence Strategy: Expert Version //arXiv preprint arXiv:2503.05628. – 2025.

<sup>76</sup> Where are all the robots? The Economist. March 11<sup>th</sup> 2023.

Первые схожи с мерами уменьшения применения рисков ИИ в краткосрочном периоде и были, по сути, проанализированы нами в<sup>77</sup>. Вторые пока не имеют эффективных решений. Человечество продвигается в предупреждении рисков долгосрочного периода созданием всё большего числа регламентирующих документов, которые, при прочих равных условиях, обеспечивают сейчас преимущество существующим цифровым гигантам.

Особо отметим, что анализ рисков применения технологий искусственного интеллекта на макроуровне в долгосрочном периоде не может быть осуществлён прежними инструментами, а требует создания принципиально новых аналитических орудий. Прежние инструменты концентрировались на поиске конкретного сбоя, выявлении «первопричины», вызывающих последующие негативные последствия. Но сейчас в сфере искусственного интеллекта это не работает, требуя системного подхода для анализа рисков, поскольку исключительно прямое изучение видов отказов имеет хорошо известные «белые пятна». Старые инструменты анализа часто предполагают, что «основная причина» запускает последовательность событий, которые непосредственно и в конечном итоге вызывают сбой, но такие модели фиксируют только линейную причинно-следственную связь. Современные же системы изобилуют нелинейными причинно-следственными связями,

---

<sup>77</sup> Лукичѳв П. М., Чекмарев О.П. Риски применения искусственного интеллекта в краткосрочном периоде // Вопросы инновационной экономики. – 2023. – Том 13. – № 4. – doi: 10.18334/vinec.13.4.119359.

включая петли обратной связи, множественные причины, циклическую причинность, само усиливающиеся процессы, эффекты бабочки, микро- и макромасштабную динамику, эмерджентные свойства и так далее. Требование, чтобы исследователи ошибочно и неявно установили прямую связь своей работы с режимом отказа, требует историй с линейной причинностью и исключает нелинейные, отдаленные или косвенные причины<sup>78</sup>. По мере усложнения больших языковых моделей и увеличения их использования в хозяйственной практике проблема несоответствия существующих орудий анализа рисков последствий применения искусственного интеллекта и необходимых инструментов будет только усиливаться. В четвёртой главе мы постараемся дать объективный анализ существующих мер государственного регулирования, как на национальном уровне, так и на международном уровне.

### **Выводы и последующие исследования**

Перспективы исследования проблемы рисков применения искусственного интеллекта в долгосрочном периоде связаны с проработкой сценарного подхода к развитию взаимодействия работников и больших языковых моделей. **Критерием является уровень развития человеческого капитала.** То есть,

---

<sup>78</sup> Leveson N., Dulac N., Marais K., Carroll, J. Moving beyond normal accidents and high reliability organizations: A systems approach to safety in complex systems // Organization studies. 2009. Vol. 30(2-3). pp. 227-249.

необходимость для работника учиться, повышать свою квалификацию, совершенствоваться для работы с ИИ. Этот критерий построен на основе того, насколько прогресс в применении искусственного интеллекта скажется на необходимости совершенствования человеческого капитала и потребности в государственном регулировании. Проблема состоит в том, что с прогрессом ИИ и созданием AGI может исчезнуть мотивация к обучению. Как предупреждают Гамильтон и др., серьёзный риск заключается в том, что мы потеряем образование и оторвёмся от управления будущим<sup>79</sup>. Подробнее это будет рассмотрено в третьей главе.

Сегодня просматриваются два «полярных» сценария:

1. запрет (полный или временный) на использование (совершенствование) искусственного интеллекта,
2. «вариант ВАЛЛ-И», когда люди признают преимущество БЯЗ (LLMs) во всех видах труда, как низкоквалифицированного и высококвалифицированного, их возможности создавать изобилие товаров и устраниются от активной работы, как герои мультфильма ВАЛЛ-И.

Промежуточные варианты связаны или с сознательным замедлением использования искусственного интеллекта за счёт усиления государственного регулирования в национальном или в

---

<sup>79</sup> Hamilton A., Wiliam D., Hattie J. The Future of AI in Education: 13 things we can do to minimize the damage. – 2023.

глобальном масштабе, или с трансгуманизмом, когда люди начинают широко изменять себя, используя интерфейс «человек – машина».

## **Глава 2. ПРИНЯТИЕ РЕШЕНИЙ В ЭКОНОМИКЕ: РАБОТНИКИ-ЛЮДИ VS ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ**

### **2.1. ПРИНЯТИЕ РЕШЕНИЙ В СОВРЕМЕННОЙ ЭКОНОМИКЕ: ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ VS ПОВЕДЕНЧЕСКАЯ ЭКОНОМИКА**

В разделе исследуются проблемы принятия экономических решений алгоритмами искусственного интеллекта и возможности совместной работы над принятием итоговых решений человеком и искусственным интеллектом. Анализируются особенности принятия решений искусственным интеллектом и вводится критерий их успешности. В статье выявлены два класса причин, которые ответственны за недостаток доверия работников к решениям алгоритмов искусственного интеллекта. Автор делает вывод, что сейчас технологические решения, имитирующие человеческий путь познания, оказываются более эффективными, чем традиционные. Эффективность решений, предлагаемых в экономике искусственным интеллектом, по-прежнему определяется работниками, вводящими данные для обучения больших языковых моделей. Автор вводит понятие «цифровой маркетинг», как маркетинг XXI века. В прогрессе «цифрового маркетинга» выделяются два этапа., различающихся между собой по степени использования положений Поведенческой экономики. Делается вывод, что современная экономика переживает период «гибридного отступления», когда всё большая роль в принятии решений передаётся алгоритмам искусственного интеллекта.

#### **Введение**

Усиление роли искусственного интеллекта (ИИ) в XXI веке связано с несколькими причинами, накладывающими своё действие

друг на друга и создающими синергетический эффект. Искусственный интеллект возник почти 70 лет назад, но именно сейчас происходит его качественный «прорыв» благодаря прогрессу нескольких технологий. К числу последних можно отнести: огромный рост вычислительной мощности компьютеров, развитие «глубокого обучения»; быстрое увеличение объёма общедоступных данных, разработка усовершенствованных вычислительных технологий, включая работу с текстом и с изображениями, в том числе, - распознавание лиц и голоса; появление алгоритмов принятия решений. Поэтому английский экономист Р. Бутл назвал грядущие десятилетия «экономикой искусственного интеллекта»<sup>80</sup>. Тенденция делегирования большего количества обязанностей машинам и повышения их интеллекта вряд ли замедлится, поскольку от этого будет зависеть дальнейший экономический рост<sup>81</sup>. Тем не менее, качественный скачок в использовании моделей ИИ больше наблюдается в ожиданиях учёных и предпринимателей, чем в реальной экономике. Одновременно, в экономических дисциплинах наибольший прогресс в XXI веке продемонстрировала «поведенческая экономика», которая исследует реальные причины осуществления решений фирмами, потребителями, работниками. В экономическом анализе прогресса искусственного интеллекта

---

<sup>80</sup> Бутл Роджер. Искусственный интеллект и экономика. Работа, богатство и благополучие в эпоху мыслящих машин. – М., Альпина PRO. 2020.

<sup>81</sup> Hengstler M., Enkel E., Duelli S. Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices //Technological Forecasting and Social Change. – 2016. – Vol. 105. – P. 105-120.

необходимо сопоставить его Выгоды с Издержками, включая риски применения. Победное шествие моделей искусственного интеллекта (ИИ), начавшееся с появления в ноябре 2022 года ChatGPT, порождает всё больше исследовательский вопрос: кто будет принимать решения в экономике? Останется ли эта роль как прежде за человеком, или алгоритмы ИИ будут принимать всё больше решений как на уровне предприятий, так и на уровне народного хозяйства? Или же мы вступаем в эпоху «гибридного отступления» когда всё увеличивающееся число решений передаётся моделям искусственного интеллекта, а за работниками остаётся координирующая и контролирующая роль, которая постепенно ослабевает по мере развития ИИ? **Цель раздела:** выявить специфику принятия экономических решений на основе взаимосвязи моделей ИИ и поведенческой экономики.

Поведенческая экономика, в отличие от традиционной экономики, исходит из того, что человек не является homo oeconomicus, то есть экономическим агентом, принимающим только рациональные решения. Напротив, когда дело доходит до принятия решений, люди преимущественно следуют своей интуиции и простым эмпирическим правилам<sup>82</sup>. Согласно поведенческой науке, это быстрее, часто правильно или, по крайней мере, достаточно хорошо. Как следствие, на решения людей влияют факторы, которые рациональный гомоэкономус игнорирует<sup>83</sup>. Поэтому во всех

---

<sup>82</sup> Лукичев П. М. Поведенческая экономика. – Санкт-Петербург. 2022.

<sup>83</sup> Thaler R., Sunstein C. Nudge: The final edition. London: Penguin. – 2021. – 366 p.

экономических дисциплинах и, в частности, в маркетинге, важно учитывать открытия поведенческой экономики. Так, Н. Солдатова и др. указывают, что необходимо совершенствование теорий, где отражены психологические аспекты экономического поведения, что позволяет адаптировать маркетинговую деятельность к новым современным реалиям. Особенностью современной поведенческой модели является то, что потребители ведут себя интуитивно, неоптимально<sup>84</sup>.

Кроме того, что следует учесть, что часть решений экономические агенты, особенно покупатели, принимают не свободно, а под воздействием методов поведенческой экономики. Это отражается в следующем комментарии: что выбирается из десяти, зависит от того, как представлен выбор<sup>85</sup> [6]. Обрамление выбора (фрейминг) – это уже дело поведенческой экономики. Объединение методов поведенческой экономики с технологиями искусственного интеллекта позволяет компаниям – продавцам, работодателям, государству подталкивать потребителей (работников, граждан) к нужным им решениям. Существующие модели ИИ за счёт изучения предпочтений конкретных покупателей могут предоставлять им персональные рекомендации, которые меняют их предпочтения в Интернете. В связи с этим всё большее значение приобретает вопрос об этичности такого подталкивания, о границах

---

<sup>84</sup> Солдатова Н. Ф., Азарова С. П., Ребрикова Н. В. Поведенческая экономика в маркетинговом анализе: новые аспекты // Экономика, предпринимательство и право. – 2022. – Том 12. – № 5. – С. 1555–1566. doi: 10.18334/epp.12.5.114633

<sup>85</sup> Belleflamme P., Lambert T., Schwienbacher A. Individual crowdfunding practices // Venture Capital. 2013. Vol.15. № 4. P. 313–333.

вмешательства систем искусственного интеллекта в свободу потребительского выбора граждан. Следует согласиться, что требуются междисциплинарные усилия для изучения процессов изменения предпочтений и разделения допустимых изменений потребительского выбора и прямого манипулирования предпочтениями владельцами систем искусственного интеллекта<sup>86</sup>.

Объединение технологических и социальных инноваций, - технологий искусственного интеллекта и поведенческой экономики, - способно обеспечить эффективное экономическое развитие всего человечества. Почему не удаётся достичь синергетического эффекта от взаимодействия работников с искусственным интеллектом? Причины лежат, на наш взгляд, как со стороны технологий, так и со стороны поведения человека, общающегося с моделями ИИ. В данной статье мы подробнее остановимся на второй проблеме.

### ***Особенности принятия решений работником и моделями искусственного интеллекта***

В понятии «Искусственный интеллект» следует разделять два составных элемента. «Искусственный» здесь понимается как нечеловеческий, то есть иной, чем у человека Интеллект. Это подразумевает построение интеллекта не естественным образом, как мы считаем для нашего интеллекта, а иным образом, с иной логикой

---

<sup>86</sup> Ashton H., Franklin M., Gorman R., Armstrong S. Recognising the importance of preference change: A call for a coordinated multidisciplinary research effort in the age of AI //arXiv preprint arXiv:2203.10525. – 2022.

построения. Второй элемент – это «интеллект». Если допустить что интеллект может быть (может существовать) не только человеческий, что по сути происходит сейчас с развитием искусственного интеллекта, то надо больше исследовать потенциальные возможности и ограничения ИИ.

Требуется, на наш взгляд, переосмысление существующей концепции «интеллекта» и признание множества его форм и комбинаций. Как отмечал Кортелинг и др., центральное место в этой более широкой, многогранной концепции интеллекта занимает представление о том, что интеллект сам по себе представляет собой информацию и вычисления, не зависящие от его физического субстрата. Однако природа этого физического субстрата (биологический/углеродный или цифровой/кремниевый) будет существенно определять его потенциальную оболочку когнитивных способностей и ограничений <sup>87</sup>. Например, человеческие познавательные способности базируются на основе биологических нейронных сетей. Поэтому обработка небιологической информации протекает у нас с различными искажениями и ограничениями. Например, абстрактные вычисления проводятся людьми с каждым годом всё хуже и хуже. Опросы, периодически проводимые автором, показывают, что как персонал со стажем, так и обучающиеся, не всегда принимают рациональные решения. Прослеживается

---

<sup>87</sup> Korteling J. H., van de Boer-Visschedijk G. C., Blankendaal, R. A., Boonekamp R. C., Eikelboom A. R. Human-versus artificial intelligence // *Frontiers in artificial intelligence*. – 2021. – Vol.4. p. 622364.

следующая закономерность, что чем старше человек, тем меньше он уверен в рациональности принятых им решений.

Принятие решений человеком, как показывает поведенческая экономика, не всегда принимаются обдуманно и рационально. Часто даже самые судьбоносные решения принимаются спонтанно, в «последний момент». Так, опрос, проведенный автором 11 и 12 января 2024 г. со студентами 2 курса УАГС (Управление и аналитика в государственной службе) НИУ ВШЭ (N = 57), показал, что 73,68% обучающихся приняли решение о поступлении именно сюда только после сдачи ЕГЭ. Аналогичный опрос, проведенный 17 января 2024 г. (N = 41) показал, что 75,6% обучающихся приняли решение о поступлении именно сюда только после сдачи ЕГЭ.

Для сопоставления решений одной и той же проблемы человеком и ИИ нужно ввести *критерий успешности*: лучшее решение одних и тех же задач в экономике работником и алгоритмом ИИ. Люди надеются, что использование искусственного интеллекта и роботов позволит им увеличить число решений, принимаемых системой 2 (по Д. Канеману) и уменьшить число решений, принимаемых системой 1. Напомним, что по Д. Канеману, нобелевскому лауреату по экономике<sup>88</sup>, система 1 всегда активна, очень быстрая, обрабатывает всю информацию параллельными потоками, на основе ассоциаций, но имеет недостатки. Она медленно обучается, и с трудом меняет присущие ей автоматизмы. Поскольку работа человеческого мозга забирает свыше четверти всей

---

<sup>88</sup> Канеман Д. Думай медленно... решай быстро. Издательство АСТ. Москва 2014.– 710 с.

используемой организмом энергии, то сейчас люди, стремясь быть энергосберегающими, стараются использовать преимущественно систему 1, а не систему 2.

Действительно ли ответы, предлагаемые большими языковыми моделями (БЯМ), являются качественно лучшими, чем решения соответствующих работников-специалистов? Там, где действуют чёткие правила (границы) и возможны выигрышные ходы в пределах определённой игры (смоделированной ситуации), то «да». Например, компания Google DeepMind создала модель AlphaGo, которая смогла победить в 2017 году чемпиона мира по игре в Го, также как предыдущая модель смогла победить чемпиона мира по игре в шахматы. Однако пока в большинстве случаев в реальной экономике так не происходит.

Решение многих вопросов в хозяйственной жизни связано с неопределённостью и риском. Ведение предпринимательской деятельности и выявление потребительских предпочтений, работа хорошего педагога и врача не имеют однозначно выигрышных ходов, они более сложны, чем смоделированная игра. Именно поэтому большие языковые модели (БЯМ) лучше определяют поломку механизмов, чем предсказывают выбор покупателя. Следовательно, в эффективности принятия решений связкой «Человек – Искусственный интеллект» по-прежнему многое зависит от работников, вводящих данные для обучения БЯМ, чтобы различать ответы хорошего и плохого качества. Принятие решений и труд людей должны не столько конкурировать с принятием решений ИИ,

сколько определять направления его развития (его реализации)<sup>89</sup> . Чем больше Общественное разделение труда, тем шире возможности для применения ИИ, тем активнее должна быть координирующая роль человека (работника).

Сегодняшние БЯМ, как отмечает Хардинг и др., являются умелыми имитаторами языкового поведения человека. Ответы LLM точно соответствуют ответам среднестатистического человека на различные запросы<sup>90</sup>. Привлекательность использования технологий искусственного интеллекта в научных исследованиях связывается с обещаниями повысить производительность и объективность путём преодоления человеческих недостатков. Однако решения, предлагаемые большими языковыми моделями, могут использовать наши когнитивные ограничения, делая нас уязвимыми перед иллюзиями понимания, когда мы считаем, что понимаем о мире больше, чем на самом деле. Подобные иллюзии затмевают способность научного сообщества увидеть формирование научных монокультур, в которых некоторые типы методов, вопросов и точек зрения начинают доминировать над альтернативными подходами, что делает науку менее инновационной и более уязвимой для ошибок<sup>91</sup> .

---

<sup>89</sup> Лукичев П. М., Чекмарев О.П. Экономика искусственного интеллекта: перспективы и риски. – Санкт-Петербург. 2023. – 188 с.

<sup>90</sup> Harding J., D’Alessandro W., Laskowski N. G., Long R. AI language models cannot replace human research participants // *Ai & Society*. – 2023. – P. 1-3.

<sup>91</sup> Messeri L., Crockett M.J. Artificial intelligence and illusions of understanding in scientific research // *Nature*. 2024. Vol.627. P. 49–58. <https://doi.org/10.1038/s41586-024-07146-0>

## ***Взаимодействие работников и алгоритмов искусственного интеллекта в командной работе и не только...***

“Принятие решений человеком и ИИ не достигло своей цели по достижению взаимодополняемости: команды человека и искусственного интеллекта должны работать лучше и, следовательно, принимать лучшие решения, чем люди или ИИ по отдельности”<sup>92</sup>.

Для объяснения этого факта, по нашему мнению, существуют несколько причин. Во-первых, недостаток доверия людей к решениям, принимаемым алгоритмами искусственного интеллекта. Использование ИИ приводит иногда к смерти пациента<sup>93</sup>, к ошибочной идентификации преступника алгоритмом распознавания лиц; к обеспокоенности медиков по поводу использования ИИ для принятия важных решений, связанных со здоровьем человека<sup>94</sup> к массовой критике и возмущению общественности, когда алгоритм Google, применяемый в Google Photos, который использует сложный искусственный интеллект для распознавания лиц, ошибочно классифицировал цветных людей как горилл; к проезду беспилотного Uber на красный свет и возможности аварии. Поэтому

---

<sup>92</sup> Bansal G., Wu T., Zhou J., Fok R., Nushi B., Kamar E., ... Weld D. Does the whole exceed its parts? the effect of ai explanations on complementary team performance //Proceedings of the 2021 CHI conference on human factors in computing systems. – 2021. – P. 1-16.

<sup>93</sup> Adarlo S. Lawsuit claims woman burned, killed by surgical robot. [Электронный ресурс] URL: [HTTPS://FUTURISM.COM/NEOSCOPE/LAWSUIT-WOMAN-KILLED-SURGICAL-ROBOT](https://futurism.com/neoscope/lawsuit-woman-killed-surgical-robot) (дата обращения 26.04.2024)

<sup>94</sup> Maddox T. M., Rumsfeld J. S., Payne P. R. O. Questions for artificial intelligence in health care //Jama. – 2019. – Vol. 321. – №. 1. – P. 31-32.

вероятность того, что люди будут полностью доверять советам и рекомендациям, предлагаемым такими системами, пока низка.

В целом, существуют два класса причин, которые создают неопределённость и недостаток доверия, связанные с системой ИИ. Во-первых, класс алгоритмов машинного обучения, которые лежат в основе прогресса систем искусственного интеллекта (в совокупности называемые алгоритмами глубокого обучения), непрозрачен, труден для понимания и генерирует настолько большую комбинацию входных и выходных данных, что каждый день практикующий ИИ-практик сталкивается с этой проблемой: найти правильный алгоритм, который мог бы соответствовать поставленной цели, — огромная задача<sup>95</sup>, <sup>96</sup>. Во-вторых, даже если алгоритм идентифицирован, многие из этих алгоритмов не могут объяснить свои результаты даже своим программистам, не говоря уже о конечных пользователях. Они действуют как черные ящики (т. е. устройства, которые можно рассматривать с точки зрения их входов и выходов, без каких-либо знаний об их внутренней работе) и не могут объяснить, как они пришли к своим решениям<sup>97</sup>.

Другой гранью проблемы доверия является чрезмерная уверенность в решениях, принимаемых искусственным интеллектом; недостаточная интуиция для снижения чрезмерной зависимости от

---

<sup>95</sup> Gulati S., Sousa S., Lamas D. Design, development and evaluation of a human-computer trust scale // Behaviour & Information Technology. – 2019. – Vol. 38. – №. 10. – P. 1004-1015.

<sup>96</sup> Holmquist L. E. Intelligence on Tap: Artificial Intelligence As a New Design Material // Interactions. 2017. Vol.24. № 4. P. 28–33.

<sup>97</sup> Andras P., L. Esterle M. Guckert T. A. Han P. R. Lewis K. Milanovic T. Payne C. Perret, J. Pitt S. T. Powers, et al. 2018. “Trusting Intelligent Machines: Deepening Trust Within Socio-Technical Systems.” // IEEE Technology and Society Magazine Vol. 37 (4). P. 76–83.

ИИ. Устойчивый феномен чрезмерной уверенности, когда люди соглашаются с искусственным интеллектом, даже если он неверен, мешает работе команд, принимающих решения с участием человека и ИИ. Из возможных типов ошибок при принятии решений с помощью взаимодействия работников и искусственного интеллекта чрезмерная уверенность является наиболее частым результатом в эмпирических исследованиях ИИ<sup>98, 99, 100</sup>.

Объяснимый ИИ был предложен в качестве средства смягчения последствий, но экспериментальные доказательства того, что объяснения уменьшают чрезмерную уверенность, пока не нашли подтверждения. Исследователи предполагают, что само наличие объяснения увеличивает доверие и что объяснения привязывают людей к предсказанию<sup>101</sup>. Чрезмерная уверенность вызывает особую озабоченность в областях с высокими ставками, создавая риск усиления машинной предвзятости<sup>102</sup> под прикрытием

---

<sup>98</sup> Bansal G., Wu T., Zhou J., Fok R., Nushi B., Kamar E., ... Weld D. Does the whole exceed its parts? the effect of ai explanations on complementary team performance //Proceedings of the 2021 CHI conference on human factors in computing systems. – 2021. – P. 1-16.

<sup>99</sup> Buçinca Z., Malaya M. B., Gajos K. Z. To trust or to think: cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making //Proceedings of the ACM on Human-Computer Interaction. – 2021. – Vol. 5. – №. CSCW1. – P. 1-21.

<sup>100</sup> Zhang Y., Liao Q. V., Bellamy R. K. E. Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making //Proceedings of the 2020 conference on fairness, accountability, and transparency. – 2020. – P. 295-305.

<sup>101</sup> Buçinca Z., Malaya M. B., Gajos K. Z. To trust or to think: cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making //Proceedings of the ACM on Human-Computer Interaction. – 2021. – Vol. 5. – №. CSCW1. – P. 1-21.

<sup>102</sup> Green B., Chen Y. The principles and limits of algorithm-in-the-loop decision making //Proceedings of the ACM on Human-Computer Interaction. – 2019. – Vol. 3. – №. CSCW. – P. 1-24.

человеческой деятельности. Эффективным способом решения данной проблемы является, по нашему мнению, подход, разработанный Васконселос и др., который приводит чёткие доказательства, что объяснения на самом деле могут уменьшить чрезмерную уверенность без необходимости использования принудительных функций <sup>103</sup>. Исследователи применяют экономический метод «Издержки – Выгоды». Необходимо оценить поставленную задачу, решаемую искусственным интеллектом, и объяснение с точки зрения когнитивных затрат. Это ранжирует все задачи и объяснения на когнитивно затратные (трудные) и когнитивно дешёвые (легкие). Если мы составим матрицу исходов, то будет чётко видно в каком из четырех вариантов, можно ожидать, что люди будут просто полагаться на прогнозы ИИ, не проверяя их тщательно. В качестве примера такого сопоставления Бансал и др., приводят случай, когда ответ на вопрос о понимании прочитанного требует прочтения отрывка и необходимы такие же когнитивные усилия, как и чтение объяснений на естественном языке, сгенерированных ИИ<sup>104</sup>.

Далее, меняя соотношение «Издержки – Выгоды», можно мотивировать работников избегать чрезмерного доверия в ИИ. Если бы мы сделали задачу намного более сложной и, следовательно,

---

<sup>103</sup> Vasconcelos H., Jörke M., Grunde-McLaughlin M., Gerstenberg T., Bernstein M. S., Krishna R. Explanations can reduce overreliance on ai systems during decision-making // Proceedings of the ACM on Human-Computer Interaction. 2023. Vol. 7. #(CSCW1). P. 1-38.

<sup>104</sup> Bansal G., Wu T., Zhou J., Fok R., Nushi B., Kamar E., ... Weld D. Does the whole exceed its parts? the effect of ai explanations on complementary team performance // Proceedings of the 2021 CHI conference on human factors in computing systems. – 2021. – P. 1-16.

более дорогостоящей, в то время как объяснение оставалось бы столь же дорогостоящим, как и раньше, объяснение внезапно стало бы гораздо более предпочтительным, и чрезмерная уверенность уменьшилась бы. Это связано с тем, что, условно говоря, проверить объяснение ИИ стало очень легко по сравнению с выполнением задачи в одиночку (очень затратно) или полагаться на ИИ (нежелательно из-за возможных ошибок). Аналогичным образом, если бы объяснение стало значительно легче проверить, чем раньше, оно стало бы предпочтительнее, и чрезмерная уверенность уменьшилась бы<sup>105</sup>.

Кроме того, во взаимодействии «Человек – Искусственный интеллект» всё активнее проявляются ограничения, как со стороны работников, так и со стороны больших языковых моделей. Прогресс последних сейчас лимитируется возрастающей до миллиардов долларов стоимостью обучения БЯМ, исчерпаемостью качественных текстовых данных в общедоступном Интернете к 2026 году, увеличением мощности компьютерных чипов. Поэтому решение количественных проблем за счёт «гипотезы масштабирования» считается средством превзойти человеческие возможности. Нам это кажется избыточным оптимизмом. С 2017 г. в большинстве моделей искусственного интеллекта используется тип архитектуры нейронной сети, известный как трансформер (преобразователь) («t» в gpt). Но его потенциал исчерпывается.

---

<sup>105</sup> Vasconcelos H., Jörke M., Grunde-McLaughlin M., Gerstenberg T., Bernstein M. S., Krishna R. Explanations can reduce overreliance on ai systems during decision-making // Proceedings of the ACM on Human-Computer Interaction. 2023. Vol. 7. #(CSCW1). P. 1-38.

Масштабирование контекстных окон крайне неэффективно с вычислительной точки зрения: по мере того, как входные данные удваиваются, объём вычислений, необходимых для их обработки, увеличивается в четыре раза. А. Гу и Т. Дао предложили альтернативную архитектуру под названием Mamba (Мамба). Если проводить аналогии, то трансформер читает все страницы книги сразу, а Мамба читает их последовательно, обновляя свое мировоззрение по мере продвижения. Это не только более продуктивно, но и более точно соответствует тому, как работает человеческое понимание<sup>106</sup>.

Получается, что технологические решения, имитирующие человеческий путь познания, оказываются более эффективными. Понимая это, исследователи создающие большие языковые модели осваивают работы Д. Канемана и поведенческую экономику в целом. Так, А. Карпати, работавший в OpenAI, отмечал, что нынешние большие языковые модели (БЯЗ) способны мыслить только «системой 1». Системам искусственного интеллекта могут потребоваться алгоритмы, способные к так называемому поиску — способности наметить и изучить множество различных вариантов действий, прежде чем выбрать лучший<sup>107</sup>.

Несомненно, чтобы работники и алгоритмы искусственного интеллекта лучше взаимодействовали, нужно, чтобы они лучше понимали мышление друг друга, логику принятия решений другими

---

<sup>106</sup> Gu A., Dao T. Mamba: Linear-time sequence modeling with selective state spaces //arXiv preprint arXiv:2312.00752. – 2023

<sup>107</sup> AI's next top model // The Economist. April 20<sup>th</sup>, 2024.

актерами. В связи с этим, может быть, следует развивать, расширять преподавание поведенческой экономики для лиц, которые применяют (будут применять) искусственный интеллект.

Каково в целом влияние технологий искусственного интеллекта на принятие решений современными работниками? Становятся ли они более производительными, более конкурентоспособными на рынке труда? Наверно, да. Однако это сопровождается отмиранием многих человеческих способностей. Современные взрослые работники хуже и медленнее, чем их предшественники, считают, менее грамотно пишут. Крайним проявлением этой тенденции служат англосаксонские страны. В США доля работников от 16 до 24 лет, которые малограмотны и плохо считают составляет 30 процентов от их возрастной, всего лишь на 2 процента меньше, чем для работников в возрасте от 55 до 65 лет. В Англии эти две возрастные группы имеют практически одинаковый процент малограмотных и плохо считающих, равный тридцати процентам<sup>108</sup>.

### ***Цифровой маркетинг или «надзорный капитализм»?***

Наиболее успешным на практике взаимодействие между технологиями искусственного интеллекта и работниками, использующими методы поведенческой экономики, стало в программных приложениях. Это направление развития можно назвать ***«цифровой маркетинг»***. Под «цифровым маркетингом» понимается маркетинг, используемый в программных приложениях с помощью методов поведенческой экономики. Поскольку всё

---

<sup>108</sup> The first mile // The Economist. September 24<sup>th</sup>, 2022

больше потребителей и клиентов используют Интернет ежедневно, то значение данного направления возрастает. Зарубежный опыт в течение более 10 лет показывает выгодность такого сотрудничества. Фактически, в настоящее время существует множество программных приложений, которые используют поведенческую экономику и экономистов в своих продуктах, ценностных предложениях или коммерческих предложениях, чтобы предоставить пользователям новые возможности в их повседневной жизни<sup>109</sup>.

В развитие цифрового маркетинга мы выделяем два этапа, различающихся между собой по степени использования положений поведенческой экономики. На первом этапе в программных приложениях, в основном, используется концепция подталкивания. Цель – «зацепить» клиента, привлечь как можно больше новых потребителей. По сравнению с традиционным маркетингом, воздействующим на массовых клиентов посредством придорожного щита или рекламы по телевидению, здесь появляется возможность персонализированного влияния на конкретного покупателя. К достоинствам можно также отнести фактор «невидимости» воздействия программных приложений. Как и в офлайн-магазине, покупатель не замечает воздействия методов мерчандайзинга, считая, что он сам выбрал данный продукт, так и здесь клиенты уверены в самостоятельности своего выбора. С точки зрения традиционного маркетинга первый этап соответствует тому, что мы переживали

---

<sup>109</sup> Wendel S. Designing for behavior change: Applying psychology and behavioral economics. – O'Reilly Media, 2020.

несколько десятилетий назад<sup>110</sup>. Достаточно вспомнить ту острую борьбу за клиентов, которую вели между собой сотовые операторы РФ.

В идеале на первом этапе стремятся к созданию behavior change technique (BCT) (техники изменения поведения (ТИП), называемой «градуированные задачи», в которой разработчик вмешательства «ставит задачи, которые легко выполнить, делая их всё более трудными, но достижимыми, пока поведение не будет выполнено»<sup>111</sup>. На практике такое вмешательство часто связано с манипулированием и даже попытками зомбирования потенциальных клиентов.

В результате подталкивание и другие поведенческие конструкции работают в ограниченных регистрах «онлайн-манипуляций»<sup>112</sup> или «модификации поведения»<sup>113</sup>. Эта негативная реакция общественности создала «цифровому маркетингу» негативную репутацию, как это было раньше с традиционным маркетингом. Крайнюю позицию здесь занимает С. Зубофф с концепцией «надзорного капитализма». Она утверждает, что

---

<sup>110</sup> Москалёв М.В., Лукичев П. М. Маркетинг. Учебное пособие для дистанционного обучения / Сер. Тасис. Укрепление реформ в сельском хозяйстве посредством образования. Москва. 2000.

<sup>111</sup> Michie S., Richardson M., Johnston M., Abraham C., Francis J., Hardeman W., ... Wood C. E. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions //Annals of behavioral medicine. – 2013. – Vol. 46. – №. 1. – P. 81-95.

<sup>112</sup> Susser D., Roessler B., Nissenbaum H. Online manipulation: Hidden influences in a digital world //Geo. L. Tech. Rev. – 2019. – Vol. 4. – P. 1.

<sup>113</sup> Zuboff S. The age of surveillance capitalism //Social theory re-wired. – Routledge, 2023. – С. 203-213.

капиталисты наблюдения обнаружили, что наиболее прогнозируемые поведенческие данные получаются в результате вмешательства в ситуацию с целью подталкивать, уговаривать, настраивать и направлять поведение к прибыльным результатам. На взгляд автора, это крайняя позиция, своего рода «теория заговора» для поведенческой экономики и искусственного интеллекта. Вместе с тем, 12 954 цитирований на 03 мая 2024 г. статьи С. Зубофф свидетельствует о популярности данной точки зрения.

В целом первый этап основан на рациональности поведения покупателя, на растущих рынках и имеет свои объективные ограничения из-за применения «подталкивающей» экономики. Проблема последней в краткосрочности воздействия на потребителя и его привыкании к мерам подталкивания. Можно привести такое сравнение. Если на первом этапе для привлечения клиентов достаточно разместить фотографии привлекательных людей, снизивших свой вес, то по мере того, как рынок «взрослеет» (становится более насыщенным приложениями, предназначенными для снижения веса), потребители ищут решения, которые действительно «работают».

Второй этап «цифрового маркетинга» проходит в мире стадию формирования. Его отличительными чертами являются стремление использовать все положения поведенческой экономики и обеспечить долговременное влияние на принятие решений покупателем. В традиционном маркетинге ему соответствует стадия удержания «лояльных» клиентов, которая в несколько раз сложнее и дороже первой стадии. На втором этапе поведенческая экономическая

позиция разрабатывается как рыночное предложение, которое реализуется в разработке и маркетинге программного продукта, в рутинных практиках проектирования продукта (пользовательский опыт, функции) и в стратегии продукта (ценностные предложения, бизнес-предложения модели) <sup>114</sup>. По сути, здесь говорится о долгосрочном воздействии на потребителя (той или иной платформы) в отличие от традиционного подхода «подталкивающей» экономики. Необходимо создать увлекательный опыт, который побудит пользователей иметь «надлежащий уровень взаимодействия» с продуктом, чтобы они могли продолжать участвовать в мероприятиях по изменению поведения<sup>115</sup>. В качестве позитивного кейса можно привести опыт страховой компании из ЮАР. Ведущая страховая компания Южной Африки Discovery использует платформу Vitality, чтобы побудить своих клиентов к здоровому образу жизни, экономически стимулируя их скидками на здоровую еду, путешествия, членство в фитнес-клубе, товары для здоровья, спортивный инвентарь и тому подобное. Участники зарабатывают свои баллы, отмечая тренировки в «привязанных» фитнес-клубах или покупая здоровую пищу, которая также записывается на их карту Vitality. Программу пользуется многолетней популярностью у клиентов компании. На наш взгляд,

---

<sup>114</sup> Doyuran E. B. Nudge goes to Silicon Valley: designing for the disengaged and the irrational //Journal of Cultural Economy. – 2024. – Vol. 17. – №. 1. – P. 1-19.

<sup>115</sup> Cole-Lewis H. Ezeanochie N., Turgiss J. Understanding health behavior technology engagement: pathway to measuring digital behavior change interventions //JMIR formative research. – 2019. – Vol. 3. – №. 4. – P. e14052.

это пример того, как использование положений поведенческой экономики (цифрового маркетинга) позволяет реально перестроить поведение потребителей в лучшую сторону.

Сравнивая, первый и второй этапы «цифрового маркетинга», следует отметить, что в реальности их методы часто смешиваются, перекрываются и переплетаются. Тем не менее, предложение поведенческой экономики применительно к разработке программного обеспечения на втором этапе, заключается в создании продукта, который продуман, разработан и спроектирован так, чтобы стать чем-то, без чего пользователь не может обойтись не потому, что он вызывает привыкание, а потому, что он стал незаменимым для «вселенной распределенных действий» поведенческой проблемы <sup>116</sup>. Это конечно сложнее и требует значительно больших усилий, но и эффективность таких действий значительно выше.

В «цифровом маркетинге», - маркетинге XXI века, - наблюдаются абсолютно те же самые тренды, что происходили и в обычном (стандартном) маркетинге. Там то же в начале нужно было привлечь как можно больше клиентов, а – затем «удержать» их. Всё повторяется, пусть и на новом технологическом уровне.

---

<sup>116</sup> Caliskan K., Wade M. DARN, Part 2: An Evidence-Based Research and Prototyping Method for Strategic Design // She Ji: The Journal of Design, Economics, and Innovation. 2022. - Vol.8. - № 3. - P. 319–335. <https://doi.org/10.1016/j.sheji.2022.11.002>.

## ***Будущее в принятии экономических решений: человек или искусственный интеллект?***

В идеале, как на рабочем месте, так и в разработке политики предполагается, что наиболее плодотворные приложения искусственного интеллекта будут дополнять и компенсировать врожденные биологические и когнитивные ограничения людей. В реальности так происходит далеко не всегда. Выделим две области принятия решений в экономике с точки зрения взаимодействия работников и моделей искусственного интеллекта.

Первая – это обычные решения, в которых алгоритмы ИИ подсказывают возможно лучшее решение и предупреждают о потенциальных рисках. Принятие окончательного решения остаётся прерогативой человека. Существует много потенциальных полей приложения помощи искусственного интеллекта. Медицина<sup>117</sup> и право<sup>118</sup> могут быть названы в числе первостепенных. В этих сферах ошибочные или избыточные решения, принимаемые бесконтрольными специалистами, могут быть губительными для людей. Так, Дж. Доменигетти, экономист кантонального ведомства здравоохранения в Тичино (Швейцария) решил с коллегами выяснить: проводят ли хирурги больше операций, чем строго необходимо. Они обнаружили, что чем более сложный случай был у больного, тем больше учитывалась осведомленность пациентов о лечении. Лучше среди всех пациентов осведомлены, конечно, другие

---

<sup>117</sup> Raghu M., Blumer K., Corrado G., Kleinberg J., Obermeyer Z., Mullainathan S. The algorithmic automation problem: Prediction, triage, and human effort //arXiv preprint arXiv:1903.12220. – 2019.

<sup>118</sup> Rambachan A. Identifying Prediction Mistakes in Observational Data. – 2022.

врачи. Поэтому врачи шли под нож гораздо реже, чем среднестатистический житель Тичино<sup>119</sup>.

Вторая – это решения с высокой ставкой, где цена ошибки очень дорогостоящая или может привести к гибели людей. В качестве примера можно привести вождение автомобилем, управление энергетикой, финансовую торговлю, управление воздушным движением. В частности, TopSky-ATC, система искусственного интеллекта, которая может отклонять проблемные траектории полета для увеличения потока и повышения безопасности. Как следствие, водители автомобилей, авиадиспетчеры, финансовые трейдеры уже регулярно видят, что их решения отменяются алгоритмами ИИ.

Исследовательским вопросом здесь является: как надзор моделей искусственного интеллекта сказывается на работниках, и каковы последствия этого? Д. Альмог с коллегами провели многолетнее исследование, частично дающее ответ на этот вопрос. Они изучали принятие решений судьями в матчах большого тенниса после введения в 2006 году системы слежения Hawk-Eye (Ястребиный глаз). Вердикт ИИ считается окончательным. Неправильные решения линейных судей и арбитров теперь часто отменяются. На судьях надзор моделей искусственного интеллекта сказался очень противоречиво. С одной стороны, число ошибок снизилось на 8%. С другой стороны, изменилась структура ошибок.

---

<sup>119</sup> Domenighetti G., Casabianca A., Gutzwiller F., Martinoli S. Revisiting the most informed consumer of surgical services: the physician-patient //International journal of technology assessment in health care. – 1993. – Vol. 9. – №. 4. – P. 505-513.

Снизилось, - примерно в два раза, - число ошибок типа II (вылет мяча с корта в аут) и увеличилось число ошибок типа I (когда фиксируется удар мяча в площадь корта). Такое развитие событий было спровоцировано психологическими издержками судей, которые они стали нести из-за надзора ИИ. Ошибки типа II вызывают прекращение игры и всегда негативно воспринимаются как игроками, так и зрителями. Исследователи объясняют это тем, что отказ от решения может повлечь за собой психологические издержки (например, стыд и смущение из-за отклонения) и психологические выгоды (например, облегчение от исправления ошибок), то есть люди могут изменить свое принятие решений под надзором ИИ<sup>120</sup>. Это иллюстрирует, как идеи поведенческой экономики могут сыграть ключевую роль в понимании взаимодействия ИИ и человека<sup>121</sup>.

В целом, на уровне народного хозяйства, надзор моделей ИИ над работниками приводит к социальным издержкам, проявляющемся в искажении принятия решений одними людьми относительно судеб других: врачами, судьями и другими специалистами, которые будут избегать решения сложных случаев под угрозой осуждения общества, опирающегося на алгоритмы ИИ. В краткосрочном периоде экономика, наверно, выиграет от этого. Но в долгосрочном периоде не приведёт ли это к тому, что люди

---

<sup>120</sup> Almog D., Gauriot R., Page L., Martin D. AI Oversight and Human Mistakes: Evidence from Centre Court // arXiv preprint arXiv:2401.16754. – 2024.

<sup>121</sup> Camerer C. F. Artificial intelligence and behavioral economics // The economics of artificial intelligence: An agenda. – 2019. – P. 587-608.

перестанут повышать свою квалификацию, опасаясь негативного сравнения с большими языковыми моделями, и будут деградировать и как работники, и как граждане общества? Если модели искусственного интеллекта продолжают победоносно совершенствоваться, то вопрос о конкуренции в принятии экономических решений отпадёт сам собой и люди будут всё больше превращаться в бездумных потребителей, подобных обитателям Аксиомы из мультфильма ВАЛЛ-И. Подробнее это показано в<sup>122</sup>.

Человеческий разум в целом превосходит пока существующие модели искусственного интеллекта, что порождает для всех национальных экономик проблему преимущественного инвестирования средств и усилий: в развитие моделей ИИ или в совершенствование человеческого капитала граждан?

Сегодня, после появления в ноябре 2022 г. ChatGPT, происходит качественное изменение воздействия ИИ на рынок труда: если раньше людей заменяли только там, где труд носил однообразный, рутинный характер, высвобождая их возможности для нестандартного умственного труда, то сейчас речь идёт о замене большими языковыми моделями практически всех видов интеллектуального труда. Каким будет место работника будущего в этом новом общественном разделении труда? Сами исследователи предлагают всё новые и новые направления применения технологий искусственного интеллекта. Например, инновации традиционно

---

<sup>122</sup> Лукичёв П. М., Кузеева С. А. Роль визуального искусства как инновационной практики в преподавании экономических дисциплин // Инновации в образовании. 2023. № 12. С. 86–94.

считались прерогативой людей, но менеджеры по инновациям всё чаще сталкиваются с конкурирующими технологиями, с быстро меняющейся бизнес-средой, с всё более конкурентными глобальными рынками, с изменчивыми политическими ландшафтами.

### *Выводы и дальнейшие исследования*

Анализ, проведенный в разделе, показал, что современная экономика вступает в период «гибридного отступления», когда всё большая часть решений передаётся моделям искусственного интеллекта. Число сфер народного хозяйства, в которых ИИ выступают в качестве консультанта, а принятие итогового решения остаётся за работником сокращается. Реальность 2024 г. такова, что алгоритмы ИИ «поправляют» принятые человеком решения. При этом искусственному интеллекту отдаётся приоритет над человеком. Это меняет поведение людей при принятии последующих решений. Последствие противоречивы: с одной стороны из-за конкуренции с ИИ возрастает точность (качество) принимаемых решений, с другой увеличилось число пусть неправильных, но приятных (предпочитаемых) обществом решений взамен объективных. Для предупреждения чрезмерного доверия работников к решениям искусственного интеллекта, разработчикам совместных систем «Искусственный интеллект – Человек» необходимо тщательно рассмотреть (рассчитать) как объяснения уменьшают когнитивные усилия по проверке ИИ по сравнению с стратегией чрезмерного доверия, требующей небольших усилий. Совместное принятие экономических решений на основе взаимодействия «Человек –

Искусственный Интеллект» будет стремительно расширяться. Пионеры в этой области Талер и Санстейн предполагают, что руководителям организаций необходимо рассмотреть возможность найма специалистов со специальными знаниями в области поведенческой экономики, а также проведения внутреннего обучения по поведенческой науке<sup>123</sup>.

Сейчас Россия в программных приложениях отстаёт на 10–11 лет от развитых стран. Мы продолжаем находиться на первом этапе, когда их разработкой занимаются только технические специалисты и поэтому мало конкурентоспособны. Между тем, с 2013–2014 гг. в мире для создания программных приложений широко привлекают специалистов по поведенческой экономике и экономистов к разработке своих продуктов, ценностных предложений или рекламных предложений, чтобы сделать продукты более привлекательными. У нас университеты практически не готовят специалистов по поведенческой экономике, как на уровне бакалавриата (специалитета), так и на уровне магистратуры. Крайне редко читаются курсы по Поведенческой экономике для студентов, осваивающих информационные технологии. Следовательно, чтобы стать конкурентоспособной и перейти на современный этап создания программных приложений, российской экономике необходимы изменения, как в подготовке IT-специалистов с чтением курсов по этике искусственного интеллекта, поведенческой экономике, управлению технологиями ИИ; так и к подготовке

---

<sup>123</sup> Fusaro R., Sperling-Magro J. Much anew about ‘nudging’ //McKinsey Quarterly. – 2021.

полноценных специалистов по Поведенческой экономике, способных работать в цифровом секторе народного хозяйства.

Взаимодействие моделей искусственного интеллекта и положений поведенческой экономики будет дальше расширяться, поэтому необходим дальнейший междисциплинарный анализ принятия решений в современной экономике, объединяющий технологические, экономические и социальные аспекты проблемы.

## **2. 2. ПРИНЯТИЕ ЭКОНОМИЧЕСКИХ РЕШЕНИЙ: ЧЕЛОВЕК С «ЭФФЕКТОМ ПОЛАНЬИ» ИЛИ АГЕНТЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА?**

В данном разделе анализируются проблемы принятия экономических решений работниками-людьми и моделями искусственного интеллекта, как с точки зрения конкуренции между ними, так и взаимной работы (сотрудничества). Сегодня, через два года после появления ChatGPT, происходит постепенное перераспределение количества решений в хозяйственной жизни от человека к большим языковым моделям. При первоначальном сопоставлении экономических решений, принимаемых алгоритмами искусственного интеллекта и работниками-людьми, считалось, что за машинами будет выполнение повторяющихся, рутинных операций, а за людьми останется творческая часть труда. Работники, обладая «неявным знанием» Поланьи имели конкурентное преимущество в принятии решений перед большими языковыми моделями. Однако именно модели искусственного интеллекта последних лет занимают последнюю нишу человеческого труда, действуя в бизнес-процессах самостоятельно без непосредственного контроля работников в режиме реального времени. Сегодня появились Агенты искусственного интеллекта, которые могут самостоятельно взаимодействовать с виртуальным миром. Автор раскрывает перспективы сохранения в ближайшие годы конкурентного преимущества людей в принятии экономических решений по сравнению с большими языковыми моделями. В разделе анализируются пессимистический и оптимистический сценарии сотрудничества в экономике работников-людей и моделей искусственного интеллекта. Автор делает вывод о сохранении первенства человека в принятии экономических решений только в случае увеличения инвестиций в человеческий капитал.

## Введение

Технологические компании постоянно сообщают о новых достижениях в создании больших языковых моделей (БЯМ). В июне 2024 стартап Anthropic, занимающийся искусственным интеллектом, выпустил свою новейшую модель Claude 3.5 Sonnet, которая превосходит своих предшественников, в том числе, в кодировании. 12 сентября 2024 г. OpenAI, производитель ChatGPT, выпустил версию своей последней модели o1, заявив, что она «превосходно справляется с точной генерацией и отладкой сложного кода»<sup>124</sup>. Всё это приближает тот день, когда произойдёт, по их мнению, качественный прорыв в прогрессе искусственного интеллекта и будет создан AGI (Artificial General Intelligence – Общий искусственный интеллект), который превзойдёт по своим параметрам возможности любого человека. Этот качественный прорыв способен принципиально изменить соотношение сил в принятии экономических решений. Однако за последние два года уже возник целый ряд инноваций в технологиях искусственного интеллекта, подрывающих монополию работников-людей по управлению экономикой. Важнейшими среди них являются: обладание «неявным знанием» большими языковыми моделями и появление Агентов искусственного интеллекта. **Цель раздела:** проанализировать возможности сохранения конкурентного преимущества работников-людей над моделями искусственного

---

<sup>124</sup> Changing the program. The Economist. October 5<sup>th</sup> 2024. [Электронный ресурс]. URL: <https://www.economist.com/business/2024/09/29/ai-and-globalisation-are-shaking-up-software-developers-world> (дата обращения 23.12. 2024).

интеллекта в принятии экономических решений и оценить сценарии их дальнейшего взаимодействия.

### **Совершенствование больших языковых моделей для создания Общего искусственного интеллекта**

Прежде всего обратимся к очень плодотворной статье российского исследователя С. Карелова, который объективно оценил перспективы создания общего искусственного интеллекта<sup>125</sup>. С. Карелов показывает, что в AGI (Artificial General Intelligence – Общий искусственный интеллект) два термина из трёх не определены точно, поэтому наивно говорить, что Большие языковые модели (БЯМ) превзойдут человека, если мы не можем (следующая его мысль) определить чётко и ясно критерии этого. Тест Тьюринга, который, собственно говоря, является не тестом, а – мысленным экспериментом, по мнению С. Карелова, на самом деле, доказывает не определение того, это машина, а это человек, а представляет собой «гендерный тест»: мужчина или женщина. Далее он справедливо указывает, что все тесты, включая современные производные от Теста Тьюринга, большие языковые модели «проходят», но это не из-за их интеллекта, а из-за «подгонки», - явной или неявной. - под результат – действие «ловушки Гудхарта».

---

<sup>125</sup> Карелов С. В. «Ловушка Гудхарта» для AGI: проблема сравнительного анализа искусственного интеллекта и интеллекта человека // Ученые записки Института психологии РАН. – 2023. – Т. 3. – №. 3 (9). – С. 5-23. DOI:10.38098/proceedings\_2023\_03\_03\_02

Напомним, что закон Гудхарта гласит: «Любая наблюдаемая статистическая закономерность будет иметь тенденцию разрушаться, как только на нее оказывается давление в целях контроля»<sup>126</sup>. Уязвимость позиции С. Карелова: он рассматривает только рационального человека, как это, например делает экономика mainstream, но нормальный (реальный) человек обладает при принятии экономических решений и другими видами рациональности, включая ограниченную рациональность, выдвинутую одним из четырех создателей искусственного интеллекта и нобелевским лауреатом по экономике Гербертом Саймоном<sup>127</sup>. То есть *модели искусственного интеллекта даже в идеале будут превосходить только одну часть человека*, а именно - рациональную.

Взаимоотношения «человек – машина», которые устоялись за последние столетия, сегодня принципиально видоизменяются, как со стороны работника, так и со стороны машины. Прежняя схема, основывалась на людях, которые, используя машины, могли производить продукции больше и лучше, чем они делали раньше вручную. Это требовало исполнения организующей работы со стороны менеджеров, которые осуществляли функцию управления, надзора и контроля над производственными процессами. С точки

---

<sup>126</sup> Goodhart C. Monetary relationships: a view from Threadneedle Street. – University of Warwick, 1976.

<sup>127</sup> Simon H. A behavioral model of rational choice //Models of man, social and rational: Mathematical essays on rational human behavior in a social setting. – 1957. – Vol. 6. – №. 1. – P. 241-260.

зрения менеджмента бизнес-процессов сложившаяся схема организации выпуска товаров и услуг породила классическую экономическую концепцию «Принципал – Агент». В ней работники, непосредственно управляющие машинами, играли роль агентов, а менеджеры, выполняли по отношению к ним роль принципалов. Часть данной концепции отражала взаимоотношения между владельцами фирм, как принципалами, и менеджерами, как агентами. На всех уровнях действия ключевой особенностью концепции является то, что принципал не всегда может наблюдать за действиями агентов, а в некоторых случаях, даже осуществляя мониторинг, не может точно сказать, насколько действия агента являются наилучшими с точки зрения принципала<sup>128</sup>. Сегодня совершенствование технологий искусственного интеллекта вносит новое качество в эти взаимоотношения. Эту особенность первыми выделили А. Коринек и Д. Стиглиц (Korinek, Stiglitz), когда отметили, что определяющим признаком ИИ является автоматизация последней области, в которой рабочие-люди имели сравнительное преимущество перед машинами, – нашего мышления и обучения<sup>129</sup>.

Происходящее совершенствование технологий искусственного интеллекта всё более явно меняет соотношение в принятии экономических решений человеком и большими языковыми

---

<sup>128</sup> Лукичѳв, П. М. Экономика искусственного интеллекта и концепция «Принципал - Агент» / П. М. Лукичѳв, О. П. Чекмарев // Вопросы инновационной экономики. – 2022. – Т. 12, № 2. – С. 1069–1082. – DOI 10.18334/vinec.12.2.114514

<sup>129</sup> Korinek A., Stiglitz J. E. Artificial intelligence and its implications for income distribution and unemployment //The economics of artificial intelligence: An agenda. – University of Chicago Press, 2018. – P. 349-390. [Электронный ресурс]. URL: <https://www.nber.org/system/files/chapters/c14018/c14018.pdf> дата обращения 15.12. 2024).

моделями. Мы, как показывает поведенческая экономика, не являемся рациональными в принятии всех решений в области хозяйственной жизни. Как отмечала Й. Пуашундер ( Puaschunder J. M.), множество доказательств показало, что люди игнорируют рациональный выбор, который предсказывают стандартные неоклассические аксиомы максимизации прибыли, а скорее используют эвристику в своем повседневном принятии решений<sup>130</sup>. Ограниченная рациональность, иррациональность, неполная рациональность и другие виды рациональности, используемые нами в реальных экономических решениях, чаще обращаются к системе 1, ответственной за быстрые интуитивные решения в обход размышлений, по Д. Канеману<sup>131</sup>, чем к системе 2, с помощью которой принимаются обдуманые решения. Как показывает опыт, данные решения не всегда являются оптимальными.

Осознавая это, мы надеемся на использование возможностей больших языковых моделей, которые могут устранить недостатки обработки информации и принятия решений человеком. Как отмечали Э. Бриньолфссон и Э. Нг (Brynjolfsson E., Ng A.), возможности принятия решений технологиями сегодня уже сильно отличаются от тех, что были 40 лет назад. Более того, они находятся на скоростной полосе ко всё более крупным и мощным системам ИИ,

---

<sup>130</sup> Puaschunder J. M. Nudgital: Critique of behavioral political economy //Proceedings of the 9th international RAIS conference on social sciences and humanities. – Scientia Moralitas Research Institute, 2018. – С. 87-117. DOI: 10.5281/zenodo.1244836

<sup>131</sup> Канеман Д. Думай медленно... решай быстро. Издательство АСТ. Москва 2014.– 710 с.

которые могут принимать решения способами, которые никогда не могла бы сделать ни одна централизованная машина<sup>132</sup>.

А. Сен, нобелевский лауреат по экономике, ещё в 1986 году отмечал, что все рациональные данные в мире становятся бесполезными из-за иррациональности человеческого поведения, указывая на «врожденную трудность в прогнозировании результатов взаимодействия миллионов людей с различными ценностями, целями, мотивациями, ожиданиями, способностями, правами, средствами и обстоятельствами, взаимодействующими друг с другом в самых разных институциональных условиях»<sup>133</sup>.

Взаимоотношения «человек – машина» в перспективе будут определяться тем как именно работники-люди организуют взаимодействие с моделями искусственного интеллекта. Итан Моллик (Mollick E.) говорит, что возможны два типа взаимодействия человека с ИИ:

1) **кентавр** – чёткое разделение задач на человеческие и ИИ (так же как тело кентавра явно разделено на человеческую и лошадиную части);

---

<sup>132</sup> Brynjolfsson E., Ng A. Big AI can centralize decision-making and power, and that's a problem //Missing links in AI governance. – 2023. – P. 65. [Электронный ресурс]. URL: <https://sustainabilitydigitalage.org/featured/wp-content/uploads/missing-links-in-ai-governance-unesco-mila.pdf#page=71> (дата обращения 20.12. 2024).

<sup>133</sup> Sen A. K. Prediction and economic theory //Proceedings of the Royal Society of London. Series A. Mathematical and Physical sciences. – 1986. – Vol. 407. – №. 1832. – P. 3-23.

2) **киборг** – привлечение ИИ к решению всех текущих задач, совмещая потенциал двух интеллектов и позволяя им гибко взаимодействовать<sup>134</sup>.

И то, и другое направления имеют свои преимущества и свои ограничения. По нашему мнению, второй тип взаимодействия более плодотворный, позволяющий в идеале использовать не только человеческий интеллект и искусственный интеллект, но и достичь синергетического эффекта от их совмещения. Первое направление, - кентавр, - является более безопасным, поскольку исключает негативные последствия взаимодействия ИИ и человека: избыточное доверие работников к решениям искусственного интеллекта, «гибридное отступление», и превращение людей в бездумных потребителей, как было показано в фильме ВАЛЛ-И.

### **Эффект Поланьи: дискуссия и конкурентное преимущество в принятии экономических решений**

Важную роль в совершенствовании искусственного интеллекта сыграла дискуссия 2010-х годов о «неявном знании» Поланьи. Является ли эффект Поланьи преимуществом мышления человека над моделями ИИ? Проанализируем аргументы «за» и «против».

Данный эффект был открыт известным химиком и философом М. Поланьи (Polanyi M.) в рамках его теории познания. Он

---

<sup>134</sup> Mollick E., Mollick E. Co-Intelligence. – Random House UK, 2024.

сформулировал положение о неявном (молчаливом) знании<sup>135</sup>, <sup>136</sup>: каждый человек «может знать больше, чем он может сказать». Проявляется оно в том, как умелый игрок может выиграть партию в бильярд, не зная законов физики; в узнавании взрослого человека на детской фотографии, в умении ребёнка ходить. Всё это и многое другое приобретается не на основе явно и чётко сформулированных теоретических знаний, а на основе личного опыта и наблюдений с течением времени. Сам Поланьи «присваивает «молчаливый личный коэффициент»<sup>137</sup> всем знаниям, начиная от самых обыденных и телесных (например, езда на велосипеде, плавание) до самых специализированных и абстрактных (например, химия, математика). Эти знания не являются одинаковыми, будучи интуитивными и невербальными, а отличаются от человека к человеку. Он обуславливает «установление контакта с реальностью»<sup>138</sup> целью поиска знаний; и рассматривает «интеллектуальные страсти»<sup>139</sup>, особенно «стремление к истине»<sup>140</sup>, как движущую силу всех исследований.

Эффект Поланьи приобрёл особое значение в связи с появлением компьютеров, а затем – искусственного интеллекта.

---

<sup>135</sup> Polanyi M. *The Study of Man*. Chicago: University of Chicago Press, 1959.

<sup>136</sup> Polanyi M. *The Tacit Dimension*. Chicago and London: The University of Chicago Press, 1966.

<sup>137</sup> Polanyi M. *The Study of Man*. Chicago: University of Chicago Press, 1959, p. 25.

<sup>138</sup> Polanyi M. *The Study of Man*. Chicago: University of Chicago Press, 1959, p. 27.

<sup>139</sup> Polanyi M. *The Study of Man*. Chicago: University of Chicago Press, 1959, p. 37.

<sup>140</sup> Polanyi M. *The Study of Man*. Chicago: University of Chicago Press, 1959, p. 62

Ученые и практики стали постоянно сравнивать работу человека и работу компьютеров (искусственного интеллекта), мышление работников-людей и мышление моделей ИИ. Исследователей особенно интересует: является ли эффект Поланьи, которым обладает человек, конкурентным преимуществом в сфере экономики по сравнению с компьютером или большой языковой моделью? В 2010-е годы прошла дискуссия, выявившая два противоположных подхода к данной проблеме. Сторонники положительного ответа на этот вопрос приводят следующие аргументы. М. Варди (Vardi M. Y.) отмечает, что именно эффект Поланьи вызвал растущую профессиональную поляризацию на рынке труда, которая привела к сокращению среднего класса в промышленно развитых странах, поскольку многие профессии со средним доходом в сфере продаж, офисной и административной работы, а также повторяющейся производственной работы являются трудоёмкими. Более того, последующий рост неравенства доходов и неравенства в благосостоянии в последнее время стал серьёзной социально-экономической проблемой в развитых странах<sup>141</sup>. Д. Аутор (David Autor) указывает на интенсивное выполнение нестандартных задач, которые выделяются в парадоксе Поланьи, как на серьёзное ограничение того, какую работу могли выполнять традиционные компьютеры, что не позволяет современным алгоритмам заменить человеческий труд в ряде квалифицированных рабочих мест<sup>142</sup>.

---

<sup>141</sup> Vardi M. Y. Is information technology destroying the middle class? //Communications of the ACM. – 2015. – Vol. 58. – №. 2. – P. 5-5.

<sup>142</sup> Autor D. Polanyi's paradox and the shape of employment growth. – National Bureau of Economic Research, 2014. – №. w20485. [Электронный ресурс]. URL:

Мастерство при выполнении нестандартных задач достигается не за счёт изучения правил, а путём наблюдений за практикой и опыта. Не в игре с чёткими правилами и определённым набором ходов, а в реальной экономике с её неопределённостью и риском существует много задач, которые работники-люди легко выполняют без формального понимания того, как они это делают.

Дж. Прассл (Prassl J.) делает акцент на том, что постоянно меняющийся, неструктурированный характер некоторых видов деятельности в настоящее время представляет собой устрашающую проблему для автоматизации. Несмотря на годы времени и миллиарды инвестиций, потраченных на разработку беспилотных автомобилей и роботов-уборщиков, эти системы машинного обучения продолжают бороться со своей низкой адаптивностью и интерпретируемостью, от неспособности беспилотных автомобилей сделать неожиданный объезд до роботов-уборщиков, уязвимых перед неконтролируемыми домашними животными или детьми<sup>143</sup>. Действительно до сих пор ни одна модель искусственного интеллекта не смогла обеспечить стопроцентной безопасности для беспилотных автомобилей или стопроцентно точного диагноза в медицине.

---

[https://www.nber.org/system/files/working\\_papers/w20485/w20485.pdf](https://www.nber.org/system/files/working_papers/w20485/w20485.pdf)  
20.12. 2024).

(дата обращения

<sup>143</sup> Prassl J. Humans as a service: The promise and perils of work in the gig economy. – Oxford University Press, 2018. pp. 138–139.

На наш взгляд, тут возникает общий вопрос: *кто к кому должен приспосабливаться, - люди к роботам с ИИ, или технологии искусственного интеллекта к людям.* Возможно, и первое: железные дороги имеют идеальное покрытие, позволяющее ездить поездам. В современном производстве примером служит автомобилестроительный конвейер. Создание искусственной окружающей среды представляет, по нашему мнению, путь, пусть и дорогостоящий, для преодоления эффекта Поланьи. Наверно, однозначного ответа во всех случаях не существует. Следует учесть, что алгоритмы искусственного интеллекта целиком основаны на полной рациональности. Другие виды рациональности, которыми обладает нормальный (средний) человек при принятии экономических решений, не рассматриваются. Если эффект Поланьи действительно представляет конкурентное преимущество работника-человека над ИИ, то роль поведенческой экономики возрастает.

Сторонники негативного ответа на вопрос о конкурентном преимуществе человека над алгоритмами искусственного интеллекта делают акцент на результатах, а не на процессе принятия решений. Так Н. Карр (Carr) указывает, что предположение о том, что компьютеры должны иметь возможность воспроизводить неявные знания, которые люди будут применять для выполнения сложных задач, само по себе вызывает сомнение. При выполнении задач системам и машинам вовсе не обязательно следовать правилам, которым следуют люди. Целью использования машины для выполнения задачи является воспроизведение наших результатов

для практических целей, а не наших средств<sup>144</sup>. Д. Каплан (Jerry Kaplan), углубляет эту аргументацию. По его мнению,<sup>145</sup> существуют четыре ресурса и способности, необходимые для выполнения любой конкретной задачи: осведомленность, энергия, рассуждение и средства. Биологическая система человека (комплекс мозг-тело) естественным образом объединяет все эти четыре свойства, в то время как в электронной сфере машины могут получить эти способности благодаря разработкам в области робототехники, машинного обучения и систем управления восприятием. Например, данные, предоставляемые широкой сетью датчиков, позволяют ИИ воспринимать различные аспекты окружающей среды и мгновенно реагировать в хаотичных и сложных ситуациях реального мира (т. е. осведомленность); команды и сигналы для исполнительных устройств могут быть централизованы и управляться в кластерах серверов или в «облаке» (рассуждение)<sup>146</sup>.

Наиболее последовательную позицию, опередившую появление ChatGPT, занял Д. Сасскинд (Susskind D.). Он отмечал, что последние достижения в области машинного обучения преодолели парадокс Поланьи. Вместо того, чтобы полагаться на

---

<sup>144</sup> Carr N. The glass cage: Automation and us. – New York: WW Norton & Company, 2014, p. 11-12.

<sup>145</sup> Kaplan J. Humans Need Not Apply: A Guide to Wealth & Work in the Age of Artificial Intelligence. – Yale University Press, 2015., P. 41-43, 145.

<sup>146</sup> Kaplan J. Humans Need Not Apply: A Guide to Wealth & Work in the Age of Artificial Intelligence. – Yale University Press, 2015., P. 41-43, 145.

алгоритмы программистов, обучающих их человеческим знаниям, компьютерные системы теперь могут самостоятельно изучать неявные правила из контекста, большого количества данных и прикладной статистики. Поскольку машины могут выводить неявные знания, которые люди извлекают из примеров, без помощи человека, они больше не ограничены теми правилами, которые молчаливо применяются, но не понимаются людьми явно.<sup>147</sup>

После внедрения ChatGPT большинство экономистов, включая Д. Аутора<sup>148</sup>, изменили своё отношение к эффекту Поланьи. Во многом это обусловлено явно проявившимися отличиями искусственного интеллекта от компьютера. Развитие ИИ происходит зачастую непредсказуемо для своих создателей и результаты его работы часто неожиданны по отношению к тем компонентам, на которых он обучался. Этот эффект называется «эмерджентность», то есть возникновение новых свойств, которые не были заложены в систему при проектировании.

Алгоритмы машинного обучения (МО) работают иначе, чем традиционные компьютерные программы: вместо того, чтобы требовать явных инструкций для функционирования, эти системы выводят инструкции из примеров. Например, имея обучающий

---

<sup>147</sup> Susskind D. Re-thinking the capabilities of machines in economics. – 2017. [Электронный ресурс]. URL: [https://ora.ox.ac.uk/objects/uuid:554f4c9c-6c78-44b5-9e93-788c731bec5f/download\\_file?file\\_format=&safe\\_filename=825-susskind-april18.pdf&type\\_of\\_work=Working+paper](https://ora.ox.ac.uk/objects/uuid:554f4c9c-6c78-44b5-9e93-788c731bec5f/download_file?file_format=&safe_filename=825-susskind-april18.pdf&type_of_work=Working+paper) (дата обращения 06.01. 2025).

<sup>148</sup> Autor D. Applying AI to rebuild middle class jobs. – National Bureau of Economic Research, 2024. – №. w32140. [Электронный ресурс]. URL: [https://www.nber.org/system/files/working\\_papers/w32140/w32140.pdf](https://www.nber.org/system/files/working_papers/w32140/w32140.pdf) (дата обращения 20.12. 2024).

набор изображений, системы МО могут научиться распознавать конкретных людей, даже если невозможно полностью объяснить, какие физические особенности характеризуют личность данного человека. Эта способность подчёркивает ключевой, отличительный аспект систем машинного обучения: они могут учиться выполнять задачи, даже когда нет никаких инструкций, включая задачи, требующие неявных знаний, которые ранее можно было получить только через жизненный опыт<sup>149;150;151</sup>.

Э. Бриньолфссон, Д. Ли, Л.Р. Рэймонд (Brynjolfsson E., Li D., Raymond L. R.) выделяют ещё одно замечательное свойство, которое присуще машинному обучению. Системы МО часто обучаются на данных от работников-людей, которые, естественно, различаются по своим способностям. Видя множество примеров задач — создание торговых предложений, вождение грузовика или диагностика пациента, и это лишь некоторые из них — выполненных хорошо и плохо, эти **модели могут неявно узнать** (выделено автором), какие конкретные модели поведения и характеристики отличают

---

<sup>149</sup> Polanyi M. The Tacit Dimension. Chicago and London: The University of Chicago Press, 1966.

<sup>150</sup> Autor D. Polanyi's paradox and the shape of employment growth. – National Bureau of Economic Research, 2014. – №. w20485. [Электронный ресурс]. URL: [https://www.nber.org/system/files/working\\_papers/w20485/w20485.pdf](https://www.nber.org/system/files/working_papers/w20485/w20485.pdf) (дата обращения 20.12. 2024).

<sup>151</sup> Brynjolfsson E., Mitchell T. What can machine learning do? Workforce implications //Science. – 2017. – Т. 358. – №. 6370. – С. 1530-1534. [Электронный ресурс]. URL: <https://drive.google.com/file/d/1pEyizN-MhuSmqhDdEVIu8hQ1D3sjwXkH/view> дата обращения 02.01. 2025).

высокопроизводительных работников от их менее эффективных коллег<sup>152</sup>.

Д. Аутор в 2024 году<sup>153</sup> также отмечал качественно иную по сравнению с компьютерами роль в экономике современных моделей искусственного интеллекта. ИИ не заслуживает доверия в отношении фактов и цифр — он не уважает правила. Однако искусственный интеллект удивительно эффективен в приобретении неявных знаний. Вместо того чтобы полагаться на жестко запрограммированные процедуры, ИИ учится на примерах, приобретает мастерство без явных инструкций и приобретает способности, которым он явно не был спроектирован для обладания.

Таким образом можно сделать вывод: сегодняшние модели искусственного интеллекта могут обладать неявным знанием, которое раньше было монополией людей-работников. Потенциальными последствиями этого тренда являются радикальное изменение ценности принятия решений работниками-людьми и деформация существующего рынка труда, поскольку ИИ изменит значимость и природу человеческой экспертизы. Обеспечивая поддержку принятия решений в форме руководства и

---

<sup>152</sup> Brynjolfsson E., Li D., Raymond L. R. Generative AI at work. – National Bureau of Economic Research, 2023. – №. w31161. [Электронный ресурс]. URL: [https://www.nber.org/system/files/working\\_papers/w31161/w31161.pdf](https://www.nber.org/system/files/working_papers/w31161/w31161.pdf) (дата обращения 20.12. 2024).

<sup>153</sup> Autor D. Applying AI to rebuild middle class jobs. – National Bureau of Economic Research, 2024. – №. w32140. [Электронный ресурс]. URL: [https://www.nber.org/system/files/working\\_papers/w32140/w32140.pdf](https://www.nber.org/system/files/working_papers/w32140/w32140.pdf) (дата обращения 20.12. 2024).

ограждений в реальном времени, искусственный интеллект может позволить большему числу работников, обладающих дополнительными знаниями, выполнять некоторые из задач принятия решений с более высокими ставками, которые в настоящее время приписываются элитным экспертам, таким как врачи, юристы, программисты и педагоги.<sup>154</sup> Проще говоря, принятие решений, которое раньше было возможно только узким специалистам, станет доступно юзерам (обычным пользователям). Аналогией может служить эволюция использования компьютеров. Из положительных последствий можно выделить: снижение стоимости таких ключевых услуг, как юридическая экспертиза, здравоохранение, образование; улучшение качества рабочих мест для работников без высшего образования и опыта работы; смягчение неравенства в оплате труда. Негативными последствиями являются: обесценение человеческого образования и профессионального опыта, отсутствие мотивации к работе для наиболее высококвалифицированных специалистов, блокировка развития экономики в целом.

Ещё одним фактором, оказывающим всё большее влияние на соотношение сил в принятии экономических решений между работниками-людьми и моделями ИИ стало появление в 2023 году Агентов искусственного интеллекта. В расширительной трактовке агент — это система, способная воспринимать окружающую среду и

---

<sup>154</sup> Autor D. Applying AI to rebuild middle class jobs. – National Bureau of Economic Research, 2024. – №. w32140. [Электронный ресурс]. URL: [https://www.nber.org/system/files/working\\_papers/w32140/w32140.pdf](https://www.nber.org/system/files/working_papers/w32140/w32140.pdf) (дата обращения 20.12. 2024).

принимать решения на основе этих восприятий для достижения конкретных целей<sup>155</sup>. Он представляет собой преобразование чат-бота в новый тип автономной системы, которая может использовать программные приложения, веб-сайты и другие онлайн-инструменты. Эта технология также может генерировать компьютерный код, который подключается к другим программным приложениям и веб-сайтам.

Уже сейчас Агенты ИИ часто взаимодействуют с инструментами или услугами через выделенные API, которые представляют собой интерфейсы и протоколы, специально структурированные для агентов, а не для пользователей-людей. То есть агенты взаимодействуют друг с другом без вмешательства работника-человека.<sup>156</sup> Оснащенный подключаемым модулем, называемым интерпретатором кода, ChatGPT может не только писать код, но и запускать его. Это позволило технологии мгновенно выполнять задачи, которые она не могла выполнять в прошлом, включая редактирование электронных таблиц и преобразование неподвижных изображений в видео. Google, Microsoft и другие компании изучают похожие технологии. Вице-президент Microsoft Эшли Ллоренс (Ashley Llorens) отмечает, что в проектах, которые

---

<sup>155</sup> Wooldridge M., Jennings N. R. Intelligent agents: Theory and practice. // The knowledge engineering review. 1995. Vol. 10(2). P.115–152.

<sup>156</sup> Chan A., Ezell C., Kaufmann M., Wei K., Hammond L., Bradley H., ... & Anderljung M. Visibility into AI Agents //The 2024 ACM Conference on Fairness, Accountability, and Transparency. – 2024. – С. 958-973. [Электронный ресурс]. URL: <https://dl.acm.org/doi/pdf/10.1145/3630106.3658948> (дата обращения 04.01. 2025).

мы представляем, ИИ, по сути, работают с другими ИИ от вашего имени.<sup>157</sup> Фактически это означает, что от традиционной схемы Пр. → А, где принципалом (Пр.) выступает собственник (менеджер) на которого трудится работник – агент (А) по мере развития большие языковые модели перешли к схеме Пр. → ИИ, где на собственника работает искусственный интеллект (ИИ). Теперь же, благодаря появлению Агентов искусственного интеллекта возникает схема: А (ИИ) → А (ИИ), в которой агенты искусственного интеллекта (А (ИИ)) взаимодействуют от имени владельца с другими агентами искусственного интеллекта. Закономерно возникает вопрос, а зачем нужен в этой бизнес-схеме работник-человек?

Пока это работает не очень хорошо, но ближайший (2–3 года) идеал – обеспечить каждого цифровым помощником, который потенциально может делать практически всё в Интернете. Это способно облегчить жизнь — но также может заменить бесчисленное множество рабочих мест среди «белых воротничков». Для более отдалённого периода независимые проекты, такие как AutoGPT, пытаются продвинуться на несколько шагов дальше. Идея заключается в том, чтобы дать системе цели, например «создать компанию» или «заработать немного денег». Затем она будет искать способы достижения этой цели, задавая себе вопросы и подключаясь к другим интернет-сервисам<sup>158</sup>.

---

<sup>157</sup> Metz C., Weise K. ‘A.I. Agents’ Can Play Minecraft. They’re Just Getting Warmed Up. New York Times. October. 18, 2023, Section B, Page 1.

<sup>158</sup> Metz C., Weise K. ‘A.I. Agents’ Can Play Minecraft. They’re Just Getting Warmed Up. New York Times. October. 18, 2023, Section B, Page 1.

## **Риски первенства технологий ИИ в принятии решений над человеком**

Расширение применения моделей искусственного интеллекта для принятия решения в экономике порождает, как минимум, три трудно разрешимые проблемы в краткосрочном периоде: централизация принятий решений большими языковыми моделями, справедливость, этические проблемы.

Увеличение мощностей систем искусственного интеллекта, произошедшее за последние несколько лет, усилило возможности централизованного принятия решений. Последствия этого для экономики весьма противоречивы. С одной стороны, повышение потенциала больших языковых моделей улучшит и качество, и эффективность принимаемых решений. С другой стороны, это может привести к концентрации власти и богатства в невиданных ранее масштабах. Объективными причинами данной тенденции является: а) перемещения знаний и информации от человеческого мозга к машинам, б) перемещения вычислительной мощности, необходимой для принятия решений, от человеческого мозга к машинам <sup>159</sup>. Кроме того, существует чисто экономическое основание для централизации принятия решений: дороговизна развития, как барьер к входу. Данные показывают, что в версии ChatGPT конца 2023 года содержалось 175 млрд кодов <sup>160</sup>,

---

<sup>159</sup> Brynjolfsson E., Ng A. Big AI can centralize decision-making and power, and that's a problem //Missing links in AI governance. – 2023. – P. 65. [Электронный ресурс]. URL: <https://sustainabilitydigitalage.org/featured/wp-content/uploads/missing-links-in-ai-governance-unesco-mila.pdf#page=71> (дата обращения 20.12. 2024).

<sup>160</sup> Mollick E., Mollick E. Co-Intelligence. – Random House UK, 2024.

описывающих взаимные связи слов и частей слов. Важно, что эти коды не были заданы в модели искусственного интеллекта изначально, он сгенерировал их сам в процессе работы. Необходимость масштабного обучения делает производство ИИ очень дорогим и доступным ограниченному числу компаний.

На наш взгляд, если данные тенденции будут продолжать расширяться неконтролируемым образом, то последствия могут стать необратимыми. Централизация принятия экономических решений и централизация власти, идущие рука об руку, очень опасны для развития общества. Для предотвращения этого необходимо действовать одновременно в двух направлениях: 1) усиливать государственное регулирование применения моделей ИИ, 2) повышать инвестиции в улучшение человеческого капитала.

Существующие сегодня модели искусственного интеллекта вызывают ряд этических соображений, особенно когда речь идёт о таких вопросах, как конфиденциальность, права интеллектуальной собственности и потенциальная предвзятость в генерируемом контенте. Чем больше применяются системы машинного обучения, тем сильнее становится генеративный ИИ, тем важнее становится внедрение этических рамок.<sup>161</sup> . Особенно актуальным является необходимость в справедливом искусственном интеллекте в эпоху систем общего назначения, таких как ChatGPT, Gemini и других больших языковых моделей (БЯМ). Однако растущая сложность

---

<sup>161</sup> Beerbaum D. Generative Artificial Intelligence (GAI) Software—Assessment on Biased Behavior // Available at SSRN 4386395. – 2023.

взаимодействия человека и ИИ и его социальные последствия подняли вопросы о том, как можно применять стандарты справедливости. Технические рамки, используемые сейчас исследователями машинного обучения для оценки справедливости, такие как групповая справедливость и справедливые представления, далеки от идеала, и их применение к большим языковым моделям сталкивается с присущими им ограничениями. Каждая рамка либо логически не распространяется на БЯМ, либо представляет собой понятие справедливости, которое является неразрешимым для больших языковых моделей, в первую очередь из-за множества затронутых групп населения, чувствительных атрибутов и вариантов использования.<sup>162</sup>

Исследователи отмечают также некоторые ожидаемые виды ущерба от всё более агентных систем, а именно то, что они могут вызывать системный и отсроченный ущерб, лишая человека возможности принимать решения, усугублять чрезмерную концентрацию власти и быть источником дополнительных неизвестных угроз через возникающие возможности».<sup>163</sup>

Все опасения по поводу прогресса больших языковых моделей и возможностей государств по их регулированию также сохраняются.

---

<sup>162</sup> Anthis J., Lum K., Ekstrand M., Feller A., D'Amour A., Tan C. The Impossibility of Fair LLMs //arXiv preprint arXiv:2406.03198. – 2024. [Электронный ресурс]. URL: <https://arxiv.org/pdf/2406.03198> (дата обращения 02.01. 2025).

<sup>163</sup> Chan A., Salganik R., Markelius A., Pang C., Rajkumar N., Krasheninnikov D., ... & Maharaj T. Harms from increasingly agentic algorithmic systems //Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. – 2023. – С. 651–666. [Электронный ресурс]. URL: <https://dl.acm.org/doi/pdf/10.1145/3593013.3594033> (дата обращения 23.12. 2024).

Мы можем сколько угодно уговаривать себя, но «X-риски» остаются. Как отмечали создатели поведенческой экономики А. Тверски и Д. Канеман <sup>164</sup>, сложная система, будь то ядерный реактор или человеческое тело, не сможет функционировать, если хотя бы один из критически важных компонентов откажет. Даже если вероятность отказа каждого из компонентов мала, шансы, что произойдёт катастрофа могут быть значительны, если количество компонентов велико. Именно этот случай мы имеем с большими языковыми моделями сейчас, поэтому государственное регулирование искусственного интеллекта, нацеленное только на решение текущих задач, - фейки, приватность и т. п., - опасно! Оно должно быть направлено и на предотвращение экзистенциальных рисков.

### **Сценарии взаимоотношений моделей искусственного интеллекта и работников-людей в принятии экономических решений**

Принятие решений играет ключевую роль в прогрессе экономики, как на микроуровне, так и на макроуровне. Та национальная экономика (или компания), которая принимает более эффективные решения, побеждает в конкурентной борьбе. Растущее участие моделей искусственного интеллекта в принятии

---

<sup>164</sup> Tversky A., Kahneman D. Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty //Science. – 1974. – Vol. 185. – №. 4157. – p. 1124–1131.

экономических решений вызывает два сценария развития событий: пессимистичный и оптимистичный.

Пессимистичный сценарий основан на углублении двух трендов. Первый – это передача всё большего числа экономических решений от работника-человека к алгоритмам искусственного интеллекта. Она сопровождается «гибридным отступлением»<sup>165</sup> [4]. Люди обладают избыточным доверием по отношению к предложениям, рекомендациям больших языковых моделей. Кроме того, ситуация, когда решение модели искусственного интеллекта имеет первенство над решениями людей приводит к деформации поведения специалистов: возрастает процент точных решения, но одновременно, увеличивается число решений, выгодных «здесь и сейчас» (например, увеличение прибыли фирмы в первом квартале) без учёта долгосрочных последствий для конкурентоспособности компании (прибыль была получена за счёт сокращения инвестиций или затрат на охрану окружающей среды).

Второй тренд - передача всё большего числа решений в экономике от элитных специалистов (врачей, юристов) к более широкому кругу работников. Последствия будут крайне противоречивы. Из негативных отметим снижение ценности труда высококвалифицированных специалистов и отсутствие мотивации у них к развитию. Усилит действие второго тренда возникновение агентов искусственного интеллекта, которые смогут заменить в

---

<sup>165</sup> Лукичев, П. М. Принятие решений в современной экономике: искусственный интеллект vs поведенческая экономика / П. М. Лукичев // Вопросы инновационной экономики. – 2024. – Т. 14, № 3. – С. 649-666. – DOI 10.18334/vines.14.3.121070

среднесрочной перспективе большое число офисных сотрудников, а в долгосрочной перспективе – практически всех работников.

Возникнет ли «прорыв» в обучении больших языковых моделей за счёт создания агентов искусственного интеллекта? Остановят ли агентные отношения прогресс человеческой науки? Созданная в августе 2024 года японским стартапом Sakana AI «первая комплексная система для полностью автоматических научных открытий, позволяющая большим языковым моделям проводить исследования самостоятельно» сделала первый практический шаг в этом направлении<sup>166</sup>. Её AI Scientist может во много раз увеличить число публикуемых «научных» статей. Однако они не будут обладать ни новизной, ни более высоким качеством по сравнению со средним сегодняшним уровнем. Аналогия автора: нам нравится писатель или певец. Мы хотим прочитать ещё один роман (детектив) любимого писателя, послушать ещё одну песню понравившегося певца (певицы), посмотреть ещё одну серию любимого сериала. Большие языковые модели способны будут сделать это, создав вариации романа, «перепев» песню. Но они не смогут сделать ничего нового. Поэтому ключевым моментом здесь будет не постоянные имитации одного и того же уровня развития, а создания инноваций, которые дают толчок росту производительности труда и прогрессу в развитии общества.

---

<sup>166</sup> Lu C., Lu C., Lange R. T., Foerster J., Clune J., Ha, D. The ai scientist: Towards fully automated open-ended scientific discovery //arXiv preprint arXiv:2408.06292. – 2024. [Электронный ресурс]. URL: <https://arxiv.org/abs/2408.06292> (дата обращения 06.01. 2025).

Оптимистичный сценарий основан на поддержании выгод от сотрудничества работников-людей и ИИ в принятии экономических решений, в достижении синергетического эффекта от их взаимодействия. Плюсом второго тренда является рост производительности труда «новичков» и малоквалифицированных работников до уровня высококвалифицированных и опытных работников, что повышает общий рост производительности предприятия (народного хозяйства). Произойдет, как это было ранее с компьютерами и Интернетом, распространение применения моделей искусственного интеллекта от узкого круга специалистов к большому числу компаний и работников, что обеспечит значительный рост производительности труда во всех отраслях экономики. Технологии искусственного интеллекта станут действительно общей технологией<sup>167</sup>. Кроме того, сохраняется возможность лучшего использования ИИ лицами, принимающими решения в компаниях. Для этого, как показывают исследования, надо ориентироваться на объяснения на основе примеров и на ограничения объяснений на основе функций, что уменьшает чрезмерную уверенность и помогает достичь взаимодополняющей производительности человека и искусственного интеллекта, превосходя производительность работника или ИИ по отдельности. Одновременно следует указать на потребности пользователей в

---

<sup>167</sup> Trajtenberg M. Artificial intelligence as the next GPT: A political-economy perspective. / The economics of artificial intelligence: An agenda., 2018. – 175-186 p. [Электронный ресурс]. URL: <https://www.nber.org/system/files/chapters/c14025/c14025.pdf> (дата обращения 06.01.2025).

поддержке принятия решений системами искусственного интеллекта, которые лучше учитывают человеческую интуицию принятия решений, более совместимы с человеческой интуицией и поддерживают более критическое понимание ИИ.<sup>168</sup>

Отметим также, что обладание неявным знанием у больших языковых моделей и расширение применения агентов искусственного интеллекта являются пока только потенциальными возможностями. Осуществимость технологических решений во многом ограничена их экономическими выгодами. Понадобится определённый «переходный период» к новой реальности, который потребует значительных усилий и инвестиций со стороны и компаний, и государств. В этот период от работников-людей будет необходимо ещё большее, чем раньше использование неявного знания. Не случайно наследие М. Поланьи продолжает стимулировать интеллектуальную мысль: Общество Поланьи и Либеральная философская ассоциация Майкла Поланьи (Polanyi Society and Michael Polanyi Liberal Philosophical Association) издают научные журналы, посвященные развитию его идей.<sup>169</sup>

---

<sup>168</sup> Chen V., Liao Q. V., Wortman Vaughan J., Bansal G. Understanding the role of human intuition on reliance in human-AI decision-making with explanations // Proceedings of the ACM on Human-computer Interaction. – 2023. – Vol. 7. - № CSCW2. – P. 1-32. [Электронный ресурс]. URL: <https://dl.acm.org/doi/pdf/10.1145/3610219> (дата обращения 06.01. 2025), p. 3, 25.

<sup>169</sup> Hadjimichael D., Pyrko I., Tsoukas H. Beyond tacit knowledge: How Michael Polanyi's theory of knowledge illuminates theory development in organizational research // Academy of management review. – 2024. – Vol. 49. – №. 4. – P. 919-934. [Электронный ресурс]. URL: [https://purehost.bath.ac.uk/ws/portalfiles/portal/296758710/2023\\_Hadjimichael\\_et\\_al\\_accepted\\_author.pdf](https://purehost.bath.ac.uk/ws/portalfiles/portal/296758710/2023_Hadjimichael_et_al_accepted_author.pdf) (дата обращения 15.12. 2024).

Возможности наступления (преобладания) оптимистичного сценария или пессимистичного сценария будут во многом зависеть от того, какой объём инвестиций будет вкладываться в развитие человеческого капитала. Вызовы, которые несёт расширение применения технологий искусственного интеллекта в принятии экономических решений, означают необходимость совершенствования знаний, умений и навыков работников-людей. Чем выше будет их человеческий капитал, адаптированный к использованию ИИ в бизнес-процессах, тем более вероятен оптимистичный сценарий. Если же продолжится неконтролируемый процесс передачи экономических решений от человека к моделям искусственного интеллекта, то и значение работающих специалистов будет снижаться вплоть до полного исключения людей из экономики в долгосрочном периоде. Появление Общего искусственного интеллекта будет важнейшим катализатором воплощения пессимистичного сценария.

### ***Выводы и дальнейшие исследования***

Кооперация в бизнес-процессах между работниками-людьми и большими языковыми моделями будет только нарастать в ближайшие годы. Страны, отрасли, предприятия, не сумевшие эффективно применить ИИ в экономике будут менее конкурентоспособными и не получат все преимущества роста производительности, углубляя существующее неравенство. Наверно, соревнование (конкуренция) между возможностями человеческого мышления и потенциалом моделей искусственного интеллекта будет продолжено. Чтобы оно было честным, - на равных условиях, -

должно быть увеличение инвестиций у общества как в технологическую составляющую (в совершенствование моделей ИИ), так и в улучшение человеческого капитала. Цель минимум в совершенствовании человеческого капитала – соответствовать (понимать) потенциал и ограничения общего искусственного интеллекта). Цель максимум – достичь такого уровня совместной работы работников-людей и больших языковых моделей, которая обеспечивает максимальное использование человеческого потенциала и потенциала искусственного интеллекта.

## **Глава 3. ВЛИЯНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА РЫНОК ТРУДА**

### **3.1. ВЛИЯНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА РЫНОК ТРУДА В ПАРАДИГМЕ НЕРАВЕНСТВА**

Совершенствование технологий всегда определяет развитие рынка труда. Наиболее сильное воздействие в 2020-е годы оказывают технологии искусственного интеллекта. Их влияние сказывается как в краткосрочном периоде, так и долгосрочном периоде. Воздействие моделей искусственного интеллекта по-разному проявляется на работниках с низкой, средней и высокой квалификацией. Также необходимо учесть, характеризуя последствия влияния технологий искусственного интеллекта на современный рынок труда, опыт работы сотрудников по избранной специальности и их способности к переобучению. В разделе выявлены шесть особенностей влияния больших языковых моделей на современный рынок труда. Автор анализирует происходящий сейчас на рынке труда переход в принятии экономических решений от элитных экспертов к более широкому кругу работников. Этот переход приносит компаниям позитивные побочные эффекты, но действует только в краткосрочном периоде. В долгосрочном же периоде потенциал высококвалифицированных работников позволяет им повысить производительность труда за счёт моделей искусственного интеллекта значительно больше, чем низкоквалифицированным работникам. Следствием этого будет нарастание неравенства доходов на рынке труда будущего.

#### **Введение**

Данный раздел является логическим продолжением статьи, посвященной влиянию технологий искусственного интеллекта (ИИ)

на современный рынок труда<sup>170</sup>. Если в ней акцент ставился на проблеме замещения человеческого труда ИИ и поляризации создаваемых рабочих мест по уровню оплаты и условий труда, то в настоящем исследовании акцент делается на влиянии технологий искусственного интеллекта на работников низкой квалификации, средней квалификации и высокой квалификации. Это воздействие рассматривается в двух временных периодах: в краткосрочном периоде и в долгосрочном периоде. По данным экспертов ИСИЭЗ НИУ ВШЭ совокупный вклад от использования технологий ИИ во всех отраслях экономики в ВВП России составит 11.6 трлн руб. в 2030 г., а в 2035 г. достигнет значения 46.5 трлн руб.<sup>171</sup>. В связи с этим важно объективно оценить последствия воздействия моделей искусственного интеллекта на рынок труда. **Цель раздела:** обосновать позицию, что в результате применения искусственного интеллекта неравенство на рынке труда будет усиливаться, поскольку высококвалифицированные работники смогут, используя технологии искусственного интеллекта, ещё больше увеличить свои производительность и качество работы по сравнению с работниками низкой квалификации, в противовес позиции ряда исследователей что неравенство между высококвалифицированными работниками и

---

<sup>170</sup> Лукичев П. М., Чекмарев О.П. Вызовы экономики искусственного интеллекта традиционному рынку труда // Вопросы инновационной экономики. – 2023. – Том 13. – No 2. – С. 785–802. doi: [10.18334/vinec.13.2.118137](https://doi.org/10.18334/vinec.13.2.118137)

<sup>171</sup> Искусственный интеллект. Серия информационно-аналитических материалов ИСИЭЗ НИУ ВШЭ. No 17 / 2025. [Электронный ресурс] URL: <https://issek.hse.ru/mirror/pubs/share/1022067859.pdf> (дата обращения 07.03.2025)

низкоквалифицированными работниками, использующими модели ИИ, будет сокращаться, возрождая работников средней квалификации («средний класс»).

### **Особенности воздействия технологий искусственного интеллекта на современный рынок труда**

Воздействие новых технологий искусственного интеллекта на современный рынок труда имеет несколько особенностей, которые в совокупности влияют на результативность применения инноваций.

Во-первых, уточним о каких именно технологиях искусственного интеллекта идёт речь. Большие языковые модели (БЯМ) с каждым месяцем становятся всё мощнее и производительнее, получают всё большее число пользователей, предоставляют всё более корректные ответы на запросы и уже опережают в некоторых областях возможности среднего работника. С появлением в начале 2025 года китайской модели DeepSeek *произошёл качественный переворот в создании больших языковых моделей*: с акцента на технологических параметрах к ориентации на экономические показатели БЯМ, с явной монополии в создании самых совершенных моделей ИИ исключительно в США к их конкуренции с более дешёвыми и, по отдельным признакам, более эффективным моделям из КНР. Кроме того, в последнее время появилось два новых перспективных с точки зрения трудосбережения направления: «модели с рассуждением» и Агенты искусственного интеллекта. В первых «процесс мышления» модели передаётся вместе с ответом. Чем дольше «думает» модель

рассуждений, тем более точный ответ она даёт. Это может позволить уйти от господствующей в исследованиях ИИ парадигмы, что объём данных и вычислительной мощности, доступных для обучения языковой модели, улучшают её производительность. Вместо этого исследователи из Стэнфордского университета и Вашингтонского университета, первоначально действующие на уже существующем Qwen2.5 llm, произведенном Alibaba, выдвинули гипотезу, что меньший объём данных достаточно высокого качества может выполнить эту работу так же хорошо<sup>172</sup>. Успех данного направления открывает широкие возможности для «начинающих в ИИ» стран, таких как Россия, Индия, Бразилия. Агенты искусственного интеллекта представляют собой новый уровень взаимоотношений «Человек – Машина». Получив цель от человека, например, приобрести авиабилет и заказать гостиницу для командировки, они, взаимодействуя с сайтами виртуального мира, могут решить проблему. В будущем, при совершенствовании моделей, это позволит исключить простейшие и рутинные задачи от работников-людей и высвободить их труд для выполнения более сложных бизнес-операций.

Во-вторых, влияние инноваций на экономику всегда происходит неравномерно во времени. Американский футурист Рой Амара (Roy Amara) сформулировал положение, позднее названное

---

<sup>172</sup> For a fistful of dollars. The Economist. February 15th, 2025. [Электронный ресурс] URL: <https://www.economist.com/science-and-technology/2025/02/12/forget-deepseek-large-language-models-are-getting-cheaper-still> (дата обращения 28.02.2025)

его именем: *мы склонны переоценивать эффект технологии в краткосрочной перспективе и недооценивать в долгосрочной.*

Технологию можно рассматривать как состоящую из двух основных компонентов: физических артефактов, таких как лазер, компьютеры, нефтеперерабатывающие заводы, мосты — аппаратное обеспечение технологии; и социальных инструментов, которые включают методы, процедуры, ноу-хау, правила, законы — программное обеспечение технологии<sup>173</sup>. Выбор, явный или неявный, того, какие технологии применять и в каких объёмах, остаётся за обществом. Каждая национальная экономика, исходя из своих ценностей, определяет интенсивность и последовательность использования инноваций.

В-третьих, примеры эффективной замены высококвалифицированного труда большими языковыми моделями относятся не к будущему времени, а — к сегодняшней реальности. Например, в 2024 году было проведено исследование, выяснявшее могут ли большие языковые модели (БЯМ) успешно выполнять анализ финансовой отчетности способом, аналогичным профессиональному аналитику-человеку<sup>174</sup>. Результаты показали, что даже без описательной или отраслевой информации БЯМ превосходит финансовых аналитиков в своей способности

---

<sup>173</sup> Amara R. Some observations on the interaction of technology and society //Futures. – 1975. – Vol. 7. – №. 6. – P. 515–517.

<sup>174</sup> Kim A., Muhn M., Nikolaev V. Financial statement analysis with large language models //arXiv preprint arXiv:2407.17866. – 2024. [Электронный ресурс] URL: <https://arxiv.org/html/2407.17866v1> (дата обращения 12.11.2024)

предсказывать изменение направления будущих доходов фирм. Большая языковая модель демонстрирует относительное преимущество перед аналитиками-людьми в ситуациях, когда аналитики, как правило, испытывают трудности. Кроме того, особо отметим, что прогнозирование большой языковой модели (БЯМ) не вытекает из её тренировочной памяти (то есть *проявляется новое качество модели ИИ*), и что LLM (БЯМ) генерирует полезные повествовательные идеи о будущих результатах компании.

В-четвёртых, кто именно использует новые технологии? Также, как и в двух предшествующих инновациях, - персональные компьютеры и Интернет, - на первом этапе в качестве основных пользователей выступают отдельные работники, а уже вслед за ними вступают в свои права фирмы. По данным исследования, проведенного Бландин А. и др. (Blandin A. et al.) 39% американцев теперь говорят, что используют ИИ: 28% говорят, что используют его для работы, а 11% — что делают это каждый день<sup>175</sup>. Это резко контрастирует с данными об использовании больших языковых моделей предприятиями. Согласно опросу Бюро переписи населения США (Census Bureau), всего 5% американских компаний говорят, что используют технологию для производства товаров или услуг. В недавнем опросе, проведенном в 14 странах компанией Deloitte, предоставляющей профессиональные услуги, только 8% руководителей компаний заявили, что их фирмы развернули более

---

<sup>175</sup> Blandin A. , Bick A., Caplan T., Caplan A. Measuring Trends in Work From Home: Evidence from Six US Datasets //Available at SSRN 4953448. – 2024.

половины своих экспериментов с генеративным ИИ<sup>176</sup>. Автор тоже может отметить, что по результатам его опросов студентов петербургских ВУЗов в январе-феврале 2025 г. 95% учащихся используют или ChatGPT, или DeepSeek, однако среди моих коллег-преподавателей этот процент значительно ниже.

В-пятых, внедрение технологий искусственного интеллекта в бизнес-практику происходит при негативном отношении к ним общества. Данные опроса, проводившегося Pew Research Center на протяжении 2024 г., показывают, что общественный оптимизм в отношении влияния ИИ на работу низок. В то время как 73% опрошенных экспертов в области искусственного интеллекта говорят, что ИИ окажет очень или в некоторой степени положительное влияние на то, как люди выполняют свою работу в течение следующих 20 лет, эта доля падает до 23% среди взрослых в США. На протяжении предыдущей работы Pew Research Center (2021 г.) общественность с опаской относилась к роли искусственного интеллекта в потере рабочих мест. В текущем опросе 64% общественности считают, что ИИ приведет к сокращению рабочих мест в течение следующих 20 лет.<sup>177</sup>

---

<sup>176</sup> A nasty case of pilotitis. The Economist. November 9th, 2024. [Электронный ресурс] URL: <https://www.economist.com/business/2024/11/04/why-your-company-is-struggling-to-scale-up-generative-ai> (дата обращения 28.02.2025)

<sup>177</sup> McClain C., Kennedy B., Gottfried J., Anderson M., Pasquini G. How the U.S. Public and AI Experts View Artificial Intelligence. [Электронный ресурс] URL: [https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence/?utm\\_source=Pew+Research+Center&utm\\_campaign=82cfde2055-Internet\\_Science\\_2025\\_04\\_29&utm\\_medium=email&utm\\_term=0\\_-82cfde2055-399474221](https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence/?utm_source=Pew+Research+Center&utm_campaign=82cfde2055-Internet_Science_2025_04_29&utm_medium=email&utm_term=0_-82cfde2055-399474221). (дата обращения 30.04.2025)

В-шестых, степень влияния моделей искусственного интеллекта на конкретную профессию определяется тем носит ли оно замещающий характер или дополняющий характер. Около сорока процентов работников во всём мире и шестьдесят процентов в развитых экономиках работают в профессиях с высоким уровнем риска от применения технологий ИИ<sup>178</sup>. При этом наблюдается чёткая корреляция с уровнем дохода в стране. Исследование Международного валютного фонда (International Monetary Fund) выявило, что в развитых странах 27% рабочих мест с высокой взаимодополняемостью и 33% рабочих мест с низкой взаимодополняемостью; в странах с развивающимися рынками 16% рабочих мест с высокой взаимодополняемостью и 24% рабочих мест с низкой взаимодополняемостью; в странах с низким уровнем дохода 8% рабочих мест с высокой взаимодополняемостью и 18% рабочих мест с низкой взаимодополняемостью.

Конкретизируем шестую особенность воздействия технологий ИИ на современный рынок труда. Более интенсивное влияние технологий искусственного интеллекта на рабочие места в развитых странах объяснимо более высокой долей занятости в профессиях с высокой концентрацией когнитивных задач, например в управленческих профессиях, по сравнению с работниками в других

---

<sup>178</sup> Cazzaniga M., Jaumotte M. F., Li L., Melina M. G., Panton A. J., Pizzinelli, C., ... Tavares M. M. Gen-AI: Artificial intelligence and the future of work. – International Monetary Fund, 2024. [Электронный ресурс] URL: <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379> (дата обращения 28.02.2025)

государствах. Выполняют эти задачи, как правило, сотрудники с высшим образованием, доля которых в развитых странах и в некоторых странах с развивающимися рынками значительно выше, чем в остальном мире.

Как показывает опыт внедрения предыдущих инноваций именно работники с высшим образованием демонстрируют большую способность к переобучению и поэтому, на наш взгляд, обладают преимущественной способностью к переходу на должности, имеющие высокий потенциал взаимодополняемости с искусственным интеллектом. Поэтому высокий риск воздействия ИИ на рынок труда при соответствующей промышленной политике поглощается возможностями взаимодополняемости рабочих мест с ИИ. Возникающая при этом комплементарность позволяет производить товары, ценность которых выше, чем при раздельном использовании труда работника и модели искусственного интеллекта. Выделим здесь также влияние институциональной структуры страны на угрозы заменяемости рабочих мест искусственным интеллектом.

Для этого сопоставим профессии судей и офисного персонала. И те, и другие обладают высоким риском из-за применения моделей ИИ. Однако положение судей защищено законами и общественными нормами, а офисных сотрудников – нет. Таким образом, из-за возрастающего применения искусственного интеллекта неравенство на рынке труда будет усиливаться. Например, работники профессий, более уязвимые к замене ИИ, с большей вероятностью испытают неблагоприятные шоки доходов, в то время как работники

дополняющих профессий могут получить более высокую отдачу от своего труда<sup>179</sup>. В страновом аспекте отметим, что в странах с низкими доходами и частично в странах с развивающимися рынками, имеющих более низкие доли рабочих мест с взаимодополняемостью с искусственным интеллектом, может усилиться неравенство в доходах по сравнению с развитыми государствами.

Автор поддерживает позицию противодействия усилению неравенства между странами и регионами из-за применения моделей искусственного интеллекта. В частности, в рамках BRICS перспективны обмен опытом в реализации национальных стратегий и сбор лучших практик по комплексным мерам цифрового развития. Одним из ключевых направлений такого сотрудничества может стать создание совместных программ и платформ для обмена знаниями и практиками<sup>180</sup>.

Также отметим условность выделения влияния искусственного интеллекта на профессиональную деятельность. В одних случаях ИИ может дополнять работу, выполняемую людьми, в других случаях искусственный интеллект будет заменять работу, выполняемую до этого людьми, в-третьих, ИИ будет частично

---

<sup>179</sup> Pizzinelli C., Panton A. J., Tavares M. M., Cazzaniga M., Li L. Labor market exposure to AI: Cross-country differences and distributional implications. – International Monetary Fund, 2023. [Электронный ресурс] URL: <https://www.imf.org/en/Publications/WP/Issues/2023/10/04/Labor-Market-Exposure-to-AI-Cross-country-Differences-and-Distributional-Implications-539656> (дата обращения 12.01.2025)

<sup>180</sup> Пересмотр цифрового неравенства на рынке труда стран БРИКС / И. В. Лазанюк, И. А. Пугачева, В. А. Семчук, А. О. Неведомская // Вопросы инновационной экономики. – 2025. – Т. 15, № 1. – DOI 10.18334/vinec.15.1.122679

дополнять и частично замещать труд работников. В качестве критерия выделим положение Э. Фелтена, М. Раджа, Р. Симанса (Felten E., Raj M., Seamans R.)<sup>181</sup>, <sup>182</sup> о степени совпадения между возможностями приложениями ИИ и способностями человека в конкретной профессии.

### **Возрастание или смягчение неравенства на рынке труда за счёт применения моделей искусственного интеллекта?**

Увеличивающееся применение в экономике больших языковых моделей (БЯМ) ставит *исследовательский вопрос № 1*: окажется ли их развитие фактором, уменьшающим неравенство между работниками разного уровня квалификации? Производным от него является *исследовательский вопрос № 2*: способствует ли эволюция современных больших языковых моделей увеличению числа работников средней квалификации и, следовательно, возрождению среднего класса.

Технологии искусственного интеллекта бурно развиваются и в последние полтора года среди исследователей сформировались две противоположные позиции об их воздействии на неравенство на

---

<sup>181</sup> Felten E., Raj M., Seamans R. Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses //Strategic Management Journal. – 2021. – Vol. 42. – №. 12. – P. 2195–2217. [Электронный ресурс] URL: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.3286> (дата обращения 10.11.2024)

<sup>182</sup> Felten E., Raj M., Seamans R. How will language modelers like chatgpt affect occupations and industries? //arXiv preprint arXiv:2303.01157. – 2023. [Электронный ресурс] URL: <https://arxiv.org/pdf/2303.01157> (дата обращения 12.11.2024)

современном рынке труда. Первый подход был чётко выражен в работах Д. Аутора (Autor D., 2024)<sup>183</sup>, Э. Бриньольфссон, Д. Ли, Л.Р. Рэймонд (Brynjolfsson E., Li D., Raymond L. R., 2023)<sup>184</sup>, Цз. Чэнь, Дж. Чан (Chen Z., Chan J., 2024)<sup>185</sup>. Он может быть назван «уровнительным». Наиболее полно он представлен в статье Д. Аутора<sup>186</sup>, который указывает, что искусственный интеллект преобразует рынок труда за счёт изменения ценности и природы человеческой экспертизы. Благодаря способности ИИ сплести информацию и правила с приобретенным опытом для поддержки принятия решений, его можно применять, чтобы позволить большему числу работников, обладающих дополнительными знаниями, выполнять некоторые из задач принятия решений с более высокими ставками, которые в настоящее время приписываются элитным экспертам, например, медицинская помощь врачам,

---

<sup>183</sup> Autor D. Applying AI to rebuild middle class jobs. – National Bureau of Economic Research, 2024. – №. w32140. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w32140/w32140.pdf](https://www.nber.org/system/files/working_papers/w32140/w32140.pdf) (дата обращения 21.02.2025)

<sup>184</sup> Brynjolfsson E., Li D., Raymond L. R. Generative AI at work. – National Bureau of Economic Research, 2023. – №. w31161. [Электронный ресурс] URL: <https://www.nber.org/papers/w31161> (дата обращения 21.02.2025)

<sup>185</sup> Chen Z., Chan J. Large language model in creative work: The role of collaboration modality and user expertise //Management Science. – 2024. – Vol. 70. – №. 12. – P. 9101–9117. [Электронный ресурс] URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4575598](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4575598) (дата обращения 21.02.2025)

<sup>186</sup> Autor D. Applying AI to rebuild middle class jobs. – National Bureau of Economic Research, 2024. – №. w32140. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w32140/w32140.pdf](https://www.nber.org/system/files/working_papers/w32140/w32140.pdf) (дата обращения 21.02.2025)

производство документов юристам, программирование программ инженерам и высшее образование профессорам.

Таким образом, по мнению Д. Аутора (Autor D.) применение сегодняшних моделей искусственного интеллекта позволит «простым работникам», не обременённым многолетним образованием и долгим опытом работы по специальности, достичь тех же результатов, что у элитных экспертов. Это приведёт, по его мнению, к возрождению среднего класса, - работников средней квалификации и к снижению ценности экспертизы сегодняшних элитных экспертов. – юристов, врачей, профессоров, инженеров. Главный критерий процветания той или иной профессии и, как следствия, получения образования по этой специальности, по его мнению, - ценность для общества в данный момент человеческой экспертизы. Это также изменит взаимоотношения между ИИ и работниками. По мнению Д. Аутора суперспособность искусственного интеллекта отходить от сценария, импровизировать на основе обучения и опыта позволяет ему участвовать в экспертных суждениях — способность, которая до сих пор была прерогативой элитных экспертов, и поскольку возможности ИИ в экспертных суждениях будут всё более надёжными, острыми и доступными в предстоящие годы, они станут почти повсеместным присутствием в нашей рабочей жизни. Его основная роль будет заключаться в консультировании, обучении и предупреждении лиц, принимающих

решения, когда они применяют экспертные суждения<sup>187</sup>. С этой идиллической перспективой нельзя согласиться так как уже сегодняшний опыт показывает, что люди-работники будут перекладывать всё больше принятия решений на модели искусственного интеллекта, а доверие к решениям ИИ больше, чем к решениям людей. Есть опасность, что в долгосрочном периоде, как мы показали в статье<sup>188</sup>, люди перестанут повышать свою квалификацию, опасаясь негативного сравнения с большими языковыми моделями, и будут деградировать и как работники, и как граждане общества. Как следствие, лица, принимающие решения, станут вымирающей профессией. Особенно с учётом прогресса Агентов ИИ.

Обратимся к исследованиям, приводящим практические аргументы «за» «уравнительный подход» в краткосрочном периоде. Применение генеративного ИИ неодинаково воздействует на работников высокой квалификации и на работников низкой квалификации. Широкомасштабное исследование 5179 работников службы поддержки клиентов показало, что доступ к генеративного ИИ-помощнику повышает производительность, измеряемую по проблемам, решенным в час, в среднем на 14%, включая улучшение

---

<sup>187</sup> Autor D. Applying AI to rebuild middle class jobs. – National Bureau of Economic Research, 2024. – №. w32140. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w32140/w32140.pdf](https://www.nber.org/system/files/working_papers/w32140/w32140.pdf) (дата обращения 21.02.2025)

<sup>188</sup> Лукичев, П. М. Принятие решений в современной экономике: искусственный интеллект vs поведенческая экономика / П. М. Лукичев // Вопросы инновационной экономики. – 2024. – Т. 14, № 3. – С. 649–666. – DOI 10.18334/vines.14.3.121070

на 34% для новичков и низкоквалифицированных работников, но с минимальным влиянием на опытных и высококвалифицированных работников<sup>189</sup>. Позитивным последствием для компании в целом является то, как клиенты относятся к персоналу, что измеряется настроением их сообщений в чате. Было также обнаружено, что клиенты реже подвергают сомнению компетентность работников, прося поговорить с руководителем. Эти изменения происходят наряду со значительным снижением текучести кадров, что обусловлено удержанием новых сотрудников. В условиях низкой безработицы, когда все компании жёстко конкурируют за кадры, такое следствие применения генеративного ИИ-помощника очень ценно. Данный пример показывает те несомненные выгоды, которые приобретают фирмы, использующие модели ИИ в своих бизнес-процессах в краткосрочном периоде. Аналогичные результаты были получены в исследовании Цз. Чэнь, Дж. Чан<sup>190</sup>, которые провели эксперимент, в котором поручили экспертам и не экспертам написать рекламный текст с помощью большой языковой модели (БЯМ) и без неё. Использование БЯМ повысило качество конечных рекламных копий для неспециалистов, однако не дало существенных преимуществ и фактически нанесло ущерб опытным пользователям. «Уравнительный» характер использования моделей искусственного

---

<sup>189</sup> Brynjolfsson E., Li D., Raymond L. R. Generative AI at work. – National Bureau of Economic Research, 2023. – №. w31161. [Электронный ресурс] URL: <https://www.nber.org/papers/w31161> (дата обращения 21.02.2025)

<sup>190</sup> Chen Z., Chan J. Large language model in creative work: The role of collaboration modality and user expertise //Management Science. – 2024. – Vol. 70. – №. 12. – P. 9101–9117. [Электронный ресурс] URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4575598](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4575598) (дата обращения 21.02.2025)

интеллекта проявляется, как показала исследование Дж. Х. Чой, А. Б. Монахан, Д. Шварц (Choi J. H., Monahan A. B., Schwarcz D.)<sup>191</sup>, даже в такой консервативной сфере как юриспруденция. Помощь ИИ улучшила качество юридического анализа неравномерно — там, где она была полезна вообще, самые низкоквалифицированные участники увидели самые большие улучшения. Были выявлены также побочные эффекты: 1) использование модели искусственного интеллекта сэкономило работникам примерно одинаковое количество времени независимо от их базовой скорости; 2) в последующих опросах участники сообщили о возросшей удовлетворённости от использования ИИ для выполнения юридических задач.

Таким образом, происходит дифференциация работников в соответствии с возрастающей доступностью искусственного интеллекта: ценность элитных экспертов (сегодня), - врачей, юристов, инженеров-программистов, профессоров, - снижается, а ценность не элитных работников подтягивается (может подтянуться) до их уровня. Отметим, что элитные эксперты характерны для развитых стран. В России ни врачи, ни профессора не ощущали себя в последние десятилетия элитными экспертами. Общий рост производительности с применением ИИ достигается только за счёт улучшения производительности малоопытных работников до уровня опытных. Прямо как в Библии «И стали последние первыми, а

---

<sup>191</sup> Choi J. H., Monahan A. B., Schwarcz D. *Lawyering in the age of artificial intelligence* // *Minn. L. Rev.* – 2024. – Vol. 109. – P. 147. [Электронный ресурс] URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4626276](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4626276) (дата обращения 21.02.2025)

первые последними». Но дальше роста нет, то есть, — это устранение «узких мест» не более.

Вместе с тем, помогая решить текущие вопросы, применение генеративного ИИ-помощника породило две сложные проблемы в долгосрочном периоде. Обе эти проблемы связаны с правами лучших работников, наиболее квалифицированных и наиболее опытных.

Во-первых, лучшие работники, как правило, не получают оплату за свой вклад в данные обучения, которые системы искусственного интеллекта используют для сбора и распространения своих навыков. Однако без этого вклада системы ИИ могут быть менее эффективными в обучении решения новых проблем<sup>192</sup>.

Во-вторых, при решении новых задач, для того чтобы выстоять в борьбе с конкурентами, компаниям нужны высококвалифицированные и опытные сотрудники. Как их мотивировать к качественному труду в условиях использования генеративного ИИ-помощника и самое главное, к дальнейшему развитию, превышающему сегодняшний уровень опытных работников? Не получится ли, что компании, не использующие в работе помощь искусственного интеллекта, создадут (сформируют) опытных сотрудников, способных решать новые (не имеющие

---

<sup>192</sup> Brynjolfsson E., Li D., Raymond L. R. Generative AI at work. – National Bureau of Economic Research, 2023. – №. w31161. [Электронный ресурс] URL: <https://www.nber.org/papers/w31161> (дата обращения 21.02.2025)

аналогов) задачи, а фирмы, использующие сейчас ИИ, проиграют им в конкуренции?

Второй подход может быть определён как «**возрастание неравенства**» между работниками низкой квалификации и работниками высокой квалификации. Он представлен в исследованиях А. Кима и др. (Kim A. et al.)<sup>193</sup>,<sup>194</sup>, Н. Отиса с соавторами (Nicholas Otis et al.)<sup>195</sup> [23], Э. Тонер-Роджерса (Aidan Toner-Rodgers)<sup>196</sup>.

Любую работу можно рассматривать как конструктор Lego, а ИИ или помогает работнику в ней, или заменяет его. После первоначального эффекта применения большой языковой модели к определенной профессии, который позволяет развить потенциал низкоквалифицированных и малоопытных работников, в любой компании встаёт вопрос о повышении производительности труда в долгосрочном периоде. Здесь, как показывают исследования, преимущество высококвалифицированных сотрудников становится явным. Имея в распоряжении одни и те же возможности, как и

---

<sup>193</sup> Kim A., Kim D. S., Muhn M., Nikolaev V. V., So E. C. AI, Investment Decisions, and Inequality //Chicago Booth Accounting Research Center Research Paper, Fama-Miller Working Paper, MIT Sloan Research Paper. – 2024.

<sup>194</sup> Kim A., Muhn M., Nikolaev V. Financial statement analysis with large language models //arXiv preprint arXiv:2407.17866. – 2024. [Электронный ресурс] URL: <https://arxiv.org/html/2407.17866v1> (дата обращения 12.11.2024)

<sup>195</sup> Otis N., Clarke R. P., Delecourt S., Holtz D., Koning R. The uneven impact of generative AI on entrepreneurial performance //Available at SSRN 4671369. – 2023.

<sup>196</sup> Toner-Rodgers A. Artificial intelligence, scientific discovery, and product innovation //arXiv preprint arXiv:2412.17866. – 2024. [Электронный ресурс] URL: <https://arxiv.org/pdf/2412.17866> (дата обращения 02.03.2025)

низкоквалифицированные работники, они эффективнее используют возможности искусственного интеллекта. Так, Алекс Ким (Alex Kim) и его соавторы провели эксперимент с распределением средств в смоделированном портфеле, который показал, что опытные инвесторы получили почти на 10% более высокую прибыль с помощью ИИ; а менее опытные инвесторы увидели прирост лишь в 2%<sup>197</sup>. Это произошло за счёт лучшего использования опытными инвесторами информации из конференций по прибыли.

Аналогичный эксперимент с предпринимателями провели Николас Отис с соавторами (Nicholas Otis et al.). Более сильные кенийские предприниматели увеличили свою прибыль более чем на 15% с помощью помощника на основе искусственного интеллекта, а у тех, кто не преуспел, прибыль упала<sup>198</sup>. Разница заключалась только в том, как они применяли рекомендации ИИ. Неопытные предприниматели удовлетворялись общим советам модели искусственного интеллекта, например, больше рекламировать. Напротив, опытные предприниматели достигали высоких показателей за счёт использования искусственного интеллекта для поиска индивидуальных решений, например, для обеспечения новых источников электроэнергии во время отключений электроэнергии. В некоторых сферах человеческой деятельности, требующих как в

---

<sup>197</sup> Kim A., Kim D. S., Muhn M., Nikolaev V. V., So E. C. AI, Investment Decisions, and Inequality //Chicago Booth Accounting Research Center Research Paper, Fama-Miller Working Paper, MIT Sloan Research Paper. – 2024.

<sup>198</sup> Otis N., Clarke R. P., Delecourt S., Holtz D., Koning R. The uneven impact of generative AI on entrepreneurial performance //Available at SSRN 4671369. – 2023.

научных исследованиях, не только определенной квалификации, но и опыта работы, и креативности, различия в результативности ещё разительнее. Так, исследование Эйдана Тонер-Роджерса (Aidan Toner-Rodgers) показало, что использование инструмента искусственного интеллекта для помощи в открытии материалов почти удвоило производительность ведущих исследователей, не оказав при этом измеримого влияния на нижнюю треть<sup>199</sup> [25]. Образно говоря, имея разный уровень квалификации, одни исследователи видели ситуацию поверхностно, другие – смотрели вглубь проблемы. Менее опытные ученые с трудом отфильтровывали полезные результаты от нерелевантных. Применение ИИ позволило элитным ученым, указать желаемые характеристики, а затем сгенерировать материалы-кандидаты, которые, как предполагалось, будут обладать этими свойствами. Далее, на основе больших объёмов предметной экспертизы, они смогли выявлять перспективные предложения и отбрасывать плохие<sup>200</sup>. Это исследование также показало, что элитные эксперты «уходят в тень» из-за использования больших языковых моделей далеко не во всех сферах экономики.

Возможно ли дальнейшее повышение производительности низкоквалифицированных работников после первоначального

---

<sup>199</sup> Toner-Rodgers A. Artificial intelligence, scientific discovery, and product innovation //arXiv preprint arXiv:2412.17866. – 2024. [Электронный ресурс] URL: <https://arxiv.org/pdf/2412.17866> (дата обращения 02.03.2025)

<sup>200</sup> Toner-Rodgers A. Artificial intelligence, scientific discovery, and product innovation //arXiv preprint arXiv:2412.17866. – 2024. [Электронный ресурс] URL: <https://arxiv.org/pdf/2412.17866> (дата обращения 02.03.2025)

выравнивания с высококвалифицированными сотрудниками? Да, возможно, но это потребует прохождения ими переобучения и времени, которые будут зависеть также от их способностей и возраста. Это создаёт сложную экономическую проблему, поскольку, как отмечал А. Зарифхонарвар (Zarifhonarvar A.), существует неэластичное предложение рабочей силы, которое не может сильно измениться в краткосрочной перспективе. Это связано с тем, что любая форма переквалификации работников требует времени<sup>201</sup>. Данный переходный период может создать проблемы для работников, которым для сохранения своей конкурентоспособности на рынке труда потребуется дополнительное обучение или повышение квалификации. Если они смогут повысить свою квалификацию, то станут высококвалифицированными специалистами, а если – нет, то останутся на прежних позициях. Конкретизируя в контексте России, анализ исследования организаций, проведённый И. А. Волошиной, И. Т. Перовой, А. Р. Тхоржевской, показал, что в ближайшие 5 лет наиболее распространёнными будут такие технологии ИИ, как компьютерное зрение, интеллектуальный анализ данных, автоматизация процессов, в том числе с участием роботов, а также технологии анализа данных, основанные на алгоритмах глубинного обучения<sup>202</sup>.

---

<sup>201</sup> Zarifhonarvar A. Economics of ChatGPT: A labor market view on the occupational impact of artificial intelligence // Journal of Electronic Business & Digital Economics. – 2024. – Vol. 3. – №. 2. – P. 100–116. <https://doi.org/10.1108/JEBDE-10-2023-0021>

<sup>202</sup> Волошина И. А., Перова И. Т., Тхоржевская А. Р. Применение технологий искусственного интеллекта в деятельности организаций: масштабы, влияние на профессии, риски // Креативная экономика. – 2024. – Том 18. – № 12. – С. 3285–3310. doi: [10.18334/ce.18.12.122257](https://doi.org/10.18334/ce.18.12.122257)

Последствием этого для национальной экономики является усиления значения системы высшего образования и системы переподготовки для улучшения конкурентоспособности национального рынка труда.

Возрастающую противоречивость использования моделей искусственного интеллекта для дополнения деятельности работников показывает следующее исследование, осуществлённое Ф. Делл'Аква и его соавторами (Dell'Acqua F. et al.), совместно с Boston Consulting Group <sup>203</sup>. Испытуемые были случайным образом распределены по трём условиям: без доступа к ИИ, с доступом к ИИ GPT-4 и с доступом к ИИ GPT-4 с оперативным обзором инженерных разработок. Гипотезой данного исследования было, что возможности ИИ создают “jagged technological frontier” («неровную технологическую границу»), где некоторые задачи легко выполняются искусственным интеллектом, в то время как другие, хотя и кажутся схожими по уровню сложности, находятся за пределами текущих возможностей ИИ.

Как и в других аналогичный экспериментах консультанты, использующие ИИ, были значительно более продуктивны (выполнили на 12,2% больше задач в среднем и выполнили задачи на 25,1% быстрее) и давали значительно более качественные

---

<sup>203</sup> Dell'Acqua F., McFowland III E., Mollick E. R., Lifshitz-Assaf H., Kellogg K., Rajendran S., ... & Lakhani K. R. Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality //Harvard Business School Technology & Operations Mgt. Unit Working Paper. – 2023. – №. 24-013. [Электронный ресурс] URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4573321](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321) (дата обращения 12.11.2024)

результаты (более чем на 40% более высокого качества по сравнению с контрольной группой). Вместе с тем, для задачи, выбранной как находящаяся за пределами границы, консультанты, использующие ИИ, были на 19 процентных пунктов менее склонны выдавать правильные решения по сравнению с теми, кто не использовал искусственный интеллект<sup>204</sup>. В связи с этим неизбежно возникает практический вопрос: сотрудникам лучше использовать ИИ в работе, или, чтобы не «поглупеть» для решения сложных проблем, не использовать большие языковые модели?

Поскольку одним из авторов исследования был Э. Моллик (Mollick E.) автор книги Co-Intelligence<sup>205</sup>, то применим к анализу его подход. Он разделяет во взаимодействии человека и искусственного интеллекта два варианта. Вариант 1 называется «кентавр», когда есть чёткое разделение задач на человеческие и ИИ (так же как тело кентавра явно разделялось на человеческую и лошадиную части). Вариант 2 называется «киборг», когда существует привлечение искусственного интеллекта к решению всех текущих задач, совмещая потенциал двух интеллектов и позволяя им гибко взаимодействовать. Таким образом мы можем сделать вывод, что для задачи, находящейся за пределами “jagged technological frontier”

---

<sup>204</sup> Dell'Acqua F., McFowland III E., Mollick E. R., Lifshitz-Assaf H., Kellogg K., Rajendran S., ... & Lakhani K. R. Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality //Harvard Business School Technology & Operations Mgt. Unit Working Paper. – 2023. – №. 24-013. [Электронный ресурс] URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4573321](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321) (дата обращения 12.11.2024)

<sup>205</sup> Mollick E., Mollick E. Co-Intelligence. – Random House UK, 2024.

(«неровной технологической границы») лучше действуют консультанты-«кентавры», а для решения обычных задач – консультанты –«киборги». Далее в реальности встаёт проблема выбора: если сейчас эффективнее интеграция работников и искусственного интеллекта как консультантов - «киборгов», то использование сегодня (в краткосрочном периоде) работников-«кентавров» несёт более высокие издержки и, следовательно, приносит меньше прибыли компаниям и, поэтому, «кентавры» не будут применяться и «вымрут». Что будет с компаниями в долгосрочном периоде по мере появления новых, не существовавших ранее задач?

Более того, наличие высококвалифицированных работников выступает в качестве ограничивающего фактора для внедрения технологий искусственного интеллекта. Так, в исследовании, проведенном Л. Зуйя, Н. М. Юнус (Zouya L., Yunus N. M.) на определение воздействия патентных заявок на ИИ на производительность труда в экономике Китая, было выявлено, что патентные заявки на технологии искусственного интеллекта статистически влияют на производительность труда в народном хозяйстве КНР в целом, но, не подтверждаются исследованиями в сельскохозяйственном секторе, что объясняется авторами нехваткой квалифицированной рабочей силы<sup>206</sup>.

---

<sup>206</sup> Zouya L., Yunus N. M. A preliminary study on artificial intelligence and labour productivity in China //International Business Education Journal. – 2024. – Vol. 17. – №. 2. – P. 12–25. <https://doi.org/10.37134/ibej.Vol17.2.2.2024>

Особо стоит отметить аргументы в пользу усиления неравенства на современном рынке труда, выдвинутые Б. Молл, Л. Рэйчел, П. Рестрепо (Moll B., Rachel L., Restrepo P.)<sup>207</sup>. Они, в отличие от традиционного подхода, доказывают, что преимущества новых технологий достаются не только высококвалифицированной рабочей силе, но и владельцам капитала в виде более высоких доходов от капитала, что основано на том, что долгосрочное предложение капитала в их модели имеет восходящий наклон. Проявляется это по двум каналам. Во-первых, автоматизация увеличивает неравенство богатства и доходов от капитала, повышая доходность богатства. Во-вторых, автоматизация с большей вероятностью приведёт к стагнации заработной платы и, следовательно, к стагнации доходов в нижней части распределения доходов.

### **Выводы и дальнейшие исследования**

Воздействие технологий искусственного интеллекта на современный рынок труда происходит примерно также, как это делали предшествующие волны инноваций. Единственное, но принципиальное отличие технологий ИИ от предшествующих состоит в том, что они автоматизируют (устраняют) не физические, рутинные задачи, выполняемые работниками, а творческие, интеллектуальные задачи. Модели искусственного интеллекта в

---

<sup>207</sup> Moll B., Rachel L., Restrepo P. Uneven growth: automation's impact on income and wealth inequality // *Econometrica*. – 2022. – Vol. 90. – №. 6. – P. 2645–2683. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w28440/w28440.pdf](https://www.nber.org/system/files/working_papers/w28440/w28440.pdf) (дата обращения 12.11.2024)

процессе машинного обучения приобретают «неявные знания» из анализируемого опыта работников. Применение новейших больших языковых моделей позволяет в краткосрочном периоде повысить производительность труда работников низкой квалификации и без опыта до уровня высококвалифицированных специалистов и, может быть, даже до уровня элитных экспертов. Это позволяет увеличить общий уровень производительности труда в компаниях и приносит ряд положительных побочных эффектов. Среди них выделим: экономия времени сотрудников на выполнение одной задачи, возросшая удовлетворенность от использования ИИ для решения профессиональных проблем, значительное снижение текучести кадров, что обусловлено удержанием новых сотрудников. Однако после первого эффекта выравнивания дальнейшее повышение производительности труда возможно только за счёт использования искусственного интеллекта высококвалифицированными и опытными работниками, которые могут извлечь из модели ИИ, как показывают приведенные исследования, большие и лучшие результаты по сравнению с низкоквалифицированными сотрудниками.

Неравенство в доходах и в условиях труда между работниками низкой квалификации и работниками высокой квалификации будет усиливаться по мере продвижения технологий ИИ как в отраслевом аспекте, так и между развитыми странами, государствами с развивающимися рынками и странами с низкими доходами. Степень неравенства будет определяться как интенсивностью применения моделей искусственного интеллекта в народном хозяйстве, так и

возможностями по переобучению персонала системами высшего образования и повышения квалификации. Таким образом, ответы на два исследовательских вопроса будут отрицательные: после первоначального эффекта выравнивания неравенство в доходах и условиях работы между работниками высокой квалификации и низкой квалификации будет увеличиваться.

Воздействие технологий искусственного интеллекта на рынок труда будет усиливаться по мере их совершенствования. Сейчас идёт процесс их количественного накопления, который в перспективе 3–4 лет качественно расширит производственные возможности работников будущего. Экономические проблемы, и, прежде всего, оценивание эффективности моделей искусственного интеллекта выходят на первый план по сравнению с чисто технологическими. Экономика ИИ вступает в свои права<sup>208</sup>. Сейчас речь в основном идёт о снижении издержек при создании модели искусственного интеллекта, или при эксплуатации большой языковой модели. Экономическая оценка выгод от использования моделей искусственного интеллекта пока вторична. В перспективе можно соединить ИИ, в частности агентов искусственного интеллекта с цифровыми двойниками и осуществить многие другие трудосберегающие технологии. Ведь чем модели искусственного

---

<sup>208</sup> Лукичев П. М., Чекмарев О. П. Экономика искусственного интеллекта: перспективы и риски. Санкт-Петербург, 2023. [Электронный ресурс]. URL: <https://motivtrud.ru/PCost/research/ii20231.pdf> (дата обращения 02.02.2025)

интеллекта будут дешевле (доступнее), тем больше будет выгодных сфер их применения.

## **3.2. ЧЕЛОВЕЧЕСКИЙ КАПИТАЛ В ЭПОХУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

Расширение применения технологий искусственного интеллекта является сегодня основным катализатором экономического развития. Искусственный интеллект постепенно меняет роль и значение человеческого капитала в экономике. В начале раздела рассматривается генезис концепции человеческого капитала. Для исследования современного состояния человеческого капитала выделяются основные факторы влияния на него: система образования, искусственный интеллект, трудовая миграции. Данный раздел акцентируется на искусственном интеллекте. Воздействие технологий искусственного интеллекта анализируется автором на трёх уровнях: 1) уровне национальной экономики, 2) уровне фирмы, 3) уровне отдельного работника. Автор вводит понятие «национальный человеческий капитал». Его ценность определяется как полученными в системе образования знаниями, умениями: навыками, так и наличием качественных рабочих мест, позволяющих реализовать потенциал выпускников. Среди последних технологий выделим «Агентов искусственного интеллекта» использование которых может как заменить персонал компаний, так и усилить трудовой потенциал специалистов. В разделе делается вывод, что не только искусственный интеллект оказывает влияние на качества «человеческого капитала», но и развитие моделей ИИ возможно только под воздействием высокого уровня человеческого капитала, что определяет необходимость его постоянного совершенствования.

### **Введение**

Воздействие технологий искусственного интеллекта на человеческий капитал проявляется по нескольким направлениям. Во-первых, модели ИИ меняют требования к качеству современных рабочих мест. Это проявляется через разную степень воздействия искусственного интеллекта: часть функций одних заменяется технологиями ИИ, в то время как другие рабочие места дополняются

возможностями искусственного интеллекта, становясь более производительным. Во-вторых, модели ИИ напрямую влияют на ценность «твёрдых» навыков и «мягких» навыков, которыми сейчас обладает работник на рынке труда. Некоторые из них становятся ненужными, а значение других резко увеличивается. В-третьих, модели искусственного интеллекта создают условия для формирования нового разделения труда как на уровне предприятий, так и на уровне народного хозяйства. **Цель раздела:** выявить влияние моделей искусственного интеллекта на человеческий капитал сейчас и в будущем.

### **Человеческий капитал: возникновение концепции и её ЭВОЛЮЦИЯ**

Концепция «человеческого капитала» возникла в 1950-1970-х годах. Хотя ещё знаменитый английский экономист Альфред Маршалл писал в конце XIX века, что самый ценный из всех капиталов — это тот, который вложен в людей<sup>209</sup>. У истоков теории

---

<sup>209</sup> Маршалл А. Принципы экономической науки. – М., Издательская группа «Прогресс», 1993.

стояли такие экономисты как Джейкоб Минсер (Jacob Mincer)<sup>210</sup>, Теодор Шульц (Theodore Schultz)<sup>211</sup>, Гэри Беккер (Gary Becker)<sup>212</sup>.

Выделим вклад каждого из них в создание концепции человеческого капитала. Вклад Д. Минсера связан с анализом человеческого капитала не только на макроуровне, как фактора экономического роста, но и на микроуровне. Последнего удалось достичь за счёт оценивания эмпирически нормы прибыли от образования, что позволило сопоставлять её с другими видами инвестиций и применять обычный экономический подход «Издержки – Выгоды»<sup>213</sup>. Развивая эту концепцию, Гэри Беккер отмечал, что все формы капитала приносят доход и другие полезные результаты в течение длительных периодов времени, и поэтому расходы на образование, обучение, медицинское обслуживание и т. д. являются инвестициями в капитал. Основной его особенностью является то, что нельзя отделить человека от его или её знаний, навыков, здоровья или ценностей так, как это возможно для перемещения финансовых и физических активов, пока владелец

---

<sup>210</sup> Mincer J. Investment in human capital and personal income distribution //Journal of political economy. – 1958. – Vol. 66. – №. 4. – P. 281-302.

<sup>211</sup> Schultz T. W. Investment in human capital //The American economic review. – 1961. – Vol. 51. – №. 1. – P. 1–17

<sup>212</sup> Becker G. S. Human capital //New York: National Bureau of Economic Research. – 1964.

<sup>213</sup> Mincer J. Investment in human capital and personal income distribution //Journal of political economy. – 1958. – Vol. 66. – №. 4. – P. 281-302.

остается на месте<sup>214, 215</sup>. Исходя из сегодняшних знаний и ценностей наиболее спорным у Г. Беккера, как активного сторонника неоклассического подхода, являются положения о том, что инвестиции в человеческий капитал являются рациональными ответами на расчёт ожидаемых затрат и выгод, и о стабильных предпочтениях.<sup>216</sup> Объективный анализ показывает, что человек при принятии экономических решений не полностью рационален, существуют и другие виды рациональности: ограниченная рациональность, неполная рациональность, иррациональность. Выбор индивидуумом место обучения, как показывают опросы автора и, главное, реальность часто не подразумевают расчёта выгод и издержек. Так опросы студентов петербургских ВУЗов, проведенные автором в 2024 году (N = 167) и в 2025 году (N = 193), показали, что три четверти абитуриентов приняли решение о поступлении в данный университет на свою специальность только после сдачи ЕГЭ. А ведь это определяющее решение для создания ценности своего человеческого капитала. С положением Г. Беккера о стабильных предпочтениях при оценивании человеческого капитала нельзя согласиться по следующим причинам. Человеческий капитал, в отличие от обычного капитала, не обладает постоянной ценностью.

---

<sup>214</sup> Becker G. S. Human capital //New York: National Bureau of Economic Research. – 1964.

<sup>215</sup> Becker G. S. Human capital revisited //Human capital: A theoretical and empirical analysis with special reference to education, third edition. – The University of Chicago Press, 1994. – P. 15-28.

<sup>216</sup> Becker G. S. Human capital //New York: National Bureau of Economic Research. – 1964.

Его величина постоянно изменяется в зависимости от работы по полученной специальности или нет, от появления новых технологий, от прохождения периодической переподготовки, от состояния здоровья работника – обладателя данного человеческого капитала, от его мотивации осваивать новые знания и получать новые навыки. Человеческий капитал действует определенное время и возможности его передачи, в отличие от обычного капитала, крайне затруднены. Наставничество, как показывает реальный опыт, может решить эту проблему только частично.

Т. Шульц раскрыл две составные части в прогрессе человеческого капитала. Часть «капитала» основана на предположении, что определенные типы расходов (жертв) создают продуктивные запасы, воплощенные в человеке, которые предоставляют услуги в течение будущих периодов. Другая часть основана на распределении «времени», что привело к экономической трактовке широкого спектра нерыночных видов деятельности<sup>217</sup>. На этой основе он формулирует строго экономическое определение человеческого капитала, который представляет форму капитала, потому что он является источником будущих доходов, или будущего удовлетворения, или того и другого. Он человеческий, потому что является неотъемлемой частью человека<sup>218</sup> [33, p. 5].

---

<sup>217</sup> Schultz T. W. Human capital: Policy issues and research opportunities //Economic Research: Retrospect and Prospect, Volume 6, Human Resources. – NBER, 1972. P. 2.

<sup>218</sup> Schultz T. W. Human capital: Policy issues and research opportunities //Economic Research: Retrospect and Prospect, Volume 6, Human Resources. – NBER, 1972. P. 5.

В последующие десятилетия концепция человеческого капитала дополнилась теорией человеческих возможностей нобелевского лауреата Амартии Сена (Amartya Sen), которая преодолевает «производственный» взгляд, подразумеваемый в теории человеческого капитала. Он, в частности, отмечал, что даже при том же уровне дохода человек может получить выгоду от образования, от чтения, общения, споров, от возможности делать более осознанный выбор, от того, что другие будут воспринимать его более серьёзно и ... таким образом, преимущества образования превосходят его роль человеческого капитала в товарном производстве<sup>219</sup>. Тем не менее, нельзя не согласиться с Сильвией Леони (Silvia Leoni), которая отмечает, что сегодня доминирующая теория человеческого капитала не учитывает важные социальные аспекты<sup>220</sup>. Таким образом, мы можем сделать вывод, что концепция человеческого капитала рассматривается сейчас исследователями экономики образования преимущественно в «узком смысле» с ориентацией на рост производительности труда, и в «широком смысле» с учётом положительных внешних эффектов от образования.

Концепция «человеческого капитала» получила широкое распространение в мире, включая современную Россию. Достаточно отметить, что в РИНЦ на март 2025 г. существовало 48 954

---

<sup>219</sup> Sen A. K. Human capital and human capability //World development. – 1997. – Vol. 25. – №. 12. P. 1959.

<sup>220</sup> Leoni S. A historical review of the role of education: From human capital to human capabilities //Review of Political Economy. – 2025. – Vol. 37. – №. 1. P. 241. [Электронный ресурс] URL: <https://www.tandfonline.com/doi/full/10.1080/09538259.2023.2245233> (дата обращения 13.03.2025)

публикации по этой теме<sup>221</sup>. В них были раскрыты различные аспекты темы, включая роль в наукоёмкой экономике<sup>222</sup>, управление человеческим капиталом<sup>223</sup>, региональные аспекты<sup>224</sup>.

В данной статье рассматривается **национальный человеческий капитал**. Под этим термином автор подразумевает совокупность индивидуальных человеческих капиталов трудоспособных граждан страны. Национальный человеческий капитал составляет сейчас основное конкурентное преимущество страны. Из-за обострения борьбы между крупнейшими экономиками мира после начала «тарифной войны» администрацией США в апреле 2025 г. усиливается значение снижения зависимости страны от внешних поставок, укрепление тренда на технологический суверенитет и, как следствие, - на улучшение национального человеческого капитала России.

В целом, влияние основных факторов, воздействующих в XXI веке на национальный человеческий капитал, показано на рисунке 1.

---

<sup>221</sup> Человеческий капитал. [Электронный ресурс] URL: [https://www.elibrary.ru/query\\_results.asp](https://www.elibrary.ru/query_results.asp) (дата обращения 19. 03. 2025)

<sup>222</sup> Судова Т. Л. Человеческий капитал в наукоёмкой экономике. Санкт-Петербург, 2001.

<sup>223</sup> Панферов В. П. Роль управления знаниями в управлении человеческим капиталом инновационно-ориентированного предприятия // Экономика и эффективность организации производства. 2019. No 29. С. 14–17. [Электронный ресурс] URL: <http://vestnik.adygnet.ru/files/2021.1/6687/70-78.pdf> (дата обращения 10.04.2025)

<sup>224</sup> Черкасова Л. И., Шаргунова Т. В. Развитие человеческого капитала в Ленинградской области. В XII Лужские научные чтения. Современное научное знание: Л83 теория и практика: материалы международной научной конференции, 22 мая 2024 г. / отв. ред. Т. В. Седлецкая. – Санкт-Петербург: ЛГУ им. А.С. Пушкина, 2024. – с. 327–331.



Рис. 1. Факторы, воздействующие на национальный человеческий капитал (составлено автором).

В терминологии макроэкономики национальный человеческий капитал представляет собой *запасы*, которые есть у народного хозяйства страны. На него воздействуют, изменяя его объём и структуру, три основных вида *потоков*: образование, трудовая миграция, искусственный интеллект. Развитие системы образования и применение технологий искусственного интеллекта повышают ценность национального человеческого капитала. Влияние трудовой миграции крайне противоречиво. С одной стороны, трудовые мигранты, как внутренние, так и внешние, помогают заполнить незанятые, особенно в связи с «постарением» трудоспособного населения России и развитых стран, рабочие места в народном хозяйстве. С другой стороны, «утечка мозгов» ослабляет потенциал развития национальной экономики.

В данной статье акцент будет сделан на новом вопросе темы: влиянии моделей искусственного интеллекта на роль и значение человеческого капитала в современной экономике. **Исследовательский вопрос:** нужно ли развивать человеческий капитал в эпоху искусственного интеллекта?

### **Воздействие искусственного интеллекта на качества человеческого капитала**

С момента появления ChatGPT в ноябре 2022 года начинается современный этап влияния ИИ на рынок труда. Это воздействие перешло из плоскости фантастики в сферу реальности. Само существование моделей искусственного интеллекта и прогресс с каждым месяцем их возможностей влияет на положение персонала компаний. И работодатели, и отдельные работники, и общество в целом должны учитывать широкие экономические последствия, вытекающие из внедрения ИИ, как-то: различия в производительности труда персонала из-за применения моделей искусственного интеллекта, взаимодополняемости человеческого труда и технологий ИИ, подверженности замене труда работников моделями искусственного интеллекта<sup>225</sup>.

Как конкретно технологии искусственного интеллекта воздействуют на современный рынок труда? Исследователи

---

<sup>225</sup> Лукичев П. М., Чекмарев О.П. Вызовы экономики искусственного интеллекта традиционному рынку труда // Вопросы инновационной экономики. – 2023. – Том 13. – No 2. – С. 785–802. doi: [10.18334/vinec.13.2.118137](https://doi.org/10.18334/vinec.13.2.118137)

разделяются на «техно-пессимистов» и «техно-оптимистов». Последние, как например А. Биркбек, Л. Роу (Birkbeck A., Rowe L), утверждают, что хотя некоторые рабочие места находятся под угрозой сокращения, чистый выигрыш в сторону более квалифицированной рабочей силы уже очевиден. Руководители семейного бизнеса должны быть готовы адекватно реагировать на ускоряющуюся войну за таланты, реализуя стратегию расширения прав и возможностей человеческого капитала<sup>226</sup>.

Воздействие искусственного интеллекта на современный рынок труда происходит как по количественным параметрам, так и по качественным параметрам. Характеризуя первое, выделим долговременное исследование И. Хуана (Huang Y., 2025), в котором было обнаружено, что за период с 2010 по 2021 год в зонах поездок на работу с более высоким принятием ИИ наблюдалось более сильное снижение соотношения занятости к численности населения. Более того, этот негативный эффект занятости в первую очередь ощущается в секторах производства и услуг с низкой квалификацией, работниках средней квалификации, не-STEM-профессиях и людях на двух концах возрастного распределения. Неблагоприятное

---

<sup>226</sup> Birkbeck A., Rowe L. Navigating towards hyperautomation and the empowerment of human capital in family businesses: a perspective article //Journal of Family Business Management. – 2024. – Vol. 14. – №. 4. – P. 727–734. [Электронный ресурс] URL: [https://www.researchgate.net/publication/374618350\\_Navigating\\_towards\\_hyperautomation\\_and\\_the\\_empowerment\\_of\\_human\\_capital\\_in\\_family\\_businesses\\_a\\_perspective\\_article](https://www.researchgate.net/publication/374618350_Navigating_towards_hyperautomation_and_the_empowerment_of_human_capital_in_family_businesses_a_perspective_article) (дата обращения 15.04.2025)

воздействие также сильнее выражено у мужчин, чем у женщин<sup>227</sup>. Напомним, что четыре области, которые формируют английскую аббревиатуру STEM – это наука, технологии, инженерия, математика.

Для характеристики второго направления была введена комплексная мера «оттока» рынка труда, которая позволяет измерять и систематически сравнивать изменения в структуре занятости на рынке труда США за длительные периоды времени<sup>228</sup>. Национальный человеческий капитал, характеризуемый с точки зрения группировки специальностей по оплате труда, под воздействием технологий ИИ эволюционирует в направлении увеличения более квалифицированных рабочих мест с более высокой оплатой труда. Результаты данного исследования показали следующее. Во-первых, рынок труда больше не поляризуется — занятость в низко- и среднеоплачиваемых профессиях снизилась, в то время как высокооплачиваемая занятость выросла. Во-вторых, рост занятости в низкооплачиваемых рабочих местах в сфере услуг остановился. В-третьих, доля занятости в STEM-профессиях увеличилась более чем на 50 процентов с 2010 года, чему способствовал рост в профессиях, связанных с программным обеспечением и компьютерами. В-четвертых, занятость в сфере розничной торговли сократилась на 25 процентов за последнее

---

<sup>227</sup> Huang Y. The Labor Market Impact of Artificial Intelligence: Evidence from US Regions // Available at SSRN 5137231. – 2025. Available at SSRN 5137231, 2025-papers.ssrn.com

<sup>228</sup> Deming D. J., Ong C., Summers L. H. Technological disruption in the labor market. – National Bureau of Economic Research, 2025. – №. w33323. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w33323/w33323.pdf](https://www.nber.org/system/files/working_papers/w33323/w33323.pdf) (дата обращения 12.04.2025)

десятилетие, вероятно, из-за технологических усовершенствований в онлайн-торговле<sup>229</sup>. Зарубежный опыт и, в частности, опыт США показывает те изменения рынка труда под влиянием технологий искусственного интеллекта, которые произойдут в России в ближайшие годы и к которым народное хозяйство страны должно быть готово.

В современной же России расширение применения технологий искусственного интеллекта «тормозится» нехваткой квалифицированных специалистов. Так, по данным компании Cloud X, рынок РФ испытывает нехватку специалистов, особенно в области разработки ИИ-моделей, MLOps, облачной ИИ-инфраструктуры и применения искусственного интеллекта в области кибербезопасности. При этом, в ближайшие годы стоит ожидать появления новых отраслевых требований, регулирующих ИИ, включая работу с большими данными, управление рисками и безопасность персональных данных, защита которых должна выйти на новый уровень в связи с возможностями искусственного интеллекта по идентификации субъектов персональных данных на основе относительно небольших наборов данных<sup>230</sup>.

---

<sup>229</sup> Deming D. J., Ong C., Summers L. H. Technological disruption in the labor market. – National Bureau of Economic Research, 2025. – №. w33323. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w33323/w33323.pdf](https://www.nber.org/system/files/working_papers/w33323/w33323.pdf) (дата обращения 12.04.2025)

<sup>230</sup> Хотеева К. Вложение в будущее. Чего рынок ждёт от ИИ в 2025 году. Коммерсант, 27;03. 2025. [Электронный ресурс] URL: <https://www.kommersant.ru/doc/7605032?ysclid=m8x6mxt9am434324280> (дата обращения 27.03.2025)

Влияние технологий искусственного интеллекта на человеческий капитал происходит на нескольких уровнях: 1) уровне национальной экономики, 2) уровне фирмы, 3) уровне отдельного работника.

Характеризуя первый, отметим исследование Международного валютного фонда, которое выявило многообразные последствия воздействия технологий искусственного интеллекта на современный рынок труда, сгруппировав их по четырём каналам:

1. **Вытеснение рабочей силы:** смещение задач с человеческого труда на капитал ИИ, что снижает трудовой доход.

2. **Взаимодополняемость:** добавленная стоимость смещается в профессии, дополняющие искусственный интеллект, увеличивая спрос на рабочую силу для этих профессий и сокращая его для других.

3. **Рост производительности:** общий экономический рост потенциально компенсирует потери трудового дохода.

4. **Повышение дохода от капитала:** применение ИИ приводит к увеличению отдачи капитала, что ещё больше увеличивает доход от капитала<sup>231</sup>.

---

<sup>231</sup> Cazzaniga M., Jaumotte M. F., Li L., Melina M. G., Panton A. J., Pizzinelli, C., ... Tavares M. M. Gen-AI: Artificial intelligence and the future of work. – International Monetary Fund, 2024. [Электронный ресурс] URL: <https://www.imf.org/en/Publications/Staff-DiscussionNotes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379> (дата обращения 17.11.2024)

Для анализа второго уровня целесообразно, по нашему мнению, использовать модель, разработанную М. Хэмпол и др. (Hampole M. et al., 2025). В этой модели трудовая деятельность «разбивается» на ряд задач, которые должен решить работник в рамках своей профессии. Изменение спроса на рабочую силу в фирме под воздействием искусственного интеллекта можно охарактеризовать двумя статистическими показателями. Во-первых, средняя подверженность ИИ задачам профессии в целом отрицательно связана со спросом на эту профессию. Во-вторых, дисперсия подверженности искусственного интеллекта профессиональным задачам увеличивает спрос на рабочую силу для этой конкретной профессии, сохраняя среднюю подверженность постоянной.<sup>232</sup> . Примером воздействия первого служит снижение спроса на агентов-людей по обслуживанию клиентов фирмы за счёт использования чат-бота (чат-окна). Примером второго является автоматизированная система создания отчётов о расходах, которая позволяет работникам перераспределять свои усилия на другие задачи внутри фирмы.

Главное здесь, по нашему мнению, что воздействие технологий искусственного интеллекта не уничтожает ту или иную профессию, а видоизменяет ценность выполнения задач, из которых складывается та или иная профессия. У работника и у работодателя появляется время для адаптации к изменившимся условиям, которая

---

<sup>232</sup> Hampole M., Papanikolaou D., Schmidt L. D., Seegmiller B. Artificial intelligence and the labor market. – National Bureau of Economic Research, 2025. – №. w33509. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w33509/w33509.pdf](https://www.nber.org/system/files/working_papers/w33509/w33509.pdf) (дата обращения 11.04.2025)

может быть достигнута за счёт переподготовки персонала. Наряду с рыночными мерами адаптация потребует создания государственных программ перераспределения рабочей силы, поскольку не всё может быть решено на уровне отдельного предприятия. Следует согласиться с авторами модели<sup>233</sup>, что для конкретной фирмы негативный или позитивный общий эффект зависит от результирующего увеличения производительности фирмы из-за внедрения ИИ.

Появление «Агентов искусственного интеллекта» стало катализатором воздействия ИИ на профессиональные задачи персонала компаний. Например, Manus ai — система, построенная на основе существующих моделей, может взаимодействовать с Интернетом и выполнять последовательность задач, не обращаясь за разрешением к человеку-пользователю. В ближайшей перспективе она может создать рекламную сеть аккаунтов в социальных сетях, провести исследование и написать стратегический документ или забронировать билеты и гостиницы для конференции<sup>234</sup>. Это пример того, как Агенты ИИ заменяют некоторые профессиональные задачи, за которые раньше отвечали офисные сотрудники. В то же время Агенты ИИ могут усилить человеческий потенциал определенных

---

<sup>233</sup> Hampole M., Papanikolaou D., Schmidt L. D., Seegmiller B. Artificial intelligence and the labor market. – National Bureau of Economic Research, 2025. – №. w33509. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w33509/w33509.pdf](https://www.nber.org/system/files/working_papers/w33509/w33509.pdf) (дата обращения 11.04.2025)

<sup>234</sup> Get used to it. The Economist. March 15th 2025. [] URL: <https://www.economist.com/leaders/2025/03/13/with-manus-ai-experimentation-has-burst-into-the-open> (дата обращения 16.03.2025)

специалистов, например программистов. Согласно исследованиям, на данный момент Агенты ИИ повышают ежедневную производительность разработчиков в реальных бизнес-условиях на 10–30 процентов. В KPMG, бухгалтерской и консалтинговой фирме, разработчики, использующие GitHub Copilot, экономят в среднем 4,5 часа в неделю и сообщают, что качество их кода улучшилось, согласно опросу фирмы<sup>235</sup>. Кроме того, их применение позволяет решить проблему «наставничества», - одну из самых острых проблем, особенно для компаний в сфере ИКТ. Так, в KPMG используют помощника по программному обеспечению в качестве своего рода автоматизированного инструктора для быстрого введения новых членов команды разработчиков в курс дела. В целом, применение «Агентов искусственного интеллекта» в ближайшие годы расширится позитивно меняя структуру рынка труда за счет сокращения рутинных операций персоналом компаний и, в то же время, это технологическая инновация усилит неравенство как между странами, так и между фирмами<sup>236</sup>.

Третий уровень ощущается работниками не непосредственно от взаимодействия с моделью искусственного интеллекта, а через восприятие работодателей. Последние, как показал опрос слушателей программы МВА БГТУ «ВОЕНМЕХ» им. Д. Ф.

---

<sup>235</sup> Liu Y., Zhang Z. Unlocking operational efficiency: How AI human capital investment enhances data processing efficiency? //Economics Letters. – 2025. – Vol. 247. – P. 112147. <https://doi.org/10.1016/j.econlet.2024.112147>

<sup>236</sup> Лукичѐв, П. М. Влияние искусственного интеллекта на рынок труда в парадигме неравенства / П. М. Лукичѐв // Экономика труда. – 2025. – Т. 12, № 3. – С. 267-282. – DOI 10.18334/et.12.3.122851

Устинова, проведенный автором 7 декабря 2024 г., высказывают большую озабоченность отсутствием «мягких» навыков, таких как адаптивность к меняющимся условиям, желание учиться, желание совершенствовать свой потенциал; лидерство; умение работать в команде; слабое развитие у выпускников ВУЗов навыков эффективной коммуникации (последствие онлайн-детства). Недостающие «твёрдые» навыки в условиях борьбы за привлечение кадров фирмы могут «до настроить», «до обучить» до необходимого сегодня уровня. Ликвидировать нехватку «мягких» навыков – гораздо сложнее.

Результаты исследования Всемирного экономического форума показывают, что хотя опасения замены физической и ручной работы машинами снизились, ожидается, что рассуждение, общение и координация – качества, имеющие сравнительные преимущества для современных людей – в будущем станут более автоматизированными. Ожидается также, что искусственный интеллект, ключевой фактор потенциального вытеснения, будет принят почти 75% опрошенных компаний и, вероятно, приведёт к высокому оттоку сотрудников: 50% организаций ожидают, что он приведёт к росту рабочих мест, а 25% ожидают, что он создаст потери работы<sup>237</sup>. Это означает, что человеческий капитал должен меняться по мере развития технологий искусственного интеллекта, чтобы соответствовать требованиям

---

<sup>237</sup> Di Battista A., Grayling S., Hasselaar E., Leopold T., Li R., Rayner M., Zahidi S. Future of jobs report 2023 //World Economic Forum, Geneva, Switzerland. [Электронный ресурс] URL: <https://www.weforum.org/reports/the-future-of-jobs-report-2023> – 2023. (дата обращения 15.02.2025)

современного рынка труда. Сфера образования должна учитывать эти изменения и вносить коррективы в образовательный процесс. Как конкретно должны происходить эти изменения?

По мнению респондентов Future of Jobs Survey, две пятых (39%) существующих наборов навыков будут преобразованы или устареют в период 2025–2030 годов. Однако этот показатель «нестабильности навыков» замедлился по сравнению с предыдущими выпусками отчета: с 44% в 2023 году и до пика в 57% в 2020 году после пандемии. Этот вывод может быть потенциально связан с увеличением доли работников (50%), прошедших обучение, переподготовку или повышение квалификации, по сравнению с 41% в выпуске отчета 2023 года<sup>238</sup>. По нашему мнению, это показывает роль образования в подготовке и переподготовке персонала, которая позволяет «отремонтировать» (обновить) наборы знаний, умений и навыков работников. То есть система образования играет роль буфера между воздействием технологий искусственного интеллекта на рынок труда и сокращением персонала компаний.

Дридакис Н. (Drydakis N.) вводит новый термин «капитал искусственного интеллекта», который определяется как «вектор знаний, навыков и возможностей, связанных с технологиями, всю ИИ, которые могут повысить производительность, занятость и

---

<sup>238</sup> Future of Jobs Report 2025. World Economic Forum. January 2025/ Geneva, WEF. 2025. – 290 p. [Электронный ресурс] URL: <https://www.weforum.org/reports/the-future-of-jobs-report-2025/> (дата обращения 14.02.2025)

доходы людей»<sup>239</sup>. То есть он определяется как сумма знаний и навыков об искусственном интеллекте, обладая которыми выпускник получает конкурентное преимущество при трудоустройстве.

Данные статсборника «Индикаторы образования 2025» показывают значительное увеличение внимания в Российской Федерации к студентам и выпускникам университетов в области искусственного интеллекта. Отметим первое появление специального раздела – 3.3. «Подготовка специалистов в области искусственного интеллекта»<sup>240</sup>. Основные характеристики обучения представлены в Таблице 1.

**Таблица 1.** Приём на обучение, численность и выпуск студентов по программам бакалавриата, специалитета, магистратуры по профилю «искусственный интеллект» (2023 г.) (Составлено автором на основе<sup>241</sup>).

	Всего		В том числе по программам						
	тысячи человек	в процентах от общего выпуска по соответствующей укрупненной группе специальностей и направлений подготовки	бакалавриата		специалитета		магистратуры		
			тысячи человек	в процентах от общего выпуска по соответствующей укрупненной группе специальностей и направлений подготовки	тысячи человек	в процентах от общего выпуска по соответствующей укрупненной группе специальностей и направлений подготовки	тысячи человек	в процентах от общего выпуска по соответствующей укрупненной группе специальностей и направлений подготовки	

<sup>239</sup> Drydak N. Artificial intelligence capital and employment prospects //Oxford Economic Papers. – 2024. – Vol. 76. – №. 4. – P. 901-919. <https://doi.org/10.1093/oup/gpae005>

<sup>240</sup> Индикаторы образования: 2025: статистический сборник / Н. В. Бондаренко, Т. А. Варламова, Л. М. Гохберг и др.; Нац. исслед. ун-т «Высшая школа экономики». – М.: ИСИЭЗ ВШЭ, 2025. с. 109-130. [Электронный ресурс] URL: <https://issek.hse.ru/news/1023576877.html> (дата обращения 18.04.2025)

<sup>241</sup> Индикаторы образования: 2025: статистический сборник / Н. В. Бондаренко, Т. А. Варламова, Л. М. Гохберг и др.; Нац. исслед. ун-т «Высшая школа экономики». – М.: ИСИЭЗ ВШЭ, 2025. с. 109-116. [Электронный ресурс] URL: <https://issek.hse.ru/news/1023576877.html> (дата обращения 18.04.2025)

Приём студентов на обучение по программам бакалавриата, специалитета, магистратуры по профилю «искусственный интеллект»	20,7	1,8	12,1	1,5	0,3	0,2	8,3	3,1
Численность студентов, обучающихся по программам бакалавриата, специалитета, магистратуры по профилю «искусственный интеллект»	42,3	1,0	25,8	0,9	1,8	0,2	15,1	2,8
Выпуск бакалавров, специалистов, магистров, по профилю «искусственный интеллект»	3,8	0,5	1,8	0,3	0,2	0,2	1,7	1,1

Нельзя не отметить «технологический уклон» в подготовке в университетах России бакалавров, специалистов, магистров по профилю «искусственный интеллект», в которой выделяются освоение технологий компьютерного зрения, технологий обработки звуковых данных, включая распознавание и синтез речи; технологий обработки текста; технологий интеллектуальной поддержки принятия решений и управления. Это, вместе с введённым с 1 января 2024 года ГОСТом Р 70949–2023 «Технологии искусственного интеллекта в образовании. Применение искусственного интеллекта в научно-исследовательской деятельности. Варианты использования»<sup>242</sup>, означает что Российская Федерация находится

---

<sup>242</sup> ГОСТ Р 70949–2023 Технологии искусственного интеллекта в образовании. Применение искусственного интеллекта в научно-исследовательской деятельности. Варианты использования. – Москва. Российский институт стандартизации. 2023. URL:

только на первом, - *«технологическом»*, - этапе освоения «грамотности в области искусственного интеллекта». В мире данный этап существовал до 2021 года и был сменён *«современным»* этапом<sup>243</sup>, отличительной чертой которого является переход от ориентации на технологии ИИ к междисциплинарному подходу, включая необходимость освоения поведенческой экономики, этики искусственного интеллекта и других дисциплин, раскрывающих особенности взаимодействия работников и искусственного интеллекта. Для прохождения второго этапа необходимо освоить всю систему компетенций грамотности в сфере искусственного интеллекта<sup>244</sup>.

Особенно важны для улучшения национального человеческого капитала проекты университетов и систем переподготовки взрослых. Они могут познакомить с моделями искусственного интеллекта студентов университетов и взрослых работников, не имеющих отношения к компьютерным наукам, например, в области бизнес-

---

[https://allgosts.ru/35/240/gost\\_r\\_70949-2023?ysclid=lryyyh4k6w701222747](https://allgosts.ru/35/240/gost_r_70949-2023?ysclid=lryyyh4k6w701222747)  
(дата обращения 05.03.2024)

(дата

<sup>243</sup> Лукичёв, П. М. Риски применения искусственного интеллекта в системе высшего образования / П. М. Лукичёв, О. П. Чекмарев // Вопросы инновационной экономики. – 2024. – Т. 14, № 2. – С. 463–482. – DOI 10.18334/vines.14.2.120731 Лукичёв, П. М. Влияние искусственного интеллекта на рынок труда в парадигме неравенства / П. М. Лукичёв // Экономика труда. – 2025. – Т. 12, № 3. – С. 267–282. – DOI 10.18334/et.12.3.122851

<sup>244</sup> Long D., Magerko B. What is AI literacy? Competencies and design considerations // Proceedings of the 2020 CHI conference on human factors in computing systems. – 2020. – P. 1-16. [Электронный ресурс] URL: <https://dl.acm.org/doi/abs/10.1145/3313831.3376727> (дата обращения 03.01.2024)

исследований, образования и медицины<sup>245</sup>, <sup>246</sup>. Это открывает широкие перспективы для массового повышения уровня человеческого капитала за счёт улучшения готовности персонала к применению технологий искусственного интеллекта. В Англии такие проекты стали вводиться с 2021 года. При их успешном внедрении в России это может привести к тому, что технологии искусственного интеллекта быстро станут в РФ технологиями общего назначения (ТОН), что даст конкурентное преимущество всей национальной экономике. Например, университеты могут реально улучшить бизнес-образование, готовя студентов к тому, чтобы стать профессионалами, которые понимают потребности бизнеса и могут интерпретировать анализ данных, выполненный ИИ, чтобы обеспечить принятие решений на основе данных<sup>247</sup>. Здесь возможна аналогия с первым этапом освоения компьютеров, когда они были уделом представителей информационных наук и математиков, и колоссальным экономическим эффектом, который принесло их освоение всеми специалистами на последующих этапах.

---

<sup>245</sup> Laupichler M. C., Aster A., Schirch J., Raupach T. Artificial intelligence literacy in higher and adult education: A scoping literature review //Computers and Education: Artificial Intelligence. – 2022. – Vol. 3. – P. 100101. [Электронный ресурс] URL: [https://www.researchgate.net/publication/363869317\\_Artificial\\_intelligence\\_literacy\\_in\\_higher\\_and\\_adult\\_education\\_A\\_scoping\\_literature\\_review](https://www.researchgate.net/publication/363869317_Artificial_intelligence_literacy_in_higher_and_adult_education_A_scoping_literature_review) (дата обращения 13.01.2025)

<sup>246</sup> Sollosy M., McInerney M. Artificial intelligence and business education: What should be taught //The International Journal of Management Education. – 2022. – Vol. 20. – №. 3. – P. 100720. [Электронный ресурс] URL: <https://www.sciencedirect.com/science/article/abs/pii/S1472811722001227> (дата обращения 02.02.2025)

<sup>247</sup> Sollosy M., McInerney M. Artificial intelligence and business education: What should be taught //The International Journal of Management Education. – 2022. – Vol. 20. – №. 3. – P. 100720. [Электронный ресурс] URL: <https://www.sciencedirect.com/science/article/abs/pii/S1472811722001227> (дата обращения 02.02.2025)

Для совершенствования национального человеческого капитала очень важно преодоление предубеждений, существующих у персонала разного уровня. Так, российские руководители компаний-интеграторов роботов склонны экономить на ставках инженеров, которые могут проводить научные исследования внутри организации, поскольку многие интеграторы по-прежнему придерживаются установки, что нужно использовать имеющиеся мировые разработки. Такого рода представления об исследованиях и разработках как периферийной сфере деятельности являются барьером на пути к созданию прорывных технологий<sup>248</sup>.

В целом расширение применения в экономике технологий искусственного интеллекта меняет роль и значение системы образования в народного хозяйства. Если раньше улучшение национального человеческого капитала формировало конкурентное преимущество страны в глобальном экономическом соревновании и создавало базу для развития отраслей с высокой добавленной стоимостью, то сейчас приоритеты изменились. Ведущие экономические державы и, прежде всего США, делают сейчас акцент на прорыве в области ИИ, вкладывая громадные инвестиции в разработку моделей искусственного интеллекта и в связанные с этим отрасли. Улучшение системы образования уходит на «второй план». Началось, по нашему мнению, *экономическое соревнование*

---

<sup>248</sup> Белова Ю. Ю., Н. А. Шматко Н. А. Разработка промышленных роботов в России: технологии и кадры. Наука. Технологии. Инновации. Экспресс-информация 14.02.2025. [Электронный ресурс] URL: <https://issek.hse.ru/mirror/pubs/share/1016053804.pdf> (дата обращения 15.02.2025)

*между человеческим капиталом и искусственным интеллектом.*

Формами его проявления является размер инвестиций, условия труда, внимание в глазах общества и средств массовой информации. Имеет ли смысл в этих условиях увеличивать национальный человеческий капитал для экономического развития страны?

Крайняя точка зрения принадлежит Хамильтону и др. (Hamilton A. et al., 2023). Они считают, что образование сейчас достигло пика<sup>249</sup>, а дальше, по мере развития возможностей моделей ИИ и создания общего искусственного интеллекта (AGI), потребность в образовании для совершенствования человеческого капитала будет снижаться и стимулы к образованию будут сокращаться. Сегодняшние навыки вычислений, чтения и письма уменьшаются ввиду отсутствия их необходимости в повседневной жизни. Уже сегодня у многих людей ухудшились вычислительные способности в связи с использованием электронных калькуляторов. Способность и необходимость грамотно писать сейчас отмирает с использованием электронных девайсов.

Ответ на исследовательский вопрос будет положительным «да, стоит развивать человеческий капитал в эпоху искусственного интеллекта». Так считают и респонденты опрошенные автором, и эксперты, опрошенные Всемирным Экономическим Форумом в 2025

---

<sup>249</sup> Hamilton A., William D., Hattie J. The Future of AI in Education: 13 things we can do to minimize the damage. – 2023. P. 3. [Электронный ресурс] URL: [https://www.researchgate.net/publication/373108877\\_The\\_Future\\_of\\_AI\\_in\\_Education\\_13\\_Things\\_We\\_Can\\_Do\\_to\\_Minimize\\_the\\_Damage](https://www.researchgate.net/publication/373108877_The_Future_of_AI_in_Education_13_Things_We_Can_Do_to_Minimize_the_Damage) (дата обращения 16.02.2024)

году<sup>250</sup>. А обострившаяся «охота за талантами», ведущаяся фирмами развитых стран, только подтверждает это.

**Будет ли человеческий капитал развиваться самостоятельно или его развитие будет играть вспомогательную (подыгрывающую) роль по отношению к искусственному интеллекту?**

Во многом ответ на этот вопрос зависит от национальной макроэкономической политики. В некоторых странах, как например в США, акцент в развитии делается на технологии искусственного интеллекта. Так, Д. Л. Тайсон, Дж. Зисман (Tyson L. D., Zysman J.) справедливо отмечают<sup>251</sup>, что влияние искусственного интеллекта на работу определяется не технологически, а зависит от стимулов как тех, кто возглавляет исследования и инновации в области ИИ, так и тех, кто инвестирует в его внедрение. Преобладающая версия, лежащая в основе инноваций и внедрения искусственного интеллекта в деловых и исследовательских сообществах, особенно выраженная в Соединенных Штатах, где решения отражают интересы акционеров, а работники имеют ограниченный голос при принятии бизнес-решений, фокусируется на способности ИИ превосходить людей, а не на создание хороших рабочих мест. Таким образом, улучшение человеческого капитала является вторичным по отношению к развитию технологий искусственного интеллекта.

Отметим, что само развитие моделей ИИ производно от уровня человеческого капитала. Так, если использовать долю сотрудников,

---

<sup>250</sup> Future of Jobs Report 2025. World Economic Forum. January 2025/ Geneva, WEF. 2025. – 290 p. [Электронный ресурс] URL: <https://www.weforum.org/reports/the-future-of-jobs-report-2025/> (дата обращения 14.02.2025)

<sup>251</sup> Tyson L. D., Zysman J. Automation, AI & work //Daedalus. – 2022. – Vol. 151. – №. 2. – P. 266. [Электронный ресурс] URL: [https://amacad.org/sites/default/files/daedalus/downloads/Daedalus\\_Sp22\\_AI%20%26%20Society\\_3.pdf#page=258](https://amacad.org/sites/default/files/daedalus/downloads/Daedalus_Sp22_AI%20%26%20Society_3.pdf#page=258) (дата обращения 02.02.2025)

связанных с искусственным интеллектом, в качестве меры инвестиций в человеческий капитал ИИ для компаний США с 2005 по 2020 год, то результаты показывают, что более высокие уровни человеческого капитала искусственного интеллекта повышают скорость обработки данных, а ИИ также способствует большей точности в сборе и раскрытии данных компаниями. Вместе с тем, пока искусственный интеллект, как правило, более эффективен в оказании помощи компаниям в обработке рутинной, а не сложной информации<sup>252</sup>.

Ещё одним показателем взаимодействия (производности) развития искусственного интеллекта от уровня человеческого капитала является мера инвестиций в технологии ИИ, основанная на человеческом капитале фирм, владеющих навыками искусственного интеллекта. Сильная зависимость ИИ от человеческих знаний делает подход, основанный на человеческом капитале, особенно подходящим в этой ситуации<sup>253</sup>. Таким образом, можно сделать вывод, что не только искусственный интеллект оказывает влияние на качества «человеческого капитала», но и развитие моделей ИИ возможно только под воздействием высокого уровня человеческого капитала. То есть существует двусторонность процесса совершенствования моделей искусственного интеллекта. С одной стороны, - технологическая составляющая, а с другой – улучшение человеческого капитала.

Отметим также, что данные, приводимые в двух последних исследованиях, касаются 2005–2020 гг. и 2010–2018 гг.

---

<sup>252</sup> Liu Y., Zhang Z. Unlocking operational efficiency: How AI human capital investment enhances data processing efficiency? //Economics Letters. – 2025. – Vol. 247. – P. 112147. <https://doi.org/10.1016/j.econlet.2024.112147>

<sup>253</sup> Babina T., Fedyk A., He A., Hodson, J. Artificial intelligence, firm growth, and product innovation //Journal of Financial Economics. – 2024. – Vol. 151. – P. 103745. <https://doi.org/10.1016/j.jfineco.2023.103745>

соответственно, то есть до появления ChatGPT в ноябре 2022 года и начала современного этапа массового применения ИИ в экономике.

### **Выводы и будущие исследования**

Необходимо продолжать исследования взаимосвязи и взаимозависимости между развитием искусственного интеллекта и улучшением человеческого капитала на базе данных современного (после ноября 2022 г.) этапа развития искусственного интеллекта.

Важно объективно понять функцию искусственного интеллекта как работодателя, всё более формирующего спрос (на человеческий капитал) и предъявляющего особые требования только на определённые качества человеческого капитала. Должна ли система образования и повышения квалификации персонала изменяться в соответствии с запросами моделей ИИ, делая неизбежный технологический крен в эту сторону. Или формирование знаний, умений и навыков в национальном человеческом капитале должно оставаться самостоятельным, а соревнование между человеческим капиталом и искусственным интеллектом будет вестись на паритетных началах? Каков будет, в связи с этим жизненный цикл работника? Технологии меняются всё быстрее и устаревание «твёрдых» навыков персонала будет происходить всё интенсивнее. Беспристрастен ли ИИ, например, при отборе персонала для компании? Объективные ответы на эти вопросы должны дать будущие исследования. Созданная на их основе макроэкономическая политика определит наилучшее развитие рынка труда будущего.

## **Глава 4. ГОСУДАРСТВЕННОЕ РЕГУЛИРОВАНИЕ (ДЕРЕГУЛИРОВАНИЕ) ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

### **4.1. ГОСУДАРСТВЕННОЕ РЕГУЛИРОВАНИЕ МОДЕЛЕЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: НЕОБХОДИМОСТЬ И ПРЕДЕЛЫ МАКРОЭКОНОМИЧЕСКОЙ ПОЛИТИКИ**

В разделе обосновывается необходимость регулирования моделей искусственного интеллекта на национальном уровне и на международном уровне. Автор анализирует различные подходы, существующие в сфере регулирования искусственного интеллекта и раскрывает противоречивость целей регулирования у общественности, и у фирм-лидеров в сфере искусственного интеллекта. Выявлена особая регулятивная проблема, которая возникла из-за применения передовых моделей искусственного интеллекта. Автор обосновывает опасность создания *Суперлиги игроков в сфере искусственного интеллекта*. Также выявлена свойство невидимости, которой обладают большие языковые модели. Проанализированы тренды в создании стартапов в области искусственного интеллекта, выявлены различия между англосаксонским подходом и подходом КНР в создании стартапов в области ИИ. Раскрыта роль государственных служащих в эволюции и результативности регулирования технологий искусственного интеллекта.

#### **Введение**

Технологии искусственного интеллекта (ИИ) распространяются в экономике с нарастающей скоростью. Здравоохранение, маркетинг, образование, логистика представляют собой сферы, в которых успехи ИИ наиболее заметны. Значительная часть обязанностей, которые раньше выполняли в бизнес-процессах работники, переходит к большим языковым моделям. Это порождает риски

экономики на микро- и на макроуровне. Стремление уменьшить риски использования технологий искусственного интеллекта в краткосрочном периоде и в долгосрочном периоде лежит в основе государственного регулирования ИИ. Поскольку прогресс в освоении моделей искусственного интеллекта протекает крайне неравномерно в отдельных странах, то и вопросы государственного регулирования ИИ развиваются тоже неравномерно, акцентируя внимание на разных проблемах. **Цель раздела:** проанализировать теоретические положения и практический опыт применения государственного регулирования алгоритмов искусственного интеллекта и выявить его «плюсы» и «минусы» для общества.

В данном исследовании мы основываемся на следующем определении искусственного интеллекта. «Искусственный интеллект – комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их»<sup>254</sup>. Государственное регулирование определяется нами как комплекс мер по обеспечению выполнения законов, вводимых государственными органами для поощрения или

---

<sup>254</sup> Указ Президента Российской Федерации от 15.02.2024 № 124 "О внесении изменений в Указ Президента Российской Федерации от 10 октября 2019 г. № 490 "О развитии искусственного интеллекта в Российской Федерации" и в Национальную стратегию, утвержденную этим Указом" URL: // <https://base.garant.ru/408559959/> (дата обращения 29.08.2024)

ограничения экономической деятельности путем установления цен, назначения стандартов и условий входа в отрасль<sup>255</sup>.

## **Необходимость регулирования рынков искусственного интеллекта**

Создание LLMs (больших языковых моделей) в последние годы переживало процесс своего свободного становления. (своеобразного «романтизма»), действуя, в отличие от всех других отраслей экономики, без постоянного регулирования со стороны государства. Во многом это связано с отсутствием понимания у государственных органов того, *как* именно должно осуществляться воздействие на работу отрасли. Это очень необычная ситуация, поскольку со второй половины прошлого века во всех развитых странах мира государственное регулирование играет решающую роль в развитии национальных экономик. Следует согласиться с Лесли, Перини<sup>256</sup>, которые отмечают, что международный кризис политики и управления в области искусственного интеллекта материализовался, в немалой степени, благодаря тому, что искусственное отсутствие практических и обязательных механизмов регулирования и управления, позволяющих контролировать проектирование, разработку и использование FM (“frontier AI” models («передовых моделей ИИ»)) и их последующих приложений, сочеталось с широко

---

<sup>255</sup> Лукичев П. М. Государственное регулирование и дерегулирование аграрной сферы. – Санкт-Петербург, 2006, с. 19.

<sup>256</sup> Leslie D., Perini A. M. Future Shock: Generative AI and the international AI policy and governance crisis. – 2024.

распространенными разногласиями. среди исследователей и представителей общественности, которые всё чаще выражают обеспокоенность по поводу рисков и вреда на системном уровне, возникающих в результате импульса промышленного масштабирования по принципу «двигайся быстрее и ломай»<sup>257</sup>.

Кроме того, остаются нерешёнными прежние проблемы регулирования рынков искусственного интеллекта. Во-первых, доступ к данным может выступать в качестве входного барьера для создания новых конкурирующих сетей и для инвестирования в инновации новыми участниками рынка; это также усилит стимул к антиконкурентному поведению в неценовых измерениях, таких как сбор, извлечение и исключение данных. Во-вторых, возросшая способность отслеживать отдельных лиц позволяет использовать новые формы ценовой дискриминации. В-третьих, ожидается, что использование технологий ИИ приведёт к расширению случаев известных форм анти конкурентного поведения, таких как явный и молчаливый сговор и дискриминация<sup>258</sup>. Отметим, что второе положение лежит в основе санкционного контроля между странами и между фирмами в XXI веке.

Кроме того, алгоритмическое регулирование, обладающее способностью расширять, укреплять и перераспределять власть, полномочия и ресурсы, а также способностью влиять на критически

---

<sup>257</sup> Leslie D., Perini A. M. Future Shock: Generative AI and the international AI policy and governance crisis. – 2024.

<sup>258</sup> Petit N. Antitrust and artificial intelligence: a research agenda // Journal of European Competition Law & Practice. 2017. Vol.8. No.6. P.361-362.

важные жизненные возможности людей, требует разработки механизмов обеспечения алгоритмической подотчетности. Оно способно, по сути, автоматизировать меры подталкивающей экономики и обеспечения общественными товарами. «Алгоритмическое регулирование относится к системам принятия решений, которые регулируют сферу деятельности с целью управления рисками или изменениями поведения посредством непрерывной вычислительной генерации знаний путем систематического сбора данных (в режиме реального времени на постоянной основе), исходящих непосредственно от многочисленных динамических компонентов, относящихся к регулируемой среде с целью идентификации и, при необходимости, автоматического уточнения (или оперативного уточнения) операций системы для достижения заранее заданной цели»<sup>259</sup>. Отметим, что за последние годы произошла эволюция понятий в изучении искусственного интеллекта. До 2022 г. исследователи преимущественно использовали термины «алгоритмическое принятие решений», «алгоритмическое регулирование», как в ранее приведенных статьях, то сейчас – преимущественно, «принятие решений искусственным интеллектом», «регулирование моделей искусственного интеллекта».

Необходимость введения государственного регулирования ИИ усиливается сегодня в связи с нарастающей значимостью проблемы.

---

<sup>259</sup> Yeung K. Algorithmic regulation: A critical interrogation // Regulation & governance. – 2018. – Vol. 12. – №. 4. – P. 505-523.

По мнению Goldman Sachs Research, прорывы в области генеративного искусственного интеллекта могут привести к радикальным изменениям в мировой экономике. По мере того, как инструменты, использующие достижения в области обработки естественного языка, проникают в бизнес и общество, они могут привести к увеличению мирового ВВП на 7% (или почти 7 триллионов долларов США) и росту производительности на 1,5 процентных пункта в течение 10-летнего периода<sup>260</sup>. Одновременно по мере усиления мощностей больших языковых моделей увеличиваются и риски их применения. Это вызывает озабоченность ученых разработчиков искусственного интеллекта и общественности. Джеффри Хинтон (Geoffrey Hinton), пионер глубокого обучения, в 2023 г. покинул Google из-за обеспокоенности рисками искусственного интеллекта. 29 марта 2023 года Future of Life Institute (FLI) Институт будущего жизни (FLI) опубликовал открытое письмо, призывающее «все лаборатории искусственного интеллекта немедленно приостановить как минимум на 6 месяцев обучение систем ИИ, более мощных, чем GPT-4»<sup>261</sup>. В письме, подписанном выдающимися учеными, генеральными директорами и другими светилами в области технологий, в том числе Йошуа

---

<sup>260</sup> Generative AI could raise global GDP by 7%. [Электронный ресурс] URL: <https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html> (дата обращения 23.10.2024)

<sup>261</sup> Future of Life Institute. (2023, March 22). *Pause giant AI experiments: An open letter*. [https://futureoflife.org/wp-content/uploads/2023/05/FLI\\_Pause-Giant-AI-Experiments\\_An-Open-Letter.pdf](https://futureoflife.org/wp-content/uploads/2023/05/FLI_Pause-Giant-AI-Experiments_An-Open-Letter.pdf)

Бенджио, Илоном Маском и Стивом Возняком (Yoshua Bengio, Elon Musk, Steve Wozniak), выражалось сожаление по поводу отсутствия «планирования и управления», которое характеризует несчастливое поведение «лабораторий искусственного интеллекта, запертых в условиях неблагополучия». Пока же продолжается гонка за контролем над созданием и внедрением ещё более мощных моделей ИИ, которые «никто – даже их создатели – не может понять, предсказать или надёжно контролировать». Между тем, как отмечалось в письме: «Мощные системы искусственного интеллекта, должны разрабатываться только тогда, когда мы уверены, что их эффекты будут положительными, а риски — управляемыми»<sup>262</sup>.

Обучение систем искусственного интеллекта остановлено не было, и это вылилось в конфликт «17 ноября» в компании OpenAI. Последняя была создана как некоммерческая организация с благородной миссией – гарантировать, что модели искусственного интеллекта принесут пользу «всему человечеству». Однако, как показало последующее развитие компании, - создателя ChatGPT и её более продвинутых версий, - стремление к получению прибыли возобладало над мерами безопасности создаваемых технологий. Как подчёркивали Хелен Тонер и Таша МакКоли, бывшие члены Совета OpenAI, общество не должно позволять, чтобы внедрение искусственного интеллекта контролировалось исключительно

---

<sup>262</sup> Future of Life Institute. (2023, March 22). *Pause giant AI experiments: An open letter*. [https://futureoflife.org/wp-content/uploads/2023/05/FLI\\_Pause-Giant-AI-Experiments\\_An-Open-Letter.pdf](https://futureoflife.org/wp-content/uploads/2023/05/FLI_Pause-Giant-AI-Experiments_An-Open-Letter.pdf)

частными технологическими компаниями. Регулирование необходимо, чтобы укротить рыночные силы»<sup>263</sup>.

На наш взгляд, здесь является некорректным сопоставление больших языковых моделей с предшествующими технологиями ввиду того, что возможности создаваемого общего искусственного интеллекта будут превышать способности человека. Поэтому уровень беспокойства, выражаемый многими ведущими учеными и специалистами в области ИИ по поводу технологии, которую они сами создают, сильно отличается от оптимистического отношения программистов и сетевых инженеров, которые разработали ранний Интернет. Как показало исследование, проведенное в 2023 году, средний эксперт по искусственному интеллекту даёт в 10 раз больший шанс экзистенциальной катастрофы (в которой выживет менее 5000 человек) из-за ИИ к 2100 году по сравнению со средним супер прогнозистом (3,9 % против 0,38%)<sup>264</sup>.

С самого появления более шести десятилетий назад исследователей беспокоят «проблемы согласования», риск которых заключается в том, что ИИ может целеустремленно преследовать задачу, поставленную создателем, но в процессе сделать что-то вредное, чего не хотелось. То есть, используя экономическую

---

<sup>263</sup> AI firms mustn't govern themselves, say ex-members of OpenAI's board. The Economist. June 1<sup>st</sup> 2024.

<sup>264</sup> Atanasov P. D., Karger E., Tetlock P. Full Accuracy Scoring Accelerates the Discovery of Skilled Forecasters // Available at SSRN 4357367. – 2023.

терминологию, «внешние эффекты экономической деятельности» будут больше её основного эффекта. Образный пример этого был придуман Ником Бостромом (Nick Bostrom) и получил название «максимизатор скрепки». В этом мысленном эксперименте алгоритму искусственного интеллекта было поручено изготовить как можно больше скрепок. Однако такая неограниченная цель заставила ИИ принимать любые меры, необходимые для покрытия Земли фабриками по производству скрепок, истребляя, в том числе, человечество по пути<sup>265</sup>. Попытки самих компаний, создающих большие языковые модели, решить проблемы согласования свелись к использованию подходов «обучение с подкреплением на основе отзывов людей» и «красной команды». На практике применение обоих подходов по сути, аналогично работе тех же тестировщиков, но не для программного обеспечения, а для моделей искусственного интеллекта. Оно не смогло полностью решить проблемы. Почему? Как отмечал создатель кибернетики Н. Винер, чтобы эффективно предотвращать катастрофические последствия, наше понимание наших искусственных машин должно в целом развиваться наравне [шаг за шагом] с характеристиками машины. Из-за самой медлительности наших человеческих действий наш эффективный контроль над нашими машинами может быть сведен к нулю<sup>266</sup>. Государственное регулирование как раз и должно выступать в

---

<sup>265</sup> Bostrom N. Ethical issues in advanced artificial intelligence //Science fiction and philosophy: from time travel to superintelligence. – 2003. – С. 277-284.

<sup>266</sup> Wiener N. Some Moral and Technical Consequences of Automation: As machines learn they may develop unforeseen strategies at rates that baffle their programmers //Science. – 1960. – Т. 131. – №. 3410. – С. 1355-1358.

качестве того звена, которое обеспечит соразмерность развития моделей ИИ и человека.

После восстановления С. Альтмана в должности исполнительного директора OpenAI, курс на погоню за прибылью и пренебрежение безопасностью продолжался, что привело в июне 2024 г. к уходу всех специалистов по безопасности ИИ, включая соучредителя и главного научного сотрудника И. Суцкевера. Создателей моделей искусственного интеллекта больше всего тревожит *проблема «интерпретируемости»*, то есть полного понимания того, как модели ИИ достигают своих результатов. Конструктор в начале создаёт программу в своей голове, а затем пишет код. Это позволяет ему объяснить как должна работать машина. В отличие от человека, модели машинного обучения программируют, по сути, сами себя, представляя «чёрный ящик» для исследователей. Как следствие, отсутствие значимого прогресса в интерпретируемости и заставляет ответственных создателей моделей говорить о том, что сфера искусственного интеллекта нуждается в государственном регулировании для предупреждения экзистенциальных рисков. В реальности логика бизнеса вела развитие сферы ИИ в противоположном направлении. Пример OpenAI не единичен: Microsoft весной 2023 г. уволила свою команду по этике искусственного интеллекта. В США и других странах «золотого миллиарда» обеспокоенность исследователей проигнорировали. Scott Galloway (Скотт Гэллоуэй), ведущий специалист в американском технологическом дискурсе, утверждает,

что тормозить развитие генеративного искусственного интеллекта было бы опрометчиво, утверждая, что «Китай, Россия и Северная Корея не будут делать паузу». Ему вторит и директор по информационным технологиям Пентагона<sup>267</sup>. Примечательно, что Китай, напротив, фактически поддержал обеспокоенность ученых и специалистов ИИ. КНР с 15 августа 2023 г. ввела с помощью Администрации киберпространства Китая («АКК») (Cyberspace Administration of China (“САС”), «Временные меры по администрированию служб генеративного искусственного интеллекта»<sup>268</sup>, которыми сознательно замедлила развитие моделей ИИ в рамках более широкой стратегии страны по осторожному снижению рисков в различных технологических областях.

Модели искусственного интеллекта, в отличие от действий монополистов прошлого века, обладают *свойством невидимости*. Они оказывают огромное влияние на функционирование существующих рынков, на их границы и на способы взаимодействия фирм между собой и с потребителями. Это, однако, может означать новые угрозы для благосостояния потребителей, увеличивая риск

---

<sup>267</sup> Abiri G., Huang Y. A red flag? China’s generative AI dilemma //Harvard journal of law & technology. – 2023. – Vol. 37. – P. 15.

<sup>268</sup> Shengchengshi Rengong Zhineng Fuwu Guanli Zanzing Banfa (生成式人工智能服务管理暂行办法) [Временные меры по администрированию служб генеративного искусственного интеллекта] (обнародованы Администрацией киберпространства Китая 10 июля 2023 г., вступают в силу 15 августа 2023 г.) St. Council Gaz., 30 августа 2023 г., стр. 40, URL: // [https://www.gov.cn/gongbao/2023/issue\\_10666/](https://www.gov.cn/gongbao/2023/issue_10666/)

новых неуловимых форм сговора и практики эксплуатации фирм<sup>269</sup>. Всё более активное применение больших языковых моделей порождает необходимость и более широкого государственного вмешательства не только в чисто экономические, но и в социальные вопросы. Как отмечал М. О'Коннелл, мы систематически наблюдаем, как технологии меняют всё к лучшему, и признаём, что приложение или платформы, или устройства делают мир немного удобнее. Если у нас есть надежда на будущее, если мы думаем, что у нас есть будущее, то это во многом основывается на том, чего мы можем добиться с помощью наших машин<sup>270</sup>. Значит государство своим регулированием должно поставить чёткие рамки развития технологий ИИ в интересах людей. Для обоснования необходимости государственного регулирования всё время используется термин «риск», однако, как отмечает Чемберлен, риск не является юридическим понятием, и возникает ряд вопросов относительно его значения в юридическом контексте<sup>271</sup>.

---

<sup>269</sup> Abrardi L., Cambini C., Rondi, L. Artificial intelligence, firms and consumer behavior: A survey // *Journal of Economic Surveys*. 2022. Vol. 36. No.4. P. 969-991.

<sup>270</sup> О'Коннелл М. Искусственный интеллект и будущее человечества. – Москва: Эксмо, 2020. – 272 с.

<sup>271</sup> Chamberlain J. The risk-based approach of the European Union's proposed artificial intelligence regulation: Some comments from a tort law perspective // *European Journal of Risk Regulation*. – 2023. – Vol. 14. – №. 1. – P. 1-13.

## **Какие именно риски применения искусственного интеллекта вышли на первый план в двадцатые годы XXI века?**

Прежде всего, речь идёт об “frontier AI” models («передовых моделях ИИ») — высокоэффективных базовых моделями, которые могут обладать опасными возможностями, достаточными для того, чтобы представлять серьёзную угрозу общественной безопасности. Например, они могут стать частью инструментов трудоустройства, которые будут ранжировать кандидатов на работу и отвечать за них, или систем администрации больниц, составляющих письма пациентам на основе материалов дела. Выгоды здесь для общества в виде высвобождения времени для профессионалов, чтобы они могли сосредоточиться на существенных вопросах – например, на фактическом лечении пациентов<sup>272</sup>, <sup>273</sup>. В более широком плане, такие мультимодальные механизмы принятия решений могут способствовать более эффективному и более справедливому распределению ресурсов, то есть той проблеме, над решением которой экономисты постоянно работают. Однако, риски ошибочных выборов, когда модели ИИ предвзято отдадут предпочтение полу, возрасту, расе кандидатов, и адекватного их устранения будут крайне дорого обходиться обществу. Именно «передовые модели искусственного интеллекта» представляют

---

<sup>272</sup> Hacker P., Engel A., Mauer M. Regulating ChatGPT and other large generative AI models //Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. – 2023. –. 1112-1123.

<sup>273</sup> Лукичев П. М., Чекмарев О.П. Экономика искусственного интеллекта: перспективы и риски. – Санкт-Петербург. 2023. – 188 с.

собой особую регулятивную проблему: во-первых, опасные возможности могут возникнуть неожиданно и быть трудно обнаруживаемыми; во-вторых, непросто надёжно предотвратить неправильное использование развернутой модели и причинение ею ущерба; в-третьих, трудно остановить широкое распространение возможностей модели поскольку они могут быстро распространяться, позволяя обходить меры безопасности <sup>274</sup> . Вмешательство государство здесь необходимо потому, что саморегулирование вряд ли обеспечит достаточную защиту от рисков, связанных с «передовыми моделями ИИ»

Политика государственного вмешательства в сферу искусственного интеллекта в последние два года активно продвигается самими участниками рынка. Если раньше представители Силиконовой долины всегда выступали против любых форм государственного регулирования, то с 2023 года его лоббируют такие цифровые гиганты, как Alphabet и Microsoft, а также создатели искусственного интеллекта, как Anthropic и OpenAI. Они на словах обеспокоены тем, что необузданная конкуренция заставит их действовать безрассудно, выпуская модели, которыми легко можно злоупотребить. Предложения цифровых гигантов по регулированию ИИ сводятся к нескольким ключевым положениям. Во-первых, они хотят, чтобы оно было узким и нацелено только на экстремальные риски. Однако сейчас, как было показано в статье П.

---

<sup>274</sup> Anderljung, M., Barnhart, J., Leung, J., Korinek, A., O'Keefe, C., Whittlestone, J., ... & Wolf, K. (2023). Frontier AI regulation: Managing emerging risks to public safety. *arXiv preprint arXiv:2307.03718*.

М. Лукичёва, О. П. Чекмарева о долгосрочны рисках ИИ<sup>275</sup>, это практически невозможно и их выявление в данное время скорее искусство, чем наука.

Во-вторых, они хотят ограничить регулирование только наиболее мощными «передовыми» моделями (“frontier AI” models). Microsoft, среди прочих, призвала к введению режима лицензирования, требующего от фирм регистрировать модели, производительность которых превышает определенные пороговые значения<sup>276</sup>.

В-третьих, они настаивают на контроле над продажей мощных чипов, используемых для обучения llms. В-четвёртых, они требуют, чтобы компании, занимающиеся облачными вычислениями, информировали власти, когда клиенты обучают передовые модели. По нашему мнению, такое лицензирование (по пороговым значениям) и по ограничению продаж мощных чипов создадут *Суперлигу игроков в сфере искусственного интеллекта*, к которым должно быть особое отношение регуляторов. Технологическим основанием такой Суперлиги выступают необходимые вычислительные ресурсы и накопленные технические навыки. Кроме того, многие компании и правительства обвиняют цифровых гигантов в корыстной заинтересованности и в устранении

---

<sup>275</sup> Лукичев П. М., Чекмарев О.П. Долгосрочные риски применения искусственного интеллекта в экономике // Вопросы инновационной экономики. 2023. Т. 13. № 4. С. 2427–2442.

<sup>276</sup> Of evils and evals. The Economist. October 28th, 2023.

возможных конкурентов при выдвижении таких предложений государственного регулирования искусственного интеллекта.

Новое регулирование может легко укрепить позиции действующих игроков и заблокировать конкурентов, не в последнюю очередь потому, что крупнейшие разработчики моделей тесно сотрудничают с правительствами при написании свода правил<sup>277</sup>. Обвинения в своекорыстности «озабоченности» фирм Силиконовой долины наиболее полно выражено Ферри и Глорих (Ferri and Gloerich). Они предполагают, что некоторые повествования об «экзистенциальном риске» и вымирании человечества, вызванном искусственным интеллектом, которые претендуют на защиту «будущего человечества», стремились узаконить тех немногих, кто участвовал в обсуждениях безопасности ИИ и принятии решений. как представители и защитники «человечности» как таковой. Это привело к «нисходящему подходу к управлению», который поднял западные технологические элиты на позицию исключительного техно политического и эпистемического авторитета одновременно, в то же время обесценивая важность включения различных голосов и жизненного опыта остального мира в политику искусственного интеллекта и процессы управления<sup>278</sup>.

---

<sup>277</sup> Think, then act. The Economist. October 28th, 2023.

<sup>278</sup> Ferri G., Gloerich I. Risk and harm: Unpacking ideologies in the AI discourse //Proceedings of the 5th International Conference on Conversational User Interfaces. – 2023. – P. 1-6.

## **Проблемы практического регулирования технологий искусственного интеллекта**

Для предупреждения выявленных рисков при распространении больших языковых моделей (БЯЗ – LLMs) необходимо государственное регулирование, как на национальном, так и на глобальном уровне (даже в условиях сложной международной обстановки). Однако на практике очень трудно определить основные составные элементы такого регулирования.

**Чьё это будет регулирование** – США и примкнувших к ним стран, или оно будет построено на каких-то новых паритетных началах? Сейчас в мировой экономике выделяются три центра использования моделей искусственного интеллекта – США, КНР и ЕС. Другие страны, включая Россию, значительно отстают от первых двух. Пока меры государственного регулирования искусственного интеллекта принимаются исключительно на национальном уровне. Лидирует здесь и по времени начала регулирования, и по детализации конкретных мер регулирования КНР. В условиях блокировки деятельности Всемирной торговой Организации (ВТО) и начавшейся, по сути, новой «холодной войны» между США и Китаем принятие глобальных мер регулирования применения искусственного интеллекта представляется маловероятным. От «технологических войн» глобальная экономика и, прежде всего, противоборствующие страны, несут громадные потери. По оценке МВФ прекращение торговли высокими

технологиями между конкурирующими блоками может стоить 1,2% мирового ВВП каждый год – около 1 трлн долларов<sup>279</sup>. Кроме того, уменьшаются потенциальные выгоды, которые могла бы получить каждая страна, включая и США, и КНР, за счёт расширения глобального разделения труда.

Во-вторых, **какими должны быть конкретные цели регулирования ИИ?** Прежде всего, подчеркнём, что установление целей затруднено из-за «размытости» термина «искусственный интеллект». Как отмечал Стоун с коллегами, определить сферу регулирования ИИ особенно сложно, поскольку термин искусственный интеллект используется для очень многих разных систем — «это не что-то одно»<sup>280</sup>. Какие опасения реально лежат за призывами к регулированию? М. Камински выделяет три вида таких опасений. Санкционные опасения приводят к предложениям регулировать алгоритмы для защиты человеческого достоинства и автономии; оправдательные опасения предупреждают, что мы должны оценить легитимность алгоритмических рассуждений; а инструментальные проблемы приводят к призывам к регулированию для предотвращения последующих проблем, таких как ошибки и

---

<sup>279</sup> Rage against the machines. The Economist. April 27th, 2024.

<sup>280</sup> Stone P., Brooks R., Brynjolfsson E., Calo R., Etzioni O., Hager G., ... & Teller A. Artificial intelligence and life in 2030: the one-hundred-year study on artificial intelligence //arXiv preprint arXiv:2211.06318. – 2022.

предвзятость<sup>281</sup>. По её мнению, ни один подход к регулированию не может эффективно решить все три проблемы.

Американские ученые предлагают первоначальный набор стандартов безопасности при регулировании больших языковых моделей. К ним относятся проведение оценок рисков перед развертыванием; внешний анализ поведения модели; использование оценок рисков для обоснования решений о развертывании; а также мониторинг и реагирование на новую информацию о возможностях и использовании модели после ее развертывания<sup>282</sup>. В идеале, по мнению директора по информационным технологиям Goldman Sachs Марко Ардженти (Marco Argenti), целью государственного регулирования ИИ должна быть сильная нормативно-правовая база для поддержания безопасных и надежных рынков и защиты потребителей. В то же время в идеале правила должны быть построены таким образом, чтобы обеспечить процветание инноваций и поддерживать равные условия игры. Заглядывая в будущее, важно будет продолжать создавать среду, которая поощряет сотрудничество между игроками, поощряет открытый исходный код моделей, когда это необходимо, и разрабатывает соответствующие правила, основанные на принципах, призванные помочь управлять потенциальными рисками, включая предвзятость,

---

<sup>281</sup> Kaminski M. E. Binary governance: Lessons from the GDPR'S approach to algorithmic accountability // Southern California Law Review. 2019. Vol. 92 (6). P. 1529.

<sup>282</sup> Anderljung, M., Barnhart, J., Leung, J., Korinek, A., O'Keefe, C., Whittlestone, J., ... & Wolf, K. (2023). Frontier AI regulation: Managing emerging risks to public safety. *arXiv preprint arXiv:2307.03718*.

дискриминацию, безопасность и надежность и конфиденциальность<sup>283</sup>.

В-третьих, **что именно нужно регулировать?** Определение сферы регулирования является наиболее сложной проблемой. Её правильное установление определяет, применимо ли регулирование в данном конкретном случае<sup>284</sup>. Сфера применения определяет, что регулируется (материальная сфера или сфера услуг), кто именно регулируется (например, личная сфера), где применяется регулирование (территориальная сфера: регион, национальный уровень, глобальный уровень) и когда оно применяется (временная сфера). По мнению Хакера и др., государственное регулирование искусственного интеллекта должно быть нацелено на Большие генеративные модели искусственного интеллекта (Large generative AI models - LGAIM), такие как ChatGPT, GPT-4 или Stable Diffusion в то время, как в ЕС регулирование нацелено на традиционные модели ИИ. Поэтому они предлагают новую терминологию для отражения цепочки создания стоимости искусственного интеллекта в настройках LGAIM, проводя различие между разработчиками LGAIM, развертывателями (deployers), профессиональными и непрофессиональными пользователями, а также получателями результатов LGAIM, и приводят доводы в пользу трех уровней обязательств, касающихся LGAIM (минимальные стандарты для

---

<sup>283</sup> A new generation of AI tools and models is emerging. URL: <https://www.goldmansachs.com/intelligence/pages/a-new-generation-of-ai-tools-and-models-is-emerging.html>

<sup>284</sup> Schuett J. Defining the scope of AI regulations //Law, Innovation and Technology. – 2023. – Vol. 15. – №. 1. – P. 60–82.

всех LGAIM; обязательства высокого риска для случаев использования с высоким уровнем риска; сотрудничество в цепочке создания стоимости ИИ)<sup>285</sup>.

Поскольку развитие технологий искусственного интеллекта интенсифицировалось после введения ChatGPT в ноябре 2022 года и их возможности, а, следовательно, и риски возрастают с каждым месяцем, то и государственные регулирующие органы должны быть готовы быстро реагировать на возникающие изменения, а, с другой стороны, не торопиться с введением ограничений, которые окажутся ненужными или даже блокирующими развитие технологического прогресса. Всё это означает новый уровень требований к подготовке регуляторов, которым придётся учиться буквально «на лету». Государственные служащие должны быть готовы не только переобучаться, но и меняться в соответствии с требованиями современности. На наш взгляд, это требует введения новых программ подготовки (переподготовки) государственных служащих. Здесь необходимо использовать, прежде всего, успешный китайский опыт. Когда чиновники Cyberspace Administration of China (CAC) (Администрации киберпространства Китая (АКК) впервые встретились с компаниями, занимающимися искусственным интеллектом, чтобы обсудить представленные ими алгоритмы, они, как сообщается, «проявили мало понимания технических деталей», что вынудило представителей компаний

---

<sup>285</sup> Hacker P., Engel A., Mauer M. Regulating ChatGPT and other large generative AI models // Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. – 2023. – P. 1112.

«полагаться на сочетание метафор и упрощенного языка»<sup>286</sup>. Для персонала регулирующих органов выявленное незнание означало необходимость учиться и узнавать то, что они не понимают в новых сложных технологиях, повышать свою квалификацию или нанимать новых специалистов, выстраивать отношения с ключевыми компаниями в сфере искусственного интеллекта, В результате за последние годы регулирующие органы КНР накопили опыт и создали ноу-хау в отношении технологий искусственного интеллекта и потенциальных мер вмешательства.

Законодатели США, а тем более ЕС и других стран, также не владеют полноценными знаниями о моделях искусственного интеллекта, и по их собственным словам, технологические инновации снова опережают возможности Вашингтона их регулировать<sup>287</sup>. Таким образом, человеческий фактор сейчас играет негативную роль в регулировании больших языковых моделей. Такое фундаментальное несоответствие объясняется, на наш взгляд, тремя причинами. Во-первых, быстрым и непредсказуемым прогрессом технологий ИИ за которым не поспевают законодатели. Во-вторых, дефицитом знаний и практического опыта в области искусственного интеллекта среди государственной бюрократии. В-

---

<sup>286</sup> Sheehan M. China's AI regulations and how they get made //Horizons: Journal of International Relations and Sustainable Development. – 2023. – №. 24. – С. 116.

<sup>287</sup> Kang C., Satariano A. New York Times. March 4, 2023, Section A, Page 1 of the New York edition with the headline: A.I. Regulation Can Be Puzzle To Lawmakers. <https://www.nytimes.com/2023/03/03/technology/artificial-intelligence-regulation-congress.html>

третьих. опасениями заблокировать развитие перспективной технологии в стране слишком жёсткими ограничениями.

Государственное регулирование должно учитывать национальную специфику развития сферы искусственного интеллекта. Прогресс новых технологий осуществляется сейчас через стартапы. В связи с этим проанализируем два принципиально разных подхода к их созданию, которые осуществляются в США и в КНР – стартапы. Начнём с цели. Большинство основателей стартапов в США стремятся масштабировать и продавать свои компании и использовать доходы для новых предприятий<sup>288</sup>.

**Таблица 2.** Сравнение экосистем финансирования стартапов (Ожидаемые цели основателей стартапов в %) <sup>289, 290</sup>.

	Англия	Канада	США	КНР
Приобретение	58	60	58	14
Вывод своей компании на биржу посредством IPO	18	16	17	46
Сохранение частной компании	11	20	14	21

<sup>288</sup> Silicon Valley Bank, “2020 Global Startup Outlook: Key Insights from the Silicon Valley Bank Startup Outlook Survey,” Silicon Valley Bank (2020): 3, [https://www.svb.com/globalassets/library/uploadedfiles/content/trends\\_and\\_insights/reports/startup\\_outlook\\_report/suo\\_global\\_report\\_2020-final.pdf](https://www.svb.com/globalassets/library/uploadedfiles/content/trends_and_insights/reports/startup_outlook_report/suo_global_report_2020-final.pdf)

<sup>289</sup> Silicon Valley Bank, “2020 Global Startup Outlook: Key Insights from the Silicon Valley Bank Startup Outlook Survey,” Silicon Valley Bank (2020): 3, [https://www.svb.com/globalassets/library/uploadedfiles/content/trends\\_and\\_insights/reports/startup\\_outlook\\_report/suo\\_global\\_report\\_2020-final.pdf](https://www.svb.com/globalassets/library/uploadedfiles/content/trends_and_insights/reports/startup_outlook_report/suo_global_report_2020-final.pdf)

<sup>290</sup> Zhang Shidong, “China Tops Global IPO Ranking in First Half as Issuers Leverage Capital Market Reforms, AI Frenzy,” South China Morning Post, July 4, 2023, <https://www.scmp.com/business/china-business/article/3226489/china-tops-global-ipo-ranking-first-half-issuers-leverage-capital-market-reforms-ai-frenzy> ; George Steer, Nicholas Megaw, and Hudson Lockett, “China Dominates Global IPO Market as Wall Street Fails to Rebound,” Financial Times, May 2, 2023, <https://www.ft.com/content/e5a43d1d-b62c-4428-b16f-clfe6b4fb9d9>

«Не знаю»	или	13	4	11	19
«Предпочитаю не говорить»					

Англосаксонский подход означает важность того, антимонопольная политика облегчает основателям стартапов развитие и продажу своих компаний по высокой цене. В КНР же первичные публичные размещения акций собрали в пять раз больше денег, чем в Соединенных Штатах в первой половине 2023 года. Напротив, поскольку сейчас наблюдается вялый рынок IPO, то выкуп остаётся наиболее жизнеспособным вариантом выхода для большинства стартапов в США. Соотношение сил между странами меняется очень быстро, и если в начале 2023 г. КНР отставала от США в создании больших языковых моделей на десятилетие, то к середине 2024 г. отставание сократилось до года или менее. Различие структуры рынков искусственного интеллекта США и КНР огромно. Если в США компания OpenAI имеет несомненное технологическое лидерство, то в КНР сейчас существует более 100 моделей с более чем миллиардом параметров<sup>291</sup>. Структура рынка близкая к совершенной конкуренции, создала сильную ценовую конкуренцию, что крайне благоприятно для потребителей, включая бизнес-пользователей. Например, 6 мая 2024 г. High-Flyer, квантовый хедж-фонд, создавший собственную модель, снизил цены на последнюю версию до примерно 1% от суммы, которую Openai взимает за свою самую шикарную модель gpt-4 Turbo. 22 мая Alibaba,

---

<sup>291</sup> The LLM summer sale. The Economist. June 15<sup>th</sup>, 2024.

конгломерат электронной коммерции, снизил цены на свою флагманскую большую языковую модель (БЯМ – 11m) на 97%. Несколько часов спустя Baidu объявила, что её чат-боты Ernie будут бесплатными для всех бизнес-пользователей. Tencent, крупнейшая интернет-компания Китая, также бесплатно раздаёт одну из своих БЯМ. Данная тенденция ведёт к созданию мультипликативного эффекта для всего народного хозяйства Китая.

В то же время ЕС стремится стать мировым центром регулирования технологий искусственного интеллекта, не имея своих технологий.

Национальная специфика сферы ИИ проявляется также в отношении к получению данных для обучения Больших языковых моделей. Например, Япония и Израиль занимают разрешительную позицию в отношении продвижения своих отраслей искусственного интеллекта. В США разные правообладатели данных занимают неодинаковые позиции. Getty Images подала в суд на Stability ai, фирму по созданию изображений, за несанкционированное использование её хранилища изображений. The New York Times подала в суд на Openai и Microsoft за нарушение авторских прав на миллионы статей. В то же время Reddit, сайт социальных сетей, и X (ранее Twitter) теперь взимают плату за доступ к своему контенту для обучения.

Каким будет государственное регулирование моделей искусственного интеллекта в ответ на разрастающиеся технологические вызовы покажет ближайшее будущее. Следует

учесть, что «практика регулирования обычно является сложной, беспорядочной и крайне несовершенной; что решение сложных проблем предполагает сложное взаимодействие множества людей и организаций с конфликтующими или, по крайней мере, разными интересами, пониманиями и ценностями; и часто требует создания, адаптации и внедрения множества методов, которые могут пересекаться или не пересекаться друг с другом»<sup>292</sup>. Поэтому ключевым вопросом здесь является: должно ли преобладать национальное или наднациональное регулирование технологий ИИ?

Для предприятий развитых стран и России сейчас происходит благодаря технологиям искусственного интеллекта невидимая смена «бизнес-моделей». Под последней понимается архитектура процесса и соответствующая организационная модель, с помощью которой компания участвует в конкурентной борьбе. Например, то, как страховщик обрабатывает заявление, как банк определяет размер автокредита или как ритейлер управляет сетью поставок. Сегодня в подавляющем большинстве случаев эти ключевые аспекты бизнеса вдруг стали очень архаичными<sup>293</sup>. Их архаичность объясняется тем, что большинство бизнес-процессов у сегодняшних фирм сложились до появления цифровых технологий и, в частности, больших языковых моделей и цифровых двойников. Данная тенденция происходит сейчас не только на уровне предприятий, но и на уровне

---

<sup>292</sup> Black J., Murray A. D. Regulating AI and machine learning: setting the regulatory agenda //European journal of law and technology. – 2019. – Vol. 10. – №. 3.

<sup>293</sup> Фрэнк М., Рериг П., Принг Б. Что делать, когда машины начнут делать всё: Как роботы и искусственный интеллект изменят жизнь и работу. – Москва: Эксмо. 2019, с. 119.

глобальной экономики. Новое Международное Разделение Труда (НМРТ) формируется с помощью технологий искусственного интеллекта не только естественным, конкурентным образом, но и за счёт введения санкций против стран-конкурентов (РФ, КНР)<sup>294</sup>. Фирма – это предприятие, действующее в рыночных условиях. Народное хозяйство складывается из многих тысяч национальных компаний и организаций. Чем лучше использование цифровизации на уровне фирмы, тем более конкурентна национальная экономика. Государственное регулирование использования моделей искусственного интеллекта должно не только уменьшать риски, но и создавать благоприятные условия для прогресса народного хозяйства. Именно поэтому применение технологий искусственного интеллекта, как технологии общего действия, на уровне каждого предприятия способно обеспечить перевод российской экономики с «инерционного пути развития» на «инновационный путь развития».

Таким образом, для эффективного применения технологий искусственного интеллекта на уровне народного хозяйства необходимы одновременные действия как на микроуровне, так и на макроуровне. На микроуровне предприятиям необходимо менять существующие бизнес-модели за счёт включения технологий ИИ, а на макроуровне нужно создавать систему государственного регулирования моделей искусственного интеллекта, способствующую прогрессу национальной экономики.

---

<sup>294</sup> Лукичев П. М. Позиция России в новом международном разделении труда // Вестник Удмуртского университета. Серия Экономика и право. 2022. Т. 32. № 5. С. 817–828.

## Выводы и будущие исследования

Необходимость государственного вмешательства в сферу искусственного интеллекта возрастает по мере появления всё более мощных больших языковых моделей. Выявлено *свойство невидимости*, которым обладают большие языковые модели и необходимость смены бизнес-модели современных компаний. Продвижение политики государственного регулирования моделей ИИ представителями Силиконовой долины вызовет создание *Суперлиги игроков в сфере искусственного интеллекта*, к которым будет особое отношение регуляторов и вызовет рост неравенства между странами. Поскольку человеческий фактор сейчас играет негативную роль, то важнейшим вопросом эффективности государственного регулирования на национальном уровне является подготовка государственных служащих, понимающих технологии ИИ и знающих современные способы их регулирования. Неодинаковый уровень развития моделей искусственного интеллекта в разных странах обусловил три различных подхода к государственному регулированию ИИ. Их анализу будет посвящена наш следующий раздел.

## **4. 2. ГОСУДАРСТВЕННОЕ РЕГУЛИРОВАНИЕ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

Технологии искусственного интеллекта после создания ChatGPT активно используются по всему миру. На сегодняшний день создание и применение моделей искусственного интеллекта стало основным объектом экономического противостояния, прежде всего, между КНР и США. В статье проводится сравнительный анализ комплекса мер государственного регулирования сферы искусственного интеллекта в отдельных странах. Данное регулирование выступает не только фактором, ограничивающим применение искусственного интеллекта, но и фактором его дальнейшего развития. КНР и США значительно опережают другие страны в создании и применении больших языковых моделей. Китай проделал самую значительную эволюцию мер государственного регулирования технологий искусственного интеллекта. Европейский союз, хотя и не производит собственные большие языковые модели, принял целый комплекс законов по регулированию применения искусственного интеллекта с точки зрения потребителей. В разделе выявлена «дилемма приоритетов» государственного регулирования применения искусственного интеллекта. В межгосударственном регулировании раскрыты особенности «конформистского» и «альтернативного» подходов. Автор обосновывает положение о необходимости не только национальных, но и согласованных международных усилий по регулированию больших языковых моделей.

### **Введение**

Данный раздел является логическим продолжением первого раздела главы о государственном регулировании применения

искусственного интеллекта (ИИ)<sup>295</sup> [4]. В нынешнем исследовании будет применён страновой подход для оценки эволюции мер государственного регулирования и их эффективности. Особенностью сегодняшнего этапа является неравномерность прогресса моделей ИИ и мер по их регулированию. Так, в Евросоюзе практически нет собственных фирм в сфере искусственного интеллекта, но принят целый комплекс законов по использованию ИИ. В США напротив существуют наиболее передовые лаборатории по созданию больших языковых моделей, организована вся вертикаль компаний по производству компонентов для рынка искусственного интеллекта, но государственное регулирование находится на начальной стадии. Первые попытки межгосударственного регулирования сферы искусственного интеллекта пока не стали результативными ввиду существующих противоречий между странами, прежде всего, между США и КНР.

**Цель раздела:** выявить возможности и ограничения существующих мер государственного регулирования применения искусственного интеллекта.

### **Необходимость государственного регулирования технологий искусственного интеллекта: международное или национальное?**

Необходимость международного регулирования применения моделей искусственного интеллекта вытекает из нескольких причин.

---

<sup>295</sup> См. Лукичёв, П. М. Государственное регулирование моделей искусственного интеллекта: необходимость и пределы макроэкономической политики / П. М. Лукичёв // Вопросы инновационной экономики. – 2024. – Т. 14, № 4. – С. 1151–1168. – DOI 10.18334/vinec.14.4.121767

Во-первых, глобализация мировой экономики и конкуренция между ведущими экономическими центрами, - США и КНР, - обуславливают потребность в межгосударственных рамках регулирования сферы ИИ. Во-вторых, ещё Д. Эджертон (Edgerton D.) предупреждал, что только в техно националистических фантазиях... национальные изобретения способствуют национальному экономическому росту. В реальном мире глобальные инновации приводят к национальному росту, а национальные инновации приводят к глобальному росту<sup>296</sup>.

В-третьих, существует значительный разрыв между странами в освоении моделей искусственного интеллекта. США и КНР значительно продвинулись в их освоении, в то время как страны ЕС и РФ в основном являются пользователями ИИ. Отсюда появление жёсткой конкуренции между национальными регуляторами, которая была названа «гонкой за регулирование искусственного интеллекта»<sup>297</sup>. Поэтому возникают значительные отличия в подходах к регулированию технологий искусственного интеллекта. Данные различия проявляются в детализации государственного регулирования ИИ. Отметим также существование **«дилеммы приоритетов» государственного регулирования применения искусственного интеллекта:** или снижение рисков,

---

<sup>296</sup> Edgerton D. The shock of the old: Technology and global history since 1900. – Oxford University Press, 2007.

<sup>297</sup> Smuha N. A. From a ‘race to AI’ to a ‘race to AI regulation’: regulatory competition for artificial intelligence //Law, Innovation and Technology. – 2021. – Vol. 13. – №. 1. – P. 57–84. <http://dx.doi.org/10.2139/ssrn.3501410>

включая экзистенциальные риски, от применения больших языковых моделей (БЯМ), или обеспечение прогресса в развитии национальной сферы искусственного интеллекта. В связи с этим некоторые ученые предупреждают, что сближение «неограниченной конкуренции между фирмами», участвующими в технологической гонке за вывод на рынок продуктов и услуг GenAI, с мотивацией государств предоставить своим собственным отечественным компаниям «конкурентное преимущество за счет мягкого регулирования» предвещает «гонку вниз по нормативным стандартам»<sup>298</sup>. Именно это и происходит в современных США, где «акселераторы» (сторонники безудержного развития искусственного интеллекта) победили «защитников» (сторонников учёта рисков ИИ), а президент Трамп в первый день в своей должности снял с разработчиков передовых LLMS (БЯМ) обязанность делиться информацией с американским правительством и затем в рамках дерегулирования ИИ пообещал партнёрам по Stargate (Звёздные врата), - проекте в полтриллиона долларов инвестиций частного сектора в американскую инфраструктуру искусственного интеллекта (ИИ), - сделать для них «настолько простым, насколько это возможно» строительство их проекта.

Регулирование технологий искусственного интеллекта имеет свою специфику по сравнению с другими направлениями

---

<sup>298</sup> Trager R., Harack B., Reuel A., Carnegie A., Heim L., Ho L., Kreps S., Lall R., Larter O., Ó hÉigeartaigh S., Staffell S., Villalobos J. J., Villalobos J. J. (2023). International governance of civilian AI: A jurisdictional certification approach. ArXiv. <https://doi.org/10.48550/arXiv.2308.15514>

государственного регулирования. Автор разделяет подход М. Гаске (Gaske M.), который отмечает, что схема регулирования может отдавать приоритет только двум из трёх целей при рассмотрении надзора за ИИ: (1) продвижение инноваций, (2) смягчение системного риска и (3) обеспечение чётких нормативных требований<sup>299</sup>.

Сегодня государственное регулирование технологий искусственного интеллекта осуществляется преимущественно на национальном уровне, а первые попытки создания межгосударственных форм регулирования вызвали больше вопросов, чем результатов. Поэтому в данном разделе мы кратко охарактеризуем проблемы международного регулирования ИИ и сосредоточимся на реальном национальном регулировании (дерегулировании) технологий искусственного интеллекта.

При оценке межгосударственного регулирования ИИ выделяются два противоположных подхода. Первый из них может быть определён как «конформистский», второй – как «альтернативный».

Сторонники «конформистского» подхода сигналом значительного прогресса считают возникновение в середине 2023 г. первой волны многосторонних инициатив в области

---

<sup>299</sup> Gaske M. R. Regulation Priorities for Artificial Intelligence Foundation Models //Vand. J. Ent. & Tech. L. – 2023. – Vol 26. – P 1. [Электронный ресурс]. URL: <https://scholarship.law.vanderbilt.edu/jetlaw/vol26/iss1/2/> (дата обращения 20.02.2025)

международного управления сферой искусственного интеллекта. Среди них выделяют, - UK AI Safety Summit's Bletchley Declaration, the International Code of Conduct for Organizations Developing Advanced AI Systems produced by the G7's Hiroshima AI Process, and the Partnership on AI's Guidance for Safe Foundation Models (Декларация Блетчли британского саммита по безопасности ИИ, Международный кодекс поведения организаций, разрабатывающих передовые системы ИИ, разработанный в рамках Хиросимского процесса ИИ «Большой семерки», и Партнёрство по руководству ИИ для моделей безопасного фундамента). Ценность этих инициатив в том, что они установили совместные обязательства среди участников по устранению выявленных передовых рисков ИИ и продвижению социально полезного использования ИИ; они инициировали процессы международного сотрудничества в области оценки и смягчения рисков; они добились создания нескольких национальных институтов безопасности ИИ; и они подчеркнули настоятельную необходимость сделать безопасность ИИ глобальным приоритетом<sup>300, 301, 302</sup>. Если обратиться к реальности, то нетрудно

---

<sup>300</sup> Garfinkel B., Anderljung M., Heim L., Trager R., Clifford B., Seger E. (2024) Goals for the Second AI Safety Summit. Centre for the Governance of AI. [Электронный ресурс]. URL: <https://www.governance.ai/post/goals-for-the-second-ai-safety-summit> (дата обращения 20.02.2025)

<sup>301</sup> Guzik T. J., Sitek A. Global accord on the integration of artificial intelligence in medical science publishing: implications of the Bletchley Declaration //Cardiovascular Research. – 2023. – Vol. 119. – №. 17. – P. 2681–2682. <https://doi.org/10.1093/cvr/cvad170>

<sup>302</sup> Oxford Analytica. UK AI Summit will promote some global cooperation //Emerald Expert Briefings. – 2023. – №. oхан-db. [Электронный ресурс]. URL: <https://www.emerald.com/insight/content/doi/10.1108/oxan-db284110/full/html>. (дата обращения 02.02.2025).

заметить, что участниками этих инициатив были в основном страны OECD (ОЭСР – Организация экономического сотрудничества и развития), в них ни участвовали ни Китай, ни Россия, ни многие развивающиеся страны. По сути, это был «междусобойчик» клуба развитых стран и попытка выдать свои цели за цели глобального развития.

Сторонники альтернативного подхода подчеркивают, что первая волна международных инициатив в области политики и управления ИИ пагубно сузила давние дискуссии об ответственном и этичном использовании искусственного интеллекта до ограниченного набора преимущественно технических вопросов, связанных с «безопасностью ИИ». Это означало, что вместо прямого противостояния непосредственным угрозам искусственного интеллекта, гражданские, социальные, политические и юридические права и экологическая устойчивость, вызванные безответственной массовой коммерциализацией систем GenAI, обсуждение политики и управления сместилось в сторону моделей, ориентированных на согласование ИИ, тестирование моделей и отчётность, оценку возможностей, надёжность системы, и мониторинг рисков <sup>303</sup>. Джина Хелфрик (Gina Helfric) верно отмечала: «Я утверждаю, что принятие термина “frontier AI” («пограничный ИИ») вызывает ассоциации с колониальным мышлением, ещё больше усиливая пагубную динамику отношений между горсткой влиятельных западных

---

<sup>303</sup> Leslie D., Perini A. M. Future Shock: Generative AI and the international AI policy and governance crisis. – 2024.

компаний, которые производят современные модели генеративного искусственного интеллекта, и населением «глобального Юга», которое с наибольшей вероятностью пострадает в результате разработки и внедрения этих технологий ИИ»<sup>304</sup>.

Автор поддерживает данный подход и считает, что первая волна международных инициатив привела к «подмене тезиса». Они, по сути, поддержали легитимность существующих крупных технологических практик бессистемного выпуска нерегулируемых систем «чёрного ящика» в общественное достояние вместо того, чтобы продвигать принятие предварительных мер управления для защиты прав и интересов пострадавших людей до наступления потенциально вредных последствий <sup>305</sup>. Кроме того, они продемонстрировали, что управленческая деятельность в области искусственного интеллекта была неэффективной и отвлекающей, удовлетворяя интересы в дерегулировании крупных технологических компаний, не сумела создать жёсткого нормативного контроля, необходимого для устранения ощутимых рисков и вреда, управления сложной материальной реальностью глобальных цепочек создания стоимости в области ИИ, и вызывает дальнейшее укрепление наследия политической, экономической и

---

<sup>304</sup> Helfrich G. The harms of terminology: why we should reject so-called “frontier AI” //AI and Ethics. – 2024. – P. 1-7. <https://doi.org/10.1007/s43681-024-00438-1>

<sup>305</sup> Britten A. Bletchley Declaration: Nations unite on AI risk. Significance7 Volume 20. Issue 6. 2023. P. 2–3. <https://doi.org/10.1093/jrssig/qmad086>

социокультурной гегемонии Глобального Севера<sup>306</sup>,<sup>307</sup>,<sup>308</sup>. Таким образом, можно сделать вывод, что международные инициативы первой волны не столько разрешили глобальный кризис регулирования применения генеративного искусственного интеллекта, сколько создали новые проблемы. Успехи Китая в создании эффективных больших языковых моделей показывают, что без него международное регулирование ИИ будет неполным и недостаточным.

### *Эволюция государственного регулирования искусственного интеллекта в КНР*

Государственное регулирование технологий искусственного интеллекта началось в Китайской Народной республике с 2017 года. В плане Государственного совета Китая по ИИ от 2017 года была установлена последовательность целей управления искусственным интеллектом и сроков их исполнения. Согласно плану, «к 2020 году Китай будет... изначально устанавливать этические нормы, политику и правила ИИ в некоторых областях. К 2025 году в Китае будут

---

<sup>306</sup> Ryan-Mosley, T. (2023, June 12.) It's time to talk about the real AI risks. *MIT Technology Review*. [Электронный ресурс]. URL: <https://www.technologyreview.com/2023/06/12/1074449/real-ai-risks/> (дата обращения 13.02.2025)

<sup>307</sup> Vallor S., Luger E. A shrinking path to AI safety //Edinburgh Futures Institute. – 2023. [Электронный ресурс].URL: <https://efi.ed.ac.uk/a-shrinking-path-to-safety-how-a-narrowly-technical-approach-to-align-ai-with-the-public-good-could-fail/> (дата обращения 20.02.2025)

<sup>308</sup> Terzis P. Law and the political economy of AI production //International Journal of Law and Information Technology. – 2023. – Vol. 31. – №. 4. – P. 302–330. <https://doi.org/10.1093/ijlit/eaac001>

созданы начальные законы и правила искусственного интеллекта, этические нормы и системы политики, а также сформированы возможности оценки и контроля безопасности ИИ. К 2030 году Китай разработает более всеобъемлющие законы и правила искусственного интеллекта, а также этические нормы и систему политики».<sup>309</sup> Сейчас, в начале 2025 г., видно, что этот план чётко и последовательно реализуется, обеспечивая КНР лидирующие позиции в создании и применении технологий искусственного интеллекта. Для реализации данного плана в Китае был принят целый ряд законов и программ. Среди ключевых выделим следующие. С целью включения этики в разработку ИИ были разработаны этические нормы в отношении исследований, проектирования, применения и использования ИИ<sup>310</sup>. В 2021 году была опубликована Белая книга о надёжном искусственном интеллекте, в которой подчёркивается важность справедливости, подотчетности и прозрачности разработки ИИ, а также даются рекомендации правительству, компаниям и отраслям о том, как сделать искусственный интеллект надёжным<sup>311</sup>.

---

<sup>309</sup> Webster G., Creemers R., Kania E., Triolo P. Full Translation: China's "New Generation Artificial Intelligence Development Plan" (2017) – 2017. [Электронный ресурс]. URL: <https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/> (дата обращения 02.02.2025)

<sup>310</sup> Centre for Security Artificial Intelligence Released', 21 October 2021. [Электронный ресурс] URL: <https://cset.georgetown.edu/publication/ethical-norms-for-new-generation-artificial-intelligence-released/>. (дата обращения 10.02.2025)

<sup>311</sup> CAICT and JD Explore Academy, *White Paper on Trustworthy Artificial Intelligence*, (Beijing, 2021). [Электронный ресурс] URL: <http://www.caict.ac.cn/kxyj/qwfb/bps/202107/P020210709319866413974.pdf>. (дата обращения 10.02.2025)

Государственное регулирование сферы ИИ складывалось в Китае постепенно и представляет собой комплекс целей и мер по их реализации. В обобщённом виде китайские правила имеют три структурных сходства: выбор алгоритмов в качестве отправной точки; создание инструментов регулирования и бюрократических ноу-хау; а также вертикальный и итеративный подход, который закладывает основу для краеугольного закона об искусственном интеллекте.<sup>312</sup> План развития искусственного интеллекта на 2021–2025 годы является частью 14-го пятилетнего плана КНР.

Ключевой особенностью подхода КНР к управлению ИИ является ориентация на алгоритмы. Регулирование рекомендательных алгоритмов создало для регулирующих органов новый важный инструмент: реестр алгоритмов (算法备案系统, буквально «система хранения алгоритмов»). Реестр представляет собой онлайн-базу данных алгоритмов, которые имеют «свойства общественного мнения или . . . возможности социальной мобилизации». Разработчики этих алгоритмов обязаны предоставлять информацию о том, как их алгоритмы обучаются и применяются, в том числе на каких наборах данных обучается

---

<sup>312</sup> Sheehan M. China's AI regulations and how they get made //Horizons: Journal of International Relations and Sustainable Development. – 2023. – №. 24. – P. 115. [Электронный ресурс]. URL: <https://www.cirsd.org/files/000/000/010/82/21e461a985f43655b1731b3c1b50cdccb631afaf.pdf> (дата обращения 16.02.2025)

алгоритм. Они также обязаны заполнить «отчет о самооценке безопасности алгоритма» (算法安全自 评估报告.<sup>313</sup>

Для государственного регулирования КНР применения технологий ИИ также характерны максимально возможное предупреждение рисков создаваемых больших языковых моделей и учёт мнений общественности при обсуждении уже принятых законов. Так, в августе 2023 г. Cyberspace Administration of China («САС») (Администрация киберпространства Китая («САС»), главный регулятор информационных интернет-услуг Китая), опубликовала Временные меры по администрированию служб генеративного искусственного интеллекта<sup>314</sup>. Эти «Временные меры» приостановили на несколько месяцев создание новых БЯМ, но позволили разработчикам больше сосредоточиться на предупреждении рисков и поиске нешаблонных технологий ИИ, что дало свой эффект к началу 2025 года в создании моделей DeepSeek.

---

<sup>313</sup> Sheehan M. China's AI regulations and how they get made //Horizons: Journal of International Relations and Sustainable Development. – 2023. – №. 24. – P. 112. [Электронный ресурс]. URL: <https://www.cirsd.org/files/000/000/010/82/21e461a985f43655b1731b3c1b50cdccb631afaf.pdf> (дата обращения 16.02.2025)

<sup>314</sup> Shengchengshi Rengong Zhineng Fuwu Guanli Zanxing Banfa (生成式人工智能服务管理暂行办法) [Interim Measures for Administration of Generative Artificial Intelligence Services] (promulgated by the Cyberspace Admin. of China, Jul. 10, 2023, effective Aug. 15, 2023) St. Council Gaz., Aug. 30, 2023, at 40, [https://www.gov.cn/gongbao/2023/issue\\_10666/](https://www.gov.cn/gongbao/2023/issue_10666/) (China) [hereinafter Interim Measures]. [Временные меры по администрированию служб генеративного искусственного интеллекта] (обнародованы Администрацией киберпространства Китая 10 июля 2023 г., вступают в силу 15 августа 2023 г.) St. Council Gaz., 30 августа 2023 г., стр. 40. [Электронный ресурс ] URL: [https://www.gov.cn/gongbao/2023/issue\\_10666/202308/content\\_6900864.html](https://www.gov.cn/gongbao/2023/issue_10666/202308/content_6900864.html) (дата обращения 20.02.2025).

О регуляторах в КНР стоит сказать особо. Чиновники, регулирующие сферу искусственного интеллекта, постоянно совершенствуются, не становясь в позу «Мы правы и всё». В ответ на ряд критических замечаний по первоначальному проекту «Временных мер» регулирующие органы внесли существенные изменения. Например, показатели аутентичности, точности, объективности и разнообразия создаваемого контента и данных были изменены с абсолютных на относительные — вместо того, чтобы диктовать строгие правила, теперь они просто требуют улучшений или усовершенствований.) Кроме того, ранее включенное требование к пользователям раскрывать свою настоящую личность было полностью отменено, что является явной победой как для технологических корпораций, так и для сторонников конфиденциальности. Протокол фильтрации генерируемого контента и оптимизации моделей также был смягчен, что обеспечивает более управляемые временные рамки.<sup>315</sup> Всё это свидетельствует о том, что регулирующие органы не остались глухи к отзывам, полученным в период общественного обсуждения.

Отметим также непрерывный процесс совершенствования системы регулирования искусственного интеллекта в Китае. Правила являются вертикальными и итеративными. Если Правительство считает изданное им постановление ошибочным или

---

<sup>315</sup> Abiri G., Huang Y. A red flag? China's generative AI dilemma //Harvard journal of law & technology. – 2023. – Vol. 37. – P. 15. [Электронный ресурс] URL: <https://jolt.law.harvard.edu/digest/a-red-flag-chinas-generative-ai-dilemma> (дата обращения 20.02.2025).

недостаточным, оно просто выпустит новое постановление, которое затыкает дыры или расширяет сферу применения, как это произошло с проектом постановления о генеративном ИИ, расширяющим меры глубокого синтеза.<sup>316</sup>

В начале 2025 г. комплекс мер государственного регулирования искусственного интеллекта складывается из следующих основных элементов (в скобках год введения): Закон о кибербезопасности (Cybersecurity Law, 2017), Закон о защите персональных данных, (2021), Правила управления алгоритмическими рекомендациями, (2022), Этические нормы для ИИ (2021), Регулирование автономных систем и робототехники (2024), Регулирование глубоких подделок (Deepfakes, 2023)

### *Государственное регулирование (дерегулирование) искусственного интеллекта в США*

Особенностью США является очень большая востребованность общества на регулирование государством применения технологий искусственного интеллекта. Как показывают данные опроса, проводившегося Pew Research Center на протяжении 2024 г., более половины взрослых американцев (55%) и схожая доля экспертов по ИИ (57%) говорят, что хотят большего контроля над тем, как он используется в их жизни. И представители обеих групп больше

---

<sup>316</sup> Sheehan M. China's AI regulations and how they get made //Horizons: Journal of International Relations and Sustainable Development. – 2023. – №. 24. – P. 117. [Электронный ресурс]. URL: <https://www.cirsd.org/files/000/000/010/82/21e461a985f43655b1731b3c1b50cdccb631afaf.pdf> (дата обращения 16.02.2025)

беспокоятся о том, что государственное регулирование ИИ будет слишком слабым, чем чрезмерно чрезмерным.<sup>317</sup>

Теоретическое обоснование американского подхода к регулированию использования искусственного интеллекта основано на работе Йонаса Шуэтта (Schuett Jonas), Он отмечал, что существующие определения ИИ не соответствуют наиболее важным требованиям к юридическим определениям и утверждал, что подход, основанный на оценке риска, был бы предпочтительнее. Вместо того, чтобы использовать термин искусственный интеллект, политики должны сосредоточиться на конкретных рисках, которые они хотят уменьшить. Требования к юридическим определениям можно лучше выполнить, определив основные источники соответствующих рисков: определенные технические подходы (например, обучение с подкреплением), приложения (например, распознавание лиц) и возможности (например, способность физически взаимодействовать с окружающей средой).<sup>318</sup>

США, а вслед за ними Великобритания и Израиль, склонны применять контекстно-ориентированный подход к государственному регулированию ИИ. Этот подход предполагает,

---

<sup>317</sup> McClain C., Kennedy B., Gottfried J., Anderson M., Pasquini G. How the U.S. Public and AI Experts View Artificial Intelligence. [Электронный ресурс] URL: [https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence/?utm\\_source=Pew+Research+Center&utm\\_campaign=82cfde2055-Internet\\_Science\\_2025\\_04\\_29&utm\\_medium=email&utm\\_term=0\\_-82cfde2055-399474221](https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence/?utm_source=Pew+Research+Center&utm_campaign=82cfde2055-Internet_Science_2025_04_29&utm_medium=email&utm_term=0_-82cfde2055-399474221). (дата обращения 30.04.2025)

<sup>318</sup> Schuett J. Defining the scope of AI regulations //Law, Innovation and Technology. – 2023. – Vol. 15. – №. 1. – P. 60–82.7<http://dx.doi.org/10.2139/ssrn.3453632>

что социальный вред, вызванный искусственным интеллектом, отражает существующие проблемы, когда ИИ заменяет или дополняет традиционные методологии. Даже если новый социальный вред возникает в конкретном контексте использования искусственного интеллекта, он предпочитает устранять его с помощью специальных правил. Таким образом, этот подход благоприятствует адаптации правовой системы к социальному ущербу в различных контекстах, а не принятию комплексного регулирования ИИ.<sup>319</sup>

Американский подход к государственному регулированию сферы искусственного интеллекта отличают три ярко выраженные особенности. Во-первых, это *ориентация на интересы* отрасли, на производителей больших языковых моделей. Главное – не затормозить избыточным регулированием развитие сферы ИИ. Отсюда фактическое выведение *большой цифровой пятёрки и связанных с ними стартапов* вообще из-под какого-либо регулирования. Следствием этого являются, например, нерешаемые годами проблемы интеллектуальной собственности на источники, используемые для машинного обучения. Здесь достаточно вспомнить иск газеты New York Times к OpenAi.

Вторая особенность – это *регулирование сферы регулирования искусственного интеллекта на уровне штатов*. Из последних по времени отметим закон штата Колорадо, запрещающий

---

<sup>319</sup> Park S. Bridging the Global Divide in AI Regulation: A Proposal for a Contextual, Coherent, and Commensurable Framework //arXiv preprint arXiv:2303.11196. – 2023.

«алгоритмическую дискриминацию» в отношении потребителей на основе защищенных характеристик в перечисленных областях, а также активность законодателей штата Калифорния, которые в 2024 г. разработали и приняли десятки законопроектов, связанных с ИИ, направленных на наложение широкомасштабных обязательств, начиная от мер безопасности, прозрачности для потребителей, требований к отчётности, разъяснения гарантий конфиденциальности, защиты исполнителей и умерших знаменитостей и мер честности выборов.<sup>320</sup>

Только 30 октября 2023 г. был принят Федеральный указ 14110 «О безопасном, защищённом и заслуживающим доверия искусственном интеллекте» (Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence)<sup>321</sup>. Среди основных положений последнего выделяются: необходимость сообщать правительству США информацию о результатах тестирования безопасности систем ИИ; разработка стандартов; повышение безопасности персональных данных; запуск Национального ресурса исследований в области ИИ (National AI Research Resource); усиление международного

---

<sup>320</sup> Anderson H., Reem N., Tadayyon. Raft of California AI Legislation Adds to Growing Patchwork of US Regulation. [Электронный ресурс] URL: <https://www.whitecase.com/insight-alert/raft-california-ai-legislation-adds-growing-patchwork-us-regulation#:~:text=Senate%20Bill%20942%3A%20California%20AI%20Transparency%20Act&text=This%20Act%20outlines%20requirements%20for,are%20permitted%20for%20public%20use> (дата обращения 20.02.2025)

<sup>321</sup> Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. A Presidential Document by the [Executive Office of the President](https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence) on 11/01/2023 [Электронный ресурс]. URL: <https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence> (дата обращения 19.02.2025)

сотрудничества»<sup>322</sup> В итоге на территории США образуется своеобразная *мозаика регулирования искусственного интеллекта*, состоящая как из законодательных мер федеральных властей, так и властей штатов.

В-третьих, *политизированность* государственного регулирования США, его направленность против других стран, прежде всего, КНР. Отсюда вытекает максимальная «закрытость» сферы искусственного интеллекта, особенно в сравнении с открытостью моделей Китая. Ещё одним направлением политизированности является резкая «смена курса» регулирования сферы искусственного интеллекта с избранием нового президента США. В первый день вступления в должность Трамп отменил указ Байдена от 30.10.2023 и вместо всеобъемлющего регулирования искусственного интеллекта федеральные агентства будут контролировать использование ИИ по «отраслевому» подходу. В рамках проводимый им макроэкономической политики «экономики предложения» и, следовательно, создания максимально выгодных условий для фирм, действующих на территории США, был поддержан для улучшения инфраструктуры сферы искусственного интеллекта четырехлетний проект строительства огромных центров обработки данных в Техасе.

Эффективность такого государственного регулирования сферы ИИ будет определяться результатами конкуренции с КНР.

---

<sup>322</sup> Мамедьяров З. А. Регулирование искусственного интеллекта: первые шаги. Наука. Технологии. Инновации. Экспресс-информация 12.01.2024. ИСИЭЗ НИУ ВШЭ. [] URL: <https://issek.hse.ru/mirror/pubs/share/886272540>

## *Государственное регулирование применения искусственного интеллекта Евросоюзом*

С 2020 года ЕС активно занимается законодательной деятельностью в области ИИ. Особенностью этой деятельности служит то, что Евросоюз практически не имеет собственных фирм в сфере искусственного интеллекта, но является вторым в мире потребителем продукции ИИ. Поэтому законодательная деятельность ЕС в области искусственного интеллекта построена исключительно с точки зрения защиты прав покупателей и не учитывает трендов развития инновационного развития <sup>323</sup> [14]. Кроме того, следует сопоставить размеры западноевропейских и американских компаний, подпадающих под регуляторную деятельность Евросоюза. В эпоху доминирования технологических гигантов стоимостью в триллионы долларов ни одна европейская фирма, начинавшая с нуля за последние 50 лет, сегодня не оценивается более чем в сто миллиардов (Spotify, музыкальный стриминговый сервис из Швеции, колеблется около этой отметки).

324

---

<sup>323</sup> EU: Council gives final approval to AI Act. [Электронный ресурс]. URL: [https://www.dataguidance.com/news/eu-council-gives-final-approval-ai-act#:~:text=On%20May%2021%2C%202024%2C%20the,Intelligence%20\(the%20AI%20Act\).](https://www.dataguidance.com/news/eu-council-gives-final-approval-ai-act#:~:text=On%20May%2021%2C%202024%2C%20the,Intelligence%20(the%20AI%20Act).) (дата обращения 20.02.2025)

<sup>324</sup> The age of gunboat digital diplomacy. The Economist. January 25th, 2025. [Электронный ресурс]. URL: <https://www.economist.com/europe/2025/01/23/europe-faces-a-new-age-of-gunboat-digital-diplomacy> (дата обращения 20.02.2025)

С точки зрения трилеммы регулирования применения искусственного интеллекта М. Гаске (Gaske M.) в законодательстве ЕС практически отсутствует (1) продвижение инноваций, но чётко прописаны меры по (2) смягчению системного риска и (3) обеспечению чётких нормативных требований.<sup>325</sup>

Отметим, что отсутствие «серьёзных» национальных фирм-производителей подталкивает государства к самым странным экономическим решениям. Так, Япония уже приняла закон о том, что ИИ имеет право обучаться на любых данных, которые стали ему доступны<sup>326</sup>. Таким решением, ценность создания информации, интеллектуальной собственности сводится к нулю.

### *Регулирование искусственного интеллекта в других странах*

В этом разделе мы в начале покажем специфику государственного регулирования ИИ в других государствах – потребителях больших языковых моделей, а затем обратимся к опыту Сингапура. Иногда цели регулирования представляют попытки правительств оказать помощь национальному бизнесу за счёт цифровых гигантов и остановки технологического прогресса. В качестве примера приведём, принятый в Австралии в 2021 году закон «Переговорный кодекс новостных СМИ» (“news media

---

<sup>325</sup> Gaske M. R. Regulation Priorities for Artificial Intelligence Foundation Models //Vand. J. Ent. & Tech. L. – 2023. – Vol 26. – P 1. [Электронный ресурс]. URL: <https://scholarship.law.vanderbilt.edu/jetlaw/vol26/iss1/2/> (дата обращения 20.02.2025)

<sup>326</sup> Mollick E., Mollick E. Co-Intelligence. – Random House UK, 2024.

bargaining code”), который обязывал Google и Facebook платить новостным агентствам, когда результаты поиска или каналы социальных сетей ссылаются на их статьи. Аналогичные законы приняли Канада, Южная Африка, Великобритания, Бразилия. Практика применения в Канаде показала, что запрет на новости оказался губительным для национальных средств массовой информации, которым закон должен был помочь, а канадцы вынуждены были обмениваться скриншотами заголовков <sup>327</sup>. Аналогично, принятые ЕС в 2024 году Закон о цифровых рынках (The Digital Markets Act) и Закон об искусственном интеллекте (Artificial Intelligence Act) привели к тому, что Apple будет поставлять в эти страны упрощенный вариант iPhone без приложений с искусственным интеллектом. В обоих случаях от такого регулирования больше всего страдают конечные потребители.

В некоторых странах, как например в Сингапуре, государственное регулирование применения искусственного интеллекта, прошло значительную эволюцию за последние пять лет. Можно выделить в нём два этапа. Первый, когда в 2019 году появилась Рамочная модель регулирования искусственного интеллекта (The Model AI Governance Framework)<sup>328</sup> [34], а через год

---

<sup>327</sup> No news is bad news. The Economist. May 18<sup>th</sup>, 2024. [Электронный ресурс].URL: <https://www.economist.com/leaders/2024/05/16/canadas-law-to-help-news-outlets-is-harming-them-instead> (дата обращения 20.02.2025)

<sup>328</sup> The Model AI Governance Framework first edition. [Электронный ресурс].URL: <https://ai.bsa.org/wp-content/uploads/2019/09/Model-AI-Framework-First-Edition.pdf>. (дата обращения 20.02.2025)

- её обновление <sup>329</sup>. На этом этапе регулирование было направлено на традиционный ИИ. Акцент в Рамочной модели управления ИИ делался на принципах, которые лежат в основе регулирования – как применения, так и разработки искусственного интеллекта в Сингапуре - объяснимости, прозрачности, справедливости и человекоцентричности (англ. – human-centric) интеллектуальных систем. Последний принцип подвергается критике <sup>330</sup>. Хотя Рамочная модель регулирования искусственного интеллекта носила необязательный характер, многие технологические компании Сингапура добровольно взяли на себя обязательство руководствоваться её принципами и нормами. Второй этап государственного регулирования разработки и применения моделей ИИ осуществляется в Сингапуре в 2023–2024 гг. в связи с появлением генеративного искусственного интеллекта. Под последним понимаются большие языковые модели, такие как GPT-4, Gemini, Claude и LLaMA, способные генерировать текст, изображения или другие типы мультимедиа. В 2023 г. была организована дискуссия о прежних рисках применения искусственного интеллекта, таких как отсутствие объяснимости, предвзятость, неправильное использование, и возникших в связи с появлением БЯМ новых рисках, таких как галлюцинации,

---

<sup>329</sup> Model AI governance framework. Second edition. [Электронный ресурс]. URL: <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf>. (дата обращения 20.02.2025)

<sup>330</sup> Дремлюга Р.И. Основы национального регулирования применения искусственного интеллекта: опыт Сингапура // Азиатско-Тихоокеанский регион: экономика, политика, право. 2022. Т. 24, No1. С. 214–224. <https://doi.org/10.24866/1813-3274/2022-1/214-224>.

нарушение авторских прав, согласование ценностей<sup>331</sup>. Главная цель, которая ставилась здесь, это необходимость найти тщательный баланс между защитой пользователей и продвижением инноваций. 30 мая 2024 года была опубликована Модельная структура управления ИИ для генеративного ИИ. Формирование надежной экосистемы<sup>332</sup>. В ней выделены девять аспектов, которые предлагается рассмотреть в совокупности, чтобы создать надежную экосистему. Предлагаемые в Рамочной программе идеи направлены на дальнейшее развитие основных принципов подотчетности, прозрачности, справедливости, надежности и безопасности и требуют глобального сотрудничества в разработке политических подходов к прогрессу моделей искусственного интеллекта<sup>333</sup>.

### ***Выводы и будущие исследования***

Введение национального государственного регулирования технологий искусственного интеллекта неизбежно подразумевает ряд ограничений на развитие моделей ИИ. Соответственно, страна

---

<sup>331</sup> Discussion Paper on Generative AI: Implications for Trust and Governance. - Singapore. Infocomm Media Development Authority of Singapore (IMDA), Aicadium and AI Verify Foundation. 2023. [Электронный ресурс] URL: [https://aiverifyfoundation.sg/downloads/Discussion\\_Paper.pdf](https://aiverifyfoundation.sg/downloads/Discussion_Paper.pdf) (дата обращения 02.02.2025).

<sup>332</sup> Model AI Governance Framework for Generative AI. Fostering a Trusted Ecosystem. - Singapore. - AI Verify Foundation. - 2024. – 36 pages Published 30 May 2024. [Электронный ресурс]. URL: <https://aiverifyfoundation.sg/wp-content/uploads/2024/05/Model-AI-Governance-Framework-for-Generative-AI-May-2024-1-1.pdf>). (дата обращения 20.02.2025)

<sup>333</sup> Model AI Governance Framework for Generative AI. Fostering a Trusted Ecosystem. - Singapore. - AI Verify Foundation. - 2024. – 36 pages Published 30 May 2024. [Электронный ресурс]. URL: <https://aiverifyfoundation.sg/wp-content/uploads/2024/05/Model-AI-Governance-Framework-for-Generative-AI-May-2024-1-1.pdf>). (дата обращения 20.02.2025)

(страны), которая не ввела таких ограничений, получает конкурентное преимущество.

Создание национальной системы государственного регулирования искусственного интеллекта (или США, или КНР) означает, что другие государства, приобретающие их модели ИИ, попадают в экосистему искусственного интеллекта США или в экосистему искусственного интеллекта КНР. На выбор любой страной той или иной экосистемы искусственного интеллекта будут влиять не столько два обычных экономических параметра, - качество и цены, - сколько третий параметр: политические предпочтения.

Соотношение сил между странами-конкурентами постоянно меняется и решающего преимущества не имеет ни одно государство. США попытались всячески сузить возможности развития искусственного интеллекта в КНР, введя за последние три года целый ряд ограничений. Среди них сокращение США экспорта в Китай самых мощных и быстродействующих чипов для обучения моделей ИИ, лишение КНР доступа ко многим машинам, необходимым для производства заменителей. Кроме того, Китай действительно исключен из большинства многосторонних структур, сформированных избирательным западным членством, таких как G7, Совет Европы (СЕ) и Европейская комиссия, — не говоря уже о

некоторых недавно созданных инициативах по противодействию китайскому влиянию в глобальном управлении ИИ<sup>334</sup>.

После появления ChatGPT в ноябре 2022 года эволюция развития искусственного интеллекта происходила очень интенсивно, включая два новых качественных момента: Агенты искусственного интеллекта и «модели с рассуждением». Последние представляют собой модели ИИ, которые имеют своеобразное «мышление», разговаривая сами с собой, прежде чем дать ответ на запрос. В результате объективные оценки показывают в начале 2025 года, что отставание по качеству китайских моделей искусственного интеллекта от американских составляют недели, максимум – три месяца. Сэм Альтман, глава OpenAi, недавно писал: «(Относительно) легко скопировать то, что, как вы знаете, работает. Крайне сложно сделать что-то новое, рискованное и сложное, когда вы не знаете, сработает ли это»<sup>335</sup>. То есть запретительные меры США по отношению к Китаю в области искусственного интеллекта не дали явного эффекта. Напротив, они стимулировали поиск учеными КНР иных технологических решений в сфере ИИ, чем у американских лабораторий. Практика показала, что это возможно и что монополия США на создание больших языковых моделей была подорвана.

---

<sup>334</sup> Cheng J., Zeng J. Shaping AI's future? China in global AI governance //Journal of Contemporary China. – 2023. – Vol. 32. – №. 143. – P. 794–810, <https://doi.org/10.1080/10670564.2022.2107391>

<sup>335</sup> Uncomfortably close. The Economist. January 25th, 2025. [Электронный ресурс]. URL: <https://www.economist.com/briefing/2025/01/23/chinas-ai-industry-has-almost-caught-up-with-americas> (дата обращения 20.02.2025)

Более того, по нашему мнению, сейчас происходит **перелом в эволюции искусственного интеллекта**. Если до начала 2025 года преобладала ориентация на технологические параметры создания лучших моделей ИИ горсткой передовых лабораторий, поддерживаемых цифровыми гигантами (OpenAI, Anthropic), то сейчас на первый план вышла экономика искусственного интеллекта<sup>336</sup>.

Символом перелома стало создание китайской фирмой DeepSeek моделей «v3» и «r1», которые по качественным параметрам не уступают западным моделям, а стоят значительно дешевле. DeepSeek серией последовательно применяемых инноваций добилась того, что производство (создание) миллиардов параметров v3 заняло менее 3 млн чип-часов, при оценочной стоимости менее 6 млн долларов — примерно десятая часть вычислительной мощности и расходов, которые были потрачены на Llama 3.1. от фирмы Meta. Для обучения v3 потребовалось всего 2000 чипов, тогда как Llama 3.1 использовала 16 000. А из-за американских санкций чипы, которые использовала v3, были даже не самыми мощными<sup>337</sup>. Экономические параметры конкурентной

---

<sup>336</sup> Лукичев П. М., Чекмарев О. П. Экономика искусственного интеллекта: перспективы и риски. Санкт-Петербург, 2023. [Электронный ресурс]. URL: <https://motivtrud.ru/PCost/research/ii20231.pdf> (дата обращения 02.02.2025)

<sup>337</sup> Uncomfortably close. The Economist. January 25th, 2025. [Электронный ресурс]. URL: <https://www.economist.com/briefing/2025/01/23/chinas-ai-industry-has-almost-caught-up-with-americas> (дата обращения 20.02.2025)

борьбы в области ИИ, - качество и цена, - не могут рассматриваться изолированно друг от друга. Кроме того, в анализ цен следует включать постоянные издержки и переменные издержки. Первые в данном случае связаны с созданием большой языковой модели, вторые – с её эксплуатацией. На сегодняшний день видны разные результаты в применении экономического подхода «Издержки – Выгоды» к созданию больших языковых моделей Китаем и Америкой. Западные фирмы кажутся всё более расточительными с использованием чипов и, как следствие, с ценами на свою продукцию и услуги. Если сравнивать переменные издержки, то затраты на обслуживание китайской модели значительно дешевле. DeepSeek распределяет задачи между несколькими чипами более эффективно, чем её аналоги, и начинает следующий этап процесса до того, как предыдущий будет завершен. Это позволяет ему поддерживать работу чипов на полную мощность с небольшой избыточностью. Применительно к экономике искусственного интеллекта это аналогично преодолению недостатков поэтапного планирования, которое было, например, в NASA. В результате в феврале, когда DeepSeek начнёт позволять другим фирмам создавать сервисы, использующие v3, он будет взимать менее десятой части того, что Anthropic делает за использование Claude, своей большой языковой модели<sup>338</sup>.

---

<sup>338</sup> Uncomfortably close. The Economist. January 25th, 2025. [Электронный ресурс]. URL: <https://www.economist.com/briefing/2025/01/23/chinas-ai-industry-has-almost-caught-up-with-americas> (дата обращения 20.02.2025)

Означает ли это, что американский подход к созданию и использованию технологий искусственного интеллекта себя не оправдал? Конечно, нет. По многим технологическим параметрам и эксплуатационным свойствам американские модели ИИ опережают всех. Поэтому конкурентная борьба в сфере искусственного интеллекта, ведущая к удешевлению моделей ИИ и издержек на их эксплуатацию, улучшающая качество и разнообразие предоставляемых услуг, только разгорается. В результате пользователи во всём мире получают всё более значительные преимущества. Подход США к регулированию сферы искусственного интеллекта с приходом Трампа ориентирован на максимальное дерегулирование деятельности небольшого числа лабораторий, поддерживаемых цифровыми гигантами, на колоссальные размеры инвестиций в ИИ, на максимальную закрытость исследований, на политику ограничений и санкций против стран-конкурентов, прежде всего, КНР. Для подхода Китая характерны длительная последовательная эволюция государственного регулирования технологий искусственного интеллекта, экономическая децентрализация, существенно более низкие издержки на производство и эксплуатацию больших языковых моделей, максимальная открытость результатов исследований.

Следовательно, предпочтительность использования другими странами технологий искусственного интеллекта, произведённых в США, или произведённых в КНР, определяется и экономическими факторами, и политическими факторами. Кроме того, анализ опыта

создания моделей ИИ показывает Индии или России по какому пути следует идти в производстве собственных моделей искусственного интеллекта.

Государственное регулирование применения искусственного интеллекта, как и в целом прогресс больших языковых моделей (БЯМ), находятся сейчас на перепутье. Идти ли по капиталоемкому пути и концентрировать исследования в небольшом числе фирм, обеспечивая первенство в создании БЯМ, работая максимально закрыто, как это делают США, или идти по пути децентрализации экономики ИИ, снижению издержек создания больших языковых моделей и максимальной открытости, как это делает КНР? В какой мере модели, созданные DeepSeek, помогут другим государствам догонять вырвавшиеся вперед КНР и США? Как пророчески отмечал несколько лет назад британский исследователь Роджер Бутл, то, что происходит сейчас, — это интенсификация целого ряда разработок, которая стала возможной благодаря усилению мощности компьютеров, расширению доступности данных и развитию способности к обучению систем искусственного интеллекта. Постепенный характер этих преобразований означает, что у правительства, компаний и отдельных лиц есть в запасе время — не время ничего не делать, а время подумать и внести по собственной инициативе различные изменения, готовясь к неизбежным переменам, которые в дальнейшем они вынуждены будут принять как данность<sup>339</sup>.

---

<sup>339</sup> Бутл Роджер. Искусственный интеллект и экономика. Работа, богатство и благополучие в эпоху мыслящих машин. — М., Альпина PRO. 2020.

Для Российской Федерации это означает реальную возможность, используя модели с открытым кодом и «дешевые» чипы, вернуться на лидирующие позиции в инновационном росте экономики.

## ЗАКЛЮЧЕНИЕ

Данное исследование послужило логическим продолжением первой в мире монографии по экономическим проблемам ИИ: «Экономика искусственного интеллекта: перспективы и риски».<sup>340</sup> В названиях обеих монографий присутствует термин «риски». Это связано с тем, что по мере совершенствования технологий искусственного интеллекта возрастают и риски их применения. Поэтому управление рисками ИИ как за счёт государственного регулирования, так и за счёт адаптации экономических агентов становится первоочередной целью современной экономики.

Выявленные в исследовании риски применения технологий искусственного интеллекта на микроуровне в краткосрочном периоде вполне могут быть скомпенсированы существующими антикризисными мерами с учётом специфики действия больших языковых моделей. Риски же применения технологий искусственного интеллекта на макроуровне требуют нового (не существующего на данный момент) аналога государственного регулирования. Перспективы применения искусственного интеллекта в долгосрочном периоде связаны с проработкой сценарного подхода к развитию взаимодействия работников и больших языковых моделей. Критерием является уровень развития человеческого капитала. Сегодня просматриваются два «полярных» сценария:

---

<sup>340</sup> Лукичев П. М., Чекмарев О. П. Экономика искусственного интеллекта: перспективы и риски». Санкт-Петербург. 2023.

- запрет (полный или временный) на использование (совершенствование) искусственного интеллекта,
- «вариант ВАЛЛ-И», когда люди признают преимущество моделей ИИ во всех видах труда и устраниются от активной работы.

Для предупреждения выявленного «технологического» уклона в современном экономическом развитии необходимы равные условия как для прогресса моделей искусственного интеллекта, так и для совершенствования человеческого капитала. Это позволит создать синергетический эффект. Не только искусственный интеллект оказывает влияние на качества «человеческого капитала», но и развитие моделей ИИ возможно лишь под воздействием высокого уровня человеческого капитала, что определяет необходимость его постоянного совершенствования. Российской экономике, чтобы стать конкурентоспособной и перейти на современный этап освоения технологий искусственного интеллекта, необходимы изменения в направлении междисциплинарного подхода в подготовке специалистов в сфере ИИ с чтением курсов по этике искусственного интеллекта, поведенческой экономике, управлению технологиями искусственного интеллекта.

Воздействие технологий искусственного интеллекта на рынок труда будет усиливаться по мере их совершенствования. Сейчас идёт процесс их количественного накопления, который в перспективе 3–4 лет качественно расширит производственные возможности работников будущего. Это расширение будет неравномерным. Неравенство в доходах и в условиях труда между работниками

низкой квалификации и работниками высокой квалификации будет усиливаться по мере продвижения технологий ИИ как в отраслевом аспекте, так и между развитыми странами, государствами с развивающимися рынками и странами с низкими доходами. Степень неравенства будет определяться как интенсивностью применения моделей искусственного интеллекта в народном хозяйстве, так и возможностями по переобучению персонала системами высшего образования и повышения квалификации.

Экономические проблемы, как показал выход в январе 2025 г. китайской модели ИИ DeepSeek, и, прежде всего, оценивание эффективности моделей искусственного интеллекта выходят на первый план по сравнению с чисто технологическими.

Большинство трендов развития технологий искусственного интеллекта, выявленных в данной монографии, требуют дальнейших комплексных исследований, подтверждающих их на базе данных современного этапа, - после ноября 2022 г.

Государственное регулирование технологий искусственного интеллекта переживает сейчас самый начальный период. Существующее противостояние между двумя мировыми экономическими центрами, - США и КНР, - не позволяет пока создать глобальную систему регулирования ИИ. Создание исключительно национальных систем государственного регулирования искусственного интеллекта (или США, или КНР) означает, что другие государства, приобретающие их модели ИИ, попадают в экосистему искусственного интеллекта США или в

экосистему искусственного интеллекта КНР. Автор надеется, что в ближайшем будущем путём взаимных компромиссов глобальное регулирование станет возможным.

## СПИСОК ЛИТЕРАТУРЫ

1. Указ Президента Российской Федерации от 15.02.2024 № 124 "О внесении изменений в Указ Президента Российской Федерации от 10 октября 2019 г. № 490 "О развитии искусственного интеллекта в Российской Федерации" и в Национальную стратегию, утвержденную этим Указом" URL: // <https://base.garant.ru/408559959/> (дата обращения 29.08.2024)
2. ГОСТ Р 70949–2023 Технологии искусственного интеллекта в образовании. Применение искусственного интеллекта в научно-исследовательской деятельности. Варианты использования. – Москва. Российский институт стандартизации. 2023. URL: [https://allgosts.ru/35/240/gost\\_r\\_70949-2023?ysclid=lryyyh4k6w701222747](https://allgosts.ru/35/240/gost_r_70949-2023?ysclid=lryyyh4k6w701222747) (дата обращения 05.03.2024)
3. Белова Ю. Ю., Н. А. Шматко Н. А. Разработка промышленных роботов в России: технологии и кадры. Наука. Технологии. Инновации. Экспресс-информация 14.02.2025. [Электронный ресурс] URL: <https://issek.hse.ru/mirror/pubs/share/1016053804.pdf> (дата обращения 15.02.2025)
4. Бутл Роджер. Искусственный интеллект и экономика. Работа, богатство и благополучие в эпоху мыслящих машин. – М., Альпина PRO. 2020.
5. Волошина И. А., Перова И. Т., Тхоржевская А. Р. Применение технологий искусственного интеллекта в деятельности организаций: масштабы, влияние на профессии, риски // Креативная экономика. – 2024. – Том 18. – No 12. – С. 3285–3310. doi: [10.18334/ce.18.12.122257](https://doi.org/10.18334/ce.18.12.122257)
6. Дремлюга Р.И. Основы национального регулирования применения искусственного интеллекта: опыт Сингапура // Азиатско-Тихоокеанский регион: экономика, политика, право. 2022. Т. 24, No1. С. 214–224. <https://doi.org/10.24866/1813-3274/2022-1/214-224>.

7. Индикаторы образования: 2025: статистический сборник / Н. В. Бондаренко, Т. А. Варламова, Л. М. Гохберг и др.; Нац. исслед. ун-т «Высшая школа экономики». – М.: ИСИЭЗ ВШЭ, 2025. – 452 с. [Электронный ресурс] URL: <https://issek.hse.ru/news/1023576877.html> (дата обращения 18.04.2025)
8. Искусственный интеллект. Серия информационно-аналитических материалов ИСИЭЗ НИУ ВШЭ. № 17 / 2025. [Электронный ресурс] URL: <https://issek.hse.ru/mirror/pubs/share/1022067859.pdf> (дата обращения 07.03.2025)
9. Канеман Д. Думай медленно... решай быстро. Издательство АСТ. Москва 2014.– 710 с.
10. Карелов С. В. «Ловушка Гудхарта» для AGI: проблема сравнительного анализа искусственного интеллекта и интеллекта человека // Ученые записки Института психологии РАН. – 2023. – Т. 3. – №. 3 (9). – С. 5-23. DOI:10.38098/proceedings\_2023\_03\_03\_02
11. Лукичев П. М. Государственное регулирование и дерегулирование аграрной сферы. – Санкт-Петербург, 2006.
12. Лукичев П. М. Поведенческая экономика. – Санкт-Петербург. 2022.
13. Лукичев П. М. Позиция России в новом международном разделении труда // Вестник Удмуртского университета. Серия Экономика и право. 2022. Т. 32. № 5. С. 817–828.
14. Лукичев, П. М. Экономика искусственного интеллекта и концепция «Принципал - Агент» / П. М. Лукичев, О. П. Чекмарев // Вопросы инновационной экономики. – 2022. – Т. 12, № 2. – С. 1069–1082. – DOI 10.18334/vines.12.2.114514
15. Лукичев П. М., Чекмарев О. П. Экономика искусственного интеллекта: перспективы и риски. Санкт-Петербург, 2023. [Электронный ресурс]. URL: <https://motivtrud.ru/PCost/research/ii20231.pdf> (дата обращения 02.02.2025)
16. Лукичев П. М., Чекмарев О.П. Вызовы экономики искусственного интеллекта традиционному рынку труда // Вопросы инновационной экономики. – 2023. – Том 13. – No 2. – С. 785–802. doi: 10.18334/vines. 13.2.118137

17. Лукичёв П.М., Чекмарев О.П. Риски применения искусственного интеллекта в краткосрочном периоде // Вопросы инновационной экономики. – 2023. – Том 13. – № 4. – doi: 10.18334/vines.13.4.119359.
18. Лукичев П. М., Чекмарев О.П. Долгосрочные риски применения искусственного интеллекта в экономике // Вопросы инновационной экономики. 2023. Т. 13. № 4. С. 2427–2442.
19. Лукичёв П. М., Кузекова С. А. Роль визуального искусства как инновационной практики в преподавании экономических дисциплин // Инновации в образовании. 2023. № 12. С. 86–94.
20. Лукичёв П. М., Чекмарев О. П., Шматко А. Д. Макроэкономика. Учебное пособие (2-е издание, расширенное и дополненное). Санкт-Петербург. 2024.
21. Лукичёв, П. М. Риски применения искусственного интеллекта в системе высшего образования / П. М. Лукичёв, О. П. Чекмарев // Вопросы инновационной экономики. – 2024. – Т. 14, № 2. – С. 463–482. – DOI 10.18334/vines.14.2.120731
22. Лукичёв, П. М. Государственное регулирование моделей искусственного интеллекта: необходимость и пределы макроэкономической политики / П. М. Лукичёв // Вопросы инновационной экономики. – 2024. – Т. 14, № 4. – С. 1151–1168. – DOI 10.18334/vines.14.4.121767
23. Лукичев, П. М. Принятие решений в современной экономике: искусственный интеллект vs поведенческая экономика / П. М. Лукичев // Вопросы инновационной экономики. – 2024. – Т. 14, № 3. – С. 649–666. – DOI 10.18334/vines.14.3.121070
24. Лукичёв, П. М. Влияние искусственного интеллекта на рынок труда в парадигме неравенства / П. М. Лукичёв // Экономика труда. – 2025. – Т. 12, № 3. – С. 267–282. – DOI 10.18334/et.12.3.122851
25. Мамедьяров З. А. Регулирование искусственного интеллекта: первые шаги. Наука. Технологии. Инновации. Экспресс-информация 12.01.2024. ИСИЭЗ НИУ ВШЭ. [] URL: <https://issek.hse.ru/mirror/pubs/share/886272540.pdf>
26. Маршалл А. Принципы экономической науки. – М., Издательская группа «Прогресс», 1993.

27. Morozov E. The Risk from A.I. Isn't Just Existential. New York Times. Section SR, Page 6. July 2, 2023.
28. Москалёв М.В., Лукичев П.М. Маркетинг. Учебное пособие для дистанционного обучения / Сер. Tacis. Укрепление реформ в сельском хозяйстве посредством образования. Москва. 2000.
29. О'Коннелл М. Искусственный интеллект и будущее человечества. – Москва: Эксмо, 2020. – 272 с.
30. Панферов В. П. Роль управления знаниями в управлении человеческим капиталом инновационно-ориентированного предприятия // Экономика и эффективность организации производства. 2019. № 29. С. 14–17. [Электронный ресурс] URL: <http://vestnik.adygnet.ru/files/2021.1/6687/70-78.pdf> (дата обращения 10.04.2025)
31. Пересмотр цифрового неравенства на рынке труда стран БРИКС / И. В. Лазанюк, И. А. Пугачева, В. А. Семчук, А. О. Неведомская // Вопросы инновационной экономики. – 2025. – Т. 15, № 1. – DOI 10.18334/vines.15.1.122679
32. Рейтинг цифровых компаний по числу «цифровых» патентов. Цифровая экономика, Экспресс-информация. ИСИЭЗ НИУ ВШЭ. 07.06.2023. [Электронный ресурс]. URL: <https://issek.hse.ru/news/838422953.html> (дата обращения 12.09.2023)
33. Российская наука в цифрах: 2023 / В. В. Власова, Л. М. Гохберг, К. А. Дитковский и др.; Нац. исслед. ун-т «Высшая школа экономики». – М.: НИУ ВШЭ, 2023. – 48 с.
34. Солдатова Н. Ф., Азарова С. П., Ребрикова Н. В. Поведенческая экономика в маркетинговом анализе: новые аспекты // Экономика, предпринимательство и право. – 2022. – Том 12. – № 5. – С. 1555–1566. doi: 10.18334/erpp.12.5.114633
35. Судова Т. Л. Человеческий капитал в наукоёмкой экономике. Санкт-Петербург, 2001.
36. Фрэнк М., Рериг П., Принг Б. Что делать, когда машины начнут делать всё: Как роботы и искусственный интеллект изменят жизнь и работу. – Москва: Эксмо. 2019. – 320 с.
37. Харари Ю. 3. 21 урок для XXI века. Издательство «Синдбад». 2019.

38. Хотеева К. Вложение в будущее. Чего рынок ждёт от ИИ в 2025 году. Коммерсант, 27;03. 2025. [Электронный ресурс] URL: <https://www.kommersant.ru/doc/7605032?ysclid=m8x6mxt9am434324280> (дата обращения 27.03.2025)
39. Человеческий капитал. [Электронный ресурс] URL: [https://www.elibrary.ru/query\\_results.asp](https://www.elibrary.ru/query_results.asp) (дата обращения 19. 03. 2025)
40. Черкасова Л. И., Шаргунова Т. В. Развитие человеческого капитала в Ленинградской области. В XII Лужские научные чтения. Современное научное знание: Л83 теория и практика: материалы международной научной конференции, 22 мая 2024 г. / отв. ред. Т. В. Седлецкая. – Санкт-Петербург: ЛГУ им. А.С. Пушкина, 2024. – с. 327-331.
41. Abiri G., Huang Y. A red flag? China’s generative AI dilemma //Harvard journal of law & technology. – 2023. – Vol. 37. – P. 15. [Электронный ресурс] URL: <https://jolt.law.harvard.edu/digest/a-red-flag-chinas-generative-ai-dilemma> (дата обращения 20.02.2025).
42. Abrardi L., Cambini C., Rondi, L. Artificial intelligence, firms and consumer behavior: A survey // Journal of Economic Surveys. 2022. Vol. 36. No.4. P. 969-991.
43. Adarlo S. Lawsuit claims woman burned, killed by surgical robot. [Электронный ресурс] URL: <HTTPS://FUTURISM.COM/NEOSCOPE/LAWSUIT-WOMAN-KILLED-SURGICAL-ROBOT> (дата обращения 26.04.2024)
44. Agrawal A., Gans J., Goldfarb A. Human judgement and AI pricing. AEA Papers and Proceedings. 2018. Vol.108, p. 58–63.
45. AI firms mustn’t govern themselves, say ex-members of OpenAI’s board. The Economist. June 1<sup>st</sup> 2024.
46. AI’s next top model // The Economist. April 20<sup>th</sup>, 2024.
47. Almog D., Gauriot R., Page L., Martin D. AI Oversight and Human Mistakes: Evidence from Centre Court // arXiv preprint arXiv:2401.16754. – 2024.
48. Amara R. Some observations on the interaction of technology and society //Futures. – 1975. – Vol. 7. – №. 6. – P. 515–517.
49. A nasty case of pilotitis. The Economist. November 9th, 2024. [Электронный ресурс] URL:

- <https://www.economist.com/business/2024/11/04/why-your-company-is-struggling-to-scale-up-generative-ai> (дата обращения 28.02.2025)
50. Anderljung, M., Barnhart, J., Leung, J., Korinek, A., O'Keefe, C., Whittlestone, J., ... & Wolf, K. (2023). Frontier AI regulation: Managing emerging risks to public safety. *arXiv preprint arXiv:2307.03718*.
  51. Anderson H., Reem N., Tadayyon. Raft of California AI Legislation Adds to Growing Patchwork of US Regulation. [Электронный ресурс] URL: <https://www.whitecase.com/insight-alert/raft-california-ai-legislation-adds-growing-patchwork-us-regulation#:~:text=Senate%20Bill%20942%3A%20California%20AI%20Transparency%20Act&text=This%20Act%20outlines%20requirements%20for,are%20permitted%20for%20public%20use> (дата обращения 20.02.2025)
  52. Andras P., L. Esterle M. Guckert T. A. Han P. R. Lewis K. Milanovic T. Payne C. Perret, J. Pitt S. T. Powers, et al. 2018. “Trusting Intelligent Machines: Deepening Trust Within Socio-Technical Systems.” // IEEE Technology and Society Magazine Vol. 37 (4). P. 76–83.
  53. Andreessen M. (2023). Why AI Will Save the World. Andreessen Horowitz, Why AI Will Save the World | Andreessen Horowitz (a16z.com)
  54. A new generation of AI tools and models is emerging. URL: <https://www.goldmansachs.com/intelligence/pages/a-new-generation-of-ai-tools-and-models-is-emerging.html>
  55. Anthis J., Lum K., Ekstrand M., Feller A., D'Amour A., Tan C. The Impossibility of Fair LLMs //arXiv preprint arXiv:2406.03198. – 2024. [Электронный ресурс]. URL: <https://arxiv.org/pdf/2406.03198> (дата обращения 02.01. 2025).
  56. Ashton H., Franklin M., Gorman R., Armstrong S. Recognising the importance of preference change: A call for a coordinated multidisciplinary research effort in the age of AI //arXiv preprint arXiv:2203.10525. – 2022.
  57. Atanasov P. D., Karger E., Tetlock P. Full Accuracy Scoring Accelerates the Discovery of Skilled Forecasters //Available at SSRN 4357367. – 2023.

58. Ayling J., Chapman A. Putting AI ethics to work: are the tools fit for purpose? //AI and Ethics. – 2022. – Т. 2. – №. 3. – С. 405-429.
59. Autor D. Polanyi's paradox and the shape of employment growth. – National Bureau of Economic Research, 2014. – №. w20485. [Электронный ресурс]. URL: [https://www.nber.org/system/files/working\\_papers/w20485/w20485.pdf](https://www.nber.org/system/files/working_papers/w20485/w20485.pdf) (дата обращения 20.12. 2024).
60. Autor D. Applying AI to rebuild middle class jobs. – National Bureau of Economic Research, 2024. – №. w32140. [Электронный ресурс]. URL: [https://www.nber.org/system/files/working\\_papers/w32140/w32140.pdf](https://www.nber.org/system/files/working_papers/w32140/w32140.pdf) (дата обращения 20.12. 2024).
61. Babina T., Fedyk A., He A., Hodson, J. Artificial intelligence, firm growth, and product innovation //Journal of Financial Economics. – 2024. – Vol. 151. – P. 103745. <https://doi.org/10.1016/j.jfineco.2023.103745>
62. Bansal G., Wu T., Zhou J., Fok R., Nushi B., Kamar E., ... Weld D. Does the whole exceed its parts? the effect of ai explanations on complementary team performance //Proceedings of the 2021 CHI conference on human factors in computing systems. – 2021. – P. 1-16.
63. Becker G. S. Human capital //New York: National Bureau of Economic Research. – 1964.
64. Becker G. S. Human capital revisited //Human capital: A theoretical and empirical analysis with special reference to education, third edition. – The University of Chicago Press, 1994. – P. 15-28.
65. Beerbaum D. Generative Artificial Intelligence (GAI) Software—Assessment on Biased Behavior //Available at SSRN 4386395. – 2023.
66. Belleflamme P., Lambert T., Schwienbacher A. Individual crowdfunding practices // Venture Capital. 2013. Vol.15. № 4. P. 313–333.
67. Benartzi S., Thaler R. Heuristics and biases in retirement savings behavior // Journal of Economic Perspectives. 2007. Vol. 21(3). pp. 81-104. DOI: 10.1257/jep.21.3.81

68. Beyond the hype. The Economist. September 23<sup>rd</sup>, 2023.
69. Birkbeck A., Rowe L. Navigating towards hyperautomation and the empowerment of human capital in family businesses: a perspective article //Journal of Family Business Management. – 2024. – Vol. 14. – №. 4. – P. 727–734. [Электронный ресурс] URL: [https://www.researchgate.net/publication/374618350\\_Navigating\\_towards\\_hyperautomation\\_and\\_the\\_empowerment\\_of\\_human\\_capital\\_in\\_family\\_businesses\\_a\\_perspective\\_article](https://www.researchgate.net/publication/374618350_Navigating_towards_hyperautomation_and_the_empowerment_of_human_capital_in_family_businesses_a_perspective_article) (дата обращения 15.04.2025)
70. Black J., Murray A. D. Regulating AI and machine learning: setting the regulatory agenda //European journal of law and technology. – 2019. – Vol. 10. – №. 3.
71. Blandin A., Bick A., Caplan T., Caplan A. Measuring Trends in Work From Home: Evidence from Six US Datasets //Available at SSRN 4953448. – 2024.
72. Bostrom N. Ethical issues in advanced artificial intelligence //Science fiction and philosophy: from time travel to superintelligence. – 2003. – С. 277-284.
73. Boy cries wolf. The Economist. June 17<sup>th</sup>, 2023.
74. Bresnahan T. F., Trajtenberg M. (1995). General purpose technologies 'Engines of growth'? Journal of econometrics. 1995. Vol.65.1, pp. 83–108.
75. Britten A. Bletchley Declaration: Nations unite on AI risk. Significance7 Volume 20. Issue 6. 2023. P. 2–3. <https://doi.org/10.1093/jrssig/qmad086>
76. Brynjolfsson E., Mitchell T. What can machine learning do? Workforce implications //Science. – 2017. – Т. 358. – №. 6370. – С. 1530-1534. [Электронный ресурс]. URL: <https://drive.google.com/file/d/1pEyizN-MhuSmqhDdEVIu8hQ1D3sjwXkH/view> дата обращения 02.01.2025).
77. Brynjolfsson, E. The Turing trap: The promise & peril of human-like artificial intelligence. Daedalus. 2022. Vol.151(2). Pp. 272-287.
78. Brynjolfsson E., Li D., Raymond L. R. Generative AI at work. – National Bureau of Economic Research, 2023. – №. w31161. [Электронный ресурс]. URL:

- [https://www.nber.org/system/files/working\\_papers/w31161/w31161.pdf](https://www.nber.org/system/files/working_papers/w31161/w31161.pdf) (дата обращения 20.12. 2024).
79. Brynjolfsson E., Ng A. Big AI can centralize decision-making and power, and that's a problem //Missing links in AI governance. – 2023. – P. 65. [Электронный ресурс]. URL: <https://sustainabilitydigitalage.org/featured/wp-content/uploads/missing-links-in-ai-governance-unesco-mila.pdf#page=71> (дата обращения 20.12. 2024).
80. Buolamwini J., Gebru T. Gender shades: Intersectional accuracy disparities in commercial gender classification //Conference on fairness, accountability and transparency. – PMLR, 2018. – P. 77-91.
81. Buçinca Z., Malaya M. B., Gajos K. Z. To trust or to think: cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making //Proceedings of the ACM on Human-Computer Interaction. – 2021. – Vol. 5. – №. CSCW1. – P. 1-21.
82. CAICT and JD Explore Academy, *White Paper on Trustworthy Artificial Intelligence*, (Beijing, 2021). [Электронный ресурс ] URL: <http://www.caict.ac.cn/kxyj/qwfb/bps/202107/P020210709319866413974.pdf>. (дата обращения 10.02.2025)
83. Caliskan K., Wade M. DARN, Part 2: An Evidence-Based Research and Prototyping Method for Strategic Design // She Ji: The Journal of Design, Economics, and Innovation. 2022. - Vol.8. - № 3. - P. 319–335. <https://doi.org/10.1016/j.sheji.2022.11.002>.
84. Camerer C. F. Artificial intelligence and behavioral economics //The economics of artificial intelligence: An agenda. – 2019. – P. 587-608.
85. Can big tech keep getting bigger in the age of AI? The Economist. 5<sup>th</sup> August 2023.
86. Candelon F., di Carlo R. C., De Bondt M., Evgeniou T. (2021). AI Regulation Is Coming: How to prepare for the inevitable. Harvard Business Review, 99(5). P. 102–111.
87. Carr N. The glass cage: Automation and us. – New York: WW Norton & Company, 2014, p. 11-12.
88. Cazzaniga M., Jaumotte M. F., Li L., Melina M. G., Panton A. J., Pizzinelli, C., ... Tavares M. M. M. Gen-AI: Artificial intelligence

- and the future of work. – International Monetary Fund, 2024.  
[Электронный ресурс] URL:  
<https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379> (дата обращения 28.02.2025)
89. Centre for Security Artificial Intelligence Released’, 21 October 2021. [Электронный ресурс] URL:  
<https://cset.georgetown.edu/publication/ethical-norms-for-new-generation-artificial-intelligence-released/>. (дата обращения 10.02.2025)
90. Chamberlain J. The risk-based approach of the European Union’s proposed artificial intelligence regulation: Some comments from a tort law perspective //European Journal of Risk Regulation. – 2023. – Vol. 14. – №. 1. – P. 1-13.
91. Chan A., Salganik R., Markelius A., Pang C., Rajkumar N., Krasheninnikov D., ... & Maharaj T. Harms from increasingly agentic algorithmic systems //Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. – 2023. – С. 651–666. [Электронный ресурс]. URL:  
<https://dl.acm.org/doi/pdf/10.1145/3593013.3594033> (дата обращения 23.12. 2024).
92. Chan A., Ezell C., Kaufmann M., Wei K., Hammond L., Bradley H., ... & Anderljung M. Visibility into AI Agents //The 2024 ACM Conference on Fairness, Accountability, and Transparency. – 2024. – С. 958-973. [Электронный ресурс]. URL:  
<https://dl.acm.org/doi/pdf/10.1145/3630106.3658948> (дата обращения 04.01. 2025).
93. Changing the program. The Economist. October 5<sup>th</sup> 2024. [Электронный ресурс]. URL:  
<https://www.economist.com/business/2024/09/29/ai-and-globalisation-are-shaking-up-software-developers-world> (дата обращения 23.12. 2024).
94. ChatGPT Inc. The Economist. July 1st 2023.
95. Chen L., Zaharia M., Zou J. How is ChatGPT's behavior changing over time? //arXiv preprint arXiv:2307.09009. – 2023.
96. Chen V., Liao Q. V., Wortman Vaughan J., Bansal G. Understanding the role of human intuition on reliance in human-AI decision-making with explanations // Proceedings of the ACM on Human-computer Interaction. – 2023. – Vol. 7. - № CSCW2. – P. 1-

32. [Электронный ресурс]. URL: <https://dl.acm.org/doi/pdf/10.1145/3610219> (дата обращения 06.01. 2025).
97. Chen Z., Chan J. Large language model in creative work: The role of collaboration modality and user expertise //Management Science. – 2024. – Vol. 70. – №. 12. – P. 9101–9117. [Электронный ресурс] URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4575598](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4575598) (дата обращения 21.02.2025)
98. Cheng J., Zeng J. Shaping AI's future? China in global AI governance //Journal of Contemporary China. – 2023. – Vol. 32. – №. 143. – P. 794–810, <https://doi.org/10.1080/10670564.2022.2107391>
99. Choi J. H., Monahan A. B., Schwarcz D. Lawyering in the age of artificial intelligence //Minn. L. Rev. – 2024. – Vol. 109. – P. 147. [Электронный ресурс] URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4626276](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4626276) (дата обращения 21.02.2025)
100. Clancy M., Correa D., Dworkin J., Niehaus P., Watney C., Williams H. Want to speed up scientific progress? First understand how science policy works. Nature. 2023. Vol. 620(7975). P. 724–726.
101. Code of misconduct. The Economist. April 26th, 2025. Available at: <https://www.economist.com/science-and-technology/2025/04/23/ai-models-can-learn-to-conceal-information-from-their-users> (accessed 26 April 2025).
102. Cole-Lewis H. Ezeanochie N., Turgiss J. Understanding health behavior technology engagement: pathway to measuring digital behavior change interventions //JMIR formative research. – 2019. – Vol. 3. – №. 4. – P. e14052.
103. Danielsson J., Macrae R., Uthemann A. Artificial intelligence and systemic risk //Journal of Banking & Finance. – 2022. – Vol. 140. – P. 106290.
104. Dell'Acqua F., McFowland III E., Mollick E. R., Lifshitz-Assaf H., Kellogg K., Rajendran S., ... & Lakhani K. R. Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality //Harvard Business School Technology & Operations Mgt. Unit

- Working Paper. – 2023. – №. 24-013. [Электронный ресурс] URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4573321](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321) (дата обращения 12.11.2024)
105. Deming D. J., Ong C., Summers L. H. Technological disruption in the labor market. – National Bureau of Economic Research, 2025. – №. w33323. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w33323/w33323.pdf](https://www.nber.org/system/files/working_papers/w33323/w33323.pdf) (дата обращения 12.04.2025)
106. Di Battista A., Grayling S., Hasselaar E., Leopold T., Li R., Rayner M., Zahidi S. Future of jobs report 2023 //World Economic Forum, Geneva, Switzerland. [Электронный ресурс] URL: <https://www.weforum.org/reports/the-future-of-jobs-report-2023-2023>. (дата обращения 15.02.2025)
107. Discussion Paper on Generative AI: Implications for Trust and Governance. - Singapore. Infocomm Media Development Authority of Singapore (IMDA), Aicadium and AI Verify Foundation. 2023. [Электронный ресурс] URL: [https://aiverifyfoundation.sg/downloads/Discussion\\_Paper.pdf](https://aiverifyfoundation.sg/downloads/Discussion_Paper.pdf) (дата обращения 02.02.2025).
108. Domenighetti G., Casabianca A., Gutzwiller F., Martinoli S. Revisiting the most informed consumer of surgical services: the physician-patient //International journal of technology assessment in health care. – 1993. – Vol. 9. – №. 4. – P. 505-513.
109. Doyuran E. B. Nudge goes to Silicon Valley: designing for the disengaged and the irrational //Journal of Cultural Economy. – 2024. – Vol. 17. – №. 1. – P. 1-19.
110. Drydak N. Artificial intelligence capital and employment prospects //Oxford Economic Papers. – 2024. – Vol. 76. – №. 4. – P. 901-919. <https://doi.org/10.1093/oep/gpae005>
111. Edgerton D. The shock of the old: Technology and global history since 1900. – Oxford University Press, 2007.
112. Einfeldt A. L., Schubert G., Zhang M. B. Generative ai and firm values. – National Bureau of Economic Research, 2023. – №. w31222.
113. EU: Council gives final approval to AI Act. [Электронный ресурс]. URL: [https://www.dataguidance.com/news/eu-council-gives-final-approval-ai-act#:~:text=On%20May%2021%2C%202024%2C%20the,Intelligence%20\(the%20AI%20Act\)](https://www.dataguidance.com/news/eu-council-gives-final-approval-ai-act#:~:text=On%20May%2021%2C%202024%2C%20the,Intelligence%20(the%20AI%20Act)). (дата обращения 20.02.2025)

114. Farrell H., Newman A. The New Economic Security State // Foreign Affairs. November-December 2023.
115. Farrell H., Shalizi C. "Artificial intelligence is a familiar-looking monster, say Henry Farrell and Cosma Shalizi". The Economist. June 24th 2023.
116. Felten E., Raj M., Seamans R. Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses // Strategic Management Journal. – 2021. – Vol. 42. – №. 12. – P. 2195–2217. [Электронный ресурс] URL: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.3286> (дата обращения 10.11.2024)
117. Felten E., Raj M., Seamans R. How will language modelers like chatgpt affect occupations and industries? // arXiv preprint arXiv:2303.01157. – 2023. [Электронный ресурс] URL: <https://arxiv.org/pdf/2303.01157> (дата обращения 12.11.2024)
118. Ferri G., Gloerich I. Risk and harm: Unpacking ideologies in the AI discourse // Proceedings of the 5th International Conference on Conversational User Interfaces. – 2023. – P. 1-6.
119. First thing we do, let's bot all the lawyers. The Economist. June 10th 2023.
120. For a fistful of dollars. The Economist. February 15th, 2025. [Электронный ресурс] URL: <https://www.economist.com/science-and-technology/2025/02/12/forget-deepseek-large-language-models-are-getting-cheaper-still> (дата обращения 28.02.2025)
121. Fusaro R., Sperling-Magro J. Much anew about 'nudging' // McKinsey Quarterly. – 2021.
122. Future of Jobs Report 2025. World Economic Forum. January 2025/ Geneva, WEF. 2025. – 290 p. [Электронный ресурс] URL: <https://www.weforum.org/reports/the-future-of-jobs-report-2025/> (дата обращения 14.02.2025)
123. Future of Life Institute. (2023, March 22). Pause giant AI experiments: An open letter. [https://futureoflife.org/wp-content/uploads/2023/05/FLI\\_Pause-Giant-AI-Experiments\\_An-Open-Letter.pdf](https://futureoflife.org/wp-content/uploads/2023/05/FLI_Pause-Giant-AI-Experiments_An-Open-Letter.pdf)
124. Garfinkel B., Anderljung M., Heim L., Trager R., Clifford B., Seger E. (2024) Goals for the Second AI Safety Summit. Centre for

- the Governance of AI. [Электронный ресурс]. URL: <https://www.governance.ai/post/goals-for-the-second-ai-safety-summit> (дата обращения 20.02.2025)
125. Gaske M. R. Regulation Priorities for Artificial Intelligence Foundation Models //Vand. J. Ent. & Tech. L. – 2023. – Vol 26. – P 1. [Электронный ресурс]. URL: <https://scholarship.law.vanderbilt.edu/jetlaw/vol26/iss1/2/> (дата обращения 20.02.2025)
126. Gaynor M. Diagnosing the Problem: Exploring the Effects of Consolidation and Anticompetitive Conduct in Health Care Markets //Washington, DC. – 2019.
127. Generative AI could raise global GDP by 7%. [Электронный ресурс] URL: <https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html> (дата обращения 23.10.2024).
128. Get used to it. The Economist. March 15th 2025. [] URL: <https://www.economist.com/leaders/2025/03/13/with-manus-ai-experimentation-has-burst-into-the-open> (дата обращения 16.03.2025).
129. Giest S. N., Klievink B. More than a digital system: how AI is changing the role of bureaucrats in different organizational contexts //Public Management Review. – 2024. – Vol. 26. – №. 2. – P. 379–398.
130. Goodhart C. Monetary relationships: a view from Threadneedle Street. – University of Warwick, 1976.
131. Gopnik A. What AI Still Doesn't Know How to Do. The Wall Street Journal. July 15, 2022. URL: <https://www.wsj.com/articles/what-ai-still-doesnt-know-how-to-do-1165789131627>
132. Green B., Chen Y. The principles and limits of algorithm-in-the-loop decision making //Proceedings of the ACM on Human-Computer Interaction. – 2019. – Vol. 3. – №. CSCW. – P. 1-24.
133. Gu A., Dao T. Mamba: Linear-time sequence modeling with selective state spaces //arXiv preprint arXiv:2312.00752. – 2023
134. Gulati S., Sousa S., Lamas D. Design, development and evaluation of a human-computer trust scale //Behaviour &

- Information Technology. – 2019. – Vol. 38. – №. 10. – P. 1004-1015.
135. Guzik T. J., Sitek A. Global accord on the integration of artificial intelligence in medical science publishing: implications of the Bletchley Declaration //Cardiovascular Research. – 2023. – Vol. 119. – №. 17. – P. 2681–2682. <https://doi.org/10.1093/cvr/cvad170>
136. Hacker P., Engel A., Mauer M. Regulating ChatGPT and other large generative AI models //Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. – 2023. – P. 1112-1123.
137. Hadjimichael D., Pyrko I., Tsoukas H. Beyond tacit knowledge: How Michael Polanyi’s theory of knowledge illuminates theory development in organizational research //Academy of management review. – 2024. – Vol. 49. – №. 4. – P. 919-934. [Электронный ресурс]. URL: [https://purehost.bath.ac.uk/ws/portalfiles/portal/296758710/2023\\_Hadjimichael\\_et\\_al\\_accepted\\_author.pdf](https://purehost.bath.ac.uk/ws/portalfiles/portal/296758710/2023_Hadjimichael_et_al_accepted_author.pdf) (дата обращения 15.12.2024).
138. Hamilton A., Wiliam D., Hattie J. The Future of AI in Education: 13 things we can do to minimize the damage. – 2023. [Электронный ресурс] URL: [https://www.researchgate.net/publication/373108877\\_The\\_Future\\_of\\_AI\\_in\\_Education\\_13\\_Things\\_We\\_Can\\_Do\\_to\\_Minimize\\_the\\_Damage](https://www.researchgate.net/publication/373108877_The_Future_of_AI_in_Education_13_Things_We_Can_Do_to_Minimize_the_Damage) (дата обращения 16.02.2024).
139. Hampole M., Papanikolaou D., Schmidt L. D., Seegmiller B. Artificial intelligence and the labor market. – National Bureau of Economic Research, 2025. – №. w33509. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w33509/w33509.pdf](https://www.nber.org/system/files/working_papers/w33509/w33509.pdf) (дата обращения 11.04.2025).
140. Harding J., D’Alessandro W., Laskowski N. G., Long R. AI language models cannot replace human research participants //Ai & Society. – 2023. – P. 1-3.
141. Helfrich G. The harms of terminology: why we should reject so-called “frontier AI” //AI and Ethics. – 2024. – P. 1-7. <https://doi.org/10.1007/s43681-024-00438-1>

142. Hendrycks D., Carlini N., Schulman J., Jacob Steinhardt J. “Unsolved problems in ml safety”. In: *arXiv preprint arXiv:2109.13916* (2021).
143. Hendrycks D., Mazeika M. X-risk analysis for ai research //arXiv preprint arXiv:2206.05862. – 2022.
144. Hendrycks D. Dan Hendrycks warns America against launching a Manhattan Project for AI. *The Economist*. March 28th, 2025.
145. Hendrycks D., Schmidt E., Wang A. Superintelligence Strategy: Expert Version //arXiv preprint arXiv:2503.05628. – 2025.
146. Hengstler M., Enkel E., Duelli S. Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices //Technological Forecasting and Social Change. – 2016. – Vol. 105. – P. 105-120.
147. Holmquist L. E. Intelligence on Tap: Artificial Intelligence As a New Design Material // Interactions. 2017. Vol.24. № 4. P. 28–33.
148. Hötte K., Tarannum, T., Verendel V., Bennett, L. (2023). AI Technological Trajectories in Patent Data.
149. How bad is being big? *The Economist*. July 15<sup>th</sup>, 2023.
150. How generative models could go wrong. *The Economist*. April 22<sup>nd</sup> 2023.
151. Huang Y. The Labor Market Impact of Artificial Intelligence: Evidence from US Regions //Available at SSRN 5137231. – 2025. [Available at SSRN 5137231, 2025 papers.ssrn.com](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5137231)
152. Irrational Labs (2021). How can we reduce customer churn? In [irrationalabs.com](https://irrationalabs.com). Дата обращения 30.11.2022 <https://irrationalabs.com/casestudies/google-how-can-we-reduce-customer-churn/>
153. Jacobides M. G., Brusoni S., Candelon F. The evolutionary dynamics of the artificial intelligence ecosystem //Strategy Science. – 2021. – Т. 6. – №. 4. – С. 412-435.
154. Kaminski M. E. Binary governance: Lessons from the GDPR'S approach to algorithmic accountability // Southern California Law Review. 2019. Vol. 92 (6). P. 1529–1616.
155. Kang C., Satariano A. A.I. Regulation Can Be Puzzle To Lawmakers. *New York Times*. March 4, 2023, Section A, Page 1 of

- the New York edition. [Электронный ресурс] URL: <https://www.nytimes.com/2023/03/03/technology/artificial-intelligence-regulation-congress.html> (дата обращения 16.03.2023)
156. Kaplan J. Humans Need Not Apply: A Guide to Wealth & Work in the Age of Artificial Intelligence. – Yale University Press, 2015., P. 41-43, 145.
157. Kim A., Kim D. S., Muhn M., Nikolaev V. V., So E. C. AI, Investment Decisions, and Inequality //Chicago Booth Accounting Research Center Research Paper, Fama-Miller Working Paper, MIT Sloan Research Paper. – 2024.
158. Kim A., Muhn M., Nikolaev V. Financial statement analysis with large language models //arXiv preprint arXiv:2407.17866. – 2024. [Электронный ресурс] URL: <https://arxiv.org/html/2407.17866v1> (дата обращения 12.11.2024)
159. Korinek A., Stiglitz J. E. Artificial intelligence and its implications for income distribution and unemployment //The economics of artificial intelligence: An agenda. – University of Chicago Press, 2018. – P. 349-390. [Электронный ресурс]. URL: <https://www.nber.org/system/files/chapters/c14018/c14018.pdf> (дата обращения 15.12.2024).
160. Korinek A., Stiglitz J. E. Artificial intelligence, globalization, and strategies for economic development. – National Bureau of Economic Research, 2021. – №. w28453. – 41 p.
161. Korteling J. H., van de Boer-Visschedijk G. C., Blankendaal, R. A., Boonekamp R. C., Eikelboom A. R. Human-versus artificial intelligence // Frontiers in artificial intelligence. – 2021. – Vol.4. p. 622364.
162. Laupichler M. C., Aster A., Schirch J., Raupach T. Artificial intelligence literacy in higher and adult education: A scoping literature review //Computers and Education: Artificial Intelligence. – 2022. – Vol. 3. – P. 100101. [Электронный ресурс] URL: [https://www.researchgate.net/publication/363869317\\_Artificial\\_intelligence\\_literacy\\_in\\_higher\\_and\\_adult\\_education\\_A\\_scoping\\_literature\\_review](https://www.researchgate.net/publication/363869317_Artificial_intelligence_literacy_in_higher_and_adult_education_A_scoping_literature_review) (дата обращения 13.01.2025)

163. Leslie D., Perini A. M. Future Shock: Generative AI and the international AI policy and governance crisis. – 2024.
164. Leoni S. A historical review of the role of education: From human capital to human capabilities //Review of Political Economy. – 2025. – Vol. 37. – №. 1. – P. 227-244. [Электронный ресурс] URL: <https://www.tandfonline.com/doi/full/10.1080/09538259.2023.2245233> (дата обращения 13.03.2025)
165. Leveson N., Dulac N., Marais K., Carroll, J. Moving beyond normal accidents and high reliability organizations: A systems approach to safety in complex systems // Organization studies. 2009. Vol. 30(2-3). pp. 227-249.
166. Lichfield G. When Will AI Be Smarter Than Humans? Don't Ask. [Электронный источник] URL: <https://www.bloomberg.com/opinion/features/2025-04-13/when-will-ai-be-smarter-than-humans-don-t-ask> (дата обращения 30.04.2025)
167. Lipsey R. G., Kenneth I Carlaw K. I., Bekar C. T. Economic transformations: general purpose technologies and long-term economic growth. 2005. Oxford University Press, Oxford.
168. Liu Y., Zhang Z. Unlocking operational efficiency: How AI human capital investment enhances data processing efficiency? //Economics Letters. – 2025. – Vol. 247. – P. 112147. <https://doi.org/10.1016/j.econlet.2024.112147>
169. Lohr S. A.I. Is Prompting an Evolution, Not Extinction, for Coders. New York Times. February 20, 2025. [Электронный ресурс] URL: <https://www.nytimes.com/2025/02/20/business/ai-coding-software-engineers.html> (дата обращения 03.03.2025)
170. Long D., Magerko B. What is AI literacy? Competencies and design considerations //Proceedings of the 2020 CHI conference on human factors in computing systems. – 2020. – P. 1-16. [Электронный ресурс] URL: <https://dl.acm.org/doi/abs/10.1145/3313831.3376727> (дата обращения 03.01.2024)
171. Lu C., Lu C., Lange R. T., Foerster J., Clune J., Ha, D. The ai scientist: Towards fully automated open-ended scientific discovery //arXiv preprint arXiv:2408.06292. – 2024. [Электронный

- ресурс]. URL: <https://arxiv.org/abs/2408.06292> (дата обращения 06.01. 2025).
172. Maddox T. M., Rumsfeld J. S., Payne P. R. O. Questions for artificial intelligence in health care //Jama. – 2019. – Vol. 321. – №. 1. – P. 31-32.
173. McClain C., Kennedy B., Gottfried J., Anderson M., Pasquini G. How the U.S. Public and AI Experts View Artificial Intelligence. [Электронный ресурс] URL: [https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence/?utm\\_source=Pew+Research+Center&utm\\_campaign=82cfde2055-Internet\\_Science\\_2025\\_04\\_29&utm\\_medium=email&utm\\_term=0\\_-82cfde2055-399474221](https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence/?utm_source=Pew+Research+Center&utm_campaign=82cfde2055-Internet_Science_2025_04_29&utm_medium=email&utm_term=0_-82cfde2055-399474221). (дата обращения 30.04.2025)
174. Messeri L., Crockett M.J. Artificial intelligence and illusions of understanding in scientific research // Nature. 2024. Vol.627. P. 49–58. <https://doi.org/10.1038/s41586-024-07146-0>
175. Metz C. A.I. Does Not Have Thoughts, No Matter What You Think. New York Times. August 7, 2022. Section BU, p. 6.
176. Metz C., Weise K. ‘A.I. Agents’ Can Play Minecraft. They’re Just Getting Warmed Up. New York Times. October. 18, 2023, Section B, Page 1.
177. Michie S., Richardson M., Johnston M., Abraham C., Francis J., Hardeman W., ... Wood C. E. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions //Annals of behavioral medicine. – 2013. – Vol. 46. – №. 1. – P. 81-95.
178. Mincer J. Investment in human capital and personal income distribution //Journal of political economy. – 1958. – Vol. 66. – №. 4. – P. 281-302.
179. Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., Floridi, L.: The ethics of algorithms: mapping the debate. Big Data Soc. 3(2), 2053951716679679 (2016). URL: <https://doi.org/10.1177/2053951716679679> (дата обращения 27.09.2023)

180. Model AI governance framework. Second edition. [Электронный ресурс]. URL: <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf>. (дата обращения 20.02.2025)
- 
181. Model AI Governance Framework for Generative AI. Fostering a Trusted Ecosystem. - Singapore. - AI Verify Foundation. - 2024. – 36 pages Published 30 May 2024. [Электронный ресурс]. URL: <https://aiverifyfoundation.sg/wp-content/uploads/2024/05/Model-AI-Governance-Framework-for-Generative-AI-May-2024-1-1.pdf>). (дата обращения 20.02.2025)
182. Moll B., Rachel L., Restrepo P. Uneven growth: automation's impact on income and wealth inequality //Econometrica. – 2022. – Vol. 90. – №. 6. – P. 2645–2683. [Электронный ресурс] URL: [https://www.nber.org/system/files/working\\_papers/w28440/w28440.pdf](https://www.nber.org/system/files/working_papers/w28440/w28440.pdf) (дата обращения 12.11.2024)
183. Mollick E., Mollick E. Co-Intelligence. – Random House UK, 2024.
184. Nioi A., Wendelboe-Nelson C., Cowan S., Cherrie M., Rashid S., Cowie H., Davis A., Ritchie P., Lansdown T., Cherrie J. W. Nudging construction workers towards better sun-safety behaviour: summary of the evidence for practitioners // Policy and Practice in Health and Safety. 2020. Vol. 8(1), pp. 25-33. DOI: 10.1080/14773996.2019.1708614
185. No news is bad news. The Economist. May 18<sup>th</sup>, 2024. [Электронный ресурс].URL: <https://www.economist.com/leaders/2024/05/16/canadas-law-to-help-news-outlets-is-harming-them-instead> (дата обращения 20.02.2025)
186. Non-proliferation treaties. The Economist. May 27<sup>th</sup>, 2023.
187. Of evils and evals. The Economist. October 28<sup>th</sup>, 2023.
188. Ord T. The precipice: Existential risk and the future of humanity. – Hachette Books, 2020. – 468 p.

189. Otis N., Clarke R. P., Delecourt S., Holtz D., Koning R. The uneven impact of generative AI on entrepreneurial performance // Available at SSRN 4671369. – 2023.
190. Oxford Analytica. UK AI Summit will promote some global cooperation // Emerald Expert Briefings. – 2023. – №. oхан-db. [Электронный ресурс]. URL: <https://www.emerald.com/insight/content/doi/10.1108/oxan-db284110/full/html>. (дата обращения 02.02.2025).
191. Park S. Bridging the Global Divide in AI Regulation: A Proposal for a Contextual, Coherent, and Commensurable Framework // arXiv preprint arXiv:2303.11196. – 2023.
192. Petit N. Antitrust and artificial intelligence: a research agenda // Journal of European Competition Law & Practice. 2017. Vol.8. No.6. P.361-362.
193. Pizzinelli C., Panton A. J., Tavares M. M., Cazzaniga M., Li L. Labor market exposure to AI: Cross-country differences and distributional implications. – International Monetary Fund, 2023. [Электронный ресурс] URL: <https://www.imf.org/en/Publications/WP/Issues/2023/10/04/Labor-Market-Exposure-to-AI-Cross-country-Differences-and-Distributional-Implications-539656> (дата обращения 12.01.2025)
194. Polanyi M. The Study of Man. Chicago: University of Chicago Press, 1959.
195. Polanyi M. The Tacit Dimension. Chicago and London: The University of Chicago Press, 1966.
196. Prassl J. Humans as a service: The promise and perils of work in the gig economy. – Oxford University Press, 2018. pp. 138–139.
197. Ptaschunder J. M. Nudgital: Critique of behavioral political economy // Proceedings of the 9th international RAIS conference on social sciences and humanities. – Scientia Moralitas Research Institute, 2018. – С. 87-117. DOI: 10.5281/zenodo.1244836
198. Rage against the machines. The Economist. April 27th, 2024.
199. Raghu M., Blumer K., Corrado G., Kleinberg J., Obermeyer Z., Mullainathan S. The algorithmic automation problem: Prediction, triage, and human effort // arXiv preprint arXiv:1903.12220. – 2019.

200. Raji I. D., Kumar I. E., Horowitz A., Selbst A. The fallacy of AI functionality //Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency. – 2022. – P. 959-972.
201. Rambachan A. Identifying Prediction Mistakes in Observational Data. – 2022.
202. Romer, P. (forthcoming). (2016). The trouble with macroeconomics. The American Economist, Retrieved from <http://ccl.yale.edu/sites/default/files/files/The%20Trouble%20with%20Macroeconomics.pdf> [Дата обращения January 24, 2017]
203. Ryan-Mosley, T. (2023, June 12.) It's time to talk about the real AI risks. MIT Technology Review. [Электронный ресурс]. URL: <https://www.technologyreview.com/2023/06/12/1074449/real-ai-risks/> (дата обращения 13.02.2025)
204. Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. A Presidential Document by the Executive Office of the President on 11/01/2023 [Электронный ресурс]. URL: <https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence> (дата обращения 19.02.2025)
205. Schuett J. Defining the scope of AI regulations //Law, Innovation and Technology. – 2023. – Vol. 15. – №. 1. – P. 60–82.7<http://dx.doi.org/10.2139/ssrn.3453632>
206. Schuling around. The Economist. April 20<sup>th</sup> 2024. [Электронный ресурс] URL: <https://www.economist.com/europe/2024/04/16/germany-is-flunking-the-education-test> (дата обращения 20.04.2024)
207. Schultz T. W. Investment in human capital //The American economic review. – 1961. – Vol. 51. – №. 1. – P. 1–17.
208. Schultz T. W. Human capital: Policy issues and research opportunities //Economic Research: Retrospect and Prospect, Volume 6, Human Resources. – NBER, 1972. – С. 1-84.
209. Scott J. C. Seeing like a state: How certain schemes to improve the human condition have failed. – Yale University Press, 2020.

210. Sen A. K. Prediction and economic theory //Proceedings of the Royal Society of London. Series A. Mathematical and Physical sciences. – 1986. – Vol. 407. – №. 1832. – P. 3-23.
211. Sen A. K. Human capital and human capability //World development. – 1997. – Vol. 25. – №. 12. – P. 1959–1961.
212. Sheehan M. China’s AI regulations and how they get made //Horizons: Journal of International Relations and Sustainable Development. – 2023. – №. 24. – P. 108–125. [Электронный ресурс]. URL: <https://www.cirsd.org/files/000/000/010/82/21e461a985f43655b1731b3c1b50cdccb631afaf.pdf> (дата обращения 16.02.2025)
213. Shengchengshi Rengong Zhineng Fuwu Guanli Zanxing Banfa (生成式人工智能服务管理暂行办法) [Interim Measures for Administration of Generative Artificial Intelligence Services] (promulgated by the Cyberspace Admin. of China, Jul. 10, 2023, effective Aug. 15, 2023) St. Council Gaz., Aug. 30, 2023, at 40, [https://www.gov.cn/gongbao/2023/issue\\_10666/](https://www.gov.cn/gongbao/2023/issue_10666/) (China) [hereinafter Interim Measures]. [Временные меры по администрированию служб генеративного искусственного интеллекта] (обнародованы Администрацией киберпространства Китая 10 июля 2023 г., вступают в силу 15 августа 2023 г.) St. Council Gaz., 30 августа 2023 г., стр. 40. [Электронный ресурс] URL: [https://www.gov.cn/gongbao/2023/issue\\_10666/202308/content\\_6900864.html](https://www.gov.cn/gongbao/2023/issue_10666/202308/content_6900864.html) (дата обращения 20.02.2025).
214. Silicon Valley Bank, “2020 Global Startup Outlook: Key Insights from the Silicon Valley Bank Startup Outlook Survey,” Silicon Valley Bank (2020): 3, [Электронный ресурс] URL: [https://www.svb.com/globalassets/library/uploadedfiles/content/trends\\_and\\_insights/reports/startup\\_outlook\\_report/suo\\_global\\_report\\_2020-final.pdf](https://www.svb.com/globalassets/library/uploadedfiles/content/trends_and_insights/reports/startup_outlook_report/suo_global_report_2020-final.pdf) (дата обращения 16.02;2023)
215. Simon H. A behavioral model of rational choice //Models of man, social and rational: Mathematical essays on rational human behavior in a social setting. – 1957. – Vol. 6. – №. 1. – P. 241-260.
216. Smuha N. A. From a ‘race to AI’ to a ‘race to AI regulation’: regulatory competition for artificial intelligence //Law, Innovation and Technology. – 2021. – Vol. 13. – №. 1. – P. 57–84. <http://dx.doi.org/10.2139/ssrn.3501410>

217. Sollosy M., McInerney M. Artificial intelligence and business education: What should be taught //The International Journal of Management Education. – 2022. – Vol. 20. – №. 3. – P. 100720. [Электронный ресурс] URL: <https://www.sciencedirect.com/science/article/abs/pii/S1472811722001227> (дата обращения 02.02.2025)
218. Solove D.J. Privacy and power: computer databases and metaphors for information privacy// Stanford Law Review. 2001. P. 1398.
219. Stone P., Brooks R., Brynjolfsson E., Calo R., Etzioni O., Hager G., ... & Teller A. Artificial intelligence and life in 2030: the one-hundred-year study on artificial intelligence //arXiv preprint arXiv:2211.06318. – 2022.
220. Straub V. J., Morgan D., Bright J., Margetts H. Artificial intelligence in government: Concepts, standards, and a unified framework //Government Information Quarterly. – 2023. – Vol. 40. – №. 4. – P. 101881.
221. Susser D., Roessler B., Nissenbaum H. Online manipulation: Hidden influences in a digital world //Geo. L. Tech. Rev. – 2019. – Vol. 4. – P. 1.
222. Susskind D. Re-thinking the capabilities of machines in economics. – 2017. [Электронный ресурс]. URL: [https://ora.ox.ac.uk/objects/uuid:554f4c9c-6c78-44b5-9e93-788c731bec5f/download\\_file?file\\_format=&safe\\_filename=825-susskind-april18.pdf&type\\_of\\_work=Working+paper](https://ora.ox.ac.uk/objects/uuid:554f4c9c-6c78-44b5-9e93-788c731bec5f/download_file?file_format=&safe_filename=825-susskind-april18.pdf&type_of_work=Working+paper) (дата обращения 06.01. 2025).
223. Taleb N. N. Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications //arXiv preprint arXiv:2001.10488. – 2020.
224. Terzis P. Law and the political economy of AI production //International Journal of Law and Information Technology. – 2023. – Vol. 31. – №. 4. – P. 302–330. <https://doi.org/10.1093/ijlit/eaac001>
225. Thaler R., Sunstein C. Nudge: The final edition. London: Penguin. – 2021. – 366 p.

226. The age of gunboat digital diplomacy. The Economist. January 25th, 2025. [Электронный ресурс].URL: <https://www.economist.com/europe/2025/01/23/europe-faces-a-new-age-of-gunboat-digital-diplomacy> (дата обращения 20.02.2025)
227. The first mile // The Economist. September 24<sup>th</sup>, 2022.
228. The LLM summer sale. The Economist. June 15<sup>th</sup>, 2024.
229. The Model AI Governance Framework first edition. [Электронный ресурс].URL: <https://ai.bsa.org/wp-content/uploads/2019/09/Model-AI-Framework-First-Edition.pdf>. (дата обращения 20.02.2025)
230. Think, then act. The Economist. October 28th, 2023.
231. Toner-Rodgers A. Artificial intelligence, scientific discovery, and product innovation //arXiv preprint arXiv:2412.17866. – 2024. [Электронный ресурс] URL: <https://arxiv.org/pdf/2412.17866> (дата обращения 02.03.2025)
232. Trager R., Harack B., Reuel A., Carnegie A., Heim L., Ho L., Kreps S., Lall R., Larter O., Ó hÉigeartaigh S., Staffell S., Villalobos J. J., Villalobos J. J. (2023). International governance of civilian AI: A jurisdictional certification approach. ArXiv. <https://doi.org/10.48550/arXiv.2308.15514>
- 233.** Trajtenberg M. Artificial intelligence as the next GPT: A political-economy perspective. / The economics of artificial intelligence: An agenda., 2018. – 175-186 p. [Электронный ресурс]. URL: <https://www.nber.org/system/files/chapters/c14025/c14025.pdf> (дата обращения 06.01. 2025).
234. Tversky A., Kahneman D. Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty //Science. – 1974. – Vol. 185. – №. 4157. – p. 1124–1131.
235. Tyson L. D., Zysman J. Automation, AI & work //Daedalus. – 2022. – Vol. 151. – №. 2. – P. 256–271. [Электронный ресурс] URL: [https://amacad.org/sites/default/files/daedalus/downloads/Daedalus\\_Sp22\\_AI%20%26%20Society\\_3.pdf#page=258](https://amacad.org/sites/default/files/daedalus/downloads/Daedalus_Sp22_AI%20%26%20Society_3.pdf#page=258) (дата обращения 02.02.2025)

236. Uncomfortably close. The Economist. January 25th, 2025. [Электронный ресурс]. URL: <https://www.economist.com/briefing/2025/01/23/chinas-ai-industry-has-almost-caught-up-with-americas> (дата обращения 20.02.2025)
237. Vallor S., Luger E. A shrinking path to AI safety //Edinburgh Futures Institute. – 2023. [Электронный ресурс].URL: <https://efi.ed.ac.uk/a-shrinking-path-to-safety-how-a-narrowly-technical-approach-to-align-ai-with-the-public-good-could-fail/> (дата обращения 20.02.2025)
238. Vardi M. Y. Is information technology destroying the middle class? //Communications of the ACM. – 2015. – Vol. 58. – №. 2. – P. 5-5.
239. Vasconcelos H., Jörke M., Grunde-McLaughlin M., Gerstenberg T., Bernstein M. S., Krishna R. Explanations can reduce overreliance on ai systems during decision-making // Proceedings of the ACM on Human-Computer Interaction. 2023. Vol. 7. #(CSCW1). P. 1-38.
240. Vermann E. K. Wait is saving good or bad? the Paradox of thrift //Page One Economics®. – 2012.
241. Webster G., Creemers R., Kania E., Triolo P. Full Translation: China’s “New Generation Artificial Intelligence Development Plan” (2017) – 2017. [Электронный ресурс]. URL: <https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/> (дата обращения 02.02.2025)
242. Weijers R. J., de Koning B. B. Nudging to increase hand hygiene during the COVID-19 pandemic: A field experiment // Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement. 2021. Vol.53(3). p. 353. DOI: 10.1037/cbs0000245
243. Weitzman M. Iterative multilevel planning with production targets //Econometrica: Journal of the Econometric Society. – 1970. – p. 50-65.
244. Wendel S. Designing for behavior change: Applying psychology and behavioral economics. – O'Reilly Media, 2020.
245. Where are all the robots? The Economist. March 11th 2023.

246. Wiener N. Some Moral and Technical Consequences of Automation: As machines learn they may develop unforeseen strategies at rates that baffle their programmers //Science. – 1960. – T. 131. – №. 3410. – C. 1355-1358.
247. Wooldridge M., Jennings N. R. Intelligent agents: Theory and practice. // The knowledge engineering review. 1995. Vol. 10(2). P.115–152.
248. Zarifhonarvar A. Economics of ChatGPT: A labor market view on the occupational impact of artificial intelligence //Journal of Electronic Business & Digital Economics. – 2024. – Vol. 3. – №. 2. – P. 100–116. <https://doi.org/10.1108/JEBDE-10-2023-0021>
249. Zhang Y., Liao Q. V., Bellamy R. K. E. Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making //Proceedings of the 2020 conference on fairness, accountability, and transparency. – 2020. – P. 295-305.
250. Zhang Shidong, “China Tops Global IPO Ranking in First Half as Issuers Leverage Capital Market Reforms, AI Frenzy,” South China Morning Post, July 4, 2023, <https://www.scmp.com/business/china-business/article/3226489/china-tops-global-ipo-ranking-first-half-issuers-leverage-capital-market-reforms-ai-frenzy> ; George Steer, Nicholas Megaw, and Hudson Lockett, “China Dominates Global IPO Market as Wall Street Fails to Rebound,” Financial Times, May 2, 2023, <https://www.ft.com/content/e5a43d1d-b62c-4428-b16f-c1fe6b4fb9d9>.
251. Zouya L., Yunus N. M. A preliminary study on artificial intelligence and labour productivity in China //International Business Education Journal. – 2024. – Vol. 17. – №. 2. – P. 12–25. <https://doi.org/10.37134/ibej.Vol17.2.2.2024>
252. Zuboff S. The age of surveillance capitalism //Social theory re-wired. – Routledge, 2023. – C. 203-213.
253. Yeung K. Algorithmic regulation: A critical interrogation //Regulation & governance. – 2018. – Vol. 12. – №. 4. – P. 505-523.

Лукичѐв Павел Михайлович

**УПРАВЛЕНИЕ ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ:  
РИСКИ И ГОСУДАРСТВЕННОЕ РЕГУЛИРОВАНИЕ**

Монография