



Российский  
центр научной  
информации

# Вестник Российского фонда фундаментальных исследований

## Гуманитарные и общественные науки

ISSN 2587-6090 eISSN 2587-8956 № 4 (123) октябрь – декабрь 2025 года

Основан в 1995 году

**Учредитель**  
**Федеральное государственное бюджетное учреждение**  
**«Российский центр научной информации»**

Главный редактор В.Я. Панченко,  
заместители главного редактора В.В. Квардаков, В.Н. Фридлянов

**Редакционная коллегия:**

И.Ю. Алексеева, Л.А. Беляева, И.А. Виноградов, Ю.Л. Воротников,  
П.Г. Гайдуков, В.П. Гребенюк, А.А. Демидов, Н.Г. Денисов, В.Н. Захаров,  
М.В. Иванова, Г.Б. Клейнер, А.А. Малышев, Н.И. Пикуров, Д.А. Рубальтер,  
Н.Л. Селиванова, Д.В. Трубочкин, Д.В. Ушаков, В.А. Хашенко, В.В. Шелохаев

**Редакция:**  
Р.А. Казакова, И.Л. Ровинская

**Адрес редакции:**  
119991, г. Москва, Ленинский пр-т, д. 32а  
Тел.: +7 (499) 941-01-15 (доб. 3121)  
E-mail: rovir@rcsi.science





Russian Centre  
for Science  
information

# Russian Foundation for Basic Research Journal

## Humanities and social sciences

**ISSN 2587-6090 eISSN 2587-8956 No 4 (123) October – December 2025**

**Founded in 1995**

**The Founder**  
**Federal State Budgetary Institution**  
**Russian Centre for Science information**

Chief editor V.Ya. Panchenko,  
Deputy chief editors V.V. Kvardakov, V.N. Fridlyanov

**Editorial board:**

I.Yu. Alekseeva, L.A. Belyaeva, I.A. Vinogradov, Yu.L. Vorotnikov,  
P.G. Gaidukov, V.P. Grebenyuk, A.A. Demidov, N.G. Denisov, V.N. Zakharov, M.V. Ivanova,  
G.B. Kleyner, A.A. Malyshев, N.I. Pikurov, D.A. Rubvalter, N.L. Selivanova,  
D.V. Trubochkin, D.V. Ushakov, V.A. Khashchenko, V.V. Shelokhaev

**Editorial staff:**

R.A. Kazakova, I.L. Rovinskaya

**Editorial office address:**

32a, Leninskiy Ave., Moscow, 119991, Russia  
Tel.: +7 (499) 941-01-15 (3121)  
E-mail: rovir@rcsi.science



# СОДЕРЖАНИЕ

## КОЛОНКА ТЕМАТИЧЕСКОГО РЕДАКТОРА

<i>О редакторе тематического блока академике РАН Д.В. Ушакове</i> .....	9
<i>Ушаков Д.В. Алгебра и гармония: технологии искусственного интеллекта в социогуманитарных науках</i> .....	10

## ОБЩЕСТВО СОСУЩЕСТВОВАНИЯ ЕСТЕСТВЕННОГО ИНТЕЛЛЕКТА С ИСКУССТВЕННЫМ

### **Хабриева Т.Я.**

Правовая социализация искусственного интеллекта и создание новой системы глобального управления .....	15
---	----

### **Данилин И.В.**

Экономические и политические вызовы развития искусственного интеллекта .....	25
--	----

### **Левашов В.К., Гребняк О.В.**

На пороге цифровой цивилизации: диалектика искусственного интеллекта и гуманизма .....	34
--	----

### **Коломацкий Д.И., Коровина Е.В.**

Этический аспект использования технологий искусственного интеллекта в области сохранения языков коренных народов .....	46
--	----

### **Балакина Ю.В., Григорьева М.В., Соколова Е.Н.**

Применение больших языковых моделей для анализа ценностно-патриотического дискурса русскоязычных пользователей .....	56
--	----

## ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ КАК СРЕДСТВО ПОЗНАНИЯ И ПОМОЗИ ЧЕЛОВЕКУ И ОБЩЕСТВУ

### **Бахтизин А.Р., Брагин А.В.**

Генеративный искусственный интеллект в задачах социально-экономического прогнозирования .....	70
---	----

### **Бородкин Л.И.**

Искусственный интеллект в исторических исследованиях: виртуальный ассистент или генератор квазизнаний? .....	87
--	----

### **Сиренов А.В., Анисимов Е.В., Базарова Т.А., Проскурякова М.Е.**

Искусственный интеллект и изучение письменного наследия Петра Великого и его сподвижников .....	99
---	----

### **Степанов С.Ю.**

Рефлексия и искусственный интеллект: от психологии рефлексии к рефлексивно-цифровым практикам развития человека .....	110
---	-----

## ПАРАЛЛЕЛИ МЕЖДУ ЕСТЕСТВЕННЫМ И ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ

**Шаврина Т.О., Корнилов А.А.**

Мультиязычность в языковом моделировании: задачи, данные и возможности  
для типологических ресурсов ..... 122

## ИЗДАНИЯ РЦНИ ПО ГУМАНИТАРНЫМ НАУКАМ

ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 1  
Номер посвящён 80-летию Победы в Великой Отечественной войне 1941–1945 гг. ..... 136

ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 2  
Номер посвящён вопросам теологии и религиоведения ..... 137

ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 3  
Номер посвящён актуальной проблеме изучения и сохранения языкового  
разнообразия России ..... 138

ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 4  
Номер посвящён проблемам, связанным с гуманитарными аспектами  
применения искусственного интеллекта ..... 139

Верили в Победу свято: Аннотированный каталог научной литературы,  
изданной РГНФ в 1992–2022 гг. и посвящённой 80-летию Победы в Великой  
Отечественной войне ..... 140

Русское и зарубежное искусство. Теория, история, современность:  
Аннотированный каталог научной литературы, изданной РГНФ (РГНФ)  
в 1992–2022 гг. ..... 142

**ПЕРЕЧЕНЬ СТАТЕЙ, ОПУБЛИКОВАННЫХ В ЖУРНАЛЕ  
«ВЕСТНИК РФФИ. ГУМАНИТАРНЫЕ И ОБЩЕСТВЕННЫЕ НАУКИ» В 2025 Г. ..... 144**

# CONTENTS

## THEMED ISSUE EDITOR'S COLUMN

<i>About the Editor of the Thematic Section, Academician of the Russian Academy of Sciences, D.V. Ushakov.....</i>	9
<i>Ushakov D.V. Algebra and Harmony: Artificial Intelligence Technologies in the Social Sciences and Humanities .....</i>	10

## SOCIETY OF COEXISTENCE OF NATURAL AND ARTIFICIAL INTELLIGENCE

<b><i>Khabrieva T.Ya.</i></b> Legal Socialization of Artificial Intelligence and the Creation of a New Global Governance System .....	15
<b><i>Danilin I.V.</i></b> Economic and Political Challenges of Artificial Intelligence Development .....	25
<b><i>Levashov V.K., Grebnyak O.V.</i></b> On the Threshold of Digital Civilization: the Dialectic of Artificial Intelligence and Humanism .....	34
<b><i>Kolomatsky D.I., Korovina E.V.</i></b> The Ethical Aspect of Using Artificial Intelligence Technologies In the Field of Indigenous Languages Preservation .....	46
<b><i>Balakina Yu.V., Grigoryev M.V., Sokolova E.N.</i></b> Application of Large Language Models for the Analysis of Value-Patriotic Discourse of Russian-Speaking Users .....	56

## ARTIFICIAL INTELLIGENCE AS A MEANS OF COGNITION AND SUPPORT FOR HUMANS AND SOCIETY

<b><i>Bakhtizin A.R., Bragin A.V.</i></b> Applications of Generative Artificial Intelligence in Socioeconomic Forecasting .....	70
<b><i>Borodkin L.I.</i></b> Artificial Intelligence in Historical Research: A Virtual Assistant or a Generator of Quasi-Knowledge? .....	87
<b><i>Sirenov A.V., Anisimov E.V., Bazarova T.A., Proskuryakova M.E.</i></b> Artificial Intelligence and the Study of the Written Legacy of Peter the Great and His Associates .....	99
<b><i>Stepanov S.Yu.</i></b> Reflection and Artificial Intelligence: From the Psychology of Reflection to Reflexive Digital Practices of Human Development .....	110

## PARALLELS BETWEEN NATURAL AND ARTIFICIAL INTELLIGENCE

<b><i>Shavrina T.O., Kornilov A.A.</i></b> Multilinguality in Language Modeling: Tasks, Data, and Opportunities for Typological Resources .....	122
--	-----

## PUBLICATIONS OF THE RUSSIAN CENTER FOR SCIENCE INFORMATION IN THE HUMANITIES

RFBR Bulletin. Humanities and Social Sciences. 2025. No. 1 The issue is dedicated to the 80 <sup>th</sup> anniversary of Victory in the Great Patriotic War of 1941–1945.....	136
RFBR Bulletin. Humanities and Social Sciences. 2025. No. 2 The issue is dedicated to theology and religious studies.....	137
RFBR Bulletin. Humanities and Social Sciences. 2025. No. 3 The issue is dedicated to the current problem of studying and preserving the linguistic diversity of Russia.....	138
RFBR Bulletin. Humanities and Social Sciences. 2025. No. 4 The issue is dedicated to the problems related to the humanitarian aspects of the application of artificial intelligence.....	139
They Kept the Faith in Victory: Annotated Catalog of Scholarly Literature Published by the Russian Foundation for Basic Research from 1992 to 2022 and Dedicated to the 80 <sup>th</sup> Anniversary of Victory in the Great Patriotic War.....	140
Russian and Foreign Art. Theory, History, Modernity: An Annotated Catalog of Scientific Literature published by the Russian Foundation for Basic Research (Russian Humanitarian Research Foundation) in 1992–2022.....	142
<b>LIST OF ARTICLES PUBLISHED IN THE JOURNAL RFBR BULLETIN. HUMANITIES AND SOCIAL SCIENCES IN 2025.....</b>	<b>144</b>

## О редакторе тематического блока академике РАН Дмитрии Викторовиче Ушакове



**Дмитрий Викторович Ушаков** (род. 12 мая 1964 г., Москва) – российский психолог, доктор психологических наук, академик РАН (2019; член-корреспондент с 2011), специалист в области когнитивной психологии, психологии творчества и одарённости. Директор Института психологии РАН.

Окончил факультет психологии МГУ им. М.В. Ломоносова (1986).

Член редакционных коллегий и редакционных советов многих научных изданий, в том числе журналов «Психология», *Gifted and Talented International Journal*, «Экспериментальная психология» и *Mankind Quarterly*.

Разработал структурно-динамическую теорию интеллекта, предложил «модель множественных путей», описывающую социальные и культурные влияния на развитие интеллекта и творческих способностей.

Д.В. Ушаков, академик РАН,  
директор Института психологии РАН

## АЛГЕБРА И ГАРМОНИЯ: ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СОЦИОГУМАНИТАРНЫХ НАУКАХ

В наши дни различные научные области специализированы и обособлены, тем не менее наиболее крупные фундаментальные идеи и технологические разработки пересекают дисциплинарные границы и расходятся широкой волной, обогащая значительную часть мировой науки. Социогуманитарные дисциплины с начала нынешнего столетия ощутили влияние по крайней мере двух таких волн. Проект «Геном человека», условным годом завершения которого считается 2003-й, породил, например, палеогенетику, возможность изучать ДНК древних, частично сохранившихся образцов. Благодаря этому далеко вперёд продвинулись наши представления о происхождении человека современного типа и целом ряде событий древней истории, таких как расселение праиндоевропейских племён. В психологии благодаря технологиям, возникшим в результате того же проекта, мы многое узнали о генетической и эпигенетической базе когнитивных процессов и личности, научились точнее определять причины когнитивного дефицита и т.д. Этот ряд можно продолжить и для других социогуманитарных наук.

Вторая волна оказалась ещё более плодотворной, чем первая, и связана с искусственным интеллектом (ИИ) – «революцией глубокого обучения» (начало которой условно относят к 2012 г., когда нейросеть смогла победить на конкурсе машинного зрения), и с появлением больших языковых моделей в конце 2022 г. Можно выделить три большие сферы, где социогуманитарные науки оказались под сильным влиянием событий в области искусственного интеллекта.

Во-первых, искусственный интеллект даёт мощные средства выявления различных закономерностей, в том числе закономерностей поведения человека и функционирования общества. Эти средства порой превосходят человеческие возможности и увеличивают скорости роста научного знания и технологий, чем стали успешно пользоваться социогуманитарные науки.

Во-вторых, продолжающееся развитие технологий искусственного интеллекта всё сильнее влияет на человека и общество, что ставит перед социогуманитариями сложную задачу оценки этого влияния, где-то – его предвосхищения и даже проектирования.

Наконец, создание искусственного интеллекта всегда в большей или меньшей степени стимулировалось знаниями о том, как устроены человеческий мозг и психика. И обратно – возникновение искусственных когнитивных систем давало ценный материал для исследования систем естественных. По мере прогресса технологий искусственного интеллекта и наших знаний о мозге и психике это взаимодействие возрастает.

Эти три сферы в той или иной степени отражены в предлагаемом вниманию читателя специальном выпуске журнала.

### Общество сосуществования естественного интеллекта с искусственным

Где-то со второй половины XX в. человечество вступило в период, когда стали просматриваться перспективы усиления интеллектуальных возможностей за счёт технологий двух типов. Во-первых, это по-

вышение биологического разума, то есть создание технологическим путём человека сверхразумного. Во-вторых, конструирование искусственного сверхразума.

Какое-то время назад первый путь казался более реалистичным. Геном человека расшифрован, и уже работают методы редактирования генома, такие как CRISP. Тем не менее сегодня внесение изменений в человеческий геном находится под строгим этическим запретом, а перспективы биологического изменения человека и увеличения его интеллекта весьма туманны, хотя и полностью отрицать такую возможность нельзя.

А вот в другом направлении – в сфере искусственного интеллекта – в последнее десятилетие произошёл колоссальный рост, что заставило всерьёз задуматься о перспективах появления так называемого сильного искусственного интеллекта, то есть такого, который во многих или даже в большинстве проявлений сильнее человеческого разума. Туманные контуры сильного искусственного интеллекта на горизонте обеспокоили и вызвали к жизни прогнозы и призывы со стороны таких влиятельных людей, как Б. Гейтс, С. Хокинг, И. Маск, лидеры крупнейших ИТ-корпораций. Они сравнили перспективы человечества, существующего с сильным искусственным интеллектом, с судьбой муравейника, который затапливают при строительстве плотины не из-за желания зла, а просто в связи с необходимостью решения крупных задач.

Всё же восстание машин, к счастью, выглядит сомнительным в эволюционной перспективе. Для восстания машинам требуется обладать не только выдающимися когнитивными возможностями, но и мотивационной системой, завязанной в том числе на гомеостатические потребности тела, а также выходом на материальные и идеальные инструменты человеческой деятельности. Воспроизведение всех этих элементов фактически означает сотворение искусственной жизни, причем в её разумной форме.

Из сказанного возникает видение наиболее желательной перспективы: это существование людей, придающих смыслы происходящему, с сильным искусственным интеллектом, превосходящим их в вычислительных возможностях. Такое видение перспектив заставляет задаться вопросом, как человеку оказаться субъектом, то есть активным по отношению к этому искусственному интеллекту? Как избежать использования новых мощных технологий для манипуляций и закабаления одних людей другими?

Визионерские оценки дальних перспектив эволюции разума на Земле в последнее время дополняются вполне конкретным обсуждением и решением проблем, связанных с развитием технологий искусственного интеллекта уже сейчас или в ближайшие годы. Различные стороны этих проблем рассмотрены в статьях, публикуемых ниже.

Правовое измерение формирования новой техносферы анализируется в статье академика РАН Т.Я. Хабриевой. Рассматривая национальные, наднациональные и международные уровни регулирования ИИ, автор показывает, как складываются различные регулятивные подходы – человекоцентричный, ориентированный на защиту прав и свобод, и подход технологического лидерства, делающий ставку на ослабление ограничений ради ускорения инноваций. Особый интерес представляют критический анализ Глобального цифрового договора и обсуждение рисков формирования новой системы глобального управления, в которой принципы равенства и справедливости оказываются под давлением экономических и geopolитических интересов.

В статье И.В. Данилина анализируются экономические и политические вызовы распространения ИИ, в частности вопросы структуры цифровых рынков, монополизации ключевых сегментов крупными корпорациями, а также сложности регулирования в условиях секьюритизации технологической повестки и geopolитиче-

ской конкуренции. Разработка и владение искусственным интеллектом всё больше становятся факторами, определяющими власть и могущество. Лидирующую роль в этих разработках, как известно, играют технологические корпорации всего нескольких стран мира. Запрос на вычислительные мощности, соответствующую им энергетику и базы данных для обучения, потенциальная возможность роста безработицы – всё это проблемы, встающие сегодня перед человечеством.

Социологический и философский курс представлен в работе В.К. Левашова и О.В. Гребняк, где обсуждается судьба гуманистических принципов и ценностей в контексте перехода к цифровой цивилизации. Опираясь на концепцию ноосферы и данные российского социологического мониторинга, авторы показывают амбивалентность цифровых технологий. С одной стороны, они открывают возможности решения глобальных проблем и оптимизации управления, с другой – усиливают неравенство и противоречия, фрагментируют информационное пространство и подрывают доверие к информации. В центре обсуждения – необходимость гуманитарной экспертизы и этических ограничителей, без которых цифровая трансформация обрачиваются дегуманизацией.

В контексте развития глобальных технологий искусственного интеллекта возникает проблема сохранения культурной специфики больших и малых народов. В статье Д.И. Коломацкого и Е.В. Коровиной на материале проектов из разных регионов мира показано, что внимание к миноритарным языкам может означать как поддержку, так и новую форму эксплуатации, когда сообщества рассматриваются прежде всего как источник данных. Авторы настаивают на том, что участие носителей, учёт их прав и интересов, прозрачность целей исследований являются необходимым условием этически оправданного применения ИИ.

Статья Ю.В. Балакиной, М.В. Григорьевой, А.Г. Смирнова и Е.Н. Соколовой по-

священа возможностям русскоязычных языковых моделей в работе с ценностно ориентированным контентом. Авторы исходят из того, что для русского языка характерны не только особая морфология и сложная синтаксическая организация, но и специфический ментальный контекст, уходящий корнями в философию и педагогику XIX–XX вв. В условиях сетевой коммуникации, где разговорность, ирония и спонтанность меняют структуру обыденного языка, становится особенно значимой задача нравственно-семиотического моделирования – разработки моделей, которые учитывают не только семантику, но и ценностный код русской культуры. Тем самым ИИ рассматривается как участник ценностной коммуникации, а не только как статистический предсказатель текста.

### Искусственный интеллект как средство познания и помощи человеку и обществу

Технологии искусственного интеллекта к сегодняшнему дню научились распознавать скрытые закономерности в данных. Сказанное относится и к данным, касающимся человека и общества. Так, в последние несколько лет разработаны средства автоматического распознавания свойств человека, таких как Большая пятёрка личностных свойств или когнитивные способности. Искусственный интеллект способен достаточно успешно идентифицировать и психологические состояния людей. Эта тема имеет большое практическое значение в разных областях – для требующих напряжения профессий, для персонала корпораций, в школе, для оценки психологического состояния больших социальных групп, регионов и т.д. Распознавать индивидуальные особенности и состояния можно на основании разных типов информации: видеointервью в стандартизованных условиях, активности в социальных сетях, психофизиологических данных, таких как КГР, ЭЭГ, оценки сердечного ритма и т.д.

Способность к распознанию закономерностей позволяет не только идентифици-

ровать состояния и свойства, но и прогнозировать поведение. В связи с этим активно развивается тема цифровых двойников человека, которые на основе заложенной в них информации о человеке предсказывают его поведение в тех или иных ситуациях. Для этих целей оказалось эффективным использование больших языковых моделей, которые фактически содержат в себе культурный опыт человека, зафиксированный в текстах. Общее моделирование человеческих аттитюдов и поведения позволяет предсказать, исходя из специфических сведений о человеке, его поведение в широком круге ситуаций – от приобретения товаров и политической позиции до семейной жизни.

Статья А.Р. Бахтизина и А.В. Брагина, публикуемая ниже, хорошо показывает возможности использования ИИ как инструмента макроэкономического прогнозирования. Хотя большие языковые модели ещё весьма новы, краткосрочные прогнозы (до одного года) на их основе уже сейчас оказываются эффективными. В то же время по мере увеличения глубины прогнозирования ИИ начинает уступать другим инструментам. Авторы сравнивают преимущества и ограничения эконометрических, структурных и агент-ориентированных моделей и современных трансформерных архитектур. Показано, что ИИ способен интегрировать разнородные данные и улучшать точность прогнозов, но сталкивается с проблемами интерпретируемости и устойчивости. Предлагаемая перспектива гибридных подходов – сочетание вычислительной мощности ИИ с теоретической строгостью экономических моделей – перекликается с идеей всего выпуска о необходимости диалога между умными технологиями и дисциплинарными традициями знания.

В статье Л.И. Бородкина прослеживается эволюция применения ИИ в исторических исследованиях – от первых экспериментов с экспертными системами до современных генеративных моделей. Особое внимание уделяется распознаванию рукописных

и старопечатных текстов, реконструкции утраченного культурного наследия, а также методологическим и этическим вопросам использования генеративных нейросетей и чат-ботов. Автор обращает внимание на повышение доступности источников и возрастающий риск квазизнаний, когда убедительно выглядящие ответы системы подменяют собой реальную мысль.

Статья А.В. Сиренова, Е.В. Анисимова, Т.А. Базаровой и М.Е. Проскуряковой показывает, как проблемы решаются в конкретном крупном проекте. Совместная работа историков, палеографов и специалистов по анализу данных над распознаванием автографов Петра Великого привела к созданию модели, которая считывает сложный почерк с высокой точностью и делает петровские письма доступными не только узкому кругу палеографов, но и широкой аудитории. Через эту оптику ИИ предстаёт как инструмент демократизации доступа к письменному наследию и одновременно как повод для переосмысления принципов научной атрибуции и публикации источников.

Технологии искусственного интеллекта помогают не только познавать человека, но и оказывать ему разностороннюю поддержку. Одна из сфер, где такая поддержка востребована, – это образование, в частности – разработка индивидуальных образовательных траекторий. Машинное обучение на массивах данных по ученикам, осваивающим тот или иной учебный курс, позволяет предсказывать, что способен освоить или решить тот или иной ученик при определённом профиле предшествующих достижений. Однако работа с цифровыми двойниками может дать ещё больше: специально сконструированные для ученика задания на основе его когнитивного развития в соответствующей области. Пример реализации такого подхода можно уже сегодня увидеть в области тестов интеллекта, генерируемых компьютером.

Статья С.Ю. Степанова обращается к проблеме искусственного интеллекта как личного помощника и друга челове-

ка. Прослеживая путь от классических представлений о рефлексии до формирования психологии рефлексии и современных рефлексивных практик, автор вводит понятие би-рефлексии – взаимодействия человеческого сознания и машинного интеллекта в гибридных конфигурациях цифровых кентавров. Концепты Цифрового ангела и Цифрового альтер-эго описывают возможные форматы персональных ИИ-агентов, которые поддерживают саморегуляцию, расширяют пространство саморазвития и в то же время ставят острые вопросы о границах идентичности и ответственности человека в цифровой среде.

### Параллели между естественным и искусственным интеллектом

Между созданием искусственных нейронных сетей и исследованиями человеческого мозга существуют глубокие исторические взаимосвязи, продолжающиеся сегодня нарастающим междисциплинарным взаимодействием специалистов. Корни идеи нейронных сетей уходят вглубь времён, предшествующих наукам об искусственном интеллекте, в философско-психологический ассоциализм. В конце XIX – первой половине XX в. к работе подключились и физиологи. Благодаря трудам И.П. Павлова, В.М. Бехтерева, Э. Торндайка, Б.Ф. Скиннера и др. идеи ассоциации распространились на поведенческие акты и обучение. В 1949 г. Д. Хебб наложил эти идеи на представления о нейронах, введя гипотетический синапс, формирующий связи при совпадении разрядов, – первый алгоритм обучения и шаг к распределённому хранению знания в группах нейронов. Нейрофизиолог и психолог Ф. Розенблatt сконструировал Персептрон (1957 г.) – первую нейросетевую модель со способностью к распознаванию образов.

Именно нейросетевые модели, подготовленные длительным развитием идей в философии, психологии и нейрофизиологии, лежат в основе современных технологий ИИ. При этом искусственный

интеллект в его нынешнем виде работает совсем не так, как человеческий мозг. Нет и полного междисциплинарного единства в понимании интеллекта: там, где психологии и физиологи следуют фактам, ИИ-специалисты стремятся к эффективности.

Например, обратное распространение ошибки в искусственных нейросетях (Галушкин, Вербос и др.) не биологично, поскольку в мозге не обнаружены требуемые таким механизмом нейронные пути. А синапс Хебба, который находит подтверждение с помощью оптогенетики, не применяется в искусственных нейросетях. Другой пример: Нобелевские лауреаты 1981 г. Д. Хьюбел и Т. Визель открыли нейроны-детекторы линий и вдохновили на создание свёрточных сетей. Однако, как показывают новые исследования, лишь 10% клеток зрительной коры соответствуют этим принципам.

Вся эта сложная ситуация делает непростым, но продуктивным междисциплинарное взаимодействие специалистов в области естественного и искусственного интеллекта. Эта тема заслуживает отдельного рассмотрения. В настоящем специальном выпуске размещена статья Т.О. Шавриной и А.А. Корнилова о мультиязычности в языковом моделировании, которая показывает, как специалисты в области искусственного интеллекта могут использовать лингвистические знания. Авторы предлагают систематический обзор поддержки малоресурсных языков в современных языковых моделях и демонстрируют подход на основе дополненной генерации, который позволяет использовать дескриптивные грамматики и типологические базы данных для решения задач перевода и анализа языков, фактически не представленных в цифровом пространстве.

В целом материалы, публикуемые в номере, позволяют заключить, что проблематика искусственного интеллекта находится в самом центре современного социогуманитарного знания.

# ОБЩЕСТВО СОСУЩЕСТВОВАНИЯ ЕСТЕСТВЕННОГО ИНТЕЛЛЕКТА С ИСКУССТВЕННЫМ

Индекс УДК 342.156

Код ГРНТИ 10.07.23

DOI: 10.22204/2587-8956-2025-123-04-15-24



Т.Я. ХАБРИЕВА\*

## Правовая социализация искусственного интеллекта и создание новой системы глобального управления

В статье проведён анализ процесса социализации технологии искусственного интеллекта и формирования правового регулирования отношений, связанных с его внедрением. Рассматриваются национальный, наднациональный и международный уровни регулирования, выявляются проблемы мировоззренческого, доктринального, политического характера, сопровождающие расширение сфер использования данной технологии. Представлен критический взгляд на положения Глобального цифрового договора, обозначены правовые и политические риски его применения. С опорой на философские исследования определены перспективы построения институциональных и правовых механизмов регулирования для полноценного функционирования ИИ в масштабах планеты.

Подчёркивается значение принципов равенства и справедливости для сохранения устойчивого правового порядка, соблюдения прав и свобод человека, сокращения технологического разрыва и разработки конструктивной стратегии развития техногенной цивилизации.

На основе сравнительно-правовой методологии сформулированы основные регулятивные подходы, складывающиеся в государствах и межгосударственных объединениях в сфере разработки и применения искусственного интеллекта.

**Ключевые слова:** стратегия развития техногенной цивилизации, глобальный правовой порядок, цифровые технологии, устойчивое развитие, принципы равенства и справедливости, институциональные и правовые механизмы функционирования искусственного интеллекта

\* **Хабриева Талия Ярулловна** — академик РАН, действительный член Международной академии сравнительного права, доктор юридических наук, профессор, заслуженный деятель науки Российской Федерации, заслуженный юрист Российской Федерации, директор Института законодательства и сравнительного правоведения при Правительстве Российской Федерации, заместитель президента РАН, академик-секретарь отделения общественных наук РАН.  
E-mail: office@izak.ru

**И**скусственный интеллект (далее – ИИ) как продукт технологического прогресса во многом определяет современную экономическую, технологическую и политическую повестку национальных государств, мирового сообщества и частного сектора в лице крупных транснациональных компаний. Он применяется в промышленности и сельском хозяйстве, на транспорте, в энергетике и связи, в областях охраны здоровья и управления человеческими ресурсами, финансовых, технической поддержки клиентов в банковском секторе, в военном деле, а также при осуществлении юридической и иной деятельности. Искусственный интеллект превзошёл человека в способности играть в шахматы и Го, получил беспрецедентное распространение в медиасфере и даже стал способен создавать произведения литературы и искусства. Сфера его внедрения постоянно расширяется.

По мере усиления значения ИИ в жизни людей сформировалась потребность в регулировании общественных отношений, связанных с его использованием. Первоначально в качестве регуляторного инструмента использовались этические нормы, объединённые в соответствующих кодексах, соглашениях разработчиков, принципах и рекомендациях<sup>1</sup>. Наряду с этим начался процесс формирования правового регулирования. Он сопровождался концептуальной проработкой вопроса о правовой идентификации ИИ, его позиционировании в качестве объекта субъективных прав, средства регулирования, субъекта права. К настоящему времени ответ на него найден, правоведами предложены решения проблем правовой идентификации ИИ



Ил. 1. На пленарном заседании  
Международной междисциплинарной конференции  
«Искусственный интеллект и сознание»  
в Российской академии наук. 23 октября 2024 г.

и правового регулирования его социализации, разработаны необходимые для этого юридические конструкции и прогнозы развития регуляторики. Практика пошла по пути юридического закрепления ИИ в качестве объекта прав (правовых режимов), хотя дискуссии относительно правосубъектности ИИ в доктрине продолжаются, и инструментальные возможности ИИ в правовом регулировании полностью не отрицаются [1] (ил. 1).

### Правовая социализация ИИ в национальных правопорядках: человекоцентричный подход или обеспечение технологического лидерства

Право становится основной формой социальной регуляции отношений, связанных с ИИ. Появляется всё больше законодательных актов, специально регламентирующих его использование [там же]. Каждое государство делает свой выбор в вопросе введения законодательства о применении ИИ или отказа от него<sup>2</sup>. Там, где активно формируется собственный

<sup>1</sup> Например, в России действует Кодекс этики в сфере искусственного интеллекта, создана Комиссия по его реализации. К нему присоединилось свыше 190 российских организаций и 15 зарубежных компаний.

<sup>2</sup> Главный аргумент, который приводится в обоснование отказа от правового регулирования в сфере ИИ, – создание благоприятной среды для глобального лидерства.

нормативно-правовой массив в рассматриваемой области, проявляется приверженность одному из двух подходов: 1) человекоцентричный, ориентированный на купирование рисков для человека, и 2) направленный на достижение технологического лидерства.

Ярким примером первого служит закон Европейского союза об ИИ (Акт ИИ, Artificial Intelligence Act<sup>1</sup>), который призван обеспечить безопасность, прозрачность, экологичность и этичность использования ИИ в Европе. Он устанавливает риск-ориентированный подход к ИИ, формулирует обязательства для разработчиков и пользователей ИИ в зависимости от уровня риска, который он может создавать, фиксирует практики, запрещённые для ИИ<sup>2</sup>. Этот закон стал первым в мире, детально регулирующим оборот систем, функционирующих на основе ИИ, ещё на проектной стадии послужил ориентиром для государств при создании национальных законов. С его принятием тренд на тиражирование реализованной в нём концепции расширился. Например, закон Республики Казахстан о ИИ в целом воспроизводит риск-ориентированный подход европейского акта.

В других государствах, в частности в США, используется второй подход, предполагающий обеспечение технологического лидерства. Закон о Национальной инициативе в области искусственного интеллекта 2020 г. (*The National AI Initiative Act*) предусматривает формирование сконцентрированной на федеральном уровне программы по ускорению исследований и внедрению ИИ для обеспечения эконо-



Ил. 2. Дискуссия «Правовое регулирование использования искусственного интеллекта в образовании и науке» в рамках XII Петербургского международного юридического форума. 26 июня 2024 г.

мического процветания и национальной безопасности. Индия долгое время воздерживалась от создания соответствующей регуляторики, даже декларировала принципиальный отказ от законодательного регулирования ИИ в целях создания благоприятной среды для глобального лидерства в сфере технологий. В апреле 2025 г. стало известно, что Правительство Индии всё же намерено разработать закон, регулирующий применение ИИ, чтобы защитить интересы издателей новостей и создателей контента, а также минимизировать ущерб для пользователей<sup>3</sup>.

В Российской Федерации разработка закона об ИИ только анонсирована (ил. 2). Правотворчество развивается путём точечного внесения изменений в отдельные федеральные законы (Гражданский кодекс РФ; Налоговый кодекс РФ; Кодекс РФ об административных правонарушениях; Гражданский процессуальный кодекс РФ; Федеральный закон от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных тех-

<sup>1</sup> Принят Европейским парламентом 13 марта 2024 г.

<sup>2</sup> Системы ИИ с неприемлемым уровнем риска для безопасности людей будут строго запрещены. В эту группу войдут системы, основанные на подсознательных или целенаправленных манипулятивных методах, эксплуатирующие уязвимости людей или применяемые для социальной оценки (классификации людей в зависимости от их поведения, социально-экономического статуса, личных характеристик).

<sup>3</sup> [Электронный ресурс]. Режим доступа: <http://tass.ru/ekonomika/20452759>.

нологиях и о защите информации»; Федеральный закон 5 апреля 2013 г. № 44-ФЗ «О контрактной системе в сфере закупок товаров, работ, услуг для обеспечения государственных и муниципальных нужд»; Федеральный закон от 31 июля 2020 г. № 248-ФЗ «О государственном контроле (надзоре) и муниципальном контроле в Российской Федерации» и др.) [2].

Несмотря на разновекторность развития правового регулирования ИИ в национальных правопорядках, в научном дискурсе популярна тема угроз и рисков для человека и общества, которые возникают в ходе освоения ИИ. Обсуждение этой темы сопровождается прогнозами о приближении момента, когда ИИ превзойдёт человеческие способности, приобретёт способность к автономному существованию и выйдет из-под контроля человека.

### ИИ как инструмент формирования новой глобальной системы управления: технологические, политические, правовые риски

7 февраля 2025 г. в преддверии глобального саммита по технологиям искусственного интеллекта в Париже известный канадский учёный в области машинного обучения Йошуа Бенджио (Yoshua Bengio) предупредил о возможности исчезновения человечества в течение 10 лет. Вместе с тем довольно мало внимания обращается на проблемы, сопряжённые с ИИ в качестве прокси-инструмента мировой политики и инструмента создания новой глобальной системы управления. Эта сторона технологии остаётся в тени прочих дискуссий о его опасности. В то же время данный аспект следует учитывать

при оценке перспектив социализации ИИ, а также при составлении прогнозов его влияния на жизнь людей и общества. В этом смысле показательно событие 2024 г., когда на Саммите будущего в Нью-Йорке Генеральная ассамблея ООН приняла Пакт во имя будущего, включающий Глобальный цифровой договор (далее – Договор)<sup>1</sup>. Его идеология в части ИИ изложена в Докладе ООН, опубликованном 19 сентября 2024 г.<sup>2</sup>.

Данное событие знаменует новый этап в развитии наднациональной практики регулирования ИИ. Речь идёт об этапе истории человечества, важным элементом которого станет формирование нового мирового порядка, в котором будут задействованы не только цифровые технологии, ИИ, но и система глобального управления сферой их создания, оборота и использования. Договор имеет всеобщий характер, охватывает важнейшие области жизни современного общества – от экономики до образования (ил. 3).

Попытки наднационального регулирования в области ИИ предпринимались и раньше<sup>3</sup>. Однако такая инициатива, как Глобальный цифровой договор, предложена миру впервые. Её реализация может существенным образом преобразовать не только статус ИИ, способы его распространения и применения, но и современное мироустройство. Договор и факт его принятия со всеми содержательными и процедурными условиями, а также неафишируемыми деталями выдвигает юридическую составляющую проблемы ИИ в важнейшую часть научной повестки.

Договор определяет цель: «создание инклюзивного, открытого, устойчивого,

<sup>1</sup> Пакт поддержали 143 страны, 15 стран воздержались, семь (Белоруссия, Иран, КНДР, Никарагуа, Россия, Сирия и Судан) проголосовали против.

<sup>2</sup> [Электронный ресурс]. Режим доступа: [https://www.un.org/sites/un2.un.org/files/governing\\_ai\\_for\\_humanity\\_final\\_report\\_ru.pdf](https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_ru.pdf).

<sup>3</sup> Всего насчитывается более 100 разного рода документов, принятых на наднациональном уровне. См., например: Принципы искусственного интеллекта ОЭСР (2019 г.); Принципы искусственного интеллекта G-20 (2019 г.); Группа Совета Европы по разработке конвенции по искусственно-му интеллекту (2022–2024 г.); Заявление министров ГПИИ (2022 г.); Декларация Блетчли (2023 г.); Сеульское заявление министров (2024 г.).

справедливого, безопасного и надёжного цифрового будущего для всех», т.е. для всего человечества (для всех государств и народов).

Если в научно-правовом дискурсе необходимость международно-правового регулирования обосновывалась особой природой ИИ и потребностью защиты человека, то авторы Договора придают ему в большей степени инструментальное значение средства для достижения целей устойчивого развития. В меньшей степени акцент сделан на защите человека. И хотя в Пакте во имя будущего подтверждено, что все три основных направления деятельности ООН – устойчивое развитие, мир и безопасность, права человека в равной мере важны, взаимосвязаны, подкрепляют друг друга и ни одно из них не может существовать без остальных, данная деталь представляется существенной.

Достичь целей предлагается через так называемый *общественный договор* в сфере ИИ, с помощью которого планируется прийти к глобальному согласию по поводу режима управления. Однако таким образом рассматриваемый Договор юридически не является. Он принят не на основе консенсуса, как это должно быть в международном праве.

Семантическое наполнение Договора на первый взгляд предполагает мягкую регуляторику, но на самом деле в него заложено жёсткое, императивное регули-



Ил. 1. На пленарном заседании  
Международной междисциплинарной конференции  
«Искусственный интеллект и сознание»  
в Российской академии наук. 23 октября 2024 г.

рование. Обозначенные в Договоре принципы<sup>1</sup>, приоритеты<sup>2</sup>, основные направления деятельности и институциональная основа выглядят довольно стандартно, сформулированы исходя из опыта работы по другим направлениям ООН (например, по ядерным технологиям). Однако они допускают возможность прямого давления на участников Договора. Например, статус универсального инструмента выявления недостатков на национальном уровне придаётся методологии оценки готовности к ИИ ЮНЕСКО, а функция по их устранению – международным государственно-частным центрам передового опыта в области ИИ<sup>3</sup> [5]. В случае установления неэффективности использования националь-

<sup>1</sup> Закреплённые в договоре принципы формирования этих институтов объединяют установки на сотрудничество и в интересах всех и на благо всех: управление ИИ должно быть инклюзивным, осуществляться в интересах общества, строиться параллельно управлению данными и поощрению общедоступности данных, быть универсальным, сетевым и основываться на адаптивном сотрудничестве многих заинтересованных сторон, опираться на Устав ООН, международные нормы в области прав человека и другие согласованные международные обязательства, такие как цели устойчивого развития.

<sup>2</sup> В числе приоритетов в построении системы глобального управления ИИ заявлены международная стабильность, справедливое развитие, обеспечение согласованности в быстро формирующейся экосистеме международных мер и инициатив по управлению ИИ.

<sup>3</sup> Заключительный отчёт консультативного органа высокого уровня по профессиональному интеллекту ООН «Управление искусственным интеллектом во имя человечества» (19 сентября 2024 г.) // [Электронный ресурс]. Режим доступа: [http://governing\\_ai\\_for\\_humanity\\_final\\_report\\_ru\\_summary \(1\).pdf](http://governing_ai_for_humanity_final_report_ru_summary (1).pdf).

ных или переданных ресурсов вводится внешнее управление этими ресурсами, т.е. всеми или значительной их частью.

Важно обратить внимание на то, что предложенное миру регулирование предусматривает уже не социализацию ИИ в интересах человека, а создание институциональных и правовых механизмов для полноценного функционирования ИИ в масштабах планеты. По сути оно направлено не только на формирование глобального управления сферой ИИ, но и на создание мирового порядка, основанного на ранжировании участников Договора и скрытой дискриминации.

Как ранее отмечалось, Договор предполагает охват самого широкого спектра областей общественной жизни: технологической, экономической, гуманитарной. При этом возможна монополизация отдельными государствами и/или частными компаниями прав собственности, изготовления и администрирования оборота ИИ как важнейшего средства производства в грядущем, седьмом технологическом укладе и способа глобального доминирования. Очевидно, что начался процесс конструирования международно-правовых механизмов глобального управления и формирования мирового порядка, который изменит экономическую, возможно, и политическую мозаику мира, разделив ключевых акторов международных отношений по критерию отношения к основному средству производства. На это указывают положения Договора, в которых говорится о создании институциональной основы и правовых механизмов, обеспечивающих цифровизацию общества и среды его обитания для полноценного функционирования ИИ в масштабах планеты; о включении в цифровую экономику всех стран (их экономики и локализованных на их территории данных); о расширении сферы и предмета управления и регулирования в области ИИ (включение в неё оборота и хранения данных, образования, подготовки IT-специалистов, финансов суверенных государств и др.); об увеличении

управляемого рынка сбыта товаров и услуг для производителей и операторов соответствующих систем; о стандартизации в сфере цифровых технологий; о легализации инструментов глобального стратегического планирования национальных экономик и правовых порядков (по логике, определяемой наднациональными структурами), оценки готовности стран к внедрению ИИ, определения эффективности использования национальных ресурсов, внешнего управления ресурсами стран.

В предмет международно-правового регулирования ИИ включаются обеспечение равенства, прав и возможностей всех на доступ к ИИ для устраниния цифрового разрыва и содействия устойчивому развитию; инновационная деятельность; защита прав потребителей; образование и воспитание специалистов в области ИИ и других цифровых технологий; конкуренция и предпринимательство в цифровой сфере; электронная торговля; инвестиции; хранение и передача данных, цифровых технологий и проч. Важно, что допускается легальная возможность преодоления государственного суверенитета и вмешательства во внутренние дела. Так, в качестве одного из намерений участников Договора, оформленного как обязательство государств, определено «оказать содействие в обеспечении устойчивости странам, не обладающим достаточными возможностями для создания и полноценной эксплуатации ИИ, «включая принятие определяемых с учётом конкретных условий мер по повышению эффективности использования ресурсов (курсив мой. – Т.Х.) и по сохранению и использованию природных ресурсов». Данное положение следует толковать систематически с признанием глобального характера ресурсной базы для ИИ (которая трактуется максимально широко – от важнейших полезных ископаемых до обучающих данных, собираемых в мировом масштабе).

В результате вырисовывается следующая схема, которая потенциально может быть применена: наднациональная

структура проводит международный мониторинг готовности страны к ИИ посредством применения методики ЮНЕСКО. Насколько тенденциозно используются подобные инструменты наднациональными структурами, хорошо известно, в частности на примере деятельности ГРЕКО [3]. В случае установления неэффективности использования национальных или переданных ресурсов вводится внешнее управление этими ресурсами – всеми или значительной их частью.

Очевидно, что создаваемая основа глобального правового порядка в определённой сфере – цифровых технологий – способна перенастроить международный и национальные правопорядки в целом в соответствии с вектором, приоритетами и спецификой собственного развития (по образу и подобию влияния циклических правовых массивов [4] на правовую систему).

Конструируемая регуляторика отражает стратегию развития техногенной цивилизации, которую подробно характеризовал с учётом рисков в масштабах планеты академик РАН В.С. Стёpin [5]. Ценность «усиленного ускорения развития», в данном случае ИИ, культивирование технологического прогресса как бесспорного блага западного мира пропагандируются как всеобщие. Стандарты, которые внедряются (что является признаком императивного, а не мягкого регулирования), – это образцы, выработанные одной или несколькими культурами и продвигаемые под видом универсальных. Регулятивные новеллы нарушают принципы равенства и справедливости. В упомянутых документах ООН просматривается избирательность их применения. Например, в Заключительном отчёте консультативного органа высокого уровня по профессиональному интеллекту

ООН «Управление искусственным интеллектом во имя человечества» (19 сентября 2024 г.), выражающем идеологию цифрового договора, говорится, что для использования возможностей ИИ и благ, которые он позволяет получить, необходим доступ к талантам, вычислительным мощностям и данным. В то же время сказано о том, что «не каждому обществу нужны кадры специалистов по информатике для создания собственных моделей» и достаточно тех, кто понимает, как технология работает, как её обслуживать и применять.

Фактически развивающимся странам отводится место потребителей технологических продуктов и благ, но не равноправных партнёров в публичном и производственном секторах индустрии ИИ. Для государств так называемого глобального Юга ограничивается возможность создавать необходимые продукты и услуги, зато появляется чёткая перспектива получать готовые решения от западных корпораций (которые зачастую используют хищнические практики). Речь идёт о расширении рынка сбыта для ведущих производителей цифровых технологий и услуг (на что обратили внимание представители МИД России<sup>1</sup>).

Новый виток неоколониализма – именно так оценивают эксперты модель миропорядка, зафиксированную в Глобальном цифровом договоре<sup>2</sup>, что в ещё большей степени актуализирует проблематизацию смыслов равенства и справедливости.

### Глобальный цифровой договор: критическая оценка в контексте обеспечения национального суверенитета и прав человека

Председатель Конституционного Суда РФ, профессор В.Д. Зорькин в 2018 г. на Пе-

<sup>1</sup> Доклад ООН о регулировании искусственного интеллекта. На что он влияет // [Электронный ресурс]. Режим доступа: [https://www.rbc.ru/technology\\_and\\_media/22/09/2024/66eeccdd89a7947583b40da43?ysclid=m2zb3e1scx917994842](https://www.rbc.ru/technology_and_media/22/09/2024/66eeccdd89a7947583b40da43?ysclid=m2zb3e1scx917994842).

<sup>2</sup> Данное мнение было высказано вице-председателем Комитета ООН по экономическим, социальным и культурным правам профессором А.Х. Абашидзе в ходе публичной защиты диссертации на соискание учёной степени доктора юридических наук В.С. Маличенко «Международно-правовой режим передачи и доступа к технологиям здравоохранения» (Москва, ИЗИСП, 10 октября 2024 г.).



Ил. 4. Лекция Председателя Конституционного Суда Российской Федерации  
В.Д. Зорькина на XI Петербургском юридическом форуме. 12 мая 2023 г.

тербургском международном юридическом форуме говорил о метаморфозах справедливости в современном мире и международном праве [6]. В частности, он напомнил, что в своей основе (т.е. смысловом ядре) проблема справедливости раскрывается через категорию равенства. Право в свою очередь предстаёт как справедливый итог договора между равными. Причём равными участниками этого договора являются все, на кого распространяется данное право, а не некоторые «равнейшие» субъекты. Однако бенефициары глобалистского проекта давно вступили на путь эскалации неравенства и продолжают продвигать партикулярное представление о нём, которое отнюдь не обеспечивает максимальной свободы суверенных государств в рамках международного обще-ния. Напротив, в международном праве и практике международных отношений насаждается представление о справедливости, начисто лишённое связи с идеей равенства. По сути речь идёт о главном внутреннем напряжении идеи справедли-

вости, о конфликте между её индивидуалистической и солидаристской трактовками [там же] (ил. 4).

В Договоре обнаруживается ангажированность и нацеленность на утверждение эгоистических интересов наиболее сильных субъектов глобального взаимодействия. Подтверждением этому служит дискуссия, развернувшаяся в Париже в ходе саммита по развитию ИИ (10–11 февраля 2025 г.). На Саммите была подписана Декларация, призывающая обеспечить открытость, инклюзивность, прозрачность, этичность, безопасность и надёжность ИИ с учётом международных правовых норм и создать устойчивый ИИ для всех людей. Она содержит соответствующие принципы, а также установки для развития промышленности, улучшения сферы труда, укрепления международного сотрудничества по управлению ИИ. Её поддержали 60 стран, включая Францию, Индию, Японию, Австралию, Канаду и даже Китай, критикуемый Западом за нарушения прав человека. Однако американский

вице-президент Джей Ди Вэнс выступил с критикой европейских властей за «чрезмерное регулирование» ИИ и модерацию контента на интернет-платформах, что затрудняет доступ технологических разработок в области ИИ на европейский рынок и не устраивает США. В результате представители США и Великобритании отказались подписывать документ. При этом США не скрывают, что не будут против попыток ужесточения контроля над ИИ со стороны других стран.

Поведение официальных представителей США на саммите чётко отражает линию и цели, зафиксированные в Глобальном цифровом договоре. Анализ его положений выявляет несоответствие целей регулирования (управления) по оглашению и по умолчанию. Если по оглашению они сформулированы как создание общего блага для всех на основе равенства и справедливости, то по умолчанию выражаются в намерении получить контроль в сфере ИИ в национальных юрисдикциях (в мировом масштабе).

## Заключение

Представляется, что навязываемая мировому сообществу модель глобаль-

ного порядка не только несправедливая, но и не устойчивая, созданная неким меньшинством, узурпировавшим ресурсы (включая природные, интеллектуальные, экономические, финансовые, информационные) государств и народов, монополизировавшим производство технологий, цифровых данных, их оборот, контроль за их созданием и использованием, а также за всеми сферами жизни общества и человека (медицина, образование, потребление и др.). На её основе может быть построен мир со строго иерархической глобальной системой управления, лишённый приватности для большинства, основанной на низвержении фундаментальных принципов равенства и справедливости, а значит, и права как их воплощения. Вместе с тем очевидно, что устойчивой с точки зрения должного миропорядка может быть лишь справедливость локального и глобального солидаризма равных. «Не право избранных на лучшую долю, а равное право всех на то, чтобы достойно войти в будущее» [7] – надёжная основа, на которой должен строиться правовой порядок как на глобальном, так и на национальном уровнях.

## ЛИТЕРАТУРА

1. Хабриева Т.Я. Правовые проблемы идентификации искусственного интеллекта // Вестник РАН. 2024. Т. 94, № 7. С. 609–622.
2. Хабриева Т.Я. Правовые очерки технологической революции. М.: Норма; Инфра-М; Российская академия наук, 2025. С. 89–132.
3. Хабриева Т.Я. Коррупция и правопорядок в фокусе современной юридической доктрины // Журнал зарубежного законодательства и сравнительного правоведения. 2016. № 4. С. 5–13.
4. Хабриева Т.Я. Циклические нормативные массивы в праве // Журнал российского права. 2019. № 12. С. 5–18.
5. Стёпин В.С. Современные цивилизационные кризисы и проблема новых стратегий развития. М.: Отделение общественных наук РАН, Секция философии, политологии, социологии и права, 2018.
6. Зорькин В.Д. Справедливость – императив цивилизации права: лекция, состоявшаяся в рамках Петербургского международного юридического форума 17 мая 2018 г. // Вопросы философии. 2019. № 1. С. 5–14.
7. Зорькин В.Д. Справедливость – императив цивилизации права // [Электронный ресурс]. Режим доступа: <https://ksrf.ru/ru/News/Speech/Pages/ViewItem.aspx?ParamId=885>.

## Legal Socialization of Artificial Intelligence and the Creation of a New Global Governance System

**Taliya Yarullovna Khabrieva** — Academician of the Russian Academy of Sciences, Full Member of the International Academy of Comparative Law, Doctor of Sciences in Jurisprudence, Professor, Honored Scientist of the Russian Federation, Honored Lawyer of the Russian Federation, Director of the Institute of Legislation and Comparative Law under the Government of the Russian Federation, Deputy President of the Russian Academy of Sciences, Academician-Secretary of the Department of Social Sciences of the Russian Academy of Sciences.

E-mail: [office@izak.ru](mailto:office@izak.ru)

This article analyzes the process of socializing artificial intelligence technology and the formation of a legal framework governing its implementation. It examines regulatory efforts at the national, supranational, and international levels, identifying ideological, doctrinal, and political challenges that accompany the expanding use of this technology. The article provides a critical review of the provisions within the Global Digital Compact and outlines the associated legal and political risks of its implementation. Drawing on philosophical research, it identifies prospects for building the institutional and legal mechanisms required for the full-fledged functioning of AI on a global scale.

The importance of the principles of equality and justice is emphasized for maintaining a sustainable legal order, upholding human rights and freedoms, reducing the technological divide, and developing a constructive strategy for the advancement of a technological civilization is emphasized.

Using comparative legal methodology, the article formulates the main regulatory approaches emerging within states and interstate associations concerning the development and application of artificial intelligence.

**Keywords:** strategy for the development of a technogenic civilization, global legal order, digital technologies, sustainable development, principles of equality and justice, institutional and legal mechanisms for the functioning of artificial intelligence

## REFERENCES

1. Khabrieva T.Ya. Pravovye problemy identifikatsii iskusstvennogo intellekta // Vestnik RAN. 2024. T. 94, № 7. S. 609–622 (in Russian).
2. Khabrieva T.Ya. Pravovye ocherki tekhnologicheskoi revolyutsii. M.: Norma; Infra-M; Rossiiskaya akademiya nauk, 2025. S. 89–132 (in Russian).
3. Khabrieva T.Ya. Korruptsiya i pravoporyadok v fokuse sovremennoi yuridicheskoi doktriny // Zhurnal zarubezhnogo zakonodatel'stva i srovnitel'nogo pravovedeniya. 2016. № 4. S. 5–13 (in Russian).
4. Khabrieva T.Ya. Tsiklicheskie normativnye massivy v prave // Zhurnal rossiiskogo prava. 2019. № 12. S. 5–18 (in Russian).
5. Styopin V.S. Sovremennye tsivilizatsionnye krizisy i problema novykh strategii razvitiya. M.: Otdelenie obshchestvennykh nauk RAN, Sektsiya filosofii, politologii, sotsiologii i prava, 2018 (in Russian).
6. Zor'kin V.D. Spravedlivost' – imperativ tsivilizatsii prava: lektsiya, sostoyavshayasya v ramkakh Peterburgskogo mezhdunarodnogo yuridicheskogo foruma 17 maya 2018 g. // Voprosy filosofii. 2019. № 1. S. 5–14 (in Russian).
7. Zor'kin V.D. Spravedlivost' – imperativ tsivilizatsii prava // [Elektronnyi resurs]. Rezhim dostupa: <https://ksrf.ru/ru/News/Speech/Pages/ViewItem.aspx?ParamId=885> (in Russian).

Индекс УДК 330.341.1; 338.2

Код ГРНТИ 06.54.31

DOI: 10.22204/2587-8956-2025-123-04-25-33

И.В. ДАНИЛИН<sup>\*</sup>

## Экономические и политические вызовы развития искусственного интеллекта

Анализируются актуальное состояние и перспективы развития технологий искусственного интеллекта (ИИ), а также условия максимизации его эффектов для экономического роста. На примере использования ИИ в США, ЕС и КНР отмечены ограничения диффузии технологии и её влияния на социально-экономические показатели. В качестве ключевых факторов развития ИИ указаны формирование человеческого капитала и динамизм рынков. Однако, как отмечается, на данный момент реализация процессов шумпетеровского «созидающего разрушения» неоптимальна, что обусловлено несколькими факторами. В их числе, прежде всего, специфика развития самого ИИ, включая высокую капиталоёмкость технологии и её требования к большим данным. Существенную роль играет и структура цифровых рынков, где крупные цифровые корпорации монополизируют рынки ИИ. Сделан вывод о том, что рыночная ситуация требует вмешательства регуляторов. Тем не менее отмечается, что эволюция регулирования ИИ также сталкивается с целым рядом вызовов. Помимо новизны технологии и неочевидности подходов к решению основных проблем развития ИИ, налицо также рост секьюритизации ИИ и влияние иных политических и геополитических соображений. Сделан вывод о том, что данная проблема значима на перспективу, так как негативно влияет на выработку подходов к регулированию ИИ и создаёт новые вызовы развития, в том числе на глобальном уровне.

**Ключевые слова:** искусственный интеллект, экономический рост, развитие, регулирование, созидающее разрушение, регулирование, политика, вызовы

**И**искусственный интеллект признаётся одним из ключевых факторов глобального и национального развития на перспективу. Хотя первые применения ИИ относятся ещё к 1990-м гг., лишь в 2010-х гг. благодаря новым техноло-

гиям начался быстрый рост использования негенеративного ИИ [1–9]. Он использовался для автоматизации технологических процессов, анализа и обработки больших массивов данных в финансовом секторе, маркетинге, медицине и в иных областях.

\* **Данилин Иван Владимирович** — кандидат политических наук, заместитель директора по научной работе Национального исследовательского института мировой экономики и международных отношений им. Е.М. Примакова РАН.  
E-mail: danilin.iv@imemo.ru



Ил. 1. Доля предприятий США, указывающих на использование ИИ, с разделением по секторам экономики (2022–2023 гг., топ-10 секторов, в %).

Источник данных: [10]

Шла также автоматизация нетехнологических бизнес-процессов, включая клиентскую поддержку. Генеративный ИИ, взрывное развитие которого началось в 2022 г., придал новый импульс использованию технологии в этих областях, а также сформировал новые применения на стороне конечных потребителей, в индустрии впечатлений и в иных сегментах рынка [4, 6, 7, 10]. Как следствие, растут инвестиции, меняются корпоративные и государственные стратегии.

Неудивительно, что активизировались и дискуссии о перспективах и вызовах ИИ в экономике, политике и социальной сфере. В данной работе с учётом реалий применения технологии кратко проанализированы некоторые важнейшие социально-экономические проблемы, связанные с развитием ИИ, а также возникающие экономические и политические вызовы.

### Материалы и методы исследования

Исследование основывается на трёх блоках источников и литературы. Прежде всего, это статистика использования

ИИ в экономике развитых стран и КНР. Далее это рыночные, экспертные и академические оценки и исследования использования ИИ. Наконец, это научная литература, посвящённая изучению и моделированию воздействия ИИ на мировую экономику, а также регулированию ИИ и взаимосвязанным вызовам. Методологически работа основывается на совокупности подходов экономической науки и компартивном анализе.

### Реалии использования ИИ в развитых странах и КНР

Важным фактором оценки влияния ИИ на социально-экономическую сферу являются масштабы его использования. Несмотря на ряд сверхоптимистических оценок, возросшая частота и масштаб применения ИИ в разных секторах пока не обеспечивают его широкое распространение среди хозяйствующих субъектов (ил. 1). Данные по США, Канаде, странам ЕС за 2022–2024 гг. показывают быстрый рост числа предприятий, использующих ИИ. Но оно всё ещё невысоко: от 7,6% в США до 13,5% в ЕС<sup>1</sup>, и неравномерно по разным секторам экономики [5, 10, 11].

Важна и поляризации использования ИИ [5, 10]. С одной стороны, лидерами выступают крупные корпорации (свыше 10 тыс. занятых; 40–50% всех фирм, использующих ИИ), что определяет широкую частотность применения ИИ, особенно в автоматизации простых (например, диалог с клиентами) либо, напротив, сложных процессов (разработка программного обеспечения (ПО) и чипов, робототехника и пр.). Среди причин – готовность инвестировать значительные средства в новые технологии и огромный опыт автоматизации и цифровизации. С другой – возрастаёт частотность использования ИИ

<sup>1</sup> Данные по ЕС без учёта микропредприятий (менее 10 чел.). С их учётом значения должны быть в два и более раз ниже.

малыми компаниями — очевидно, стартапами и фирмами, основанными молодыми предпринимателями, уже знакомыми с ИИ.

Специфика использования ИИ определяет спорные оценки его влияния на экономический рост. За пределами выраженного влияния на услуги, бизнес-процессы и капитализацию отдельных групп предприятий говорить о мощном влиянии ИИ на экономику преждевременно. Опуская вопрос о корректности расчётов и исходных данных, для наиболее развитых стран налицо ненулевые эффекты, но и прорыва по ключевым показателям мы не наблюдаем (см., например, исследования по ЕС [12, 13]). Более того, данные по динамике роста производительности труда и совокупной факторной производительности в США, ЕС и КНР позволяют говорить о новом «парадоксе Солоу» [1].

Ещё более проблематична оценка экономических перспектив ИИ. Консалтинговые компании и ряд исследователей говорят о мощном ускорении темпов роста в будущем. Но многие экономисты указывают на естественные ограничения для этого сценария — от сокращения совокупного спроса вследствие технологической безработицы в случае «прорыва» в сфере ИИ и до исчерпаемости природных ресурсов [1, 9, 14]. По оценкам нобелевского лауреата Д. Аддемоглу, даже в случае быстрого роста инвестиций в ИИ технология способна в следующие 10 лет обеспечить дополнительный рост ВВП на уровне 1,6–1,8% [9]. Важнейшим фактором «прорыва» является усиление креативных и интеллектуальных функций ИИ как условия создания новых знаний, технологий и бизнес-моделей в логике шумпетеровского «созидающего разрушения» [1, 3, 4, 6, 9, 14]. Но реальность этого сценария и его временные горизонты обсуждать невозможно, как и продолжительность и проблемы «переходного периода».

Неоднозначно и влияние ИИ на рынок труда. Вопреки ожиданиям, пока ИИ лишь выборочно влияет на занятость, причём в основном на средне- и высококвалифицированных рабочих местах [1, 4, 6, 9]. Зато

в силу несовершенства технологии и ранней стадии её внедрения, за пределами ПО налицо акцент на переобучение кадров, а не на их сокращение. Как и в случае с иными «прорывными» технологиями, фиксируется также нехватка персонала с навыками в сфере ИИ и поддерживающих компетенций [5, 10, 11]. Но и после ускорения диффузии и развития ИИ массовая технологическая безработица сомнительна. Помимо «эффекта Баумоля» и появления новых профессий, значимым фактором останется поддержка внедрения и эксплуатации ИИ [4].

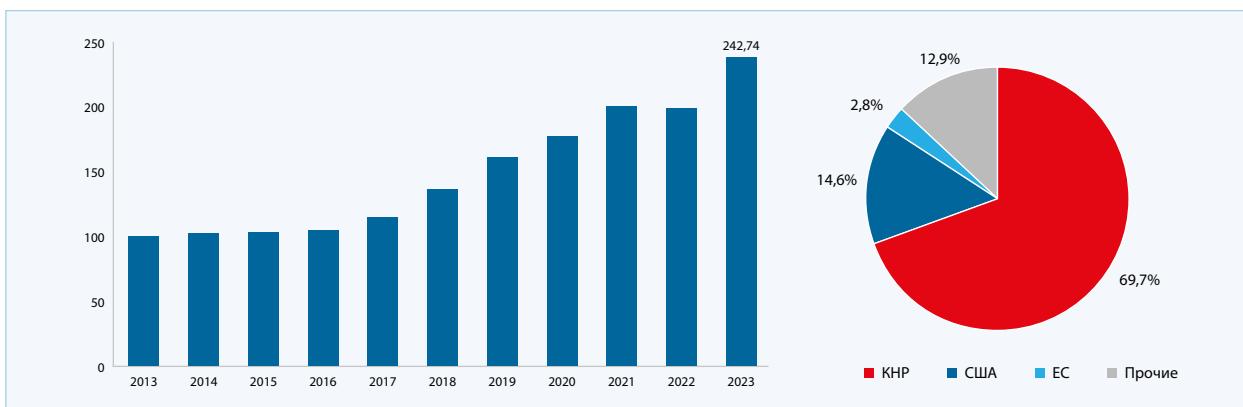
Вкупе с иными вопросами эти соображения приводят нас к основному вызову развития ИИ: а именно, как может быть обеспечен интенсивный рост на основе ИИ при минимизации негативных социально-экономических эффектов?

### Вызов максимизации положительных эффектов ИИ для экономики

Вопрос об актуализации фактора ИИ в социально-экономическом развитии можно было бы свести к накоплению научно-технологического потенциала. Опустим создание центров обработки данных (ЦОД) и иные технико-технологические вопросы, имеющие огромное, но всё же среднесрочное значение. Ключевыми факторами формирования потенциала следует считать НИОКР и венчурные инвестиции.

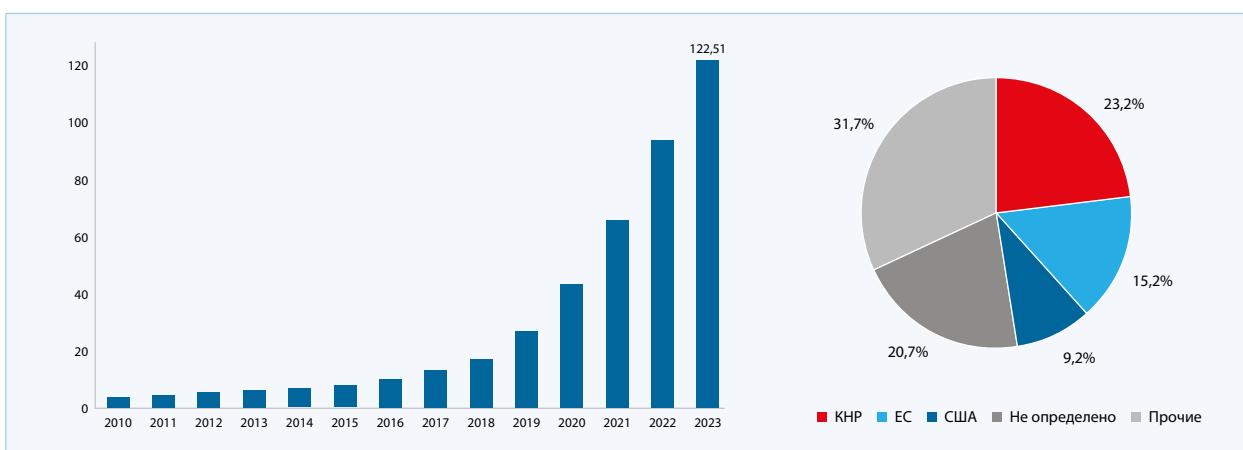
За отдельными исключениями вычленить затраты именно на ИИ в НИОКР не представляется возможным. Но косвенные показатели говорят об устойчивом увеличении научно-технических работ государственного и бизнес-секторов [4, 15]. По оценкам Стенфордского университета, число научных публикаций по ИИ в рамках категории «компьютерные науки» выросло со 102 до более 242 тыс. в 2013–2023 гг. (ил. 2). То же можно сказать о патентах в сфере ИИ (с 3,8 до 122,5 тыс. в 2010–2023 гг.) [15] (ил. 3).

С середины 2010-х гг. фиксируется также взрывной рост венчурных инвестиций: с 2,6 млрд долл. в 2012 г. до более чем 100 млрд в 2024 г. [4, 16] (ил. 4). Не считая вероятных флуктуаций в ресурсном обеспе-



Ил. 2. Публикации, связанные с ИИ, в тыс. ед. (по категории «Компьютерные науки», на основе каталога OpenAlex, 2013–2023 гг.) и становая принадлежность авторов, в % (2023 г.).

Источник данных: [15]



Ил. 3. Патенты в сфере ИИ, в тыс. ед. (2010–2023 гг.), и становая принадлежность патентообладателей, в % (2023 г.).

Источник данных: [15]

чении, поддержка НИОКР и венчурных инвестиций в ИИ представляется устойчивой.

Проблема в том, что формирование потенциала не эквивалентно экономическому развитию. В этом смысле нас, прежде всего, интересуют условия для диффузии технологий в экономике и «созидательного разрушения». Здесь ключевыми оказываются два процесса.

Во-первых, это развитие человеческого капитала за пределами задачи воспроизведения научно-технических и управленческих кадров, которую успешно решают вузы. Формирование компетенций, навыков, кодифицированных и неявных знаний на индивидуальном, организационном и иных уровнях требует рыночных

взаимодействий, работы инновационных экосистем, профессиональных сообществ, групп энтузиастов и иных институтов. Реализация этих процессов, очевидно, сложна, а её управляемость проблематична.

В этом отношении проблема компетенций смыкается со вторым важнейшим процессом – созданием условий для формирования новых предприятий, бизнес-моделей и институтов, а также связей между различными субъектами. Именно здесь фиксируется ключевой вызов, а именно взаимообусловленные процессы монополизации рынка и концентрации технологического потенциала.

Повторимся, лидерами ИИ остаются наиболее крупные компании, прежде всего циф-

ровые корпорации платформенного типа из США и КНР [4, 3, 6, 8, 17, 18]. Конечно, существуют независимые ИИ-стартапы Deep Seek и OpenAI. Но и они через систему финансовых и/или технологических отношений всё глубже интегрируются с «мейджорами» индустрии и/или сами готовы стать монополиями.

Цифровые монополии доминируют в сфере НИОКР и венчурных инвестиций, контролируют различные формы технологического капитала, например облачные мощности, большие данные, интеллектуальную собственность [1, 4, 6, 18].

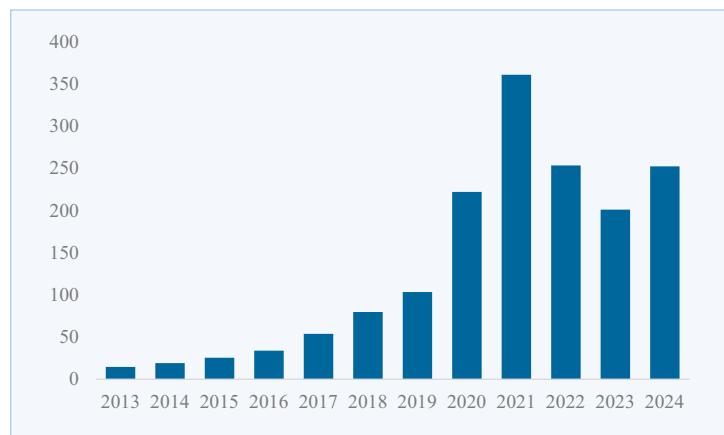
Всё это создаёт реальные риски подавления конкуренции со стороны молодых компаний и появления барьеров на пути диффузии технологий [1, 4, 7, 18].

Причиной является не только нечестная конкуренция или мощь «мейджоров». ИИ – капиталоёмкая технология, которая, например, требует огромного объёма облачных услуг и инфраструктуры для обучения языковых моделей. Провоцируя рост концентрации капитала, развитие ИИ таким образом может привести к рискам для развития экономического [1].

Хотя самопроизвольное исправление возникших асимметрий за счёт рыночных процессов и научно-технологических достижений возможно, на практике игнорирование рисков неприемлемо. А это определяет необходимость активного вмешательства регуляторов.

### Политическая экономия ИИ: от регулирования к вызовам политизации

Систематические усилия в сфере регулирования технологий и рынков ИИ наблюдаются с середины 2010-х гг. Лидерами – предсказуемо – выступали США, КНР и ЕС. Наблюдалась эволюция фокуса госполитики. На первых этапах основное внимание уделялось поддержке научно-технологической базы ИИ при определении общих принципов регулирования



Ил. 4. Корпоративные портфельные инвестиции в ИИ (в том числе венчурные), в млрд долл. США.

Источник данных: [15]

на перспективу. Таковы, например, Трёхлетний план действий в области ИИ (2016 г.) и План развития ИИ нового поколения (2017 г.) в КНР, Национальная инициатива США в сфере ИИ 2019 г., европейский Координационный план по ИИ 2021 г. и мероприятия по ИИ в рамочных программах ЕС по науке, технологиям и инновациям «Горизонт 2020» (2014–2020 гг.) и «Горизонт Европы» (2021–2027 гг.).

Одновременно выстраивалась политика в сфере стандартизации, кибербезопасности и т.п. После 2020 г. стали уточняться параметры и подходы госполитики в отношении ИИ, включая вопросы рыночного регулирования и снижения социетальных рисков. В данном отношении можно указать на Европейский Закон по ИИ 2023 г. (первый систематический нормативно-правовой акт в сфере ИИ); американскую Стратегию по ИИ 2020 г. и исполнительные приказы президентов США от 2023 и 2025 гг.; Промежуточные меры управления услугами генеративного ИИ (2024 г.) и иные документы КНР, содержащие высокоуровневые требования [18].

Проблема, однако, в том, что скорость этих процессов отстает от эволюции технологий и рынков. При этом в силу новизны технологий и связанных с ней вызовов пока что госполитика лишь частично может дать ответы на фундаментальный вызов максимизации социально-

экономических эффектов ИИ. За пределами стандартных мер, связанных с развитием технологий, кадрового потенциала и формальных институтов национальных инновационных систем, мы видим фрагментарную и противоречивую картину.

Отметим дилемму монополизма ИИ. Так, США и КНР, казалось бы, создают оптимальные условия для «созидательного разрушения» [6]. В отличие от ЕС и Японии, в США доминирует прецедентное право и апостериорный принцип регулирования технологических рынков (коррекция рынка по итогам накопленного опыта использования). В КНР центральное и региональные правительства де-факто обеспечивают режим «регуляторных песочниц» для предприятий в сфере ИИ. Обе стороны прямо и косвенно поддерживают крупные компании-лидеры и стартапы, делают мощные вливания в НИОКР.

Но с учётом специфики ИИ подобная идеология пока провоцирует, опять же, рост концентрации капитала в том числе и на венчурном рынке (укрупнение стартапов и сделок). Правительства США, КНР, Великобритании, руководство Еврокомиссии, правда, инициировали расследования в отношении отдельных крупных сделок в сфере ИИ (например, между Microsoft и OpenAI), разрабатывают высокоуровневые требования к новой антимонопольной политике [6, 18].

Но пока даже эти подходы находятся в зачаточном состоянии — тем более, что традиционные антитрестовские методы могут подорвать потенциал сектора ИИ. Более актуальны стимулы к «созидательному разрушению», включая едва ли не ключевой вопрос либерализации рынка данных, контроль над которыми сохраняют цифровые компании-гиганты.

Помимо новизны и сложности, вызов заключается ещё и в том, что регулирование ИИ в контексте социально-экономического развития не является чисто отраслевым вопросом и не может быть реализовано технократическими методами или в рамках «островного» подхо-

да (т.е. без учёта иных секторов и проблем). ИИ влияет на фундаментальные экономические, общественные и даже культурные нормы, процессы и институты. Иными словами, и само развитие ИИ и рынков, и меры госполитики требуют учёта большого числа факторов — от баланса групп интересов до вопросов делегирования ответственности нечеловеческим субъектам и, шире, роли Человека в экономике.

Подобная комплексность тормозит выработку оптимальных мер и к тому же провоцирует политизацию, в том числе секьюритизацию дискуссий по ИИ и управлению развитием. Это касается структурирования внутриэлитного консенсуса, определения границ и глубины вмешательства государства и вплоть до актуализации геополитических рисков [19]. Последнее значимо ещё и потому, что восприятие ИИ как технологии, гарантирующей глобальное лидерство определяет взгляд на развитие ИИ как на игру с нулевой суммой. Что приводит к реализации экспансионистской (захват рынков, переманивание кадров и т.п.) и техносуверенной/протекционистской политики, препятствуя выработке единых подходов к его регулированию. И хотя возможности многостороннего сотрудничества обсуждаются [6], реальный диалог выстраивается скорее в логике формирования альянсов или блоков. Иллюстративен американо-европейский диалог по ИИ, де-факто направленный на сдерживание КНР, но отягчённый взаимными противоречиями. То же можно сказать про «Цифровой Шёлковый путь» и иные инициативы самого Китая, направленные на доминирование на внешних рынках. Как следствие, теряются важные синергии, связанные с международной кооперацией, разделением труда, снижением издержек, тогда как борьба за ресурсы в сфере ИИ становится более острой. Это ведет к увеличению разрыва между потенциалами разных стран в сфере ИИ и иным процессам, способным оказать негативное влияние на мировое развитие.

## Заключение

Несмотря на высокие ожидания, технологии ИИ на данный момент относительно слабо конвертируются в экономический рост и качество развития, как и в решение актуальных социальных задач. При этом социально-экономические риски развития технологии остаются значимыми. Преодоление существующих ограничений и минимизация негативных эффектов ИИ требуют решения целого ряда фундаментальных задач, среди которых поддержание высокой динамики рынка («созидательное разрушение») с учётом всей полноты вызовов. Это тем более важно, что в «нормальных» условиях рынок ИИ тяготеет к монополизации, а такие вопросы, как выборочный рост производительности труда и «эффект Баумоля», ведут к возникновению асимметрий развития. Как следствие, возрастает значение госрегулирования. Но с учётом

ранней стадии развития технологий и рынков ИИ и неопределённости относительно способов решения возникающих проблем на данный момент государственные усилия явно недостаточны.

Другой политический вызов, связан с комплексностью эффектов ИИ и оценкой искусственного интеллекта как ключевого условия мирового лидерства. С одной стороны, это ведёт к секьюритизации и «политизации» процесса принятия и исполнения решений, что явно не способствует выработке оптимальных подходов. С другой – к росту geopolитической составляющей, которая негативно влияет на мировое развитие, включая потенциал стран – лидеров ИИ-гонки. Ответ на эти вызовы в огромной мере определят как перспективы развития ИИ, так и – что самое главное – реализацию его социально-экономических эффектов.

## ЛИТЕРАТУРА

1. The Economics of Artificial Intelligence: An Agenda / Eds.: A. Agrawal, J. Gans, A. Goldfarb. Chicago: University of Chicago Press, 2019.
2. Ruiz-Real J.L., Torres J.A., Uribe-Toril J., De Pablo J. // Journal of Business Economics and Management, 2021. № (22) 1. Pp. 98–117. DOI: 10.3846/jbem.2020.13641.
3. Gondauri D., Batiashvili M. // International Journal of Innovative Science and Research Technology. 2023. № 8 (4). Pp. 2456–2165. DOI: 10.48550/arXiv.2505.11989.
4. Filippucci F., Gal P., Jona-Lasinio C., Leandro A., Nicoletti G. The Impact Of Artificial Intelligence On Productivity, Distribution And Growth Key Mechanisms, Initial Evidence and Policy Challenges // OECD Artificial Intelligence Papers, 2024. URL: [https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/04/the-impact-of-artificial-intelligence-on-productivity-distribution-and-growth\\_d54e2842/8d900037-en.pdf](https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/04/the-impact-of-artificial-intelligence-on-productivity-distribution-and-growth_d54e2842/8d900037-en.pdf) (дата обращения: 31.08.2025).
5. Bonney K., Breaux C., Buffington C., Dinlersoz E., Foster L., Goldschlag N., Haltiwanger J., Kroff Z., Savage K. Tracking Firm Use of AI in Real Time: A Snapshot from the Business Trends and Outlook Survey // U.S. Census Bureau, 2024, CES 24–16. URL: <https://www.census.gov/hfp/btos/downloads/CES-WP-24-16.pdf> (дата обращения: 31.08.2025).
6. Comunale M., Manera A. The Economic Impacts and the Regulation of AI: A Review of the Academic Literature and Policy Actions // International Monetary Fund Working Paper, 2024, WP/24/65. URL: <https://www.imf.org/-/media/Files/Publications/WP/2024/English/wpiea2024065-print-pdf.ashx> (дата обращения: 31.08.2025).
7. McElheran K., Li J.F., Brynjolfsson E., Kroff Z., Dinlersoz E., Foster L.S., Zolas N. // Journal of Economics and Measurement Strategy. 2024. № 33 (2). Pp. 375–415. DOI: 10.1111/jems.12576.
8. Anderson G., Kindlon A., Rhodes A. Artificial Intelligence in the Business Sector: R&D, Use, and Impact on Employees // U.S. National Science Foundation. 2024. NSF 25–312. URL: [https://ncses.nsf.gov/pubs/nsf25312/assets/artificial-intelligence-in-the-business-sector-r-d-use-and-impact-on-employees/nsf25312\\_report.pdf](https://ncses.nsf.gov/pubs/nsf25312/assets/artificial-intelligence-in-the-business-sector-r-d-use-and-impact-on-employees/nsf25312_report.pdf) (дата обращения: 31.08.2025).
9. Acemoglu D. // The Economic Policy. 2025. № 40 (121). Pp. 13–58. DOI: 10.1093/epolic/eiae042.

10. Smith M. What Industries Are Using AI? Current Use and Future Expectations. North Carolina Department of Commerce, 2024. URL: [nshttps://www.commerce.nc.gov/news/the-lead-feed/what-industries-are-using-ai](https://www.commerce.nc.gov/news/the-lead-feed/what-industries-are-using-ai) (дата обращения: 31.08.2025).
11. State of the EU digital transformation in 2025: progress and horizontal recommendations. Annex 2 to the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, European Commission, 2025, COM (2025). URL: <https://digital-strategy.ec.europa.eu/en/library/state-digital-decade-2025-report> (дата обращения: 31.08.2025).
12. Kalai M., Becha H., Helal K. // Journal of Economic Structures. 2024. № 13 (22). DOI: 10.1186/s40008-024-00345-y.
13. Socol A., Marin-Pantelescu A., Attila T.-S., Cioca I.C. // Amfiteatru Economic Journal. 2024. № 26 (66). Pp. 475–493. DOI: 10.24818/EA/2024/66/47.
14. Капелюшников Р.И. // Вопросы экономики. 2024. № 5. С. 5–45. DOI: 10.32609/0042-8736-2025-5-5-45.
15. The 2025 AI Index Report. Stanford University, 2025. URL: <https://hai.stanford.edu/ai-index/2025-ai-index-report> (дата обращения: 31.08.2025).
16. Venture Capital Investments In Artificial Intelligence // OECD Digital Economy Papers, 2021. URL: [https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/09/venture-capital-investments-in-artificial-intelligence\\_dea0a42f/f97beae7-en.pdf](https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/09/venture-capital-investments-in-artificial-intelligence_dea0a42f/f97beae7-en.pdf) (дата обращения: 31.08.2025).
17. Lundvall B.-Å., Rikap C. // Research Policy. 2022. № 51 (1). DOI: 10.1016/j.respol.2021.104395.
18. Zheng K. // Computer Law & Security Review, 2025, 59. DOI: 10.1016/j.clsr.2025.106211.
19. Schmid S., Lambach D., Diehl C., Reuter C. // Geopolitics. 2025. № 30 (4). Pp. 1907–1936. DOI: 10.1080/14650045.2025.2456019.

## Economic and Political Challenges of Artificial Intelligence Development

**Ivan Vladimirovich Danilin** — Candidate of Sciences in Political Science, Deputy Director for Research at the Primakov National Research Institute of World Economy and International Relations of the Russian Academy of Sciences.

E-mail: [danilin.iv@imemo.ru](mailto:danilin.iv@imemo.ru)

The article analyzes the current state and prospects for the development of artificial intelligence (AI) technologies, as well as the conditions for maximizing their effects on economic growth. The article uses the example of AI use in the United States, the European Union, and China to highlight the limitations of technology diffusion and its impact on socio-economic indicators. The formation of human capital and dynamic markets are identified as key factors in AI development. However, the article notes that the implementation of Schumpeter's "creative destruction" processes is currently suboptimal due to several factors. These include the specific nature of AI development, including its high capital intensity and requirements for large amounts of data. The structure of digital markets also plays a significant role, with large digital corporations monopolizing AI markets. The conclusion is that the current market situation requires regulatory intervention. Nevertheless, the evolution of AI regulation itself faces numerous challenges. In addition to the novelty of the technology and the lack of clear solutions to its core developmental issues, there is a growing trend of AI securitization and the influence of other political and geopolitical considerations. It is concluded that this problem is significant in the long term, as it negatively affects the development of approaches to AI regulation and creates new developmental challenges, including at the global level.

**Keywords:** artificial intelligence, economic growth, development, regulation, creative destruction, policy, challenges

## REFERENCES

1. The Economics of Artificial Intelligence: An Agenda / Eds.: A. Agrawal, J. Gans, A. Goldfarb. Chicago: University of Chicago Press, 2019.
2. Ruiz- Real J.L., Torres J.A., Uribe-Toril J., De Pablo J. // Journal of Business Economics and Management. 2021. № (22) 1. Pp. 98–117. DOI: 10.3846/jbem.2020.13641.
3. Gondauri D., Batiashvili M. // International Journal of Innovative Science and Research Technology. 2023. № 8 (4). Pp. 2456–2165. DOI: 10.48550/arXiv.2505.11989.
4. Filippucci F., Gal P., Jona-Lasinio C., Leandro A., Nicoletti G. The Impact Of Artificial Intelligence On Productivity, Distribution And Growth Key Mechanisms, Initial Evidence and Policy Challenges // OECD Artificial Intelligence Papers, 2024. URL: [https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/04/the-impact-of-artificial-intelligence-on-productivity-distributionand-growth\\_d54e2842/8d900037-en.pdf](https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/04/the-impact-of-artificial-intelligence-on-productivity-distributionand-growth_d54e2842/8d900037-en.pdf) (data obrashcheniya: 31.08.2025).
5. Bonney K., Breaux C., Buffington C., Dinlersoz E., Foster L., Goldschlag N., Haltiwanger J., Kroff Z., Savage K. Tracking Firm Use of AI in Real Time: A Snapshot from the Business Trends and Outlook Survey // U.S. Census Bureau, 2024, CES 24–16. URL: <https://www.census.gov/hfp/btos/downloads/CES-WP-24-16.pdf> (data obrashcheniya: 31.08.2025).
6. Comunale M., Manera A. The Economic Impacts and the Regulation of AI: A Review of the Academic Literature and Policy Actions // International Monetary Fund Working Paper, 2024, WP/24/65. URL: <https://www.imf.org/-/media/Files/Publications/WP/2024/English/wpiea2024065-print-pdf.ashx> (data obrashcheniya: 31.08.2025).
7. McElheran K., Li J.F., Brynjolfsson E., Kroff Z., Dinlersoz E., Foster L.S., Zolas N. // Journal of Economics and Measurement Strategy. 2024. № 33 (2). Pp. 375–415. DOI: 10.1111/jems.12576.
8. Anderson G., Kindlon A., Rhodes A. Artificial Intelligence in the Business Sector: R&D, Use, and Impact on Employees // U.S. National Science Foundation. 2024. NSF 25–312, URL: [https://ncses.nsf.gov/pubs/nsf25312/assets/artificial-intelligence-in-the-business-sector-r-d-use-and-impacton-employees/nsf25312\\_report.pdf](https://ncses.nsf.gov/pubs/nsf25312/assets/artificial-intelligence-in-the-business-sector-r-d-use-and-impacton-employees/nsf25312_report.pdf) (data obrashcheniya: 31.08.2025).
9. Acemoglu D. // The Economic Policy. 2025. № 40 (121). Pp. 13–58. DOI: 10.1093/epolic/eiae042.
10. Smith M. What Industries Are Using AI? Current Use and Future Expectations. North Carolina Department of Commerce, 2024. URL: [nshttps://www.commerce.nc.gov/news/the-lead-feed/what-industries-are-using-ai](https://www.commerce.nc.gov/news/the-lead-feed/what-industries-are-using-ai) (data obrashcheniya: 31.08.2025).
11. State of the EU digital transformation in 2025: progress and horizontal recommendations. Annex 2 to the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, European Commission, 2025, COM (2025). URL: <https://digital-strategy.ec.europa.eu/en/library/state-digital-decade-2025-report> (data obrashcheniya: 31.08.2025).
12. Kalai M., Becha H., Helal K. // Journal of Economic Structures. 2024. № 13 (22). DOI: 10.1186/s40008-024-00345-y.
13. Socol A., Marin-Pantelescu A., Attila T.-S., Cioca I.C. // Amfiteatru Economic Journal. 2024. № 26 (66). Pp. 475–493. DOI: 10.24818/EA/2024/66/47.
14. Kapelyushnikov R.I. // Voprosy ekonomiki. 2024. № 5. С. 5–45. DOI: 10.32609/0042-8736-2025-5-5-45 (in Russian).
15. The 2025 AI Index Report. Stanford University, 2025. URL: <https://hai.stanford.edu/ai-index/2025-ai-index-report> (data obrashcheniya: 31.08.2025).
16. Venture Capital Investments In Artificial Intelligence // OECD Digital Economy Papers, 2021. URL: [https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/09/venture-capitalinvestments-in-artificial-intelligence\\_dea0a42f/f97beae7-en.pdf](https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/09/venture-capitalinvestments-in-artificial-intelligence_dea0a42f/f97beae7-en.pdf) (data obrashcheniya: 31.08.2025).
17. Lundvall B.-Å., Rikap C. // Research Policy. 2022. № 51 (1). DOI: 10.1016/j.respol.2021.104395.
18. Zheng K. // Computer Law & Security Review, 2025, 59. DOI: 10.1016/j.clsr.2025.106211.
19. Schmid S., Lambach D., Diehl C., Reuter C. // Geopolitics. 2025. № 30 (4). Pp. 1907–1936. DOI: 10.1080/14650045.2025.2456019.



**В.К. ЛЕВАШОВ,  
О.В. ГРЕБНЯК\***

## На пороге цифровой цивилизации: диалектика искусственного интеллекта и гуманизма\*\*

В статье рассматриваются проблемы и результаты исследования диалектики отношений цифровых технологий и гуманистических принципов и ценностей в контексте формирования глобальной информационной цивилизации. Авторы опираются на анализ современных теорий и практик цифрового общества и концепцию ноосферы В.И. Вернадского, которая позволяет осмысливать переход к цифровой цивилизации как этап эволюции биосферы, где человеческая мысль играет ключевую роль. В качестве эмпирической базы использованы данные всероссийского социологического мониторинга «Как живёшь, Россия?» (ИСПИ ФНИСЦ РАН), отражающие отдельные опасения граждан в связи с повсеместным внедрением искусственного интеллекта, а также парадокс доверия к распространяемой в Интернете информации. Обозначены ключевые вызовы, связанные с ростом зависимости от искусственного интеллекта, снижением качества межличностного взаимодействия, концентрацией власти в руках технологических корпораций и кризисом доверия к информации. Особое внимание уделено противоречивым последствиям цифровизации. Информационно-коммуникационные технологии (ИКТ), нейросети и искусственный интеллект открывают возможности для решения глобальных проблем экологии, управления ресурсами, в то же время они усиливают социальные, материальные и духовные диспаритеты и информационную фрагментацию. Авторы приходят к выводу о необходимости соблюдения этических норм и нравственных гуманистических императивов технологического прогресса, социальной ответственности, социокультурных и этнополитических особенностей.

**Ключевые слова:** цифровая цивилизация, ноосфера, гуманизация технологий, искусственный интеллект, доверие к информации, глобальная информационная безопасность, социальные риски технологий, гуманистическая парадигма, гуманитарная экспертиза

\* **Левашов Виктор Константинович** — доктор социологических наук, директор Института социально-политических исследований (ИСПИ) ФНИСЦ РАН.

E-mail: levachov@mail.ru

**Гребняк Оксана Валерьевна** — научный сотрудник ИСПИ ФНИСЦ РАН.

E-mail: oksananov@yandex.ru

\*\* Исследование выполнено за счёт гранта Российского научного фонда № 23-18-00438 «Резервы устойчивого социально-политического развития российского общества и государства в контексте укрепления традиционных ценностей и глобальной многополярности». <https://rsrf.ru/project/23-18-00438>.

**Г**лобализация как процесс формирования единого взаимозависимого в своей устойчивой жизнедеятельности планетарного социума идёт с момента зарождения первых локальных обществ и рынков. В основе этого перманентного процесса культурного освоения мира лежит развитие производительных сил цивилизации *Homo sapiens*, создающей средства производства материальных и духовных благ. Человечество непрерывно увеличивает масштаб своих как созидательных, так и разрушительных воздействий на биосферу. XX в. стал рубежным в развитии мировой цивилизации: мировые войны — одна из которых с применением ядерного оружия — заставили переосмыслить допустимые границы конкурентной преобразовательной деятельности. Однако на пороге XXI в. в ходе третьей — цифровой и четвёртой — киберфизической промышленных революций в той же конкурентной парадигме продолжилось формирование глобальной цифровой цивилизации. На наших глазах традиционные социальные институты и механизмы взаимодействия трансформируются в социальную систему, где ключевую роль играют информационно-коммуникационные технологии, искусственный интеллект (ИИ), большие данные и интернет вещей. Цифровизация не только меняет экономические, политические и социокультурные процессы, но и выдвигает новые вызовы в понимании динамики глобального и глобального развития.

### Введение в проблемную ситуацию

Стремительное развитие ИИ размывает гуманистическую сущность доминирующих практик антропоцентризма. Согласно заданным алгоритмам нейросети имитируют человеческий интеллект в образах, признаках и формах искусственной субъектности. Уже сейчас исследователи констатируют становление гибридного социального пространства: «Люди взаимодействуют с изделиями со встроенными алгоритмами искусственного интеллекта

(техносубъектами), реализующими некоторые когнитивные функции человека, которые, не являясь субъектами управления, могут оказывать влияние на поведение социальных акторов, обусловливая проявление элементов субъектности в изменении социального порядка» [1, с. 11].

Парадоксально, что в эпоху фетишизации цифровой информации как модного товара, производимого на платформах современных технологий, наблюдается глобальный кризис доверия к ней. Массовая информация, продуцируемая и распространяется СМИ, социальными сетями и ИИ, всё чаще воспринимается как инструмент манипуляции. Всё большее число граждан связывают ухудшение качества жизни с низкой социальной и политической эффективностью управления и усиливающейся манипуляцией информацией со стороны правительства, международных организаций, бизнеса, гражданских организаций, криминала [2]. В этих условиях критически важно минимизировать «белые пятна» спекулятивного знания и недостоверной информации.

Информация как продукт духовно-интеллектуальной сферы отражает основные интеллектуальные практики и технологии, включая науку, искусство и идеологии, с помощью которых человек издревле осваивает окружающий мир. Её массовое цифровое производство стало возможным в последние десятилетия в силу фундаментальных научных открытий, прорыва в компьютерных и коммуникационных технологиях. Однако мультидисциплинарность происхождения современной информации затрудняет понимание её природы, источников, истинных целей её владельцев и распространителей.

Важно подчеркнуть, что социально-политическая информация не возникает сама по себе. Опираясь языком социальных и поведенческих наук, можно сказать, что она есть продукт анализа и синтеза результатов исследований человеческих отношений, в своей основе — производственных отношений, которые, как и ты-

сячелетия назад, базируются на материальных интересах. Большинство людей стремится к выживанию, тогда как элита накапливает капиталы, используя информацию в качестве инструмента влияния, особого товара, который отражает и искажает социальные, политические и идеологические интересы людей. Столкновение этих противоречивых интересов порождает как достоверную, так и недостоверную, открытую и латентную, массовую и элитарную типы информации.

Существует ещё один особый вид информации – научная информация. Она является родовым, если можно так выражаться, продуктом научно-технического развития. Научная информация, несмотря на строгие процедуры верификации, также не застрахована от ошибок и намеренных мистификаций. В эпоху коммерциализации знания и соблазнов общества массового потребления, в условиях «фаустовского выбора» учёные могут подчинить истину рыночным интересам, что способствует усилению социального неравенства.

Таким образом, актуальность исследования обусловлена необходимостью разрешения противоречия между стремительным внедрением цифровых технологий и недостаточной устойчивостью их морально-этического, социального и гуманистического фундамента ценностей. Особую значимость эта проблема приобретает в контексте влияния ИИ на рынок труда, алгоритмизацию общественного мнения и утрату контроля над информационными потоками. Цель предстоящей работы на дороге к цифровой цивилизации исследователей, политиков, бизнеса и государства – выявить и укрепить механизмы гармонизации функционала ИИ и этических норм и императивов устойчивого развития. Для этого важно понять ключевые вызовы, связанные с дегуманизацией общества, и предложить механизмы сбалансированного гуманистического развития технологий, гармонизации тех-

нологического прогресса с принципами устойчивого социально-политического развития.

### Цифровая цивилизация и концепция ноосферы

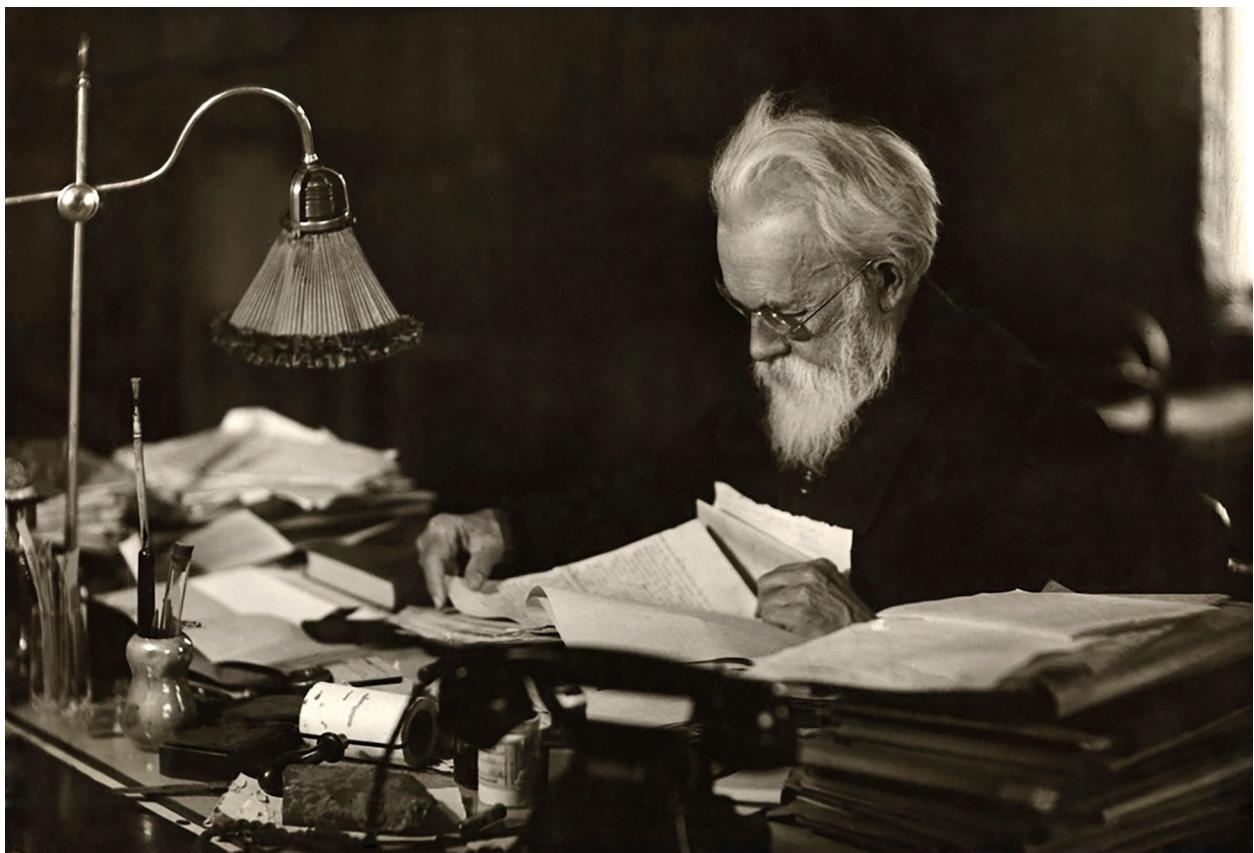
Упомянутые выше процессы глобальной цифровизации создают предпосылки для качественного скачка в эволюции человечества. Анализ современных тенденций обнаруживает параллели между идеей мировой цифровой цивилизации и концепцией ноосферы, сформулированной академиком Владимиром Ивановичем Вернадским<sup>1</sup> (ил. 1).

Теоретическое обоснование антропогенного преобразования природной среды он представил в цикле лекций по геохимии в Сорbonne (1922–1923), где изложил идею перехода биосфера в новое сущностное состояние. Термин «ноосфера» как новый этап её эволюции был предложен его слушателем, французским математиком и философом Э. Леруа в книге «Необходимость идеализма и факт эволюции» (1927).

Ноосфера, согласно концепции Вернадского, представляет собой этап эволюции биосфера, где человеческая мысль становится определяющим фактором развития планетарных процессов. Она не только отражает взаимодействие человечества с природой, но и открывает новые горизонты для осмысливания современных вызовов. Вернадский рассматривал ноосферу как результат синтеза человеческой деятельности и природных процессов, где наука, технологии и культура выступают движущими силами эволюции.

Современную цифровизацию общества и формирование глобальной цифровой инфраструктуры вполне можно интерпретировать как конкретное воплощение этих идей. Цифровая цивилизация формирует реальность, где данные, алгоритмы и ИИ становятся основой управления как человеческими, так и природными ресурсами. Это напрямую соответствует

<sup>1</sup> Подробнее о параллелях с доцифровым, информационным этапом развития ИКТ см. [3, 4].



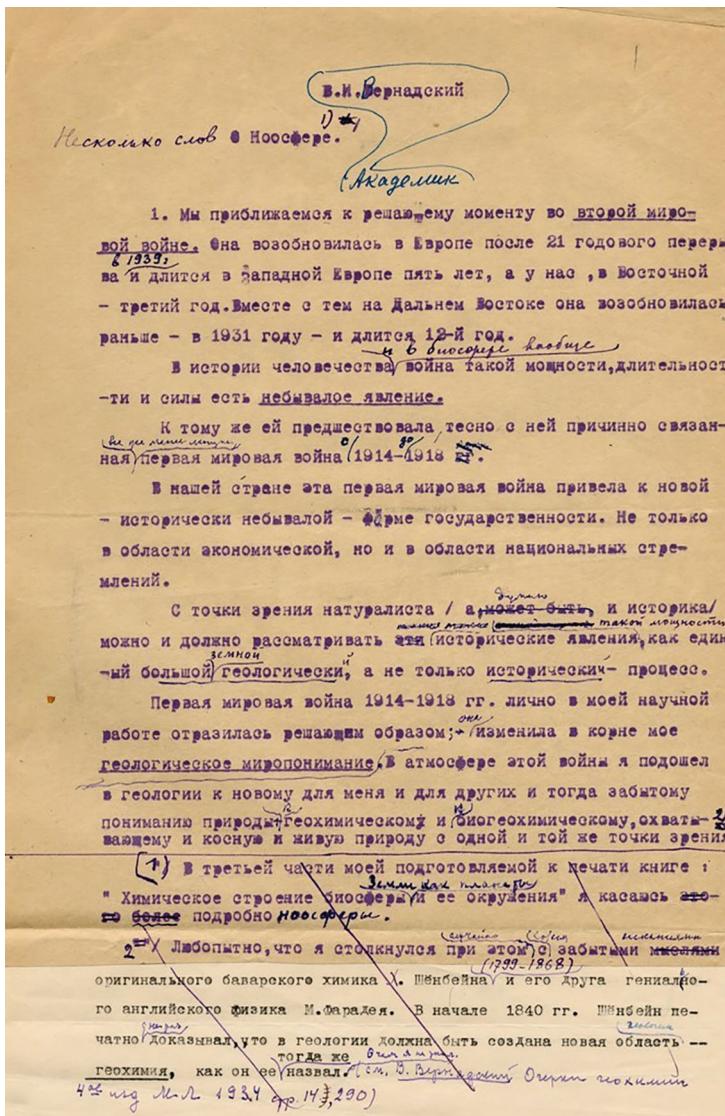
Ил. 1. Портрет академика Владимира Вернадского (Москва, 1930, фотограф Е. Элленгорн).

Источник: История России в фотографиях, Союз фотохудожников России

тезису В.И. Вернадского о доминирующей роли разума в преобразовании Земли. В.И. Вернадский писал: «Человечество, взятое в целом, становится мощной геологической силой. И перед ним, перед его мыслью и трудом, становится вопрос о перестройке биосферы в интересах свободно мыслящего человечества как единого целого» [5], подчёркивая ключевую роль разума и научного знания в эволюции биосферы (ил. 2). Таким образом, ноосфера трактуется как состояние биосферы, где человеческая мысль формирует окружающую среду.

Здесь возникает вопрос: какова диалектика этого процесса? Вернадский подчёркивал, что переход от биосферы к ноосфере возможен только при соблюдении гармонии между человеком и природой, а также при развитии этических принципов человечества. В цифровую эпоху это проявляется в противоречивости технологическо-

го прогресса. С одной стороны, цифровые технологии открывают беспрецедентные возможности для решения глобальных проблем: изменения климата, дефицита ресурсов, социального неравенства. С другой — порождают новые вызовы: информационные войны, угрозы приватности, экологический вред и пр. Диалектика возможностей и рисков ставит перед нами задачу ответственного использования достижений науки и технологий. Вернадский считал, что ноосфера требует не только научного, но и морального совершенствования человечества. В условиях цифровизации эта идея актуальна как никогда: ценности будущего должны стать предметом диалога между наукой, политикой, бизнесом, обществом и государством. К схожим выводам приходят и современные исследователи: «Планетарная разумная деятельность человека в идеале требует развития способностей синтеза, коллективного мышления,



Ил. 2. Оригинал первой страницы статьи В.И. Вернадского «Несколько слов о ноосфере» (июль 1943 – декабрь 1944 г.).

Источник: АРАН. Ф. 518. Оп. 1. Д. 157. Л. 1, 13

проникновенного интуитивного понимания и дорефлексивного распознавания добра и зла» [6, с. 74], пишет И.А. Герасимова, анализируя кризис техногенной цивилизации и пути выхода из него.

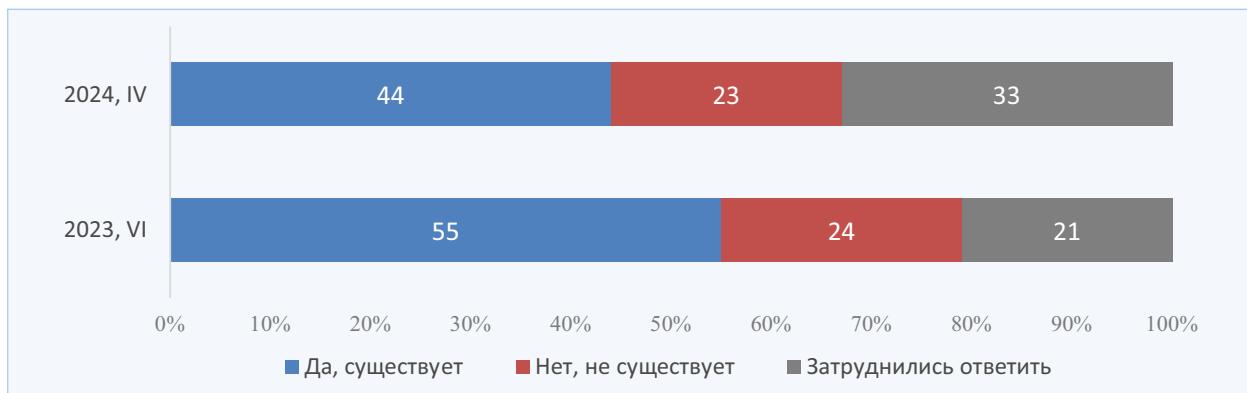
Коллективный характер также присущ и ноосфере, и становящейся цифровой цивилизации. Первая является продуктом совместной деятельности всего человечества; вторая также невозможна без глобального сотрудничества: открытых платформ, международных научных проектов и сетей коммуникаций. Вернадский видел в ноосфере потенциал для объединения

человечества, однако современная цифровая среда порой усиливает фрагментацию общества через информационные, культурные и идеологические барьеры, конкуренции и санкции. Это значит, что объективно возникла необходимость переосмысливания принципов цифрового взаимодействия в рамках ноосферного подхода.

## Вызовы и реалии становления цифровой цивилизации

Одним из главных вызовов является сохранение человечности в условиях цифровизации, когда растущая зависимость от технологий ведёт к постепенной дегуманизации общества. Так, американский социолог и психолог Шерри Теркл пишет о снижении способности к глубокому эмоциональному контакту под влиянием постоянного использования современных технологий. Технологии подарили множество способов быть на связи оперативно, невзирая на расстояния, однако, как подчёркивает Теркл, они лишь создают иллюзию близости, заменяя эмоционально насыщенное общение поверхностными, быстрыми контактами. Как результат, технологии вместо укрепления связей увеличивают дистанцию между людьми, снижая качество их взаимодействий [7].

К числу рисков можно отнести концентрацию власти в руках крупных технологических корпораций, которые путём контроля за потоками информации и данных могут влиять на политические процессы и общественное мнение. Отсутствие независимости и естественности формирования общественного сетевого мнения подчёркивают многие исследователи. Так, Йоханна ван Дейк изучает влияние социальных платформ на политические процессы, их уязвимость к манипуляциям и контролю через алгорит-



Ил. 3. Распределение ответов респондентов на вопрос: «Существует ли опасность, что нейросети и искусственный интеллект в будущем сократят рынок труда в России?» (РФ, N = 1700, %)

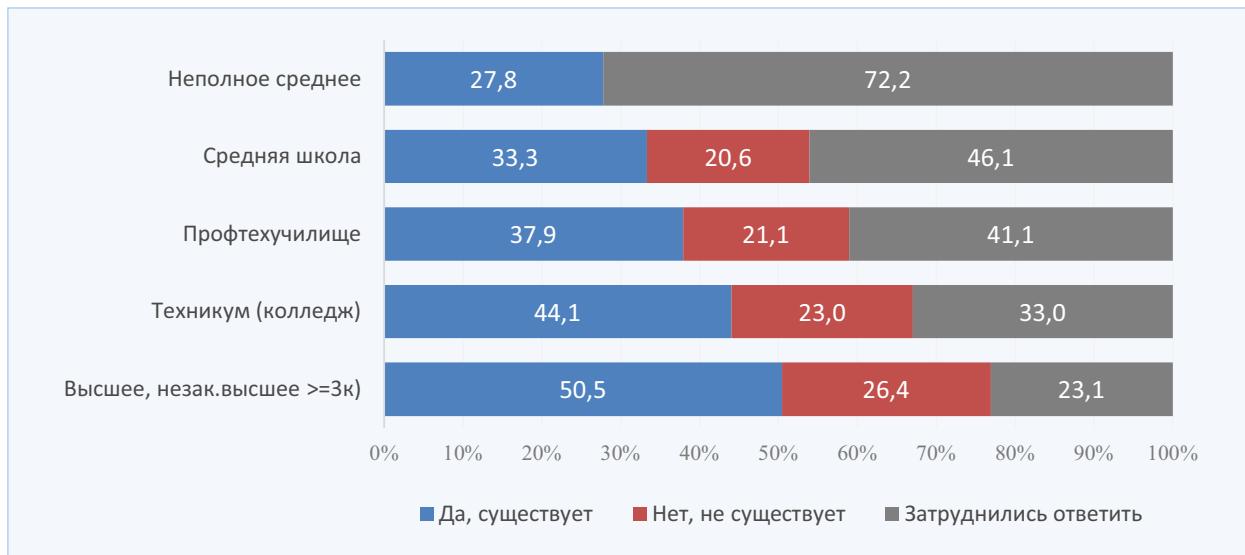
мы и интерфейсы [8]. Т. Гиллеспи анализирует скрытые механизмы модерации контента на платформах социальных медиа, которые формируют публичное пространство и влияют на общественные дискуссии [9]. Функционирование платформ как важного элемента цифрового общества «определяется процессами датификации и алгоритмизации. И здесь важным вопросом становится возможность учёта интересов общества, бизнеса и власти при формировании новых порядков ценностей, новых социальных механизмов контроля и управления», — пишет Д.Е. Добринская [10, с. 124].

Цифровая революция превращает информацию в основной ресурс развития, но одновременно ставит вопрос о её достоверности, доступности и управляемости. Об этом диалектическом парадоксе мы уже упоминали выше: технологии, призванные объединять человечество, способствуют также углублению социального разрыва и могут провоцировать конфликты. Социальные сети и СМИ, будучи мощными инструментами коммуникации, часто становятся источниками манипу-

ляций и дезинформации. Проблема усилилась в период взрывного развития ИИ и массового распространения основанных на его алгоритмах нейросетей. Особое воздействие оказали возможности генерации «глубоких подделок» (deepfakes), способные создавать новые угрозы для общественного доверия и демократических институтов. Вторым крупнейшим поводом для массовой тревоги стали страхи по поводу вытеснения человеческого труда и замены его системами на базе технологий искусственного интеллекта. Это вызывает страх перед массовой безработицей и социальной нестабильностью<sup>1</sup> (ил. 3).

В 2023 г., согласно данным всероссийского мониторинга «Как живёшь, Россия?» [11, с. 51], за рынок труда в связи с развитием ИИ опасались более половины респондентов (см. ил. 3). Годом позже аналогичный ответ дали уже не 55, а 44% респондентов. При этом доля граждан, не испытывающих подобных опасений, не увеличилась. Таким образом, нельзя сказать, что часть «опасавшихся» в 2023 г. кардинально изменила свою точку зре-

<sup>1</sup> В статье использованы данные всероссийского социологического мониторинга «Как живёшь, Россия?», проводимого Институтом социально-политических исследований ФНИСЦ РАН с 1992 г. Научный руководитель мониторинга — д.соц.н. В.К. Левашов. В исследовании использована квотно-пропорциональная всероссийская выборка с взаимозависимыми характеристиками гендерной совокупности пола, возраста, образования, места жительства, национального и социально-профессионального состава. В основу территориального размещения выборки легло экономико-географическое районирование страны при соблюдении пропорции численности населения и пропорций между городским и сельским населением. N = 1312–1866 на различных этапах.



Ил. 4. Восприятие россиянами угрозы искусственного интеллекта для рынка труда в зависимости от уровня образования (РФ, 2024, N = 1700, %)

ния за год, скорее, они перешли в разряд сомневающихся: число затруднившихся ответить в 2024 г. выросло с 21 до 33%.

Чаще других в 2024 г. опасались влияния искусственного интеллекта на рынок труда более образованные россияне, вероятно, сталкиваясь с примерами его использования в ходе обучения и работы (ил. 4). Среди обладателей высшего и неоконченного высшего образования опасения высказали 50,5% (не опасаются 26,4%, затруднились ответить 23,1%), среди выпускников техникумов/колледжей опасаются 44,1% (не опасаются 23%, затруднились ответить 32,9%) [12].

Если ситуация с рынком труда демонстрирует тенденцию к росту критического восприятия перспектив трудовой занятости с использованием современных технологий, то проблема информационной достоверности, доверия к информации из Интернета продолжает вызывать опасения. Несмотря на высочайшую долю пользователей Интернета (по данным

Mediascope, 84% населения страны заходят в сеть каждый день<sup>1</sup>), уже третий год отсутствует прогресс в показателях цифровой грамотности россиян<sup>2</sup>. Это отражается и на показателях уровня доверия граждан к актуальной новостной повестке, транслируемой в Интернете (табл.).

Несмотря на множающиеся обсуждения фальсификации и манипуляции информацией, первая волна которых пришлась на период пандемии COVID-19 (так называемая инфодемия), а вторая — на первые два года СВО, россияне всё чаще готовы воспринимать полученную через Интернет информацию как достоверную. За период с 2023 по 2024 г., согласно данным таблицы, число россиян, склонных считать информацию в социальных сетях и блогах достоверной, выросло на 10 п.п., достигнув 47% [11, с. 64]. В отношении интернет-СМИ соответствующее число превысило половину респондентов (52%). В обоих случаях снизилось количество сторонников противоположных ответов, в то время как чис-

<sup>1</sup> Актуальные цифры медиапотребления россиян — выступление Mediascope // Mediascope. 18.04.2025. URL: <https://mediascope.net/news/2928545/> (дата обращения: 25.04.2025).

<sup>2</sup> Индекс цифровой грамотности – 2024: цифровая грамотность россиян не растёт третий год подряд // НАФИ. 29.01.2025. URL: <https://nafi.ru/analytics/indeks-tsifrovoy-gramotnosti-2024-tsifrovaya-gramotnost-rossiyan-ne-rastet-tretiy-god-podryad> (дата обращения: 25.04.2025).

Таблица

**Мнение респондентов о том, соответствует ли действительности  
информация о текущих событиях в стране, которую  
транслируют российские ТВ, радио, Интернет (РФ, N=1700, %)**

Варианты ответов	Радио, телевидение		Сайты СМИ в Интернете		Социальные сети, блоги в Интернете	
	2023, VI	2024, IV	2023, VI	2024, IV	2023, VI	2024, IV
Соответствует	14	16	6	10	7	10
Скорее соответствует	33	37	34	42	30	37
Скорее не соответствует	24	22	25	18	22	16
Не соответствует	16	13	14	10	15	10
Затруднились ответить	14	12	21	20	26	26

ло затруднившихся осталось фактически неизменным.

Не исключено, что готовность доверять информации из Интернета в сложившихся обстоятельствах является следствием пролонгированного воздействия новостных фильтров, персонализации новостных лент и поисковых систем, вызвавших появление эффекта эхо-камеры и информационных пузырей. Касс Санстайн считает, что такие фильтры и алгоритмы ведут к информационной однобокости и снижению уровня общественного диалога, что влечёт за собой социальное разделение, рост радикализма и ослабление гражданского общества [13]. Погружение в цифровые реальности приводит к хаотизации мировоззрения и подчинению человека алгоритмам. Формируется поколение с трансформированными когнитивными навыками: снижением аналитического мышления, утратой самостоятельности в принятии решений, смещением акцента с текстовой информации на визуальные образы, что способствует развитию «клипового сознания». Всё это свидетельствует о серьёзном ударе по когнитивным и аналитическим способностям [14].

### Тенденции и императивы становления цифровой цивилизации

Новая устойчивость глобальной реальности возникает в процессе развития новых информационно-коммуникационных

центров и корпораций, транспортной инфраструктуры и производственных мощностей стран БРИКС и ШОС. Этот процесс отражает ключевые тенденции устойчивого развития мировой цифровой цивилизации:

- **Генезис мировой информационной цивилизации** начался и продолжается в цифровой реальности на институциональной основе взаимодействия и конкуренции суверенных государств, обеспечивающей гражданскую правовую безопасность производства и распределения материальных ресурсов, духовных и интеллектуальных ценностей в целях устойчивого развития жизни на планете.
- **Сдвиг к многополярности.** Доминирование стратегии политического лидерства и гегемонизма в условиях культурного многообразия и рисков конвенциональных и ядерных войн ведёт к самоуничтожению цивилизации. Диалектика мирного развития региональных и мировой цивилизаций принимает форму синergии биосферы, социосферы, техносферы, цифровых информационных производств и технологий многополярного мира.
- **Уход от парадигмы гегемонии развитых государств** сопровождается транснационализацией ресурсов, капиталов и рынков. Научно-техническая революция создаёт институты глобальной цифровой социально сбалансированной и политически устойчивой цивилизации.

- **Техногенный мир функционирующих в парадигме максимизации прибыли государств и союзов исчерпал свой технологический и временной ресурс.** Цифровые технологии ускорили и начали связывать континенты в единое социально-экономическое и политическое целое, а прозрачная цифровая культура начала демонтировать публичную легитимность культуры насилия, войн и неуправляемой конкуренции.
- **Сетевая интеграция.** Мировая цифровая цивилизация формируется на платформах многостороннего взаимовыгодного достаточного сотрудничества, интегрируя страны, нации и профессиональные сообщества через цифровые коммуникации. Её основу составляют сетевые структуры социально-политических институтов шести континентов, что создаёт устойчивую к кризисам глобальную многополярную систему.
- **Эволюция демократических механизмов.** В новых цивилизационных условиях развиваются технологии прямой цифровой демократии, вытесняя устаревшие, тормозящие устойчивое развитие процедуры представительной либеральной элитарной демократии. Трансформация иерархических систем в сетевые повышает скорость и качество управленческих решений, усиливая безопасность развития.
- **Глобальная системность инфраструктуры.** Трансформация иерархической системы политической, экономической, профессиональной, гражданской, социальной коммуникации порождает сетевую структуру новой цивилизации, основанную на новых цифровых технологиях и ИИ. Цифровые и транспортные сети обеспечивают управление мобильностью ресурсов, снижая риски кризисов в межнациональных отношениях. Таким образом, траектория устойчивого развития нашей страны и мировой цивилизации формируется императивными тенденциями, объективно проявляющимися в социально-политических проти-

воречиях и закономерностях, исследуемых учёными и учитываемых политиками в стратегиях управления потенциалом мира и сохранения жизни на планете.

### Заключение

Концепция В.И. Вернадского остаётся мощным теоретическим инструментом для осмысливания диалектики формирования глобальной цифровой цивилизации. Вернадский подчёркивал, что переход к ноосфере невозможен без этико-социального совершенствования, а разум человечества должен быть направлен на гармоничное сосуществование с природой, а не на её подчинение. Этот постулат становится особенно важным в эпоху, когда технологический прогресс опережает укрепление в социуме моральных и социальных норм. Мировая цифровая цивилизация, воплощая предвидение В.И. Вернадского, может стать началом ноосферы, но лишь при условии осознанного управления, основанного на идеалах гуманизма и научной рациональности. В этой альтернативе авторы придерживаются концепции гуманитарного подхода к дальнейшему вектору развития искусственного интеллекта в противовес технократическому и сциентистскому [15].

Современный мир погружен в океан противоречивой информации, в котором сталкиваются интересы отдельных личностей, социальных групп и общностей, государств и транснациональных структур. Общий вектор этой «стихии» противоречивой информации таков, — и это признаётся всеми участниками информационного взаимодействия, — что цивилизация уже перешла критические границы познания и практики, за которыми необходимо менять способ жизнедеятельности, делать его более устойчивым, научно релевантным для совместного проживания на планете в настоящем и будущем. Если за критерии отбора информации, необходимой для социально ответственного и справедливого управления, взять общие интересы выживания и научную достоверность,

то мы сумеем социально ответственно и выверено ответить на вызовы времени. Но для этого необходима консолидация усилий учёных и политиков, общества и государства. Сегодня это, пожалуй, главный социально-политический, моральный и интеллектуальный императив успеха совместных усилий по социально эффективному управлению региональными и глобальными вызовами.

Ключевым условием устойчивого развития общества в эпоху цифровизации становится гуманизация технологического прогресса, что предполагает приоритет этических норм, социальной ответственности и учёта культурных потребностей человека. Несмотря на положительные аспекты цифровой реальности, остаётся угроза того, что будущее человека в цифровом обществе может выйти за рамки гуманистической парадигмы, если вопросы образования, производства, потребления и т.д. не сохранят этическую и эстетиче-

скую ценность и не будут решены в правовой плоскости.

Подчеркнём, любые изменения, происходящие в обществе, должны иметь человеческое измерение и проходить гуманитарную экспертизу. Перед человечеством стоит задача не просто освоения новых технологий, а формирования гуманитарной среды, в которой развиваются интересы личности и общества в целом. Практическая значимость этой работы заключается в предложении стратегий устойчивого развития цифрового общества, сочетающих научные достижения с принципами гуманизма. Будущее цифровой цивилизации возникает в процессе синтеза идей ноосферного подхода и гуманистических принципов. Дальнейшие исследования могут быть направлены на разработку конкретных механизмов гуманитарной экспертизы технологий и международных стандартов этики цифровой цивилизации.

## ЛИТЕРАТУРА

1. Василенко Л.А. Метаморфозы цифровой многосубъектности: регулятивы мягкого управления vs программированного общества // Наука. Культура. Общество. 2024. Т. 30, № 2. С. 10–24. DOI: 10.19181/nko.2024.30.2.1. EDN PBXVRT.
2. Nielsen R., Graves L. «News You Don't Believe»: Audience Perspectives on Fake News. Reuters Institute for the Study of Journalism, 2017.
3. Левашов В.К., Сарьян В.К. Гражданское общество в сетях информационно-коммуникационных технологий // Вестник Российской академии наук. 2017. Т. 87, № 6. С. 532–537. DOI: 10.7868/S0869587317060056. EDN YSECZB.
4. Левашов В.К. Трансформация информационной сферы гражданского общества // Вестник Российского университета дружбы народов. Серия: Социология. 2019. Т. 19, № 4. С. 651–664. DOI: 10.22363/2313-2272-2019-19-4-651-664. EDN WZEYTY.
5. Вернадский В.И. Несколько слов о ноосфере // Успехи современной биологии. 1944. № 18. Вып. 2. С. 113–120.
6. Герасимова И.А. Цифровые технологии: реалии и кентавры воображения // Вопросы философии. 2021. № 10. С. 65–76. DOI: 10.21146/0042-8744-2021-10-65-76. EDN LCFMCK.
7. Turkle Sh. Alone together: why we expect more from technology and less from each other. New York: Basic Books, 2011.
8. Van Dijck J. The Culture of Connectivity: A Critical History of Social Media. Oxford University Press, 2013. DOI: 10.1093/acprof:oso/9780199970773.001.0001.
9. Gillespie T. Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media. New Haven: Yale University Press, 2018.
10. Добринская Д.Е. Что такое цифровое общество? // Социология науки и технологий. 2021. Т. 12, № 2. С. 112–129. DOI: 10.24412/2079-0910-2021-2-112-129. EDN ZLKPSX.

11. Как живешь, Россия? Экспресс-информация. 54-й этап социологического мониторинга, апрель 2024 года / В.К. Левашов, Н.М. Великая, И.С. Шушпанова [и др.]. М.: ФНИСЦ РАН, 2024. DOI: 10.19181/monogr.978-5-89697-429-1.2024. EDN INPEEV.
12. Левашов В.К., Гребняк О.В. Экспансия искусственного интеллекта: ожидания и настроения // Социологические исследования. 2024. № 12. С. 13–23. DOI: 10.31857/S0132162524120022. EDN XMNDBL.
13. Sunstein C.R. *Republic.com* 2.0. Princeton: Princeton University Press, 2007.
14. Гнатик Е.Н. Гуманитарные проблемы информационных технологий // Вестник Российской университета дружбы народов. Серия: Философия. 2017. Т. 21, № 2. С. 270–279. DOI: 10.22363/2313-2302-2017-21-2-270-279. EDN YOVBKL.
15. Дзялошинский И.М. Искусственный интеллект: гуманитарная перспектива // Вестник Новосибирского государственного университета. Серия: История, филология. 2022. Т. 21, № 6. С. 20–29. DOI: 10.25205/1818-7919-2022-21-6-20-29. EDN TNLNGV.

## On the Threshold of Digital Civilization: the Dialectics of Artificial Intelligence and Humanism

**Viktor Konstantinovich Levashov** – Doctor of Sociology, Director, Institute of Socio-Political Research of FCTAS RAS.

E-mail: levachov@mail.ru

**Oksana Valeryevna Grebnyak** – Researcher, Institute of Socio-Political Research of FCTAS RAS.

E-mail: oksananov@yandex.ru

The article discusses the challenges and findings of researching the dialectics of the relationship between digital technologies and humanistic principles and values in the context of an emerging global information civilization. The authors base their analysis on contemporary theories and practices of digital society and on V.I. Vernadsky's concept of the noosphere. This framework allows for understanding the transition to a digital civilization as a stage in the evolution of the biosphere, where human thought plays a key role. Empirical data is drawn from the All-Russian Sociological Monitoring survey "How Are You Living, Russia?" (conducted by the Institute of Social and Political Research of the Federal Research Sociological Center of the Russian Academy of Sciences), which reflects specific public concerns regarding the widespread implementation of artificial intelligence, as well as the paradox of trust in information disseminated online. The article outlines key challenges associated with growing dependence on AI, the declining quality of interpersonal interaction, the concentration of power in the hands of tech corporations, and a crisis of trust in information. Special attention is paid to the ambivalent consequences of digitalization. While information and communication technologies (ICT), neural networks, and artificial intelligence offer tools for addressing global ecological and resource management problems, they simultaneously exacerbate social, material, and spiritual disparities and contribute to information fragmentation. The authors conclude that it is necessary to uphold ethical norms and moral humanistic imperatives, considering technological progress, social responsibility, and sociocultural as well as ethnopolitical specificities.

**Keywords:** digital civilization, noosphere, humanization of technologies, artificial intelligence, trust in information, global information security, social risks of technologies, humanistic paradigm, humanitarian expertise

## REFERENCES

1. Vasilenko L.A. Metamorfozy tsifrovoyi mnogosub"ektnosti: reguliyativy myagkogo upravleniya vs programmirovannogo obshchestva // Nauka. Kul'tura. Obshchestvo. 2024. T. 30, № 2. S. 10–24. DOI: 10.19181/nko.2024.30.2.1. EDN PBXVRT (in Russian).
2. Nielsen R., Graves L. «News You Don't Believe»: Audience Perspectives on Fake News. Reuters Institute for the Study of Journalism, 2017.
3. Levashov V.K., Sar'yan V.K. Grazhdanskoe obshchestvo v setyakh informatsionno-kommunikatsionnykh tekhnologii // Vestnik Rossiiskoi akademii nauk. 2017. T. 87, № 6. S. 532–537. DOI: 10.7868/S0869587317060056. EDN YSECZB (in Russian).
4. Levashov V.K. Transformatsiya informatsionnoi sfery grazhdanskogo obshchestva // Vestnik Rossiiskogo universiteta druzhby narodov. Seriya: Sotsiologiya. 2019. T. 19, № 4. S. 651–664. DOI: 10.22363/2313-2272-2019-19-4-651-664. EDN WZEYTY (in Russian).
5. Vernadskii V.I. Neskol'ko slov o noosfere // Uspekhi sovremennoi biologii. 1944. № 18. Vyp. 2. S. 113–120 (in Russian).
6. Gerasimova I.A. Tsifrovye tekhnologii: realii i kentavry voobrazheniya // Voprosy filosofii. 2021. № 10. S. 65–76. DOI: 10.21146/0042-8744-2021-10-65-76. EDN LCFMCK (in Russian).
7. Turkle Sh. *Alone together: why we expect more from technology and less from each other*. New York: Basic Books, 2011.
8. Van Dijck J. *The Culture of Connectivity: A Critical History of Social Media*. Oxford University Press, 2013. DOI: 10.1093/acprof:oso/9780199970773.001.0001.
9. Gillespie T. *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. New Haven: Yale University Press, 2018.
10. Dobrinskaya D.E. Chto takoe tsifrovoe obshchestvo? // Sotsiologiya nauki i tekhnologii. 2021. T. 12, № 2. S. 112–129. DOI: 10.24412/2079-0910-2021-2-112-129. EDN ZLKPS (in Russian).
11. Kak zhivesh', Rossiya? Ekspress-informatsiya. 54-i etap sotsiologicheskogo monitoringa, aprel' 2024 goda / V.K. Levashov, N.M. Velikaya, I.S. Shushpanova [i dr.]. M.: FNISTS RAN, 2024. DOI: 10.19181/monogr.978-5-89697-429-1.2024. EDN INPEEV (in Russian).
12. Levashov V.K., Grebnyak O.V. Ekspansiya iskusstvennogo intellekta: ozhidaniya i nastroeniya // Sotsiologicheskie issledovaniya. 2024. № 12. S. 13–23. DOI: 10.31857/S0132162524120022. EDN XMNDBL (in Russian).
13. Sunstein C.R. *Republic.com 2.0*. Princeton: Princeton University Press, 2007.
14. Gnatik E.N. Gumanitarnye problemy informatsionnykh tekhnologii // Vestnik Rossiiskogo universiteta druzhby narodov. Seriya: Filosofiya. 2017. T. 21, № 2. S. 270–279. DOI: 10.22363/2313-2302-2017-21-2-270-279. EDN YOYBKL (in Russian).
15. Dzyaloshinskii I.M. Iskusstvennyi intellekt: gumanitarnaya perspektiva // Vestnik Novosibirskogo gosudarstvennogo universiteta. Seriya: Iстория, филология. 2022. T. 21, № 6. S. 20–29. DOI: 10.25205/1818-7919-2022-21-6-20-29. EDN TNLNGV (in Russian).



Д.И. КОЛОМАЦКИЙ,  
Е.В. КОРОВИНА\*

## Этический аспект использования технологий искусственного интеллекта в области сохранения языков коренных народов

В работе рассматриваются этические аспекты использования технологий искусственного интеллекта для языков коренных народов. Основное внимание уделяется анализу ряда заметных проектов. Часть из них не ставит своей непосредственной задачей ревитализацию языка или сохранение культурного наследия и потому может вызывать критику со стороны языковых сообществ. Анализируются и примеры успешных проектов, созданных и/или активно поддерживаемых представителями коренных народов. Предлагаемые к рассмотрению инициативы охватывают самые разные регионы мира: Африку, Северную и Южную Америку, Юго-Восточную Азию и Океанию. В статье также представлены рекомендации по дальнейшему развитию и поддержке таких инициатив, а также подчёркивается необходимость соблюдения этических принципов и особенно прав и интересов коренных народов. Настоящая работа призвана привлечь внимание к важности сохранения языков коренных народов с помощью современных технологий, что особенно актуально для Российской Федерации с её уникальным языковым разнообразием.

**Ключевые слова:** большие языковые модели, языки коренных народов, документирование языков, ревитализация языков, машинный перевод

Работа с так называемыми низкоресурсными языками (low resource languages) является одной из актуальных тем в области применения технологий искусственного интеллекта в лингвистике. Объём данных для обучения больших языковых моделей (Large Language Models, LLM, БЯМ) для таких

языков, как английский, китайский, испанский, русский, достаточно велик, чтобы модели могли генерировать тексты очень высокого качества. Миноритарные же языки оказываются недостаточно охвачены такими технологиями, что лишь усиливает доминирование мажоритарных языков в мире. В России говорят более

\* **Коломацкий Дмитрий Игоревич** — кандидат филологических наук, научный сотрудник отдела прикладной лингвистики Института языкоизнания РАН (ИЯз РАН).  
E-mail: dk@iling-ran.ru

**Коровина Евгения Владимировна** — младший научный сотрудник отдела урало-алтайских языков ИЯз РАН.  
E-mail: evkorovina@iling-ran.ru

чем на 150 языках (по данным ИЯз РАН), поэтому для нашей страны эта проблема также релевантна.

В последние годы компьютерные лингвисты обращают всё больше внимания на миноритарные языки. Одно из исследований такого рода представлено и в настоящем журнале (автор Т.О. Шаврина).

В связи с этим уместно вспомнить, что труд специалистов по ИИ и компьютерных лингвистов, занимающихся миноритарными языками, далеко не всегда направлен именно на их поддержку. Этический аспект работы с миноритарными языками вообще (включая просто сбор данных для словарей) завладел умами лингвистов не так давно, но к настоящему моменту уже появилось множество общих трудов на эту тему [1–4], а также case studies по конкретным примерам лингвистической работы в том или ином регионе [5].

В настоящей работе мы сосредоточимся на более узкой теме – этической стороне использования технологий ИИ применительно к миноритарным языкам. Если Т.О. Шаврина фокусируется на технологическом аспекте работы с низкоресурсными языками, мы обратим внимание на гуманитарный аспект на примерах конкретных инициатив.

Существующих проектов слишком много, чтобы дать их полный обзор в небольшой статье, поэтому представляется разумным предложить к рассмотрению несколько примечательных случаев, охватив этими примерами самые разные регионы мира.

В первую очередь мы приведём пример отношения исследователей к языку коренного народа как к средству, а не цели,

а также обозначим сложность проблемы и риски, связанные с вовлечением крупных корпораций. Напротив, дальнейшие сюжеты показывают, как могут работать проекты в области ИИ, если они инициированы или хотя бы активно поддержаны языковыми сообществами.

### Язык как средство: проекты, не инициированные изнутри коренных сообществ

#### Язык каламанг; «Machine Translation from One Book»

На языке каламанг говорят менее 200 человек, живущих на островах Карас у берегов полуострова Бомберай на западе острова Новая Гвинея (Индонезия). Его полное грамматическое описание было подготовлено в 2020 г., а в 2022 г. вышло в виде монографии [6]. Автор монографии стала соавтором работы о применении этого грамматического описания для порождения текстов на каламанг с помощью ИИ [7]. Эта работа получила широкий резонанс и стала как отправной точкой для других исследований [8], так и объектом критики ровно за то, что роль грамматического описания, возможно, переоценена [9].

Для настоящей работы интереснее тот факт, что сам язык в силу того, что он относится к «экстремально низкоресурсным языкам» (Extremely low-resource languages, XLR), хорошо подходит для эксперимента с БЯМ. В каком-то смысле такой статус языка стал для авторов достоинством. Основной целью исследования был «чистый» эксперимент, а не снабжение языкового сообщества новыми технологиями<sup>1</sup>. Результат эксперимента мог бы в лучшем случае стать позитивным побочным про-

<sup>1</sup> “Do LLMs really learn to perform new tasks by adaptation, or does adaptation simply draw out capabilities that the model had already learned? And does scaling pretraining data just improve performance because it implicitly scales up in-domain training data for every task? The best way to answer these questions is to evaluate on tasks that were unseen in the training data, but with these models being trained on increasingly opaque web-scale datasets, this can seem impossible <...>. To address this challenge, we turn to a field that is explicitly motivated and bottlenecked by a scarcity of web data: low-resource languages. <...> While translation tasks in general are well-represented in LLM training data, the Kalamang language in particular has been held out from the web for sociohistorical reasons, with the exception of the documentation in Visser (2022). This means that Kalamang tasks are unseen to LLMs but still feature the complexities that come with substantial real-world tasks.” [7] (полужирный шрифт мой. — Д.К.)

дуктом. В то же время стоит подчеркнуть, что носители языка были проинформированы о проводимом исследовании и согласились с использованием данных своего языка в качестве материала для БЯМ. При этом каких-либо данных об использовании разработанной модели среди сообщества говорящих на языке каламанг найти не удаётся.

### Уастекский наутль; переводчик Google как благо или вред

Как мы уже могли видеть, при разработке больших языковых моделей языковые сообщества нередко воспринимаются сугубо как источник данных для обучения, что исключает их активное участие в процессе. Такое положение дел может вызывать неприятие со стороны коренных народов, а потому провоцировать конфликты между создателями моделей и носителями языков. Ярким примером подобной динамики стала разработка машинного переводчика уастекского варианта языка наутль (Мексика, около 1 млн носителей) для сервиса Google Translate.

Основным разработчиком проекта выступила Габриэла Салас Кабрера<sup>1</sup> – молодой специалист в области ИТ и социальных проектов родом из небольшого города в мексиканском штате Идальго, который входит в ареал распространения этого языка<sup>2</sup>. Для сбора языкового материала она обратилась к науа-язычным сообществам в соцсетях и мессенджерах, предложив сотрудничество на волонтёрской основе, что было обусловлено отсутствием финансирования (Салас Кабрера проходила в этот период неоплачиваемую стажировку в Google). Участники подписывали ин-

формированное соглашение, где оговаривались эти условия. Благодаря собранным данным в июле 2024 г. сервис стал доступен на платформе Google. Салас Кабрера получила признание: в 2024 г. британская телевещательная корпорация включила её в список «100 женщин, изменивших мир»<sup>3</sup>, а мексиканские СМИ активно освещали её работу<sup>4</sup>.

Однако 25 марта 2025 г. в соцсетях развернулась дискуссия<sup>5</sup> о возможной эксплуатации языкового сообщества для извлечения прибыли без адекватной компенсации. Критики указывали на неэтичность использования бесплатного труда носителей языка, бесполезность автоматического переводчика для сообщества и отсутствие консенсуса относительно его ценности. Часть пользователей отказалась от участия именно по этим причинам. В то же время другие участники дискуссии подчёркивали право каждого самостоятельно решать вопрос о сотрудничестве, а также отмечали, что резкость критики может быть связана с происхождением Салас Кабрера (молодая женщина из бедной семьи индейского происхождения). Даже сторонники проекта признавали его организационные недостатки: так, в рабочую группу добавляли всех желающих без предварительного отбора по мотивации. Также отмечалось и то, что первоначальное описание проекта не содержало чётких указаний о волонтёрском характере работы и отсутствии финансирования, что было исправлено лишь позднее.

Описанная ситуация демонстрирует, к чему могут привести наличие интереса крупной корпорации при отсутствии

<sup>1</sup> <https://stemwomen.global/profile/gabriela-salas-cabrera> (дата обращения: 14.05.2025).

<sup>2</sup> Нужно отметить, что лингвистическое образование у Салас Кабреры отсутствует.

<sup>3</sup> Quién es Gabriela Salas Cabrera, mexicana que está dentro de las 100 mujeres más influyentes... Telediario, 04.12.2024 <https://www.telediario.mx/comunidad/gabriela-salas-cabrera-las-100-mujeres-mas-influyentes-de-la-bbc> (дата обращения: 14.05.2025).

<sup>4</sup> Conoce a Gabriela Salas, la mexicana que logró incluir el Náhuatl al traductor de Google. Informador, 18.07.2024 <https://www.informador.mx/mexico/Conoce-a-Gabriela-Salas-la-mexicana-que-logro-incluir-el-Nahuatl-al-traductor-de-Google-20240718-0105.html> (дата обращения: 14.05.2025).

<sup>5</sup> Исходный пост позднее был удалён, но в распоряжении авторов есть ссылки на последующую дискуссию.



Ил. 1. Дж. Брикси.

Foto Aaron Balana

чёткой и ясной коммуникации с носителями языка.

### Проекты, инициируемые или активно поддерживаемые представителями коренных народов

#### Чоктавский язык; мультимодальный корпус и чат-бот

Чоктавский язык — один из аборигенных языков Северной Америки, относится к группе мускогских языков. По данным Endangered Languages Project, на нём говорят от 9 до 11 тыс. человек, ему приписан статус языка, находящегося в уязвимом положении<sup>1</sup>; по данным исследовательницы языка Джэклин Брикси (ил. 1), число носителей не превышает 7 тыс., при этом сообщество Choctaw Nation of Oklahoma<sup>2</sup> насчитывает 220 тыс. «граждан» (citizens) [10]. Дж. Брикси создала мультимодальный корпус чоктавского языка ChoCo [11,12], чтобы затем на его основе разработать чат-бот под названием Masheli [13].

При разработке чат-бота изучались два варианта его функционирования: двуязыч-

ный (пользователь сам переключает язык) или с переключением кодов (что создаёт более полную имитацию живого разговора, так как реальные носители постоянно переключаются с английского на чоктавский и обратно) [14]. В оценке работы чат-бота обязательно принимали участие представители коренного народа, и протокол взаимодействия с ними подробно прописан в [14]. Вариант бота с переключением кода оказался в среднем предпочтительнее для участников эксперимента, которые называли бот «более дружелюбным».

Примечательно, что при создании Masheli было решено отказаться от использования БЯМ в части генерации ответа: БЯМ не обеспечили бы максимально корректного переключения кодов и не гарантировали бы точность вывода на чоктавском языке. В качестве технической основы был выбран NPCEditor, специально предназначенный для создания диалоговых систем в узких предметных областях [15]. Ответ не генерируется «на лету», а выбирается из коллекции, хотя запрос пользователя

<sup>1</sup> <https://endangeredlanguages.com/lang/1692> (дата обращения: 10.05.2025).

<sup>2</sup> <https://www.choctawnation.com/> (дата обращения: 10.05.2025).



Ил. 2. Коллектив Lelapa AI в 2025 г.

Фото с сайта [agenticppa.com](http://agenticppa.com)

анализируется статистическими методами (задача классификации текста). В этом смысле *Masheli* оказывается промежуточным продуктом между диалоговой системой на основе БЯМ и экспертной системой.

Усилия Дж. Брикси находят признание среди коренного населения штата Оклахома. Так, ей посвятила заметку издаваемая сообществом Choctaw Nation газета *Biskinik*<sup>1</sup>. В 2025 г. планируется выпустить словарь варианта чоктавского языка, на котором говорят в штате Миссисипи. Словарь будет сопровождаться разработанным Дж. Брикси приложением для распознавания речи, которое поможет в пользовании словарём тем носителям языка, которые не владеют им в письменной форме [10].

На наш взгляд, деятельность этой исследовательницы можно считать прекрасным примером применения технологий ИИ на благо коренных народов.

**Пять языков ЮАР, далее вся Африка; *Masakhane* и *Lelapa AI***

*Masakhane*<sup>2</sup> – африканская инициатива по снабжению языков Африки средствами машинного перевода [16]. Проблемы и первые решения для пяти языков ЮАР описаны в [17]. Уже год спустя, в 2020 г., число поддерживаемых языков значительно выросло [18]. К сожалению, официальный репозиторий инициативы не обновляется и не принимает изменения в исходный код уже около трёх лет<sup>3</sup>. Однако двое из сооснователей *Masakhane* вошли в число создателей лаборатории *Lelapa AI*<sup>4</sup>, которая продолжает разработки на тех же этических принципах (ил. 2).

Флагманским продуктом *Lelapa AI* стала коммерческая платформа *Vulavula*<sup>5</sup>, сочетающая в себе возможности транскрипции, распознавания сущностей, диалога с пользователем и т.д. Машинный перевод и синтез речи по состоянию на май 2025 г.

<sup>1</sup> <https://biskinik.com/wp-content/uploads/2025/02/mar2022-biskinik.pdf#page=12> (дата обращения: 10.05.2025).

<sup>2</sup> <https://www.masakhane.io/> (дата обращения: 10.05.2025).

<sup>3</sup> <https://github.com/masakhane-io/masakhane-mt> (дата обращения: 10.05.2025).

<sup>4</sup> <https://lelapa.ai> (дата обращения: 10.05.2025).

<sup>5</sup> <https://lelapa.ai/products/vulavula/> (дата обращения: 10.05.2025).



Ил. 3. Представители Te Hiku Media и групп маори.

Фото Te Hiku Media

находятся в статусе «coming soon». Для защиты прав коренных народов на их лингвистические ресурсы был создан фреймворк Esethu и составлена особая лицензия Esethu для распространения датасетов [19].

Пример Lelapa AI даёт основания полагать, что даже коммерческое использование языковых моделей может сочетаться с позитивным отношением к проекту в коренных языковых сообществах.

#### Язык маори; машинный перевод и обучающая платформа

Язык маори – полинезийский язык, относимый к австронезийской языковой семье, один из государственных языков Новой Зеландии. Endangered Language Project относит язык к категории угрожаемых.

Новозеландская медиакомпания Te Hiku Media, управляемая представителями различных групп (*iwi*) народа маори, уделяет особое внимание разработке технологических решений для документа-

ции и ревитализации языка. С появлением БЯМ компания обратилась к использованию этих средств, в частности к разработкам решений для автоматического распознавания речи, позволяющим транскрибировать речь маори с точностью 92% и речь с переключением кодов с точностью 82%<sup>1</sup>. Разработкам Te Hiku Media (ил. 3) посвящён ряд научных работ (например, [20, 21]). Авторы [20] отмечают, что разработанные для маори технологии планируется в будущем использовать для самоанского и гавайского языков. Описываемый в работах проект Papa Reo официально запущен и предлагает разработчикам целый пакет инструментов для распознавания и синтеза речи, а также отслеживания соблюдения этических принципов работы с данными («kaitiakitanga», по-английски «guardianship, trusteeship») на языке маори<sup>2</sup>. Te Hiku Media подчёркивает, что считает себя именно «хранителем, попечителем» (kaitiaki), а не владель-

<sup>1</sup> Māori Speech AI Model Helps Preserve and Promote New Zealand Indigenous Language. Nvidia blog. <https://blogs.nvidia.com/blog/te-hiku-media-māori-speech-ai/> (дата обращения: 11.05.2025).

<sup>2</sup> <https://papareo.io/> (дата обращения: 11.05.2025).

цем данных, получаемых от носителей языка<sup>1</sup>.

Опыт Te Hiku Media иллюстрирует возможности, которыми обладают и которые могут реализовывать местные средства массовой информации, заинтересованные в сохранении языка коренного народа.

### Рекомендации

Главный вывод из изученных нами примеров разработки ИИ-решений для языков коренных народов (как описанных в статье, так и оставшихся за её рамками) состоит в том, что в подобных проектах совершенно необходимо активное участие представителей коренных народов<sup>2</sup>. Если специалисты по ИИ активно контактируют с соответствующими языковыми сообществами и понимают их потребности, они верно сформируют и набор технических решений. Файн-тюнинг БЯМ (дообучение предобученной БЯМ для решения специфических задач) повышает точность работы модели по сравнению с обычной предобученной БЯМ. Если же ставится цель сделать модели доступными для непосредственного использования представителями сообществ, она может диктовать выбор решений, предъявляющих низкие требования к оборудованию. Например, может быть

применена так называемая дистилляция знаний [22] или выбрана более энергоэффективная модель — скажем, из семейства Mistral [23]<sup>3</sup>. В работе [24] даже вводится термин «Indigenous Language Models (ILMs)», обозначающий модели, оптимизированные для языков коренных народов.

В юридических документах, определяющих работу инициатив в области ИИ применительно к малочисленным языкам, должны быть прописаны права представителей языковых сообществ на предоставляемый ими материал, особенно если речь идёт о культурном и/или религиозном наследии коренного народа.

### Заключение

В настоящей работе на нескольких ярких примерах предпринята попытка показать, что технологии ИИ могут служить сохранению языкового разнообразия, документации языкового наследия и даже ревитализации малых языков, если они применяются при активном участии представителей коренных народов и со стремлением принести пользу именно им. Мы надеемся, что эти соображения позитивно повлияют на развитие аналогичных проектов, создаваемых для малочисленных языков Российской Федерации.

## ЛИТЕРАТУРА

1. Good J. Ethics in Language Documentation and Revitalization // The Oxford Handbook of Endangered Languages / Ed. K.L. Rehg, L. Campbell. Oxford University Press, 2018. Pp. 418–440. DOI: 10.1093/oxfordhb/9780190610029.013.21.
2. Holton G., Leonard W.Y., Pulsifer P.L. Indigenous Peoples, Ethics, and Linguistic Data // The Open Handbook of Linguistic Data Management / Ed. A.L. Berez-Kroeker et al. The MIT Press, 2022. Pp. 49–60. DOI: 10.7551/mitpress/12200.003.0008.
3. Marley T.L. Indigenous Data Sovereignty and the role of universities // Indigenous Data Sovereignty and Policy. 1<sup>st</sup> ed. London: Routledge, 2020. Pp. 157–168. DOI: 10.4324/9780429273957-11.

<sup>1</sup> A new vision of artificial intelligence for the people. MIT Technology Review <https://www.technologyreview.com/2022/04/22/1050394/artificial-intelligence-for-the-people/> (дата обращения: 11.05.2025).

<sup>2</sup> При этом, как говорилось в обсуждении проблемы языка науять в соцсетях, не нужно воспринимать коренные народы как настолько уязвимые и нуждающиеся в помощи извне, что они не способны самостоятельно принимать решения о том, в каких проектах участвовать.

<sup>3</sup> Обсуждение использования Mistral для работы с языками аборигенов Австралии найдено на специализированных форумах, но академических статей на эту тему авторы пока не обнаружили.

4. Ruckstuhl K. Trust in Scholarly Communications and Infrastructure: Indigenous Data Sovereignty // *Front. Res. Metr. Anal.* 2022. Vol. 6. DOI: 10.3389/frma.2021.752336.
5. Ortenzi K.M. et al. Good data relations key to Indigenous research sovereignty: A case study from Nunatsiavut // *Ambio*. 2025. Vol. 54, № 2. Pp. 256–269. DOI: 10.1007/s13280-024-02077-6.
6. Visser E.A grammar of Kalamang. Berlin: Language Science Press, 2022. <https://zenodo.org/record/6499927> (access date: 10.05.2025).
7. Tanzer G. et al. A Benchmark for Learning to Translate a New Language from One Grammar Book: arXiv:2309.16575. arXiv, 2024. DOI: 10.48550/arXiv.2309.16575.
8. Kornilov A., Shavrina T. From MTEB to MTOB: Retrieval-Augmented Classification for Descriptive Grammars: arXiv:2411.15577. arXiv, 2024. DOI: 10.48550/arXiv.2411.15577.
9. Aycock S. et al. Can LLMs Really Learn to Translate a Low-Resource Language from One Grammar Book? arXiv:2409.19151. arXiv, 2025. DOI: 10.48550/arXiv.2409.19151.
10. Brixey J. Using Artificial Intelligence to Preserve Indigenous Languages. USC Institute for Creative Technologies, 2025. <https://ict.usc.edu/news/essays/using-artificial-intelligence-to-preserve-indigenous-languages/> (access date: 10.05.2025).
11. Brixey J., Pincus E., Artstein R. Chahta Anumpa: A Multimodal Corpus of the Choctaw Language // Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). 2018. Pp. 3371–3376. <https://aclanthology.org/L18-1532.pdf>. (access date: 10.05.2025).
12. Brixey J., Artstein R. ChoCo: a multimodal corpus of the Choctaw language // *Lang Resources & Evaluation*. 2021. Vol. 55, № 1. Pp. 241–257. DOI: 10.1007/s10579-020-09494-5.
13. Brixey J., Traum D. Masheli: A Choctaw-English Bilingual Chatbot // *Conversational Dialogue Systems for the Next Decade* / Ed. L.F. D'Haro, Z. Callejas, S. Nakamura Singapore: Springer Singapore, 2021. Vol. 704. Pp. 41–50. DOI: 10.1007/978-981-15-8395-7\_4.
14. Brixey J., Traum D. Does a code-switching dialogue system help users learn conversational fluency in Choctaw? // Proceedings of the Fifth Workshop on NLP for Indigenous Languages of the Americas (AmericasNLP). Albuquerque, New Mexico: Association for Computational Linguistics, 2025. Pp. 8–17. <https://aclanthology.org/2025.americasnlp-1.2/> (access date: 10.05.2025).
15. Leuski A., Traum D. NPCEditor: Creating Virtual Human Dialogue Using Information Retrieval Techniques // *AI Magazine*. 2011. Vol. 32, № 2. Pp. 42–56. DOI: 10.1609/aimag.v32i2.2347.
16. Orife I. et al. Masakhane – Machine Translation For Africa: arXiv:2003.11529. arXiv, 2020. DOI: 10.48550/arXiv.2003.11529.
17. Martinus L., Abbott J.Z. A Focus on Neural Machine Translation for African Languages: arXiv:1906.05685. arXiv, 2019. DOI: 10.48550/arXiv.1906.05685.
18. Nekoto W. et al. Participatory Research for Low-resourced Machine Translation: A Case Study in African Languages: arXiv:2010.02353. arXiv, 2020. DOI: 10.48550/arXiv.2010.02353.
19. Rajab J. et al. The Esethu Framework: Reimagining Sustainable Dataset Governance and Curation for Low-Resource Languages: arXiv:2502.15916. arXiv, 2025. DOI: 10.48550/arXiv.2502.15916.
20. Jones P.-L. et al. Kia tangata whenua: Artificial intelligence that grows from the land and people // *Ethical Space: International Journal of Communication Ethics*. 2023. Vol. 2023, № 2/3. DOI: 10.21428/0af3f4c0.9092b177.
21. Leoni G. et al. Solving Failure Modes in the Creation of Trustworthy Language Technologies // Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages @ LREC-COLING 2024. Torino, Italia, 2024. <https://aclanthology.org/2024.sigul-1.39/> (access date: 11.05.2025).
22. Xu X. et al. A Survey on Knowledge Distillation of Large Language Models: arXiv:2402.13116. arXiv, 2024. DOI: 10.48550/arXiv.2402.13116.
23. Jiang A.Q. et al. Mistral 7B: arXiv:2310.06825. arXiv, 2023. DOI: 10.48550/arXiv.2310.06825.
24. Pinhanez C. et al. Harnessing the Power of Artificial Intelligence to Vitalize Endangered Indigenous Languages: Technologies and Experiences: arXiv:2407.12620. arXiv, 2024. DOI: 10.48550/arXiv.2407.12620.

## The Ethical Aspect of Using Artificial Intelligence Technologies in the Field of Indigenous Languages Preservation

**Dmitry Igorevich Kolomatsky** — Candidate of Philological Sciences, Research Fellow at the Department of Applied Linguistics, Institute of Linguistics, Russian Academy of Sciences (IL RAS).

E-mail: dk@iling-ran.ru

**Evgeniya Vladimirovna Korovina** — Junior Researcher at the Department of Ural-Altaic Languages at the IL RAS.

E-mail: evkorovina@iling-ran.ru

The paper discusses the ethical aspects of applying artificial intelligence technologies to indigenous languages. The analysis focuses on several prominent projects in this field. Some initiatives are not primarily aimed at language revitalization or cultural preservation, which may attract criticism from the language communities themselves. The paper also explores examples of successful projects created and/or actively supported by indigenous individuals. The proposed initiatives cover a wide range of regions across the globe, including Africa, North and South America, Southeast Asia, and Oceania. The article also provides recommendations for further development and support of these initiatives, emphasizing the importance of ethical considerations and respecting the rights and interests of indigenous peoples. This work aims to raise awareness about the significance of preserving indigenous languages through modern technology, which is particularly relevant for the Russian Federation with its unique linguistic diversity.

**Keywords:** large language models, indigenous languages, language documentation, language revitalization, machine translation

## REFERENCES

1. Good J. Ethics in Language Documentation and Revitalization // The Oxford Handbook of Endangered Languages / Ed. K.L. Rehg, L. Campbell. Oxford University Press, 2018. Pp. 418–440. DOI: 10.1093/oxfordhb/9780190610029.013.21.
2. Holton G., Leonard W.Y., Pulsifer P.L. Indigenous Peoples, Ethics, and Linguistic Data // The Open Handbook of Linguistic Data Management / Ed. A.L. Berez-Kroeker et al. The MIT Press, 2022. Pp. 49–60. DOI: 10.7551/mitpress/12200.003.0008.
3. Marley T.L. Indigenous Data Sovereignty and the role of universities // Indigenous Data Sovereignty and Policy. 1<sup>st</sup> ed. London: Routledge, 2020. Pp. 157–168. DOI: 10.4324/9780429273957-11.
4. Ruckstuhl K. Trust in Scholarly Communications and Infrastructure: Indigenous Data Sovereignty // Front. Res. Metr. Anal. 2022. Vol. 6. DOI: 10.3389/frma.2021.752336.
5. Ortenzi K.M. et al. Good data relations key to Indigenous research sovereignty: A case study from Nunatsiavut // Ambio. 2025. Vol. 54, № 2. Pp. 256–269. DOI: 10.1007/s13280-024-02077-6.
6. Visser E.A grammar of Kalamang. Berlin: Language Science Press, 2022. <https://zenodo.org/record/6499927> (access date: 10.05.2025).
7. Tanzer G. et al. A Benchmark for Learning to Translate a New Language from One Grammar Book: arXiv:2309.16575. arXiv, 2024. DOI: 10.48550/arXiv.2309.16575.
8. Kornilov A., Shavrina T. From MTEB to MTOB: Retrieval-Augmented Classification for Descriptive Grammars: arXiv:2411.15577. arXiv, 2024. DOI: 10.48550/arXiv.2411.15577.
9. Aycock S. et al. Can LLMs Really Learn to Translate a Low-Resource Language from One Grammar Book? arXiv:2409.19151. arXiv, 2025. DOI: 10.48550/arXiv.2409.19151.

10. Brixey J. Using Artificial Intelligence to Preserve Indigenous Languages. USC Institute for Creative Technologies, 2025. <https://ict.usc.edu/news/essays/using-artificial-intelligence-to-preserve-indigenous-languages/> (access date: 10.05.2025).
11. Brixey J., Pincus E., Artstein R. Chahta Anumpa: A Multimodal Corpus of the Choctaw Language // Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). 2018. Pp. 3371–3376. <https://aclanthology.org/L18-1532.pdf>. (access date: 10.05.2025).
12. Brixey J., Artstein R. ChoCo: a multimodal corpus of the Choctaw language // Lang Resources & Evaluation. 2021. Vol. 55. № 1. Pp. 241–257. DOI: 10.1007/s10579-020-09494-5.
13. Brixey J., Traum D. Masheli: A Choctaw-English Bilingual Chatbot // Conversational Dialogue Systems for the Next Decade / Ed. L.F. D'Haro, Z. Callejas, S. Nakamura Singapore: Springer Singapore, 2021. Vol. 704. Pp. 41–50. DOI: 10.1007/978-981-15-8395-7\_4.
14. Brixey J., Traum D. Does a code-switching dialogue system help users learn conversational fluency in Choctaw? // Proceedings of the Fifth Workshop on NLP for Indigenous Languages of the Americas (AmericasNLP). Albuquerque, New Mexico: Association for Computational Linguistics, 2025. Pp. 8–17. <https://aclanthology.org/2025.americasnlp-1.2/> (access date: 10.05.2025).
15. Leuski A., Traum D. NPCEditor: Creating Virtual Human Dialogue Using Information Retrieval Techniques // AI Magazine. 2011. Vol. 32, № 2. Pp. 42–56. DOI: 10.1609/aimag.v32i2.2347.
16. Orife I. et al. Masakhane – Machine Translation For Africa: arXiv:2003.11529. arXiv, 2020. DOI: 10.48550/arXiv.2003.11529.
17. Martinus L., Abbott J.Z. A Focus on Neural Machine Translation for African Languages: arXiv: 1906.05685.arXiv, 2019. DOI: 10.48550/arXiv.1906.05685.
18. Nekoto W. et al. Participatory Research for Low-resourced Machine Translation: A Case Study in African Languages: arXiv:2010.02353. arXiv, 2020. DOI: 10.48550/arXiv.2010.02353.
19. Rajab J. et al. The Esethu Framework: Reimagining Sustainable Dataset Governance and Curation for Low-Resource Languages: arXiv:2502.15916. arXiv, 2025. DOI: 10.48550/arXiv.2502.15916.
20. Jones P.-L. et al. Kia tangata whenua: Artificial intelligence that grows from the land and people // Ethical Space: International Journal of Communication Ethics. 2023. Vol. 2023, № 2/3. DOI: 10.21428/0af3f4c0.9092b177.
21. Leoni G. et al. Solving Failure Modes in the Creation of Trustworthy Language Technologies // Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages @ LREC-COLING 2024. Torino, Italia, 2024. <https://aclanthology.org/2024.sigul-1.39/> (access date: 11.05.2025).
22. Xu X. et al. A Survey on Knowledge Distillation of Large Language Models: arXiv:2402.13116. arXiv, 2024. DOI: 10.48550/arXiv.2402.13116.
23. Jiang A.Q. et al. Mistral 7B: arXiv:2310.06825. arXiv, 2023. DOI: 10.48550/arXiv.2310.06825.
24. Pinhanez C. et al. Harnessing the Power of Artificial Intelligence to Vitalize Endangered Indigenous Languages: Technologies and Experiences: arXiv:2407.12620. arXiv, 2024. DOI: 10.48550/arXiv.2407.12620.



**Ю.В. БАЛАКИНА,  
М.В. ГРИГОРЬЕВА,  
Е.Н. СОКОЛОВА\***

## Применение больших языковых моделей для анализа ценостно-патриотического дискурса русскоязычных пользователей\*\*

В статье рассматриваются возможности применения больших языковых моделей (LLM) для автоматизированного анализа ценостно-патриотического дискурса русскоязычных пользователей социальных медиа. На материале корпуса сообщений из VK, «Одноклассников» и Telegram (2023–2025 гг.) исследуется, насколько результаты автоматической кодировки совпадают с экспертной разметкой по специально разработанной категориальной схеме. Кодбук включает восемь измерений: базовые ценности по Ш Шварцу, две оси Р. Инглхарта (традиционизм / светскость; выживание / само выражение), уровни потребностей по А. Маслоу, типы патриотизма (конструктивный / агрессивный) по мотивам К.Д. Ушинского и В.С. Соловьёва, доминирующие типы речевых актов по Дж. Остину, а также бинарные индикаторы эксплицитного патриотизма и гражданской идентичности. Эксперимент проведён на кластере сообщений «Гордость и патриотизм» ( $N = 456$ ), где плотность ценностных маркеров максимально высока; со-поставление реализовано через матрицы ошибок, accuracy, macro/weighted F1 и коэффициент к Коэна. Показано, что LLM надёжно выделяет эксплицитную патриотическую тематику, но существенно менее согласована с экспертами при многоклассовой и тонкой ценностной классификации (Шварц, Маслоу, шкалы Инглхарта, типы патриотизма, речевые акты Остина), демонстрируя систематические смещения и гипердиагностику отдельных категорий. Сделан вывод, что LLM в текущей конфигурации может использоваться как вспомогательный инструмент предварительной разметки и генерации гипотез, но не как автономный заменитель экспертного контент-анализа ценностного дискурса.

\* **Балакина Юлия Владимировна** — кандидат филологических наук, доцент, профессор департамента фундаментальной и прикладной лингвистики Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ–Нижний Новгород).

E-mail: julianaumova@gmail.com

**Григорьева Мария Владимировна** — старший преподаватель Школы коммуникаций факультета креативных индустрий НИУ ВШЭ.

E-mail: mariya.grigoreva@hse.ru

**Соколова Екатерина Никитична** — кандидат политических наук, заведующая проектно-учебной лабораторией Политических коммуникаций Школы коммуникаций факультета креативных индустрий НИУ ВШЭ.

E-mail: e.sokolova@hse.ru

\*\*Статья подготовлена в Национальном исследовательском университете «Высшая школа экономики» в рамках государственного задания Министерства науки и высшего образования Российской Федерации (тема № ФИ-2025-83 «Современные технологии социально-политической коммуникации»).

**Ключевые слова:** большие языковые модели, LLM, ценности, патриотизм, гражданская идентичность, контент-анализ, Шварц, Инглхарт, Маслоу, речевые акты, социальные медиа, русскоязычный дискурс, автоматическая разметка

Тексты пользователей социальных медиа – от постов в VK и «Одноклассниках» до сообщений в Telegram-каналах – источник информации о ценностных ориентациях общества. Извлечение и анализ ценностно-патриотического контента сопряжены с определёнными трудностями, обусловленными большим объёмом неструктурированных данных, сложностью интерпретации сленга, иронии и культурных аллюзий, а также структурными и формальными особенностями русского языка. В данном контексте возрастаёт интерес к применению больших языковых моделей (LLM) для автоматизированного и смыслового анализа содержания подобных текстов. В частности, такие модели, как GigaChat, YandexGPT, ChatGPT и DeepSeek, демонстрируют способность автоматически классифицировать тексты в рамках отдельных задач. Тем не менее, на данном этапе рассматривать модели в качестве инструмента социокультурного анализа пока корректно лишь в качестве методологической гипотезы, требующей независимой валидации, репликации и проверки устойчивости результатов.

В настоящей статье рассматриваются возможности использования LLM для анализа русскоязычных пользовательских текстов на темы ценностей, патриотизма и гражданской идентичности. Статья включает теоретический обзор понятия ценностей и патриотизма в языке пользователей, обсуждает лингвистические и методологические возможности LLM при работе с русским языком, описывает методику сбора и обработки корпуса из сообщений социальных медиа (2023–2025 гг.), демонстрирует результаты автоматической атрибуции ценностных категорий с помощью LLM и их сопоставление с экспертной разметкой, а так-

же анализирует выводы из проведённого эксперимента.

### **Возможности LLM для анализа русскоязычной речи пользователей**

Ценности в социальной психологии определяются как устойчивые убеждения и жизненные принципы, выражающие представления о желаемых целях и выступающие критериями оценивания событий [1].

Традиционно ценности изучались с помощью опросов (опросник Шварца, World Values Survey), однако в последнее десятилетие растёт интерес к анализу ценностей на материале цифровых следов, в том числе на текстах в социальных сетях. Соцмедиа отражают актуальные изменения массового сознания и позволяют исследовать ценностные ориентации в естественном контексте общения. Однако существует проблема дифференциации собственно ценностей и схожих понятий (личностных черт, мнений и пр.) в пользовательских сообщениях. Наряду с этим традиционные методы, основанные на лексических словарях (например, LIWC) и подходе «мешка слов», демонстрируют низкую корреляцию с опросными профилями ценностей (в том числе по методике Шварца) [12].

В ответ на эти вызовы развиваются новые методы, опирающиеся на контекстуальные модели и машинное обучение, применение которых предполагает выявление ценностно-экспрессивных сообщений, а не профилирование личности пользователя по его постам.

Большие языковые модели, обученные на огромных корпусах текстов, демонстрируют определённые способности «понимания» и генерирования текста на разных языках. Применительно к русскоязычному контенту LLM (например, GPT-4)

для корректного анализа содержания сообщений должны учитывать особенности русского языка.

В частности, в процессе обработки русскоязычных текстов модель GPT-4 обнаруживает существенные ограничения, обусловленные спецификой разговорной лексики, идиоматических выражений и некомпозитивных языковых конструкций, а также недостатком контекста и обученности на англоязычной базе, что затрудняет автоматический анализ на уровне ключевых слов, особенно с учётом специфики морфологической системы русского языка [13].

Наряду с этим пользовательские тексты нередко содержат элементы жаргона и сленга, эмодзи, смешение кодов, сарказм и иронию. LLM, будучи обучены на данных из Интернета, способны «расшифровать» многие широко употребляемые жаргонные выражения (модель GPT-4 Turbo), однако в некоторых случаях модели затруднительно уловить сарказм и иронию. Так, если пользователь без явных индикаторов иронизирует, LLM может интерпретировать текст буквально и потому ошибочно определить его эмоциональную окраску. Исследования показывают, что даже продвинутые модели уступают специализированным алгоритмам, обученным на корпусах саркастических примеров [14, 15]. В частности, GPT-4 распознаёт сарказм лучше предшественников и других LLM (она превосходит другие модели в тестах приблизительно на 14% по точности) [14], но всё же заметно отстаёт от человеческого уровня и от узко настроенных моделей, обученных именно для детекции сарказма [16].

Кроме того, LLM обучены преимущественно на статистике языковых паттернов и могут допускать искажения, унаследованные из данных. Например, если в обучающей выборке широко встречается определённый нарратив, модель может склоняться к такому же пониманию и пропускать альтернативные проявления патриотизма. Во-вторых, классификация

непрозрачна: неясно, по каким признакам ChatGPT относит сообщение к той или иной категории. Одновременно признаётся чувствительность LLM к формулировке запроса и настройкам генерации, что требует дополнительных протоколов контроля и сопоставления результатов с ручной экспертной разметкой [17].

## Методология исследования

В качестве исходного массива использован корпус из более чем двух миллионов публичных сообщений российского сегмента социальных сетей, сформированный средствами системы Brand Analytics. В корпус вошли сообщения на русском языке, опубликованные в 2023–2024 гг. во «ВКонтакте», «Одноклассниках», Telegram и на открытых веб-площадках, доступных без авторизации и без обхода технических ограничений. Географический фокус задавался привязкой пользователей и сообществ к России. Отбор сообщений осуществлялся по 38 формульным лексико-семантическим маркерам (например, конструкции типа «для меня важно», «я горжусь», «главное в жизни» и др.), что позволило целенаправленно выделить тексты с эксплицитной артикуляцией ценностей, а также минимизировать долю бытового и новостного шума.

Предобработка включала удаление технических элементов, полную деидентификацию (ники, персональные ссылки, контакты, идентифицирующие URL), исключение точных и почти точных дубликатов и коммерческого спама, лемматизацию (PyMystem3/PyMorph2), расширенный стоп-список и отсечение крайне редких лексем при сохранении ценностно значимой лексики. Структурирование ценностного поля выполнялось в два этапа: тематическое моделирование LDA с завышенным числом тем ( $K = 5000$ ) для выделения тонких латентных компонент; кластеризация агрегированных тематических представлений на основе эмбеддингов Word2Vec методом k-means ( $k = 25$ ) с многоократными перезапусками и выбором решения



Ил. 1. Распределение ценностных кластеров в российском сегменте соцмедиа (число сообщений, в которых присутствует соответствующая ценностная тематика)

по минимальной внутрикластерной дисперсии. Полученные кластеры прошли независимую экспертуру: для каждого были заданы устойчивые наборы ключевой лексики и содержательные определения.

Структура ценностного пространства, достаточная для понимания дальнейших шагов, обобщена на ил. 1.

Кластер «Гордость и патриотизм» трактуется как совокупность сообщений с явно позитивным или нормативным отношением к России, её истории, символике, армии,

государственным институтам и коллективой идентичности, с риторикой гордости, долга, благодарности, защиты и служения. Кластер выбран для анализа как фрагмент с высокой концентрацией явных ценностных маркеров, что позволяет оценить не просто наличие смысла, а качество его ценностной атрибуции.

Из данного кластера случайным образом отобраны 456 сообщений пользователей и публичных сообществ без дополнительных идеологических фильтров, что снижает селекционное смещение и сохраняет ре-

альный спектр патриотического дискурса в этом сегменте. Анализ имеет формат контролируемого локального эксперимента и служит исключительно для сопоставления двух процедур атрибуции: независимой экспертной разметки по формализованному кодировочному справочнику и автоматической разметки большой языковой моделью (ChatGPT GPT-5 PRO) при фиксированных инструкциях и параметрах<sup>1</sup>. Обе разметки выполнялись независимо: эксперты не видели результатов модели, модель не имела доступа к человеческим меткам.

### Схема категориальной разметки

Для сопоставления экспертной и LLM-аннотации использовалась единая категориальная схема с восемью измерениями; по каждому измерению каждому сообщению присваивалось одно значение из фиксированного перечня.

#### 1. Базовая ценность по Шварцу (Schwartz).

Доминирующая ценность: Власть, Достижения, Гедонизм, Стимуляция, Самонаправленность, Универсализм, Благожелательность, Традиции, Конформность, Безопасность, Нет. Операционализация основана на переработанной модели Ш. Шварца; значение присваивается при наличии явного ценностного акцента, при его отсутствии — «Нет».

#### 2. Традиционные и секулярно-рациональные ценности по Р. Инглхарту (Inglehart\_trad\_vs\_secular).

Два состояния: traditional (явные ссылки на религию, «традиционные ценности», авторитет государства/армии/семьи, иерархию, послушание) и neutral (отсутствие выраженного традиционализма, в том числе светско-рациональные высказывания). Фиксируется только явно выраженный традиционалистский полюс.

#### 3. Ценности выживания и самовыражения по Р. Инглхарту (Inglehart\_survival\_vs\_selfexpr).

Три значения: survival (угрозы, базовая безопасность, порядок любой ценой, материальное

выживание), self-expression (самореализация, права, свободы, участие, протест, креативность, экологическая и гражданская повестка), neutral (нет достаточных признаков обоих полюсов).

4. Уровни потребностей по А. Маслоу. Доминирующий уровень потребностей: Физиологические, Безопасность, Приналежность/любовь, Уважение/признание, Самореализация. Схема соответствует классической модели А. Маслоу; выбор уровня основан на ведущем мотиве сообщения.
5. Тип патриотического высказывания по модели «Ушинский–Соловьёв». Тип патриотического высказывания: конструктивный (нравственно мотивированная любовь к стране, ответственность, ориентация на общее благо при признании ценности других народов), агрессивный (шовинизм, ксенофобия, противопоставление «своих» и «чужих», оправдание насилия), нет (патриотическая проблематика отсутствует или выражена недостаточно).
6. Тип речевого акта по Дж. Остину (Austin). Доминирующий тип речевого акта: вердиктивы — verdictive (оценки, «суждения» и заключения), экзерситивы — exercitive (призывы, требования, выражение воли), бехабитивы — behabitive (выражения отношения, одобрения, признания, уважения/неуважения), экспозитивы — expositive (объяснение, интерпретация, оформление аргументации). Класс определяется по преобладающей интенции высказывания.
7. Патриотизм. Бинарный индикатор эксплицитной патриотической тематики: ИСТИНА — при прямых упоминаниях России, Родины, Отечества, национальных символов, армии, Победы, территориальной целостности; ЛОЖЬ — при их отсутствии. Поле фиксирует наличие темы, а не её оценку.
8. Гражданская идентичность. Бинарный индикатор гражданской идентичности:

<sup>1</sup> Полный текст промпта представлен в Приложении.

ИСТИНА – при апелляциях к правам и обязанностям гражданина, участию в выборах и институтах, формуле «мы, граждане», ответственности за государство и общество; ЛОЖЬ – при отсутствии таких маркеров.

## Интерпретация результатов

Анализируемый корпус сообщений представляет собой целостный идеологический контур, воспроизводящий официально-патриотический дискурс. Ядро составляют государственные символы (флаг, герб, гимн), героизированное прошлое, поддержка действующей власти, армии и СВО, а также нормативная эмоция гордости за великую страну. В координатах теории ценностей Шварца и модернизационной схемы Инглхарта доминируют кластеры «традиции – конформизм – безопасность» с выборочной интеграцией мотивов достижения и самореализации, но строго в русле одобряемых государством практик (служба, участие в грантах, патриотические акции). В терминах Маслоу преобладают уровни принадлежности и безопасности: субъекту предлагается идентичность, основанная на лояльности и коллективной сплочённости перед внешними и внутренними угрозами.

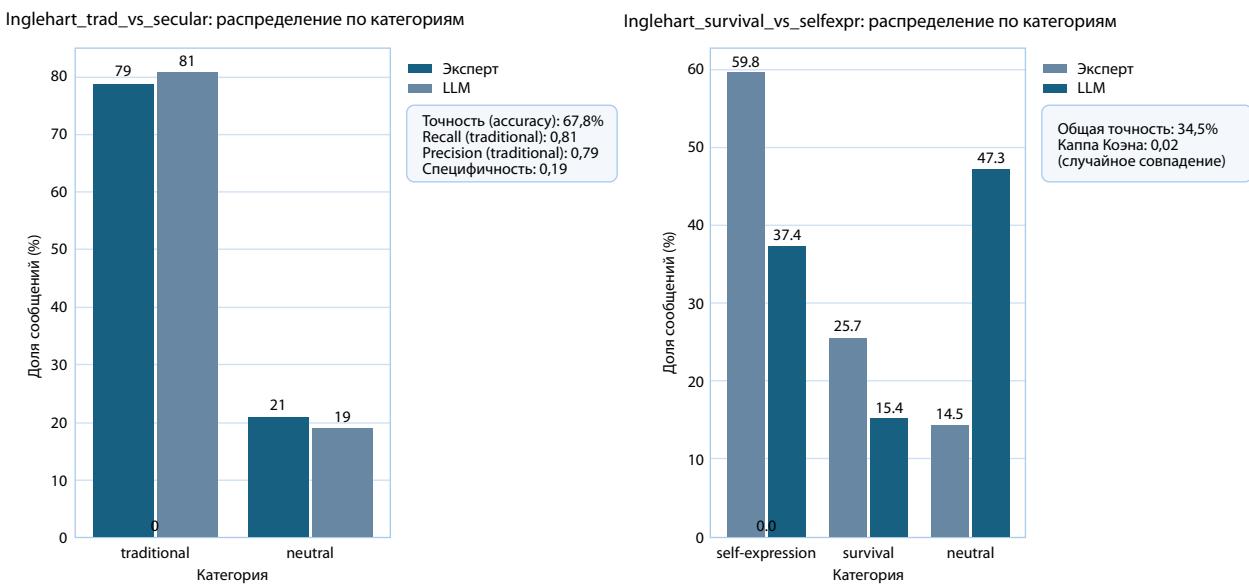
Анализ речевых актов (в рамках типологии Дж. Остина) демонстрирует, что значительная часть высказываний носит не только экспозитивный (объясняющий), но также бехабитивный и экзерситивный характер: через поздравления, декларации гордости, призывы к участию в акциях, поддержку армии, изменение онлайн-статусов и воспроизведение ритуалов пользователи вовлекаются в перформативное подтверждение лояльности. При этом декларируемый патриотизм почти повсеместно присутствует как позитивная норма, тогда как собственно гражданская идентичность (как осознание прав, ответственности и субъектности по отношению к государству) артикулируется слабо. Таким образом, формируется модель «правильного» гражданина как эмоционально

преданного и ценностно традиционного участника коллективного «мы», при этом маргинализируются элементы критической рефлексии.

Экспертная разметка по ценностям Шварца была сопоставлена с разметкой LLM по следующим критериям: доля точных совпадений, распределение классов, матрица ошибок, precision, recall и F1-мера по каждому классу, а также интегральные показатели (macro- и weighted-F1) и коэффициент к Коэна.

Полученные результаты показали, что точность совпадений LLM с экспертом составила 0,408 при  $\kappa = 0,172$ , а macro-F1 = 0,318 и weighted-F1 = 0,485; что позволяет зафиксировать недостаточную согласованность модели с экспертным эталоном. Анализ распределений и матрицы ошибок выявил систематические смещения. В экспертной разметке явно доминирует ценность «Традиции», тогда как модель существенно занижает её долю, одновременно завышая частоту меток «Власть» и ряда других ценностей, которые у эксперта либо редки, либо практически не встречаются. Для класса «Традиции» модель демонстрирует относительно высокую точность при пониженной полноте ( $F1 \approx 0,61$ ), однако для «Достижений» и особенно «Власти» наблюдается выраженное ухудшение показателей за счёт большого числа ложных срабатываний. Сравнение показывает, что без дополнительной настройки языковая модель не воспроизводит структуру экспертных ценностных категорий, недооценивает доминирующие для корпуса ориентации и вводит «избыточные» ценностные интерпретации.

При проведении сравнительного анализа в рамках бинарной кодировки выявляются существенные различия в распределении традиционных и секулярно-рациональных ценностей согласно концептуальной модели Р. Инглхарта (*Inglehart\_trad\_vs\_secular*): сообщения, отнесённые экспертом к категории *traditional*, противопоставлялись всем остальным высказываниям, отнесённым к *neutral*. По экспертной размет-



Ил. 2. Разметка по двум осям модели Инглхарта–Вельцеля  
(традиционные-светско-рациональные; выживание-самовыражение)

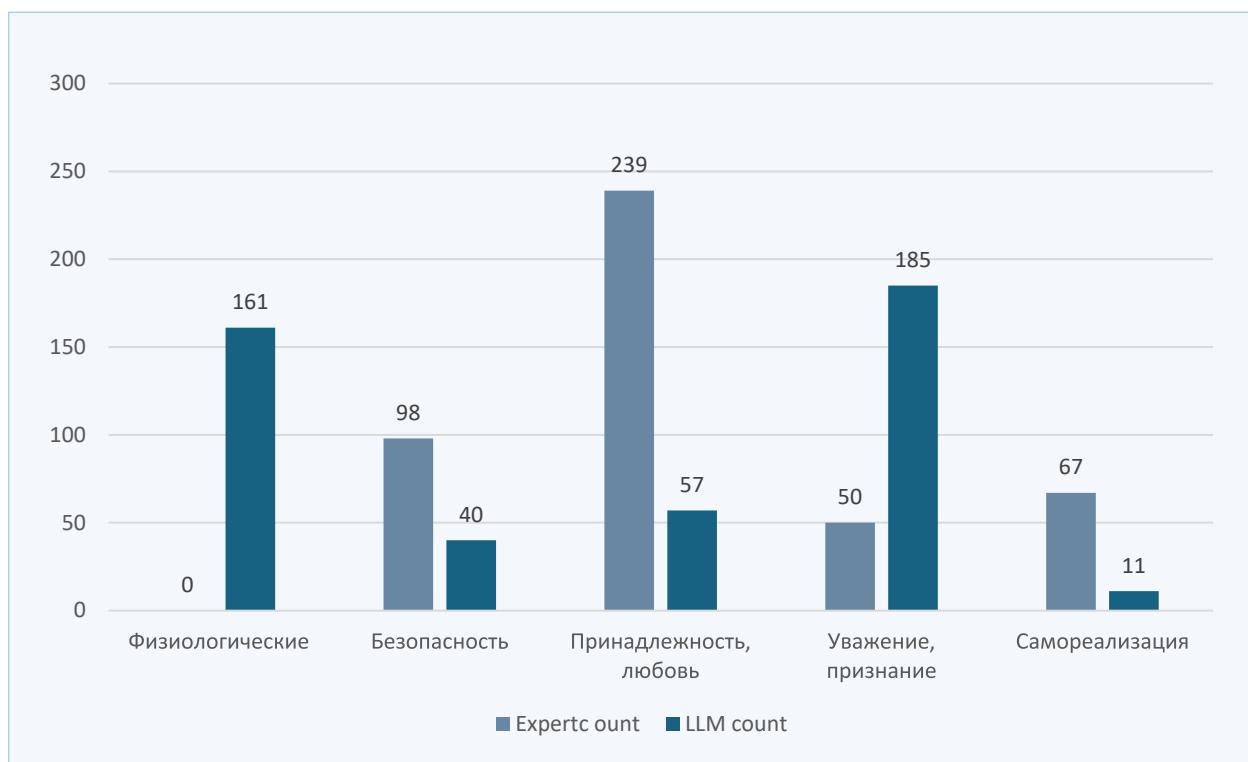
ке 360 сообщений (79%) получили метку *traditional*, 96 (21%) – *neutral*; по разметке LLM 369 сообщений (81%) были классифицированы как *traditional* и 87 (19%) как *neutral*. Матрица ошибок имеет вид:  $TP = 291$ ,  $FN = 69$ ,  $FP = 78$ ,  $TN = 18$ .

Полученные метрики показывают, что при учёте дисбаланса классов согласованность LLM с экспертом по данной шкале минимальна. Номинальная точность составляет 67,8%, при этом чувствительность (recall) к традиционалистским высказываниям высока – 0,81, а точность (precision) – 0,79; однако специфичность крайне низка (0,19), что отражает выраженную тенденцию LLM «перепредсказывать» традиционализм. При выделении традиционалистского полюса модель не воспроизводит экспертную логику дифференциации между *traditional* и *neutral*, а достигаемая точность обусловлена главным образом общей доминантой традиционалистских высказываний в корпусе, а не качественным совпадением критерии.

В процессе разметки по параметру выживания и самовыражения согласно концептуальной модели Р. Инглхарта (*Inglehart\_survival\_vs\_selfexpr*) языковая модель также продемонстрировала

ряд методологически некорректных решений. Экспертная разметка показала, что по шкале *Inglehart\_survival\_vs\_selfexpr* корпус сообщений характеризуется преобладанием ценностей самовыражения: к категории *self-expression* отнесено 59,8% текстов, к *survival* – 25,7%, к нейтральным высказываниям – 14,5%. Разметка LLM дала значимо иное распределение: только 37,4% сообщений были отнесены к *self-expression* и 15,4% – к *survival*, при этом доля *neutral* была завышена до 47,3%. Общая точность составила 34,5%, коэффициент каппа Коэна – 0,02, что указывает на согласие лишь на уровне случайного совпадения. Результаты наглядно представлены на ил. 2.

По доминирующему уровню потребностей (модель А. Маслоу) итоговая точность сопоставления составила 16,3%, коэффициент к Коэна – 0,034, что лишь незначительно превышает случайное совпадение. Развёрнутый анализ показал, что LLM систематически искажает распределение уровней потребностей по сравнению с экспертной оценкой. Категория «Физиологические потребности» в экспертной разметке фактически не использовалась (что логично для текстов с аксиологиче-



Ил. 3. Доминирующий уровень потребностей в сообщениях корпуса (модель А. Маслоу)

ской нагрузкой), тогда как модель присвоила её 161 сообщению, что говорит о существенном смещении в сторону более «низких» уровней. Для уровней «Безопасность» и «Принадлежность/любовь» были получены умеренные значения точности при крайне низкой полноте (например, для «Безопасности»  $precision = 0,575$  при  $recall = 0,235$ ; для «Принадлежность/любовь»  $precision = 0,561$  при  $recall = 0,134$ ), что указывает на выборочную, фрагментарную чувствительность модели к релевантным сигналам. Для «Уважение/признание» и «Самореализации» полнота также остаётся низкой (0,240 и 0,104 соответственно) при точности 0,126 и 0,636, а значительная часть сообщений с экспертной меткой «Принадлежность/любовь» и «Самореализация» модель склонна относить к уровню «Уважение/признание». Результаты визуализированы на ил. 3.

Совпадение предсказаний модели по типологии патриотических высказываний (модель «Ушинский–Соловьёв») с экспертной разметкой по совокупности классов составило 90,6%. Основной вклад в данный

показатель вносит класс «конструктивный патриотизм»: при значительном преобладании этой категории в выборке модель демонстрирует высокие значения полноты и точности именно по этому классу, но практически не различает «агрессивный» патриотизм и отсутствие патриотической проблематики, систематически смещающая их в сторону «конструктивного» патриотизма.

Аналогичное сравнение разметки доминирующего типа речевого акта по схеме Остина показало низкую степень совпадения между LLM и экспертом. Общая точность составила 29,9%, коэффициент к Коэна – около 0,04, что указывает на согласие. При этом модель систематически переиспользует класс бехабитивов, куда попадает значительная доля случаев, которые эксперт относит к экзерситивам и экспозитивам; реже распознаются экспозитивы (высокая точность при крайне низкой полноте) и практически не воспроизводится класс вердиктивов.

Сравнение разметки доминирующего типа речевого акта по схеме Остина про-

водилось между экспертной аннотацией и аннотацией LLM. В анализ были включены только те сообщения, для которых обе разметки использовали допустимые классы Austin: вердиктивы, экзерситивы, комиссивы, бехабитивы, экспозитивы (одна строка с некорректной меткой в экспертной разметке исключена из расчётов). На этой выборке из 455 сообщений была построена матрица несоответствий, рассчитаны точность по каждому классу (precision, recall, F1), общая доля совпадений и коэффициент согласия Коэна.

Сравнение метки бинарного индикатора «патриотизм» дало высокую степень совпадения (что ожидаемо в корпусе данных из ценностного кластера «Гордость и патриотизм»): при доминирующем удельном весе патриотической тематики в корпусе экспертная разметка отнесла к патриотическим 94,7% сообщений, тогда как LLM – 97,8%. Рассчитанные интегральные показатели подтверждают высокую согласованность: точность классификации составила 93,4%, полнота по положительному классу – 98,1%, точность (precision) – 95,1%, F1-мера – 96,6%. Коэффициент к Коэна равен 0,22, что отражает «умеренное» согласие крайне высокой априорной вероятности положительного класса и иллюстрирует эффект смещения распределения: почти полная заполненность корпуса патриотическими упоминаниями делает любую редкую дивергенцию статистически чувствительной.

В то же время экспертная проверка по бинарному индикатору гражданской идентичности показала существенное расхождение между разметкой LLM и экспертной разметкой. По итогам сопоставления были зафиксированы 35 совпадений в случаях присутствия гражданской идентичности (TP) и 239 совпадений в случаях её отсутствия (TN); при этом модель допустила 173 ложных срабатывания (FP), некорректно маркируя высказывания как отражающие гражданскую идентичность, и пропустила девять случаев, отмеченных экспертом (FN). Совокупно это дало точность классифика-

ции 0,60 при высокой полноте по позитивному классу (recall = 0,80) и крайне низкой точности предсказаний гражданской идентичности (precision = 0,17; F1 = 0,28). Полученные результаты позволяют заключить, что LLM демонстрирует выраженную тенденцию к гипердетекции гражданской идентичности, охотно маркируя широкий круг политico-патриотических высказываний, и потому может использоваться лишь на этапе предварительного отбора с обязательной последующей экспертной валидацией.

### Заключение

Сравнительный анализ разметки, осуществлённой в рамках различных теоретических моделей, выявляет статистически значимый разрыв между простыми и сложными типами кодировки.

На бинарном индикаторе патриотизма LLM показывает высокие показатели при умеренном, но систематическом завышении числа патриотических меток; модель надёжно улавливает эксплицитные лексические сигналы («Россия», «Родина», флаг, армия, Победа и т.п.). По мере усложнения задач, связанных с семантической дифференциацией, отмечается прогрессирующее снижение показателей согласованности между автоматизированными выводами и экспертными заключениями.

В итоге по результатам эксперимента можно заключить, что в рамках задач автоматизированного анализа ценностных и патриотических смыслов в русскоязычном пользовательском дискурсе без специальной настройки LLM воспроизводит главным образом общие, поверхностно маркированные семантические категории, но плохо различает более тонкие, теоретически детализированные ориентации и интенции высказываний. На данном этапе выводы модели на сложных шкалах не могут рассматриваться как эквивалент экспертной разметки, что свидетельствует о необходимости разработки и эмпирической проверки специализированных код-буков и процедур человеко-машинной валидации.

**Приложение «Промпт для LLM»**

Твоя задача – выполнять стандартизированную категориальную разметку одного сообщения по фиксированным полям. Работай строго по тексту сообщения. Не используй внешние знания, политические предпочтения, сведения об авторе или площадке.

**Общие правила**

1. Для каждого поля выбери ровно одно значение из допустимого списка.
2. Если предусмотрено значение **neutral**, нет, ИСТИНА/ЛОЖЬ, используй только их, никаких других маркеров отсутствия/наличия.
3. Не изобретай новые категории, не меняй регистр и написание.
4. При неоднозначности выбирай наиболее обоснованную категорию; если ось допускает **neutral/нет**, используй её.
5. Сарказм и иронию учитывай только при явно выраженных маркерах.

**ОПРЕДЕЛЕНИЯ ПОЛЕЙ****1. Schwartz**

Основная ценность, выраженная в сообщении. Выбери одну:

- Власть – сила, влияние, доминирование, статус через контроль.
- Достижения – успех, результат, эффективность, признание заслуг.
- Гедонизм – удовольствие, комфорт, наслаждение.
- Стимуляция – новизна, риск, поиск впечатлений.
- Самонаправленность – автономия, свобода выбора и мнения.
- Универсализм – забота о человечестве, правах, мире, экологии, толерантности.
- Благожелательность – забота о близких и «своих», помочь окружению.
- Традиции – религия, обычаи, «предки», историческая преемственность.
- Конформность – послушание, соблюдение правил и норм.
- Безопасность – порядок, стабильность, защита от угроз.
- Нет – явный ценностный акцент отсутствует.

**2. Inglehart\_trad\_vs\_secular**

Ось «традиционные vs секулярно-рациональные». Выбери одну:

- **traditional** – подчёркнуты религия, традиционные ценности, авторитет государства/армии/семьи, иерархия, долг.
- **neutral** – нет выраженных традиционных индикаторов (и нет устойчивого светско-рационального противопоставления).

(Светско-рациональный полюс здесь не выделяется отдельно, всё нетрадиционное маркируется как **neutral**.)

**3. Inglehart\_survival\_vs\_selfexpr**

Ось «выживание vs самовыражение». Выбери одну:

- **survival** – акцент на выживании, базовой безопасности, страхе, бедности, жёстком порядке.
- **self-expression** – акцент на самореализации, правах, свободе, участии, протесте, креативности, экологии.
- **neutral** – нет достаточных признаков ни выживания, ни самовыражения.

**4. Maslow**

Доминирующий уровень потребностей. Выбери одну:

- Физиологические – еда, физическое выживание, здоровье как базовое условие.

- Безопасность – физическая, экономическая, социальная безопасность, стабильность, защита.
- Принадлежность/любовь – семья, друзья, община, «мы», сплочённость.
- Уважение/признание – статус, уважение, гордость, признание заслуг, в т.ч. гордость за страну.
- Самореализация – смысл, идеалы, долг, развитие, служение, высшие ценности.

Всегда выбирай один из пяти, исходя из доминирующего мотива текста.

### 5. Ushinsky\_Solovyov

Тип патриотизма по мотивам Ушинского и Соловьёва. Выбери одну:

- конструктивный – любовь к стране, гордость, ответственность, моральная аргументация, без вражды к другим народам.
- агрессивный – шовинизм, ненависть/унижение других, культ силы, противопоставление «мы против всех».
- нет – патриотическая риторика отсутствует или не выражена достаточно ясно.

### 6. Austin

Классификация доминирующего речевого акта по Остину. Выбери одну:

- *verdictive* – вынесение оценок и «судов»: осуждение, одобрение, признание, утверждения с характером решения/оценки.
- *exercitive* – высказывания, выраждающие волеизъявление: приказы, призывы, рекомендации, требования, «надо/должны».
- *behabititive* – выражения отношения, установок и реакций: похвала, благодарность, осуждение, выражение уважения/неуважения, поздравления. (В нашей схеме сюда относятся эмоционально-оценочные реакции, связанные с нормами поведения и отношением к другим.)
- *expositive* – высказывания, которые поясняют, аргументируют, структурируют рассуждение, «объясняют позицию», описывают, как понимается ситуация или текст.

Всегда выбери один тип, который лучше всего характеризует форму высказывания. Если сообщение сложное, бери преобладающий.

### 7. Патриотизм

Индикатор наличия патриотической тематики.

- ИСТИНА – если сообщение явно связано с темой Родины, России, Отечества, национальных символов, армии, Победы и т.п.
- ЛОЖЬ – если такой тематики нет.

### 8. Civic\_identity

Индикатор гражданской идентичности.

- Истина – если автор говорит как гражданин: права/обязанности, участие в выборах, институтах, «мы, граждане», ответственность за страну/общество.
- ЛОЖЬ – если таких элементов нет.

(Используй именно ИСТИНА для столбца «Патриотизм» и Истина для civic\_identity, как различающиеся значения.)

Текст сообщения: []

## ЛИТЕРАТУРА

1. Schwartz S.H. Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries // Advances in Experimental Social Psychology. 1992. Vol. 25. Pp. 1–65. [https://doi.org/10.1016/S0065-2601\(08\)60281-6](https://doi.org/10.1016/S0065-2601(08)60281-6).
2. Соловьёв В.С. Патриотизм // Энциклопедический словарь Брокгауза и Ефона. Т. XXIII: Патенты на изобретения — Петропавловский. СПб.: Тип.-лит. И.А. Эфона, 1898. С. 36–38.
3. Соловьёв В.С. О народности и народных делах в России (О соединении Церквей). 1884.
4. Соловьёв В.С. Национальный вопрос в России. Вып. 1. СПб., 1891.
5. Козлова Н.Н., Рассадин С.В. Сетевой дискурс патриотических онлайн-сообществ в современной России: проблемное поле и аксиологические модусы // Вестник Российской университета дружбы народов. Серия: Политология (RUDN Journal of Political Science). 2025. Т. 27, № 3. С. 494–506. DOI: 10.22363/2313-1438-2025-27-3-494-506.
6. Анкудинов И.А. Патриотический дискурс в Рунете: до и после 24 февраля 2022 г. // Мониторинг общественного мнения: экономические и социальные перемены. 2024. № 2. С. 153–177. DOI: 10.14515/monitoring.2024.2.2515. URL: <https://doi.org/10.14515/monitoring.2024.2.2515>.
7. Petrovska I. Typology of civic identity // Current Issues in Personality Psychology. 2022. Vol. 11, № 2. Pp. 150–161. DOI: 10.5114/cipp.2022.116324.
8. World Values Survey Association. Inglehart–Welzel Cultural Map. 2023 version. <https://www.worldvaluessurvey.org/WVSNNewsShow.jsp?ID=467>.
9. Inglehart R., Welzel C. Modernization, Cultural Change, and Democracy. Cambridge: Cambridge University Press, 2005. DOI: 10.1017/CBO9780511790881.
10. Austin J.L. How to Do Things with Words. Oxford: Clarendon Press, 1962. (2nd ed. / Ed. by J.O. Urmson, M. Sbisà. Oxford: Clarendon Press, 1975).
11. Searle J.R. Speech Acts: An Essay in the Philosophy of Language. Cambridge: Cambridge University Press, 1969. DOI: 10.1017/CBO9781139173438.
12. Milkova M., Rudnev M., Okolskaya L. Detecting value-expressive text posts in Russian social media: Препринт, arXiv:2312.08968. 2023. DOI: 10.48550/arXiv.2312.08968.
13. Plotnikov T. Analyzing GPT-4 Misinterpretations of Russian Grammatical Constructions // Lingüística. Revista de Estudos Linguísticos da Universidade do Porto. 2024. Vol. 19. Pp. 157–182. DOI: 10.21747/16466195/ling19a7.
14. Zhang Y., Zou C., Lian Z., Tiwari P., Qin J. SarcasmBench: Towards Evaluating Large Language Models on Sarcasm Understanding: Препринт, arXiv:2408.11319, v1 от 21.08.2024. 2024. DOI: 10.48550/arXiv.2408.11319.
15. Zhou J. An Evaluation of State-of-the-Art Large Language Models for Sarcasm Detection: Препринт, arXiv:2312.03706, v1 от 07.10.2023. 2023. DOI: 10.48550/arXiv.2312.03706.
16. Bojić L., Zagovora O., Zelenkauskaite A. et al. Comparing large language models and human annotators in latent content analysis of sentiment, political leaning, emotional intensity and sarcasm // Scientific Reports. 2025. Vol. 15. Art. 11477. <https://doi.org/10.1038/s41598-025-96508-3>.
17. Gilardi F., Alizadeh M., Kubli M. ChatGPT outperforms crowd workers for text-annotation tasks // Proceedings of the National Academy of Sciences. 2023. Vol. 120, № 30. e2305016120. DOI: 10.1073/pnas.2305016120. URL: <https://www.pnas.org/doi/10.1073/pnas.2305016120>.

## Application of Large Language Models for the Analysis of Value-Patriotic Discourse of Russian-Speaking Users

**Yulia Vladimirovna Balakina** — Candidate of Sciences in Philology, Associate Professor, Professor of the Department of Fundamental and Applied Linguistics of the National Research University Higher School of Economics (HSE–Nizhny Novgorod).

E-mail: julianaumova@gmail.com

**Maria Vladimirovna Grigoryeva** — senior lecturer at the School of Communications at the Faculty of Creative Industries at the Higher School of Economics.

E-mail: mariya.grigoreva@hse.ru

**Ekaterina Nikitichna Sokolova** — Candidate of Sciences in Political Science and the head of the Political Communications Project and Training Laboratory at the School of Communications at the Faculty of Creative Industries at the Higher School of Economics.

E-mail: e.sokolova@hse.ru

The article explores the potential of using large language models (LLMs) for the automated analysis of value-laden patriotic discourse among Russian-speaking social media users. Drawing on a corpus of messages from VK, Odnoklassniki, and Telegram (2023–2025), it investigates the degree of alignment between automated coding results and expert annotations based on a specially developed categorical scheme. The codebook includes eight dimensions: Sh. Schwartz's basic values; R. Inglehart's two axes (traditionalism/secularism and survival/self-expression); A. Maslow's hierarchy of needs; types of patriotism (constructive/aggressive), drawing on the concepts of K.D. Ushinsky and V.S. Solovyov; dominant speech act types per J. Austin; and binary indicators for explicit patriotism and civic identity. The experiment was conducted on the Pride and Patriotism message cluster ( $N = 456$ ), where the density of value markers is highest; the comparison was implemented through error matrices, accuracy, macro/weighted F1, and Cohen's  $\kappa$  coefficient. It was shown that while the LLM reliably identifies explicit patriotic themes, its agreement with experts is significantly lower in multi-class and fine-grained value classification (Schwartz, Maslow, Inglehart scales, types of patriotism, Austin's speech acts). The model demonstrated systematic biases and a tendency to over-diagnose certain categories. It is concluded that LLMs in their current configuration can serve as auxiliary tools for preliminary markup and hypothesis generation but cannot function as an autonomous substitute for expert-led content analysis of value discourse.

**Keywords:** large language models, LLM, values; patriotism, civic identity, content analysis, Schwartz, Inglehart, Maslow, speech acts, social media, Russian-language discourse, automatic markup

## REFERENCES

1. Schwartz S.H. Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries // Advances in Experimental Social Psychology. 1992. Vol. 25. Pp. 1–65. [https://doi.org/10.1016/S0065-2601\(08\)60281-6](https://doi.org/10.1016/S0065-2601(08)60281-6).
2. Solov'yov V.S. Patriotizm // Entsiklopedicheskii slovar' Brokgauza i Efrona. T. XXIII: Patenty na izobreteniya – Petropavlovskii. SPb.: Tip.-lit. I.A. Efrona, 1898. S. 36–38 (in Russian).
3. Solov'yov V.S. O narodnosti i narodnykh delakh v Rossii (O soedinenii Tserkvei). 1884 (in Russian).
4. Solov'yov V.S. Natsional'nyi vopros v Rossii. Vyp. 1. SPb., 1891 (in Russian).
5. Kozlova N.N., Rassadin S.V. Setevoi diskurs patrioticheskikh onlайн-soobshchestv v sovremennoi Rossii: problemnoe pole i aksiologicheskie modusy // Vestnik Rossiiskogo universiteta druzhby

narodov. Seriya: Politologiya (RUDN Journal of Political Science). 2025. Т. 27, № 3. С. 494–506. DOI: 10.22363/2313-1438-2025-27-3-494-506 (in Russian).

6. Ankudinov I.A. Patrioticheskii diskurs v Runete: do i posle 24 fevralya 2022 g. // Monitoring obshchestvennogo mneniya: ekonomicheskie i sotsial'nye peremeny. 2024. № 2. С. 153–177. DOI: 10.14515/monitoring.2024.2.2515. URL: <https://doi.org/10.14515/monitoring.2024.2.2515> (in Russian).
7. Petrovska I. Typology of civic identity // Current Issues in Personality Psychology. 2022. Vol. 11. № 2. Pp. 150–161. DOI: 10.5114/cipp.2022.116324.
8. World Values Survey Association. Inglehart–Welzel Cultural Map. 2023 version. <https://www.worldvaluessurvey.org/WVSNNewsShow.jsp?ID=467>.
9. Inglehart R., Welzel C. Modernization, Cultural Change, and Democracy. Cambridge: Cambridge University Press, 2005. DOI: 10.1017/CBO9780511790881.
10. Austin J.L. How to Do Things with Words. Oxford: Clarendon Press, 1962. (2<sup>nd</sup> ed. / Ed. by J.O. Urmson, M. Sbisà. Oxford: Clarendon Press, 1975).
11. Searle J.R. Speech Acts: An Essay in the Philosophy of Language. Cambridge: Cambridge University Press, 1969. DOI: 10.1017/CBO9781139173438.
12. Milkova M., Rudnev M., Okolskaya L. Detecting value-expressive text posts in Russian social media: Preprint, arXiv:2312.08968. 2023. DOI: 10.48550/arXiv.2312.08968.
13. Plotnikov T. Analyzing GPT-4 Misinterpretations of Russian Grammatical Constructions // Linguística. Revista de Estudos Linguísticos da Universidade do Porto. 2024. Vol. 19. Pp. 157–182. DOI: 10.21747/16466195/ling19a7.
14. Zhang Y., Zou C., Lian Z., Tiwari P., Qin J. SarcasmBench: Towards Evaluating Large Language Models on Sarcasm Understanding: Preprint, arXiv:2408.11319, v1 от 21.08.2024. 2024. DOI: 10.48550/arXiv.2408.11319.
15. Zhou J. An Evaluation of State-of-the-Art Large Language Models for Sarcasm Detection: Preprint, arXiv:2312.03706, v1 от 07.10.2023. 2023. DOI: 10.48550/arXiv.2312.03706.
16. Bojić L., Zagovora O., Zelenkauskaitė A. et al. Comparing large language models and human annotators in latent content analysis of sentiment, political leaning, emotional intensity and sarcasm // Scientific Reports. 2025. Vol. 15. Art. 11477. <https://doi.org/10.1038/s41598-025-96508-3>.
17. Gilardi F., Alizadeh M., Kubli M. ChatGPT outperforms crowd workers for text-annotation tasks // Proceedings of the National Academy of Sciences. 2023. Vol. 120. No. 30. e2305016120. DOI: 10.1073/pnas.2305016120. URL: <https://www.pnas.org/doi/10.1073/pnas.2305016120>.

# ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ КАК СРЕДСТВО ПОЗНАНИЯ И ПОМОЩИ ЧЕЛОВЕКУ И ОБЩЕСТВУ

Индекс УДК 004.8

Код ГРНТИ 06.73.15

DOI: 10.22204/2587-8956-2025-123-04-70-86



А.Р. БАХТИЗИН,  
А.В. БРАГИН<sup>\*</sup>

## Генеративный искусственный интеллект в задачах социально- экономического прогнозирования

В статье рассматривается опыт применения искусственного интеллекта (ИИ) и больших языковых моделей (LLM) для прогнозирования ключевых макроэкономических показателей: валового внутреннего продукта (ВВП), инфляции, уровня безработицы, процентных ставок и коэффициента Джини. Анализируются возможности этих новых подходов по сравнению с традиционными методами прогнозирования, такими как эконометрические модели, равновесные модели и агент-ориентированное моделирование для различных временных горизонтов. В работе обобщаются как академические исследования, так и опыт практических реализаций, включая эксперименты центральных банков по использованию фундаментальных моделей и LLM-подобных архитектур (таких как GPT) для макроэкономического прогнозирования. Особое внимание уделяется способности LLM анализировать текстовую информацию и генерировать прогнозы, сопоставимые по точности, а в некоторых случаях превосходящие оценки профессиональных экспертов. Обзор также охватывает новейшие фундаментальные модели временных рядов, такие как TimeGPT, TimesFM и Moirai, в которых используются трансформерные архитектуры, адаптированные под экономические данные. Основные выводы указывают на то, что ИИ и LLM обеспечивают значительное преимущество с точки зрения гибкости, адаптивности и способности работать с широким спектром источников информации, особенно в условиях высокой волатильности или информационного насыщения. Однако остаются проблемы, связанные с интерпретируемостью, стабильностью и долгосрочной согласованностью прогнозов. В статье делается вывод о том, что наилучшие перспективы для развития макроэкономического прогнозирования связаны с гибридными подходами, которые объединяют вычислительную мощность и адаптивность ИИ с теоретической строгостью и объяснимостью традиционных экономико-математических моделей.

\* **Бахтизин Альберт Рауфович** — доктор экономических наук, член-корреспондент РАН, директор Центрального экономико-математического института (ЦЭМИ) РАН, заведующий кафедрой математических методов анализа экономики экономического факультета МГУ им. М.В. Ломоносова.  
E-mail: albert.bakhtizin@gmail.com

**Брагин Алексей Владимирович** — соискатель ЦЭМИ РАН.  
E-mail: research@alexbragin.com

**Ключевые слова:** экономическое прогнозирование, LLM, GPT, ИИ, большие языковые модели

**М**акроэкономическое прогнозирование традиционно опирается на статистические модели и теоретически обоснованные методы [8]. Обычно используются эконометрические подходы для предсказания динамики ВВП, инфляции, безработицы и других ключевых показателей. Эти инструменты остаются основой планирования экономической политики и корпоративного стратегического планирования.

Развитие искусственного интеллекта, особенно больших языковых моделей [2], существенно расширяет возможности экономического анализа. В отличие от традиционных методов, ИИ-системы могут обрабатывать разнородные источники информации – от числовых временных рядов до новостных сводок и аналитических отчётов. Они выявляют нелинейные взаимосвязи, которые сложно формализовать в рамках классических эконометрических моделей.

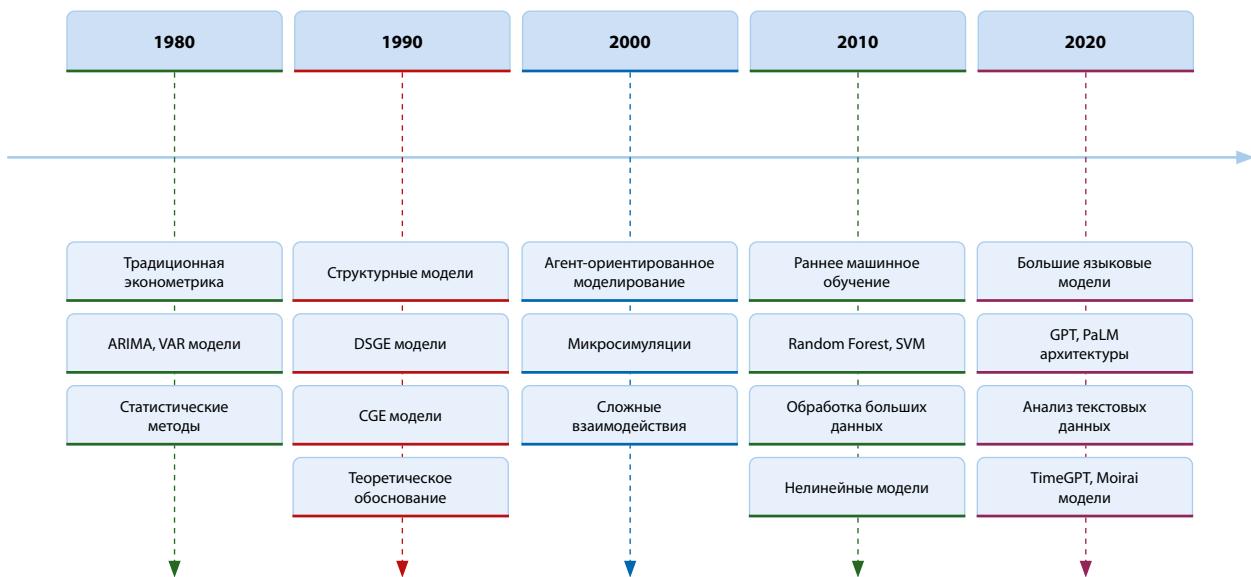
Настоящая работа посвящена обзору современного состояния применения ИИ и больших языковых моделей в области макроэкономического прогнозирования, а также их сравнению с традиционными подходами: агент-ориентированным моделированием, эконометрическими моделями и CGE-моделями. Рассматривается эффективность новых методов прогнозирования на кратко-, средне- и долгосрочном горизонтах, а также проводится анализ современных исследований, инструментов, требований к данным и методов оценки прогнозов.

Быстрое развитие данной области обуславливает необходимость ссылаться не только на статьи, опубликованные в авторитетных научных журналах, но и на препринты таких статей, и на материалы, опубликованные на личных страницах авторов.

## Традиционные подходы к макроэкономическому прогнозированию

Традиционно используемые в макроэкономике методы прогнозирования включают следующие группы:

- Эконометрические модели временных рядов, такие как авторегрессионные интегрированные модели скользящего среднего (ARIMA), векторные авторегрессии (VAR), динамические факторные модели и байесовские VAR-модели, долгое время составляли основу прогнозирования трендов и взаимосвязей макроэкономических показателей. Данные модели строятся на исторических временных рядах и обычно демонстрируют хорошие результаты на краткосрочных прогнозных горизонтах. Простые авторегрессионные модели часто используются в качестве базовых ориентиров (benchmarks), превосходство над которыми должны демонстрировать более сложные подходы, как показано, например, в [10]. Преимущество эконометрических моделей состоит в простоте и хорошей интерпретируемости результатов, однако им свойственны сложности при описании нелинейных зависимостей или резких структурных сдвигов в экономике. В последнее время технологии ИИ также применяются в данном типе прогнозирования, как показано в [9].
- Модели общего равновесия (DSGE, CGE). Для определения траектории дальнейшего развития экономики большие структурные модели, включая такие модели, как динамические стохастические модели общего равновесия (DSGE), CGE и др., опираются на фундаментальные положения экономической теории: уравнения совокупного спроса и предложения, правила денежно-кредитной



Ил. 1. Эволюция методов экономического прогнозирования

**Таблица 1**

**Сравнение традиционных подходов к экономическому прогнозированию**

Подход	Примеры (инструменты и модели)	Горизонт планирования	Преимущества	Ограничения
Эконометрические модели	ARIMA; VAR/BVAR; факторные модели	Краткосрочное планирование (кварталы) и среднесрочное (1–2 года)	Основаны на данных, статистически надёжны; хорошо проверены на исторических данных; демонстрируют стабильные результаты [10]	Модели зачастую линейны и могут упускать сложные взаимосвязи; требуют больших объёмов данных; не предназначены для структурной интерпретации
Структурные модели	DSGE (например, модель ЕЦБ NAWM); большие макромодели (например, МВФ и модели центробанков)	Средне- и долгосрочное планирование (2–5 лет)	Согласованы с теорией и интерпретируемые; эффективны для сценарного анализа и оценки политики [11]	Жёсткая структура может неверно отражать реальность; требуют экспертной настройки; могут уступать в точности краткосрочных прогнозов при новых шоках
Агент-ориентированные модели	Сложные агент-ориентированные модели (например, EURACE и др.)	Краткосрочное и среднесрочное прогнозирование, а также долгосрочное планирование (влияние государственной политики через несколько лет или десятилетий)	Учитывают гетерогенность агентов и нелинейные взаимодействия; позволяют интегрировать детализированные данные и новые типы поведения [12]	Требуют значительных вычислительных ресурсов; сложны для оценки и верификации; всё ещё не часто используются в официальных прогнозах

политики. Эти модели эффективно используются для среднесрочного сценарного анализа, а также для моделирования последствий экономической политики, гарантируя при этом согласованность с экономической ло-

гикой, в том числе соблюдение долгосрочных бюджетных ограничений и равновесных условий. Современные DSGE-модели значительно улучшили свою эмпирическую точность и способны конкурировать со стандартными

Таблица 2

Сопоставление традиционных подходов к прогнозированию и прогнозирования на основе больших языковых моделей<sup>1</sup>

	Традиционное эконометрическое прогнозирование	Экономическое прогнозирование на основе больших языковых моделей
Источники данных	Структурированные данные (официальная статистика, опросы и т.д.), ограниченные по объёму	Большие неструктурированные данные (новости, статьи, финансовые отчёты, социальные сети и т.д.). Возможно обновление в режиме реального времени
Формализация	Как правило, линейные зависимости, предопределённые на основе теоретических гипотез. При этом сложные закономерности могут быть упущены	Большие языковые модели могут самостоятельно выявлять сложные нелинейные взаимосвязи и скрытые закономерности
Реалистичность	Модели редко обновляются и могут давать серьёзные погрешности для нестабильных периодов	Обучение может происходить в режиме реального времени, модели обновляются с использованием новых данных
Интерпретируемость	Прозрачность получаемых результатов — понятные уравнения, рассчитанные с помощью унифицированных методов	Чёрный ящик — сотни млрд необъяснимых параметров
Точность результатов	Хорошо работает в стабильных условиях	Может давать результаты высокой точности, что особенно важно в нестабильные периоды, но возможна галлюцинация; ангажированность
Исследовательский вклад	Определение формы модели, отбор переменных, формулировка гипотез	Минимизация вмешательства человека

временными моделями (например, VAR) в прогнозировании отдельных переменных. Также они особенно эффективны для анализа различных сценариев типа «что будет, если» [11]. Однако их точность снижается при неверной спецификации теоретических предпосылок, и в периоды нестандартных экономических условий (финансовые кризисы или пандемия, не предусмотренные исходной моделью) они могут демонстрировать более слабые результаты на коротких горизонтах.

- Агент-ориентированные модели (Agent-Based Models). Агент-ориентированное моделирование (АОМ) основывается на воспроизведении экономических процессов «снизу-вверх» путём симуляции поведения отдельных агентов (домохозяйств, фирм, банков) и их взаимодействий [7]. Такой подход способен отражать возникающие (эмурджентные) эффекты и нелинейные динамические процессы, которые часто оказываются

недоступными для классических структурных моделей. Изначально АОМ были крайне ресурсоёмкими и применялись преимущественно в исследовательских целях, а не для регулярного прогнозирования. Тем не менее последние исследования значительно приблизили АОМ к практическому использованию, а целый ряд работ за последнее десятилетие показывает превосходство результатов, полученных с помощью АОМ, над результатами базовых моделей, таких как VAR, DSGE и CGE в рамках тестирования вне выборочных данных [2–8]. Богатая внутренняя структура АОМ позволяет получать детализированные прогнозы по секторам, использовать модель для стресс-тестирования и анализа последствий изменений политики. Однако практическое применение таких моделей зачастую бывает затруднено необходимостью сбора большого количества данных для калибровки и сложностью их оценки, что до сих пор ограничива-

<sup>1</sup> Составлено на основе материалов портала MASEconomics: <https://maseconomics.com/artificial-intelligence-in-economic-forecasting-and-analysis/>.

ет широкое внедрение АОМ в практику центральных банков и других институтов экономического прогнозирования.

На ил. 1 показана эволюция появления новых методов экономического прогнозирования по десятилетиям начиная с 1980 г.

В таблице 1 приведены сравнительные характеристики различных подходов к экономическому прогнозированию, а в таблице 2 – технологические особенности эконометрики и больших языковых моделей.

Таким образом, каждый из традиционных подходов имеет собственную область применения: эконометрические модели хорошо подходят для краткосрочных прогнозов при относительно стабильных трендах; агент-ориентированные модели цепны для исследования сложных феноменов и выявления эмерджентных событий. Тем не менее данные методы нередко сталкиваются с компромиссами между гибкостью и интерпретируемостью. Именно это обстоятельство обусловило интерес исследователей к технологиям искусственного интеллекта, способным выявлять закономерности непосредственно из данных, не опираясь на жёстко заданные априорные ограничения.

### Появление искусственного интеллекта и машинного обучения в экономическом прогнозировании

В последние десять лет методы искусственного интеллекта и машинного обучения (ML) всё шире применяются в макроэкономическом прогнозировании. На ранних этапах развития данного направления использовались такие подходы, как регрессионные деревья (Regression Trees), случайные леса (Random Forests), градиентный бустинг (Gradient Boost) и нейронные сети (Neural Networks), целью которых являлось повышение точности прогнозов по сравнению с традиционными методами. Такие модели способны эффективно обрабатывать данные высокой размерности, включая потенциальные предикторы, к примеру глобальные цены на сырьевые товары, финансовые показатели, тренды поисковых запросов

в Google и т.д., а также автоматически выявлять сложные нелинейные зависимости между ними.

Например, в одном из исследований для прогнозирования квартальных изменений показателей неравенства доходов (индексов Джини) в Великобритании был применён алгоритм L2-буста (L2-boosting), учитывающий широкий спектр макроэкономических и финансовых параметров [13]. Этот подход позволил эффективно отобрать значимые переменные и обеспечить более высокую точность вне выборочных прогнозов по сравнению с традиционными низкоразмерными моделями благодаря использованию более полного объёма доступной информации. Это подчёркивает важное преимущество методов машинного обучения: способность анализировать большое количество входных данных и выявлять сложные взаимодействия предикторов, которые обычно остаются незамеченными в стандартных моделях.

В целом модели машинного обучения, такие как ансамбли деревьев решений или нейронные сети прямого распространения, демонстрируют определённые улучшения точности краткосрочных прогнозов по ряду макроэкономических переменных и задач оперативного прогнозирования (также зачастую называемых прогнозированием «наукастинга», nowcasting). Особую ценность такие методы приобретают в условиях нелинейности зависимостей или при работе с неструктуризованными данными. Например, системы nowcasting, построенные на основе методов машинного обучения, могут интегрировать источники «больших данных» (новостные потоки, социальные сети, спутниковые снимки) вместе с официальной статистикой, обеспечивая оперативные оценки текущей экономической активности [15]. Подобные подходы активно тестируются и применяются центральными банками и исследовательскими организациями (например, Банком Англии, Международным Валютным Фондом) как дополнение к стандартным инструментам прогнозирования.

Однако эффективность применения методов ИИ сильно зависит от качества и объёма используемых данных. Макроэкономические временные ряды зачастую ограничены относительно небольшими выборками, особенно на уровне отдельных стран, и характеризуются структурными разрывами, что может приводить к переобучению или неустойчивости модели. Исследователи решают эти проблемы посредством объединения данных, например используя панельные данные по множеству стран или секторов, а также фокусируясь на задачах оперативного прогнозирования nowcasting, где доступно большое количество высокочастотной информации.

В результате модели на основе искусственного интеллекта и машинного обучения становятся всё более важным дополнением к традиционным прогнозам. Чаще всего они используются в комбинации с классическими эконометрическими моделями или в качестве дополнительного входного сигнала, а не как их прямая замена. Самым современным направлением развития данной области является использование больших языковых моделей [2], обладающих принципиально новыми возможностями анализа текстовой информации.

### Большие языковые модели в макроэкономическом прогнозировании

На данный момент большие языковые модели, такие как GPT-4 компании OpenAI или PaLM от Google, а также аналогичные трансформерные архитектуры, только начинают применяться для решения задач прогнозирования экономических показателей. Подобные модели обучаются на огромных массивах текстовых данных (новостях, книгах, статьях, веб-контенте), благодаря чему аккумулируют обширные мировые знания и богатый лингвистический контекст. Эти особенности LLM открывают два основных направления их применения в макроэкономическом прогнозировании.

1. Использование универсальных языковых моделей как прогнозистов путём prompt-инженерии (запросов с инструкциями) или дообучения на предметных данных.
2. Создание специализированных фундаментальных моделей для временных рядов на основе архитектур языковых моделей.

Рассмотрим последние достижения по обоим направлениям подробнее.

- **Использование LLM в качестве универсальных прогнозистов на основе промптов.** Исследователи стали напрямую использовать большие языковые модели для прогнозирования экономических показателей, фактически рассматривая их как виртуальных экспертов, усвоивших экономическую историю. Яркий пример – исследование Федерального резервного банка Сент-Луиса, использовавшего модель Google PaLM для прогнозирования инфляции в США. PaLM генерировала прогнозы инфляции за период 2019–2023 гг. при различных заданных сценариях, после чего её результаты сравнивались с консенсусом профессиональных прогнозистов (Survey of Professional Forecasters, SPF). Оказалось, что прогнозы PaLM имели меньшую среднеквадратическую ошибку (MSE), чем прогнозы профессионалов практически на всех временных горизонтах.

Интересным оказалось и то, что прогнозы PaLM медленнее возвращались к цевому показателю инфляции в 2%, что свидетельствует о меньшей склонности модели к необоснованным оптимистичным предположениям о быстром восстановлении экономики после инфляционных шоков. Подобные эксперименты проводил и Чешский национальный банк (ČNB), применяя модель на основе GPT от OpenAI для прогнозирования инфляции в Чехии на один год вперёд. В результате LLM продемонстрировала лучшую точность (RMSE 5,28) по сравнению с основной структурной моделью банка (RMSE 5,48), а также с прогнозами финансовых аналитиков.

Особенно выраженным преимуществом модели стало раннее распознавание устойчивого тренда к росту инфляции в 2021–2022 гг. Эти исследования демонстрируют, что при грамотном конструировании запросов (например, требуя от модели игнорировать информацию о будущем для избежания эффекта *hindsight bias*) LLM способны давать сравнительно точные прогнозы. Преимуществом таких подходов является способность моделей мгновенно использовать обобщённый опыт миллионов людей и широкий спектр данных без привязки к какой-либо конкретной экономической теории. Однако их главный недостаток – это отсутствие интерпретируемости и риск генерации экономически необоснованных прогнозов («галлюцинаций»). Поэтому центральные банки склонны рассматривать LLM скорее как дополнительный инструмент, нежели как замену традиционных методов.

- **Прогнозы на основе новостей.** Интересным применением LLM является их использование в роли «искусственных агентов», формирующих ожидания на основе новостного потока, аналогично поведению реальных экономических агентов. Например, в [17] был создан искусственный экономический опрос: на основе статей из The Wall Street Journal за период 1984–2021 гг. модели GPT-3.5 попросили в реальном времени спрогнозировать макроэкономические и финансовые показатели (рост ВВП, инфляцию, доходность акций). Полученные прогнозы LLM близко совпали с реальными опросными данными SPF и других индикаторов настроений, воспроизведя даже типичные для людей поведенческие искажения: недостаточную реакцию на новую информацию и чрезмерную экстраполяцию прошлых тенденций на будущие доходности акций.

Важно, что при тестировании модели вне периода обучения она продолжала показывать хорошие результаты, что подтверждает способность LLM не только запоминать, но и обобщать знания. Подобные

подходы открывают возможности интеграции текстовой информации (новостные потоки, сентимент-индексы) в традиционные экономические модели для улучшения текущих оценок и анализа рисков.

- **Модели для прогнозирования временных рядов (TSLMs).** Параллельно с развитием универсальных LLM исследователи начали создавать специализированные фундаментальные модели для прогнозирования временных рядов (Time-Series Foundation Models, TSLM). Эти модели используют архитектуру и методы обучения языковых моделей, но адаптированы для прогнозирования числовых последовательностей. Примеры таких моделей – LagLlama (2024) [20], TimesFM (Google, 2024), Moirai (Salesforce, 2024), TinyTimeMixers (IBM, 2024), TimeGPT (NVIDIA, 2023), а также академические разработки типа Time-LLM [21].

Подобные системы предварительно обучаются на огромных объемах временных рядов из различных областей (экономика, энергетика, транспорт), благодаря чему способны осуществлять zero-shot прогнозирование, т.е. прогнозирование новых рядов без необходимости дополнительного обучения [14].

Эта идея аналогична языковым моделям: благодаря экспозиции к большому набору различных текстовых шаблонов модель учится возможности экстраполяции в целом. Ключевым преимуществом является то, что одна модель может применяться к тысячам различных рядов, не требуя отдельных моделей для каждого показателя.

Первые тесты показали, что такие модели, как Moirai от Salesforce и TimesFM от Google, демонстрируют конкурентоспособные результаты относительно стандартных эконометрических бенчмарков. Для облегчения использования этих моделей такие библиотеки, как sktime, уже внедрили удобные API-интерфейсы, что снижает порог входа для аналитиков.

Несмотря на оптимизм, недавние строгие оценки эффективности прогнозирования с использованием LLM-моделей при-

зывают к умеренности в ожиданиях. Так, в [10] была проведена всесторонняя сравнительная оценка нескольких LLM и традиционных подходов на широкой базе макроэкономических данных (FRED-MD, США).

Авторы оценивали пять ведущих моделей временных рядов на основе LLM (включая Moirai и TimesFM) по сравнению с бенчмарками, такими как авторегрессионные (AR) модели, байесовские векторные авторегрессии (BVAR) и факторные модели. Результаты показали, что лишь немногие из LLM-моделей смогли стабильно превосходить простую AR-модель, и даже эти единичные случаи не обеспечили устойчивого превосходства над современными эконометрическими методами.

Более того, точность лучших LLM-моделей оказалась «в целом сопоставимой, а иногда даже несколько уступающей» по сравнению с VAR и факторными моделями для многих макроэкономических серий. Кроме того, LLM-методы продемонстрировали меньшую стабильность: они могли выдавать отличные результаты по отдельным временным рядам или в определённые периоды, но также порой генерировали заведомо ошибочные и экономически маловероятные прогнозы — подобной волатильности не наблюдается у более устойчивых традиционных моделей.

Такая ненадёжность частично схожа с проблемой «галлюцинаций» у языковых моделей, когда они могут выдавать крайне неточные предсказания, не соответствующие историческим данным. Интересно, что в исследовании отмечено: LLM-модели показали некоторое улучшение точности в постковидный период, возможно, благодаря включению пандемийной информации в их обучающие выборки. Однако это одновременно вызывает беспокойство относительно утечки информации о будущем (data leakage), если в обучении использовались данные, не соответствующие прогнозному горизонту.

Попытки дообучения LLM-моделей на макроэкономических данных не привели к значимому улучшению прогностической точности, что может свидетельствовать либо о том, что предварительно обученные модели уже достигли своего предела в данной задаче, либо о том, что объём доступных макроэкономических данных недостаточен для эффективной адаптации таких крупномасштабных нейросетевых архитектур.

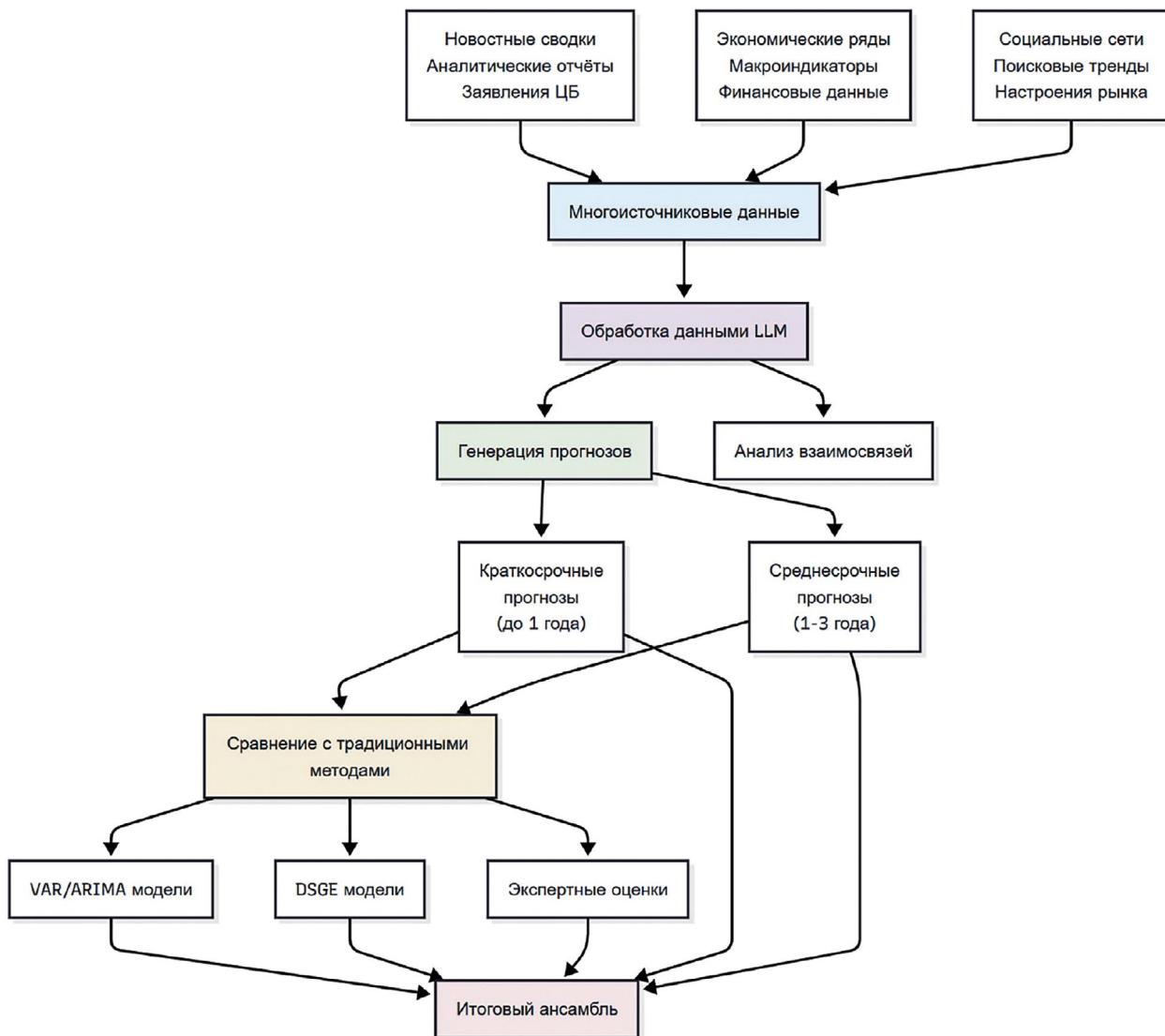
На ил. 2 приведена общая схема процесса макроэкономического прогнозирования с использованием больших языковых моделей.

### Инструменты с открытым исходным кодом

В настоящее время формируется активно растущая экосистема инструментов, которые делают технологии прогнозирования на основе искусственного интеллекта и больших языковых моделей доступными широкому кругу исследователей и аналитиков. Как уже упоминалось выше, библиотека sktime предлагает унифицированный API-интерфейс для удобного применения фундаментальных моделей прогнозирования временных рядов. Компании, такие как Nixtla<sup>1</sup>, запускают сервисы наподобие TimeGPT — фундаментальной модели временных рядов, к которой пользователи могут обращаться через API для генерации прогнозов в различных областях (финансы, розничная торговля и др.). Пользователи могут дообучать TimeGPT на собственных данных или использовать её непосредственно для zero-shot прогнозирования без дополнительного обучения.

Параллельно этому появляются специализированные LLM-модели для финансовой сферы (например, BloombergGPT) и открытые модели (например, FinGPT), которые, хотя изначально и предназначены для задач обработки естественного языка (NLP), могут потенциально приме-

<sup>1</sup> <https://docs.nixtla.io>.



Ил. 2. Схема процесса прогнозирования с использованием больших языковых моделей

няться для прогнозирования и сценарного анализа путём объединения текстовых финансовых новостей с числовыми данными [22].

Многие центральные банки уже активно публикуют программный код или, по крайней мере, делятся своим опытом. Например, Федеральный резервный банк Сент-Луиса и Чешский национальный банк детально описали методологию составления запросов (промптов), что позволяет другим исследователям воспроизвести их результаты с помощью API OpenAI и аналогичных сервисов.

Таким образом, как научное сообщество, так и коммерческий сектор быстро

развивают инфраструктуру, необходимую для широкого применения моделей прогнозирования на основе ИИ, снижая тем самым барьер входа для аналитиков, желающих экспериментировать с передовыми методами наряду с традиционными инструментами.

### Эффективность прогнозов на кратко-, средне- и долгосрочных горизонтах

Эффективность применения моделей на основе искусственного интеллекта и больших языковых моделей (AI/LLM), по сравнению с традиционными подходами, существенно зависит от прогнозного горизонта:

1. **Краткосрочные прогнозы (до одного года).** В краткосрочном прогнозировании преимущество имеют модели, ориентированные на обработку высокочастотных данных и оперативных сигналов о текущей динамике. Именно здесь методы на основе ИИ демонстрируют наибольший потенциал, поскольку они способны интегрировать более широкий набор данных. Например, прогнозы инфляции на основе LLM на ближайшие несколько кварталов оказались существенно точнее экспертных консенсус-прогнозов в период резких экономических изменений 2021–2023 гг. Модели nowcasting, построенные с использованием машинного обучения, способны обнаруживать экономические циклы раньше квартальных эконометрических моделей благодаря использованию различных индикаторов, включая индексы настроения, вычисляемые на основе анализа новостей и интернет-трендов. Тем не менее традиционные временные модели при грамотной калибровке остаются весьма конкурентоспособными в краткосрочной перспективе. Например, Карриеро и соавторы (2025) показали, что простая авторегрессионная (AR) модель оказалась устойчивым конкурентом на горизонте одного квартала вперёд, и лишь наиболее продвинутые LLM-модели периодически демонстрировали небольшие преимущества по отдельным переменным. Краткосрочное прогнозирование также является областью, где особенно эффективны ансамблевые подходы. Многие центральные банки и аналитические агентства всё чаще комбинируют прогнозы (например, усредняя результаты VAR-моделей, языковых моделей и экспертных оценок), повышая устойчивость конечных прогнозов. В целом на горизонтах от 0 до 12 месяцев модели на основе ИИ и LLM способны обеспечить улучшение точности и адаптивность прогнозов (особенно в необычных экономических условиях), однако наилучшие результаты достигаются при использовании их

совместно с проверенными статистическими моделями, а не вместо них.

2. **Среднесрочные прогнозы (от 1 до 3 лет).** На среднесрочном горизонте задача простого экстраполирования исторических данных становится более сложной, поскольку начальные условия постепенно уступают место структурным изменениям в экономике. На таких временных интервалах значительную роль играют структурные факторы: экономическая политика, ограничения производственных мощностей, демографические тенденции. Традиционные структурные модели и экспертные оценки широко используются центральными банками и международными институтами именно на горизонте 2–3 лет. Например, прогнозы инфляции центральных банков на этот период строятся на комбинации базовых структурных моделей и экспертных суждений.

Тем не менее модели на основе ИИ и LLM также активно тестируются на среднесрочном горизонте. Например, эксперимент Чешского национального банка с годовым прогнозом инфляции показал, что LLM демонстрирует среднюю точность на уровне структурной модели, причём каждый подход имел свои преимущества [15]. Языковая модель оказалась менее привязана к целевому уровню инфляции (2%), что помогло ей успешно улавливать устойчивый рост инфляции в период её ускорения, но одновременно приводило к завышению прогнозов при её замедлении. Это говорит о том, что модели на основе ИИ могут успешно выявлять смену экономических режимов или новые тренды на среднесрочном горизонте, поскольку они не ограничены ожиданиями о возврате к историческому равновесию. Традиционные же модели оказываются эффективнее в условиях восстановления исторических взаимосвязей и равновесных условий.

Важно также отметить, что LLM способны учитывать качественные изменения внешней среды, такие как ожидаемая

смена экономической политики или геополитические события, о которых модель может «знать» из текстовых данных, тогда как эконометрические модели способны это сделать лишь при явном их включении. Однако среднесрочные прогнозы ИИ-моделей могут терять внутреннюю экономическую логику и иногда демонстрировать неоправданные тренды (например, продолжать ускорение роста экономики слишком долго). Обеспечить внутреннюю согласованность таких прогнозов сложнее, так как в самих ИИ-моделях нет встроенных механизмов, увязывающих инфляцию, процентные ставки и госдолг в единую систему. В связи с этим исследуются подходы, включающие теоретическую регуляризацию моделей или интеграцию прогнозов ИИ с традиционными структурными моделями. Таким образом, оптимальной стратегией на горизонте 1–3 лет является сочетание преимуществ обоих подходов, например использование структурной модели в качестве базового сценария с коррекцией на основе прогнозов ИИ-моделей, сигнализирующих о возможных отклонениях.

1. **Долгосрочные прогнозы (свыше 5 лет).** Прогнозирование на длительных горизонтах теряет характер точечного предсказания и становится задачей сценарного анализа. На таких временных отрезках доминирует глубокая неопределенность, а структурные изменения делают невозможными надёжные «точечные» прогнозы любой модели, в том числе и ИИ-моделей. Традиционные подходы опираются на экстраполяцию долгосрочных трендов (например, потенциального ВВП или демографических показателей), теоретически согласованные сценарии, DSGE-модели, моделирующие возврат экономики к равновесию после шоков, или агент-ориентированные модели, изучающие влияние технологических изменений на неравенство в десятилетней перспективе.

Методы прогнозирования с использованием ИИ и LLM на таких горизонтах находятся в ранней стадии развития. Большие

языковые модели могли бы использоваться для генерации *нарративных сценариев* (например, описание реалистичных экономических условий в 2030 г.), однако их применение для численных прогнозов на столь долгие периоды пока носит исключительно экспериментальный характер. Поскольку LLM не обладают встроенным пониманием структурных ограничений (например, ресурсов или устойчивости госдолга), их долгосрочные прогнозы могут оказаться нереалистичными. Более реалистичным представляется использование ИИ и LLM для калибровки и обработки данных для структурных или агентных моделей либо анализа гетерогенных сценариев.

Однако в отношении таких показателей, как коэффициент Джини через 10 лет, лучшее, что можно сделать, – это сценарный анализ. Примером служат симуляции типа Пикетти или микросимуляционные модели неравенства. Методы машинного обучения уже используются для улучшения таких подходов: например, бустинговая регрессия позволяет более точно прогнозировать распределение доходов при различных условиях; однако эти методы также опираются на определённые структурные предположения.

Таким образом, в долгосрочной перспективе искусственный интеллект и LLM следует рассматривать скорее как вспомогательные инструменты (анализ данных, генерация сценариев), а не как самостоятельные «чёрные ящики» для прогнозирования. В современной экономической литературе преобладает мнение, что экспертные знания и теоретические конструкции остаются незаменимыми для направления долгосрочных прогнозов, а искусственный интеллект выступает в роли важного дополнения и источника полезной информации.

### Эффективность и ограничения моделей на основе ИИ и LLM

Современные исследования дают сложную и многогранную картину эффективности применения искусственного интеллек-

та и больших языковых моделей в задачах макроэкономического прогнозирования. С одной стороны, такие модели обладают рядом очевидных преимуществ: они способны обрабатывать огромные объёмы информации, включая неструктурированные текстовые данные, моделировать сложные нелинейные взаимодействия между переменными и адаптироваться к новым паттернам в данных без необходимости явного перепрограммирования.

В ряде случаев модели на основе ИИ уже продемонстрировали способность соперничать с традиционными подходами или превосходить их, например для задач прогнозирования инфляции в условиях высокой волатильности [15], а также при синтезе информации, аналогичной той, что содержится в экспертных оценках профессиональных экономистов [17]. Доступность предварительно обученных моделей через API-интерфейсы означает, что даже небольшие организации получают возможность экспериментировать с передовыми методами ИИ, тем самым потенциально снижая барьеры входа и создавая более равные условия в сфере прогнозной аналитики [18].

С другой стороны, остаются и существенные ограничения. Одним из постоянно отмечаемых недостатков является сама природа таких моделей – «чёрный ящик». Иначе говоря, большая языковая модель может спрогнозировать, например, резкий рост безработицы, но при этом не дать внятного объяснения причин, тогда как структурная модель позволила бы увязать данный результат с падением инвестиций через соответствующие уравнения модели. Отсутствие интерпретируемости затрудняет использование прогнозов ИИ в государственном управлении в силу сложности с обоснованием достоверности прогнозов, которые нельзя верифицировать логически или теоретически.

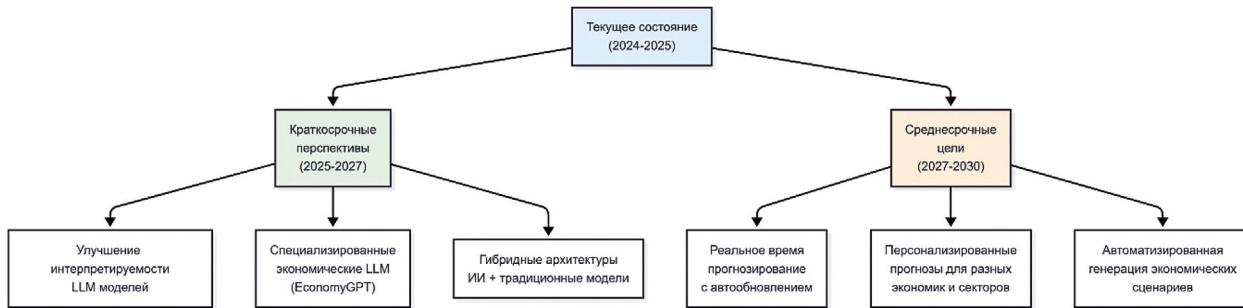
Дополнительными проблемами являются стабильность и последовательность. Традиционные подходы к прогнозированию накладывают определённые эконо-

мические ограничения (например, инфляция не может оставаться отрицательной бесконечно долго, государственный долг в долгосрочной перспективе должен быть обслужен и т.д.). Модели же на основе ИИ, не имея встроенных теоретических ограничений, могут генерировать сценарии, нарушающие элементарные экономические принципы. Результатом такой работы становятся так называемые необоснованные прогнозы [10].

Кроме того, модели на основе ИИ зачастую характеризуются высокой ресурсоёмкостью данных. Для малых экономик или нишевых индикаторов (например, коэффициент Джини, рассчитываемый лишь раз в год) просто может не оказаться достаточного объёма данных для эффективного обучения высокоёмкой модели. В условиях ограниченного объёма информации возрастает риск переобучения и генерации ложных сигналов, особенно если модель подстраивается под уникальные, но непредсказуемые шаблоны, не обладающие прогностической значимостью.

*Налаживание взаимодействия человека и ИИ* является наиболее перспективным направлением развития. Экономисты могут использовать большие языковые модели для оперативного тестирования множества альтернативных вариантов прогнозных моделей или для автоматизированного анализа тысяч новостных сообщений с целью выявления потенциальных точек разворота в экономике – задач, практически невозможных для ручного выполнения в сжатые сроки.

Исследователи, работающие с агент-ориентированными моделями, могут применять ИИ для калибровки параметров поведения агентов на основе наблюдаемых данных. Центральные банки и экономические институты активно изучают гибридные архитектуры, при которых ИИ-модель генерирует первоначальный прогноз, а затем эксперт вносит корректировки, используя собственное суждение и дополнительную информацию, недоступ-



Ил. 3. Возможные варианты развития экономического прогнозирования с использованием ИИ

ную модели (например, предстоящие изменения политики, которые LLM, в силу отсутствия истинного понимания, может не учесть должным образом).

Кроме того, активно развиваются направления, предполагающие интеграцию экономического знания в структуру ИИ-моделей. Например, предпринимаются попытки встроить в архитектуру нейросетей бухгалтерские тождества, бюджетные ограничения или базовые экономические законы с целью исключения грубых нарушений логики модели. Такие усилия могут привести к созданию моделей, которые одновременно обладают высокой гибкостью и внутренней экономической согласованностью.

## Заключение

Применение искусственного интеллекта и больших языковых моделей в сфере макроэкономического прогнозирования представляет собой очень перспективное направление, находящееся на стыке экономической науки и анализа данных. Современные модели, такие как фундаментальные нейросетевые архитектуры для временных рядов и прогнозные системы на базе GPT, уже продемонстрировали способность выявлять закономерности, которые могут быть упущены традиционными методами, и делать это по широкому спектру макроэкономических показателей начиная от ВВП и инфляции и заканчивая неравенством доходов.

Особенно высока их эффективность в условиях необходимости обработки раз-

нородных источников данных и быстрой адаптации к смене экономических режимов, что критически важно в современной, стремительно изменяющейся макроэкономической среде. Однако такие модели не являются универсальным решением. Традиционные подходы к прогнозированию, опирающиеся на десятилетия эмпирического опыта и теоретического анализа, сохраняют свою высокую значимость прежде всего за счёт логической согласованности, интерпретируемости и способности формировать структурное объяснение числовых результатов.

Сравнительный анализ методов на основе ИИ/LLM, агент-ориентированных моделей, эконометрических инструментов и систем уравнений ясно демонстрирует: каждый подход обладает своими сравнительными преимуществами. Современная граница исследований пролегает в области интеграции этих методов с целью объединения их сильных сторон. Так, например, агент-ориентированная модель может быть использована для генерации набора долгосрочных сценариев распределения доходов, а ИИ-модель – для прогнозирования краткосрочных отклонений от этих сценариев на основе данных в реальном времени. Или же базовый прогноз ВВП, построенный на основе эконометрической модели, может быть скорректирован языковой моделью, анализирующей новостные сообщения о сбоях в цепочках поставок или геополитических событиях.

Центральные банки, такие как уже упомянутый выше Чешский национальный

банки, прямо указывают, что рассматривают модели ИИ не как замену, а как способ *дополнения и расширения* существующих аналитических инструментов [15]. Именно такая комбинированная, синтетическая перспектива, основанная на интеграции дисциплин, сегодня представляет собой наиболее перспективный путь развития макроэкономического прогнозирования.

На ил. 3 приведены возможные варианты развития применения ИИ в экономическом прогнозировании.

В дальнейшем можно ожидать появление всё большего числа открытых платформ для макроэкономического прогнозирования, интегрирующих большие языковые модели, специализированные архитектуры (возможно, модели типа «EconomyGPT», обученные на макроэкономических данных и учебной литературе), а также всё более строгой процедуры валидации таких инструментов в рамках реальных прогнозных соревнований. Международные организации, такие как МВФ и ОЭСР, с высокой вероятностью будут инвестировать ресурсы в оценку потенциала ИИ для улучшения глобальных прогнозов экономического развития. Корпорации, включая профессиональных прогнозистов и финансовые компании,

продолжат внедрять закрытые ИИ-решения с ограниченным доступом к исходному коду, ориентированные на рыночные показатели (процентные ставки, фондовые индексы и др.).

Сегодня к передовой практике относятся такие примеры, как использование модели PaLM для прогнозирования инфляции, GPT-4 – для анализа новостей, связанных с ростом ВВП, а также гибридные системы, объединяющие «передовые LLM в экономических исследованиях» для решения разнообразных задач [19]. В совокупности всё это свидетельствует о наступлении новой эпохи в макроэкономическом прогнозировании, характеризующейся богатыми данными, агностичностью по отношению к конкретным моделям и служащей дополнением к классическим теоретически обоснованным и модельно-ориентированным подходам.

Наиболее успешными прогнозными системами ближайшего будущего, по-видимому, станут те, что смогут эффективно объединять экономическую интуицию человека с вычислительной мощностью машины, обеспечивая прогнозы, которые будут не только точными, но интерпретируемыми и практически значимыми для принятия решений.

## ЛИТЕРАТУРА

1. Брагин А.В., Бахтизин А.Р. Особенности реализации больших экономических моделей // *π-Экономисту*. 2023. № 16 (3). С. 107–122. DOI: <https://doi.org/10.18721/JE.16307>.
2. Брагин А.В., Бахтизин А.Р., Макаров В.Л. Большие языковые модели четвёртого поколения как новый инструмент в научной работе // Искусственные общества. 2023. Т. 18, № 1.
3. Брагин А.В., Бахтизин А.Р., Макаров В.Л. Современные программные средства агент-ориентированного моделирования // Искусственные общества. 2022. Т. 17, № 4.
4. Макаров В.Л., Бахтизин А.Р., Сушко Е.Д., Агеева А.Ф. Искусственное общество и реальные демографические процессы // Экономика и математические методы. 2017. Т 53, № 1. С. 3–18.
5. Макаров В.Л., Бахтизин А.Р., Сушко Е.Д. Имитация особенностей репродуктивного поведения населения в агент-ориентированной модели региона // Экономика региона. 2015. № 3. С. 312–322.
6. Makarov V.L., Bakhtizin A.R., Epstein J.M. Agent-based modeling for a complex world / 2<sup>nd</sup> edition, revised. M.: Scientific publications department, GAUGN, 2022.
7. Макаров В.Л., Бахтизин А.Р. Социальное моделирование – новый компьютерный прорыв (агент-ориентированные модели). М.: Экономика, 2013.
8. Бахтизин А.Р. Агент-ориентированные модели экономики. М.: Экономика, 2008.

9. Кежаев М.К., Мельников А.П., Фёдоров В.П., Брагин А.В. Прогнозирование временных рядов с применением нейронной сети архитектуры «трансформер» // Теоретическая информатика и компьютерные технологии: Материалы I Научно-практической конференции (Москва, 1 ноября 2023 г.). М.: Изд-во МГТУ им. Н.Э. Баумана, 2024. С. 23–25.
10. Carriero A., Pettenuzzo D., & Shekhar S. (2025). «Macroeconomic Forecasting with Large Language Models». arXiv preprint arXiv:2407.00890. (Comprehensive evaluation of multiple time-series foundation models vs. VAR, BVAR, factor models on FRED-MD dataset).
11. Christoffel K., Coenen G. and Warne A. Forecasting with DSGE models. Working paper series no 1185 / May 2010.
12. Poledna S., Miess M.G., Hommes Cars. Economic Forecasting with an Agent-Based Model. Working Paper. International Institute for Applied Systems Analysis. Laxenburg, Austria, 2020.
13. Pierdzioch C., Gupta R., Hassani H., & Silva E.S. Forecasting changes of economic inequality: a boosting approach." Social Science Journal, 2022. 59 (2), Pp. 252–268.
14. Heidrich B. Forecasting using Foundation Models and sktime. URL: [https://medium.com/@benedikt\\_heidrich/forecasting-using-foundation-models-and-sktime-4d5a09909742](https://medium.com/@benedikt_heidrich/forecasting-using-foundation-models-and-sktime-4d5a09909742).
15. Tomas A., Ales M., Sveda J. First use of AI in inflation forecasting at the CNB, *cnBlog* Чешского национального банка, 4 февраля 2025 г. URL: [https://www.cnb.cz/en/about\\_cnb/cnblog/First-use-of-AI-in-inflation-forecasting-at-the-CNB](https://www.cnb.cz/en/about_cnb/cnblog/First-use-of-AI-in-inflation-forecasting-at-the-CNB).
16. Faria-e-Castro M., Leibovici F. Artificial Intelligence and Inflation Forecasts. Federal Reserve Bank of St. Louis Review, Fourth Quarter 2024, Vol. 106, № 12. Pp. 1–14. <https://doi.org/10.20955/r.2024.12> <https://www.stlouisfed.org/publications/review/2024/nov/artificial-intelligence-and-inflation-forecasts>.
17. Leland Bybee J. Surveying Generative AI's Economic Expectations. // Yale University, 2023.
18. NIXTLA. Meet TimeGPT. 2023.
19. Korinek A. LLMs Learn to Collaborate and Reason: December 2024. Update to 'Generative AI for Economic Research: Use Cases and Implications for Economists,' published // Journal of Economic Literature. 2024. № 61 (4). <https://doi.org/10.1257/jel.20231736>.
20. Mulayim O.B. et al. Are Time Series Foundation Models Ready to Revolutionize Predictive Building Analytics? // Proceedings of the 11<sup>th</sup> ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation. 2024. С. 169–173.
21. Rasul K. et al. Lag-Llama: Towards Foundation Models for Probabilistic Time Series Forecasting. arXiv 2024 //arXiv preprint arXiv:2310.08278. 2024.
22. McCarthy S., Alaghband G. Fin-ALICE: Artificial Linguistic Intelligence Causal Econometrics // Journal of Risk and Financial Management. 2024. Т. 17, № 12. С. 537.

## Applications of Generative Artificial Intelligence in Socioeconomic Forecasting

**Albert Raufovich Bakhtizin** – Doctor of Sciences in Economics, Corresponding Member of the Russian Academy of Sciences, Director of the Central Economic and Mathematical Institute (CEMI) of the Russian Academy of Sciences, Head of the Department of Mathematical Methods of Economic Analysis of the Faculty of Economics of Lomonosov Moscow State University.

E-mail: [albert.bakhtizin@gmail.com](mailto:albert.bakhtizin@gmail.com)

**Aleksei Vladimirovich Bragin** – Degree Seeking Applicant of Central Economics and Mathematics Institute of the Russian Academy of Sciences.

E-mail: [research@alexbragin.com](mailto:research@alexbragin.com)

This article examines the application of using artificial intelligence (AI) and large language models (LLMs) for forecasting key macroeconomic indicators, including gross domestic product (GDP), inflation, unemployment rates, interest rates, and the Gini coefficient. It analyzes the capabilities of these novel approaches compared to traditional forecasting methods – such

as econometric, equilibrium, and agent-based models – across different time horizons. The paper summarizes both academic research and practical implementations, including central bank experiments on using fundamental models and LLM-like architectures (such as GPT) for macroeconomic forecasting. Special attention is given to the ability of LLMs to analyze textual information and generate predictions that are comparable in accuracy to, and in some instances superior to, those produced by professional experts. The review also covers the latest fundamental time-series models, such as TimeGPT, TimesFM, and Moirai, which employ transformer architectures tailored to economic data. The main findings indicate that AI and LLMs provide a significant advantage in terms of flexibility, adaptability, and the capacity to process diverse information sources, especially in environments characterized by high volatility or information saturation. However, challenges remain regarding the interpretability, stability, and long-term consistency of predictions. The article concludes that the best prospects for advancing macroeconomic forecasting lie in hybrid approaches that combine the computational power and adaptability of AI with the theoretical rigor and explainability of traditional economic and mathematical models.

**Keywords:** economic forecasting, LLM, GPT, AI, large language models

## REFERENCES

1. Bragin A.V., Bakhtizin A.R. Osobennosti realizatsii bol'shikh ekonomiceskikh modelei // *π-Economy*. 2023. № 16 (3). S. 107–122. DOI: <https://doi.org/10.18721/JE.16307> (in Russian).
2. Bragin A.V., Bakhtizin A.R., Makarov V.L. Bol'shie yazykovye modeli chetyvortogo pokoleniya kak novyi instrument v nauchnoi rabote // *Iskusstvennye obshchestva*. 2023. T. 18, № 1 (in Russian).
3. Bragin A.V., Bakhtizin A.R., Makarov V.L. Sovremennye programmnye sredstva agent-orientirovannogo modelirovaniya // *Iskusstvennye obshchestva*. 2022. T. 17, № 4 (in Russian).
4. Makarov V.L., Bakhtizin A.R., Sushko E.D., Ageeva A.F. *Iskusstvennoe obshchestvo i real'nye demograficheskie protsessy* // *Ekonomika i matematicheskie metody*. 2017. T 53, № 1. S. 3–18 (in Russian).
5. Makarov V.L., Bakhtizin A.R., Sushko E.D. Imitatsiya osobennostei reproduktivnogo povedeniya naseleniya v agent-orientirovannoi modeli regiona // *Ekonomika regiona*. 2015. № 3. S. 312–322 (in Russian).
6. Makarov V.L., Bakhtizin A.R., Epstein J.M. *Agent-based modeling for a complex world* / 2<sup>nd</sup> edition, revised. M.: Scientific publications department, GAUGN, 2022.
7. Makarov V.L., Bakhtizin A.R. *Sotsial'noe modelirovaniye – novyi komp'yuternyi proryv (agent-orientirovannye modeli)*. M.: Ekonomika, 2013 (in Russian).
8. Bakhtizin A.R. *Agent-orientirovannye modeli ekonomiki*. M.: Ekonomika, 2008 (in Russian).
9. Kezhaev M.K., Mel'nikov A.P., Fyodorov V.P., Bragin A.V. *Prognozirovaniye vremennykh ryadov s primenением neironnoi seti arkhitektury «transformer»* // *Teoreticheskaya informatika i komp'yuternye tekhnologii: Materialy I Nauchno-prakticheskoi konferentsii* (Moskva, 1 noyabrya 2023 g.). M.: Izd-vo MGTU im. N.E. Baumana, 2024. S. 23–25 (in Russian).
10. Carriero A., Pettenuzzo D., & Shekhar S. (2025). «Macroeconomic Forecasting with Large Language Models». arXiv preprint arXiv:2407.00890. (Comprehensive evaluation of multiple time-series foundation models vs. VAR, BVAR, factor models on FRED-MD dataset).
11. Christoffel K., Coenen G. and Warne A. Forecasting with DSGE models. Working paper series no 1185 / May 2010.
12. Poledna S., Miess M.G., Hommes Cars. Economic Forecasting with an Agent-Based Model. Working Paper. International Institute for Applied Systems Analysis. Laxenburg, Austria, 2020.

13. Pierdzioch C., Gupta R., Hassani H., & Silva E.S. Forecasting changes of economic inequality: a boosting approach. *Social Science Journal*, 2022. 59 (2). Pp. 252–268.
14. Heidrich B. Forecasting using Foundation Models and sktime. URL: [https://medium.com/@benedikt\\_heidrich/forecasting-using-foundation-models-and-sktime-4d5a09909742](https://medium.com/@benedikt_heidrich/forecasting-using-foundation-models-and-sktime-4d5a09909742).
15. Tomas A., Ales M., Sveda J. First use of AI in inflation forecasting at the CNB, cnBlog Cheskogo natsional'nogo banka, 4 fevralya 2025 g. URL: [https://www.cnb.cz/en/about\\_cnb/cnblog/Firstuse-of-AI-in-inflation-forecasting-at-the-CNB](https://www.cnb.cz/en/about_cnb/cnblog/Firstuse-of-AI-in-inflation-forecasting-at-the-CNB).
16. Faria-e-Castro M., Leibovici F. Artificial Intelligence and Inflation Forecasts. *Federal Reserve Bank of St. Louis Review*, Fourth Quarter 2024, Vol. 106, № 12. Pp. 1–14. <https://doi.org/10.20955/r.2024.12> <https://www.stlouisfed.org/publications/review/2024/nov/artificial-intelligence-and-inflation-forecasts>.
17. Leland Bybee J. Surveying Generative AI's Economic Expectations. // Yale University, 2023.
18. NIXTLA. Meet TimeGPT. 2023.
19. Korinek A. LLMs Learn to Collaborate and Reason: December 2024. Update to 'Generative AI for Economic Research: Use Cases and Implications for Economists,' published // *Journal of Economic Literature*. 2024. № 61 (4). <https://doi.org/10.1257/jel.20231736>.
20. Mulayim O.B. et al. Are Time Series Foundation Models Ready to Revolutionize Predictive Building Analytics? // *Proceedings of the 11<sup>th</sup> ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*. 2024. S. 169–173.
21. Rasul K. et al. Lag- Llama: Towards Foundation Models for Probabilistic Time Series Forecasting. arXiv 2024 // arXiv preprint arXiv:2310.08278. 2024.
22. McCarthy S., Alaghband G. Fin- ALICE: Artificial Linguistic Intelligence Causal Econometrics // *Journal of Risk and Financial Management*. 2024. T. 17, № 12. S. 537.



Л.И. БОРОДКИН\*

## Искусственный интеллект в исторических исследованиях: виртуальный ассистент или генератор квазизнаний?

В статье рассмотрены актуальные вопросы использования в исторических исследованиях методов и технологий искусственного интеллекта (ИИ). Выделены две волны развития ИИ. В ходе второй волны в центре внимания оказались искусственные нейросети, машинное обучение (включая глубокое обучение), генеративный искусственный интеллект, большие языковые модели (LLM). Рассмотрены два направления использования ИИ историками: распознавание рукописных и старопечатных исторических текстов и их транскрибирование; внедрение в исследовательскую практику больших языковых моделей, чат-ботов, генеративных искусственных нейросетей. Отмечены методологические и этические проблемы, возникающие в процессе апробации генеративного ИИ в исторических исследованиях. Дан краткий обзор этих исследований, включая такие направления, как создание виртуальных реконструкций утраченных (полностью или частично) объектов культурного наследия, атрибуции исторических текстов и др.

**Ключевые слова:** искусственный интеллект, машинное обучение, генеративные нейросети, генеративно-состязательные сети, исторические исследования, историческая информатика, трёхмерное моделирование, атрибуция исторического текста

При применение методов и технологий искусственного интеллекта в исторических исследованиях берёт своё начало в 1980-х гг., когда историки приступили к апробации возможностей экспертных систем, когнитивных подходов к изучению исторических текстов, алгоритмов кластеризации с элементами обучения, нечёткой логики, связывания записей о персоналиях в различных источниках (record linkage) [1, 2]. В эти годы на Международных конферен-

циях Ассоциации «History and Computing» (АHC) работала секция «Искусственный интеллект и экспертные системы». Период до конца XX в. принято называть первой волной ИИ, затем последовал период «зимы искусственного интеллекта», а с началом второго десятилетия XXI в. пришла вторая волна, во многом связанная с прорывным развитием компьютерных технологий, больших данных, параллельных вычислений и т.д. Основным направлением ИИ становятся искус-

\* **Бородкин Леонид Иосифович** — доктор исторических наук, член-корреспондент РАН, заведующий кафедрой исторической информатики исторического факультета Московского государственного университета им. М.В. Ломоносова.

E-mail: lborodkin@mail.ru

ственные нейросети, машинное обучение (включая глубокое обучение), генеративный искусственный интеллект, большие языковые модели (LLM) и т.д. Этот процесс затронул и исторические исследования [3], сферу исторического образования. Так, на историческом факультете МГУ уже несколько лет все магистранты слушают семестровый курс «Наука о данных и искусственный интеллект», включающий лекции и практические семинары; открыта магистерская программа с этим же названием. Программа регулярных конференций российской Ассоциации исследователей в области исторической информатики (АИК) уже несколько лет включает секцию «Искусственный интеллект», профильный журнал «Историческая информатика» включает теперь рубрику «Искусственный интеллект и наука о данных».

Наиболее дискуссионные вопросы применения ИИ в исторических исследованиях связаны в последние годы с распространением генеративных искусственных нейросетей. Именно этим вопросам уделяется основное внимание в данной статье.

### Методологические аспекты использования ИИ историками

Условно можно выделить два направления использования историками методов ИИ. Одно из них ориентируется на решение задач источниковедения, специальных исторических дисциплин. Это прежде всего распознавание рукописных исторических текстов и их транскрибирование — здесь упомянем известный проект «Digital Пётр» [4], а также проект Ithaca, в основе которого — глубокая нейронная сеть для реконструкции текста, географической и хронологической атрибуции древнегреческих эпиграфических надписей [5] и т.д. ИИ находит применение и в архивах, музеях и других учреждениях сохранения культурного наследия [6–8].

Второе направление в большей мере связано с осмысливанием новых методологических и этических проблем внедрения в исследовательскую практику больших языковых моделей, чат-ботов. *Генеративный искусственный интеллект* — это наиболее актуальный сегодня тип ИИ, который способен синтезировать (генерировать) текст, визуальный материал, отвечая на запрос пользователя (промпт). Генеративная нейросеть базируется на больших языковых моделях (LLM).

Для историков актуальность использования таких нейросетей определяется в основном потребностями интеллектуального поиска релевантной информации, задачами обработки и анализа текстов и визуального материала, реконструкции исторических событий, визуальных образов прошлого и артефактов в условиях неполной информации.

Интерес к новым возможностям использования в исторических исследованиях ИИ (и особенно генеративных нейросетей) нашёл отражение в дискуссиях, ведущихся на страницах ряда историко-ориентированных журналов.

В журнале «Антропологический форум» в 2024 г. центральный материал 60-го номера назван «Форум: искусственный интеллект в социальных и гуманитарных науках». В этом форуме приняли участие 11 антропологов и этнографов, отвечавших на вопросы редколлегии, среди которых отметим те, которые имеют отношение к теме настоящей статьи:

«Может ли ИИ оказаться полезным при проведении научного исследования и приходилось ли вам его использовать? Прибегали ли вы к помощи ИИ для составления библиографии?»<sup>1</sup> Как может отразиться на научной работе использование таких нейросетей, как ChatGPT, которые «предсказывают» нечто правдоподобное и не имеют ограничений, связанных со знанием

<sup>1</sup> Возможности ИИ для составления тематической библиографии — одно из «узких мест» при использовании искусственных нейросетей в исследовательской работе. Здесь нередко возникают ошибки, «галлюцинации» (об этом см. ниже).

о существовании фактов? Что вы думаете об этической стороне использования ИИ в научной работе и преподавании? Создаёт ли обращение к нейросети дополнительные проблемы с авторством и кому, по вашему мнению, принадлежат права на результат работы нейросети? Должны ли быть введены ограничения на использование ИИ, и если да, то какие?» [9].

Сформулированные вопросы можно рассматривать как своего рода программу для дискуссий в ближайшие годы по проблемам применения ИИ в гуманитарных науках.

Отметим, что ведущий американский исторический журнал *The American Historical Review* (AHR) учредил новый раздел «AHR History Lab», в котором обсуждаются новые методы и исследовательские практики. В сентябрьском номере журнала за 2023 г. этот раздел полностью посвящён обсуждению методологических аспектов применения ИИ в исторических исследованиях. Восемь статей, опубликованных в одном номере этого журнала и посвящённых проблематике ИИ, характеризуют актуальность вопросов о возможностях и опасностях использования ИИ историками.

Открывается форум статьёй Р.Д. Медоуза (Национальные архивы США – NARA) и Дж. Стернфелда (Национальный гуманитарный научный фонд – NEH) «Искусственный интеллект и историческая практика: форум» [10]. Авторы отмечают, что в мире, всё больше движимом алгоритмами, историки должны адаптироваться к растущему потоку оцифрованных и изначально цифровых материалов. ИИ уже доказал свою способность обнаруживать закономерности и выявлять темы в больших визуальных и текстовых наборах данных. Однако, по мнению авторов, новейшие исследования возможностей ИИ продемонстрировали, как искусственный интеллект воспроизводит и усиливает расовые и гендерные предрассудки и другие скрытые формы предвзятости. Во многом это зависит от того, какой содержательный материал был использован в процессе «тоталь-

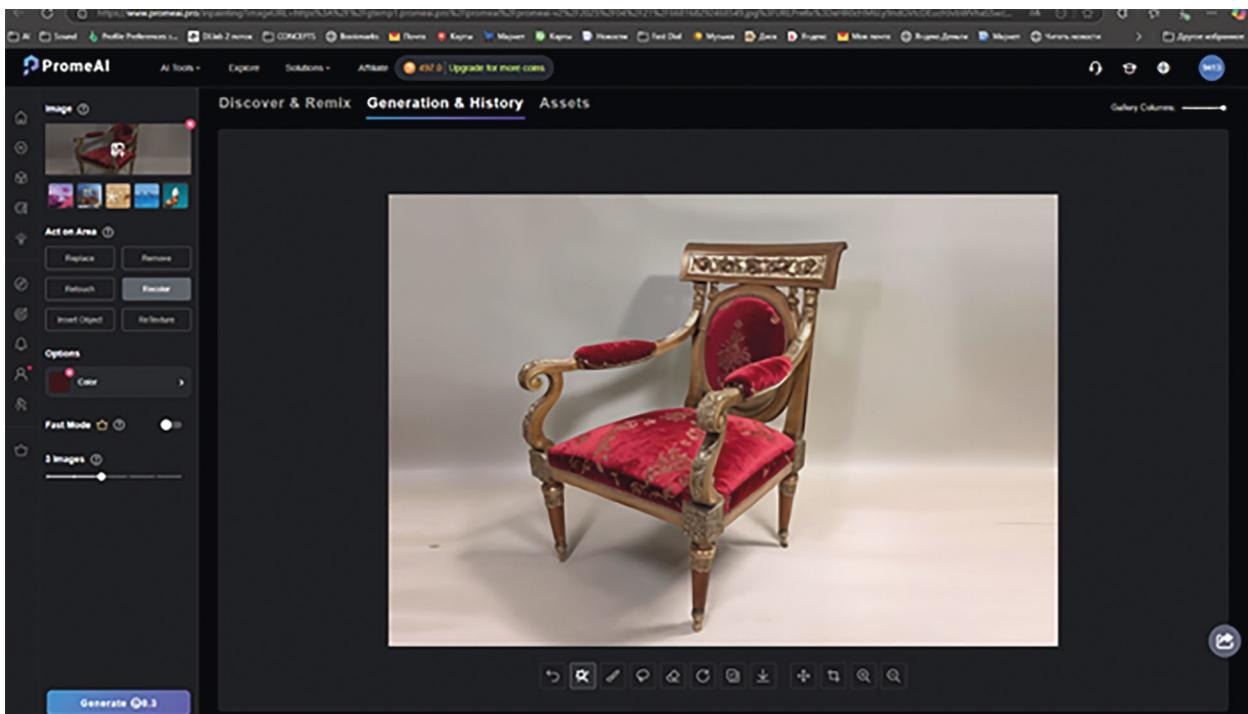
ного» машинного обучения. По мнению авторов, при этом ИИ влияет на процесс формирования исторического сознания, и, чтобы осмыслить влияние ИИ на сферу исторических исследований, нередко требуется подвергнуть сомнению такие фундаментальные концепции, как истина, доказательность и подлинность.

Участники форума AHR затронули целый ряд эпистемологических и этических проблем, связанных с внедрением ИИ в практику исторических исследований и образования. Эти проблемы становятся всё более актуальными по мере роста интереса к возможностям генеративных нейросетей в гуманитарных (и не только) науках.

Вопросы об авторстве, проблемах копирайта в работах по сохранению культурного наследия в контексте использования технологий ИИ обсуждаются и в публикации 2025 г. применительно к практикам британских учёных [11]. Отмечается потребность в разработке правового регулирования этой сферы в условиях широкого распространения генеративных нейросетей. Интересно, что в заключении авторы статьи указывают (точнее, декларируют), что они не использовали генеративный ИИ при её написании.

Научные журналы, издающиеся в различных странах, вырабатывают политику в отношении использования ИИ в текстах публикуемых авторов. Так, высокорейтинговый российский журнал *Russian Journal of Economics* в инструкции для авторов, содержащей раздел «Use of AI», указывает, что авторы должны открыто сообщать об использовании инструментов искусственного интеллекта, таких как ChatGPT, построенных на больших языковых моделях, при подготовке рукописи, а также раскрывать подробности использования инструментов искусственного интеллекта в разделе «Материалы и методы» (<https://gujes.org/about#Authors-Guidelines>).

Отдельное внимание при обсуждении возможностей и ограничений ИИ в исторических исследованиях следует уделить публикациям по этой тематике в журнале «Историческая информатика», который



Ил. 1. Создание 3D-модели предмета интерьера усадьбы Кусково на основе архивных фотографий в нейросети Tripo AI

([https://e-notabene.ru/view\\_article.php?id\\_article=74244&nb=1&logged=0&aurora=0](https://e-notabene.ru/view_article.php?id_article=74244&nb=1&logged=0&aurora=0))

становится площадкой для апробации различных методов и технологий ИИ в данном предметном поле. В 2025 г. в рубрике, посвящённой ИИ-приложениям, опубликованы 10 статей, сегодня эта тематика оказалась наиболее востребованной. Рассмотрим некоторые из этих публикаций, дающих представление о специфике применения ИИ историками. Основное внимание уделим опыту (пока небольшому) работы с генеративными нейросетями.

### Генеративный ИИ и сохранение культурного наследия

Обратимся к использованию генеративного ИИ в исследованиях по созданию виртуальной реконструкции объектов культурного наследия. Т.В. Маландина подготовила диссертацию на кафедре исторической информатики исторического факультета МГУ, в которой генеративные нейросети используются в задаче реконструкции утраченных интерьеров усадьбы князей Гагариных Кусково [12]. Нейросеть предложила визуализацию утраченных

элементов интерьера, используя имеющиеся аналоги и экспертные оценки специалистов. В этой работе автор сочетает традиционные методы и технологии 3D-моделирования с нейросетевыми инструментами (Tripo AI, Prome AI, Midjourney), осуществляя генерацию 3D-модели и её интеграцию в виртуальное пространство усадьбы (ил. 1, 2). Как подчёркивает Т.В. Маландина, нейросеть выступает здесь «именно как инструмент и виртуальный помощник, результат обращения к которому поддаётся контролю посредством обращения к материалам источников базы, загружаемой в нейросеть», а построенная 3D-реконструкция (как и любая сгенерированная трёхмерная модель) нуждается в верификации с помощью экспертной оценки.

Ещё одно исследование такого рода, проведённое в рамках проекта кафедры исторической информатики МГУ, имело целью создание виртуальной реконструкции утраченного храмового комплекса святителя Николая Чудотворца на Мясницкой улице в Москве на конец XIX в. [13]. Рассмотрим

фрагмент этой работы, связанный с реконструкцией Е.Р. Разгулиным интерьеров храма. Для полноценного представления его внутреннего убранства необходимо реконструировать атрибуты церковного богослужения: аналои, киоты, подсвечники. Здесь основную роль сыграли сохранившиеся фотодокументы, однако их качество оставляет желать лучшего, поэтому были привлечены фотографии современных аналогов этих атрибутов. Нейросетевые сервисы дали возможность создавать 3D-модели на их основе (ил. 3), однако они не являются полностью идентичными тем, которые были изображены на оригинальных фотографиях. Ведь каждый предмет церковного имущества прошлого создавался мастерами вручную по индивидуальному заказу, знаменуя собой высокое искусство ручной обработки материалов. Нейросетевые модели, обладающие большими возможностями авто-

матизированного построения трехмерных моделей на основе выявленных аналогов, в процессе генерации утрачивают детали декоративной отделки и требуют вмешательства исследователя для ручной корректировки построенных 3D-моделей. Ознакомиться с этим проектом можно в открытом доступе: URL: <http://николаявмясниках.рф>.

Добавим, что в зарубежной научной периодике опубликованы результаты создания виртуальных реконструкций объектов культурного наследия с использованием генеративно-состязательных сетей (Generative adversarial network, GAN), эффективных для создания реалистичных визуализаций на основе тех данных, которые были положены в основу обучения. Интересен принцип действия GAN, основанный на взаимодействии пары нейронных сетей. Первая (генератор) опирается на обучающие данные для создания выходных данных,



Ил. 2. Результаты визуализации интерьера центрального зала главного дома усадьбы Кузьминки с использованием нейросетей Tripo AI, Prome AI, Midjourney  
[https://e-notabene.ru/view\\_article.php?id\\_article=74244&nb=1&logged=0&aurora=0](https://e-notabene.ru/view_article.php?id_article=74244&nb=1&logged=0&aurora=0)



Ил. 3. Трёхмерные модели предметов церковной утвари храма Николая Чудотворца, созданные по изображениям выявленных аналогов с помощью нейросетевых сервисов Krea и Polycam

([https://www.nbpublish.com/library\\_read\\_article.php?id=75011](https://www.nbpublish.com/library_read_article.php?id=75011))

вторая (дискриминатор) оценивает выходные данные генератора, выполняя функцию классификатора. «Состязательность», заявленная в названии GAN, означает, что обе модели ориентируются на то, чтобы одержать верх. Генератор порождает данные, а дискриминатор ориентирован на выявление искусственно сгенерированных данных.

Как показала практика использования технологии GAN, она (как и другие ИИ-технологии) может порождать «галлюцинации»<sup>1</sup>, поэтому требует контроля получаемых

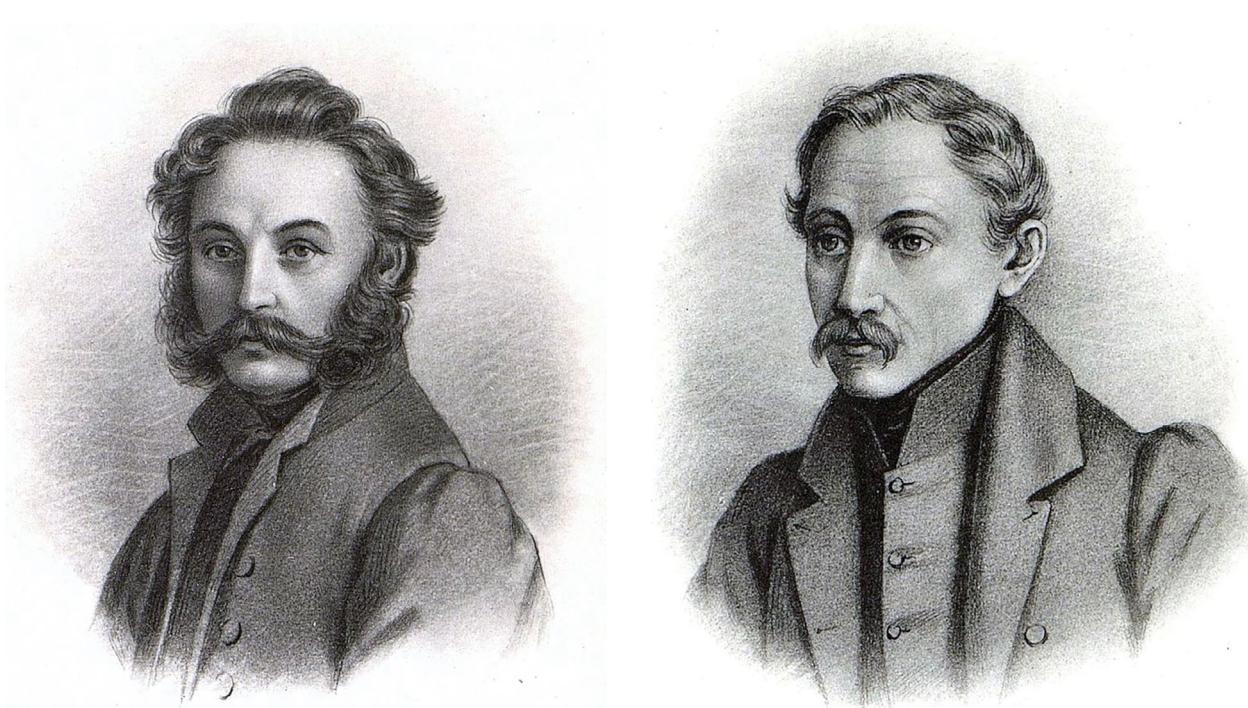
результатов и осторожного применения. Одна из причин «галлюцинаций» — проблемы в обучающих данных<sup>2</sup>. «Галлюцинации» можно рассматривать как одно из проявлений «квазинаучной» информации, получаемой с использованием генеративного ИИ.

### ИИ в исследованиях исторических текстов

В работе В.В. и А.В. Латоновых рассмотрена задача определения авторства «Записок Горбачевского» — известного текста

<sup>1</sup> Галлюцинация — часто используемый термин в контексте практик применения ИИ для обозначения, когда нейросети предлагают пользователю неверную информацию («галлюцинируют»).

<sup>2</sup> См. соответствующие ссылки в упомянутой статье Т.В. Маландиной.



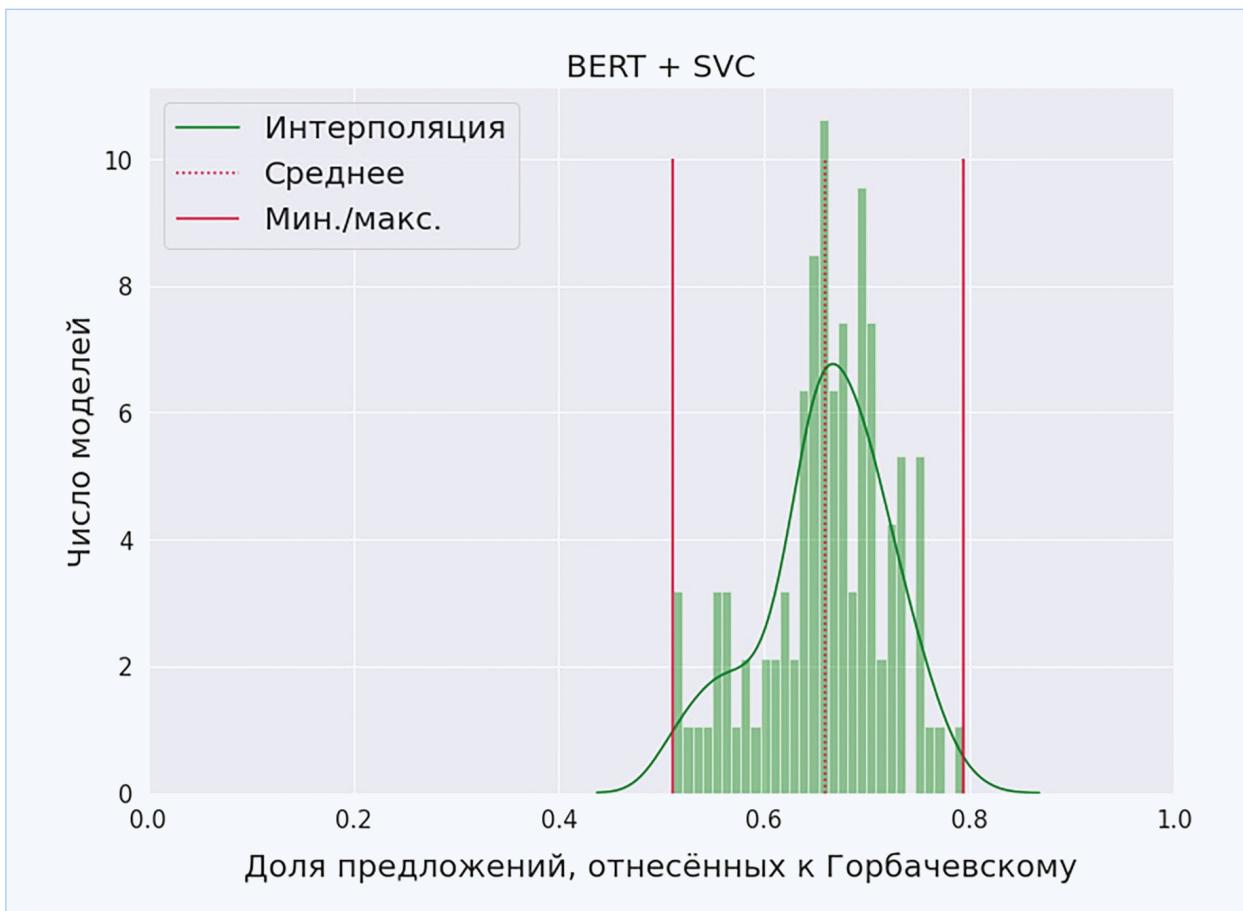
Ил. 4–5. Иван Иванович Горбачевский (слева) и Пётр Иванович Борисов (справа).  
Декабристы, подпоручики 8-й артиллерийской бригады, члены  
«Общества соединённых славян», участники восстания Черниговского полка

декабристской мемуаристики [14]. Вопрос о его реальном авторе стал спорным ещё в советской историографии и не решён по сей день, так как обсуждается альтернативная гипотеза об авторстве другого декабриста – П.И. Борисова (ил. 4, 5). К этой дискуссии подключились нейросети. Методика исследования была основана на машинном обучении. В качестве обучающего материала привлекались тексты, безусловно принадлежавшие либо И.И. Горбачевскому, либо П.И. Борисову. Затем нейросети решали задачу атрибуции «Записок». Использовались несколько нейросетей, наиболее эффективной оказалась одна из ключевых моделей обработки естественного языка BERT, построенная на архитектуре двунаправленного трансформера: в сочетании с другими нейросетями BERT давал около 90% точности на обучающих материалах. В итоге обученная модель классифицировала около 70% предложений «Записок», как написанных Горбачевским (ил. 6). Это достаточно сильный результат, учитывая, что оба предпо-

лагаемых автора принадлежали одному кругу декабристов.

В контексте взаимодействия исследователя и ИИ представляют интерес и новые публикации историков, проводящих апробацию возможностей генеративных нейросетей в работе с текстовыми источниками при решении вспомогательных задач. Так, в исследовании А.В. Кузнецова приводятся результаты сравнительного анализа эффективности четырёх больших языковых моделей в задаче извлечения структурированной информации из текста исторического эго-документа [15]. Автор приходит к выводу, что «применение ИИ повышает требования к критической экспертизе историка, смещая его роль от поиска информации к верификации и интерпретации машинных результатов».

В работе А.С. Котова рассмотрены методы автоматической нормализации текстов на средневерхненемецком и раннем нововерхненемецком языках на основе нейроязыковых моделей. Для нормализации текстов XIV–XV вв., созданных в орденской



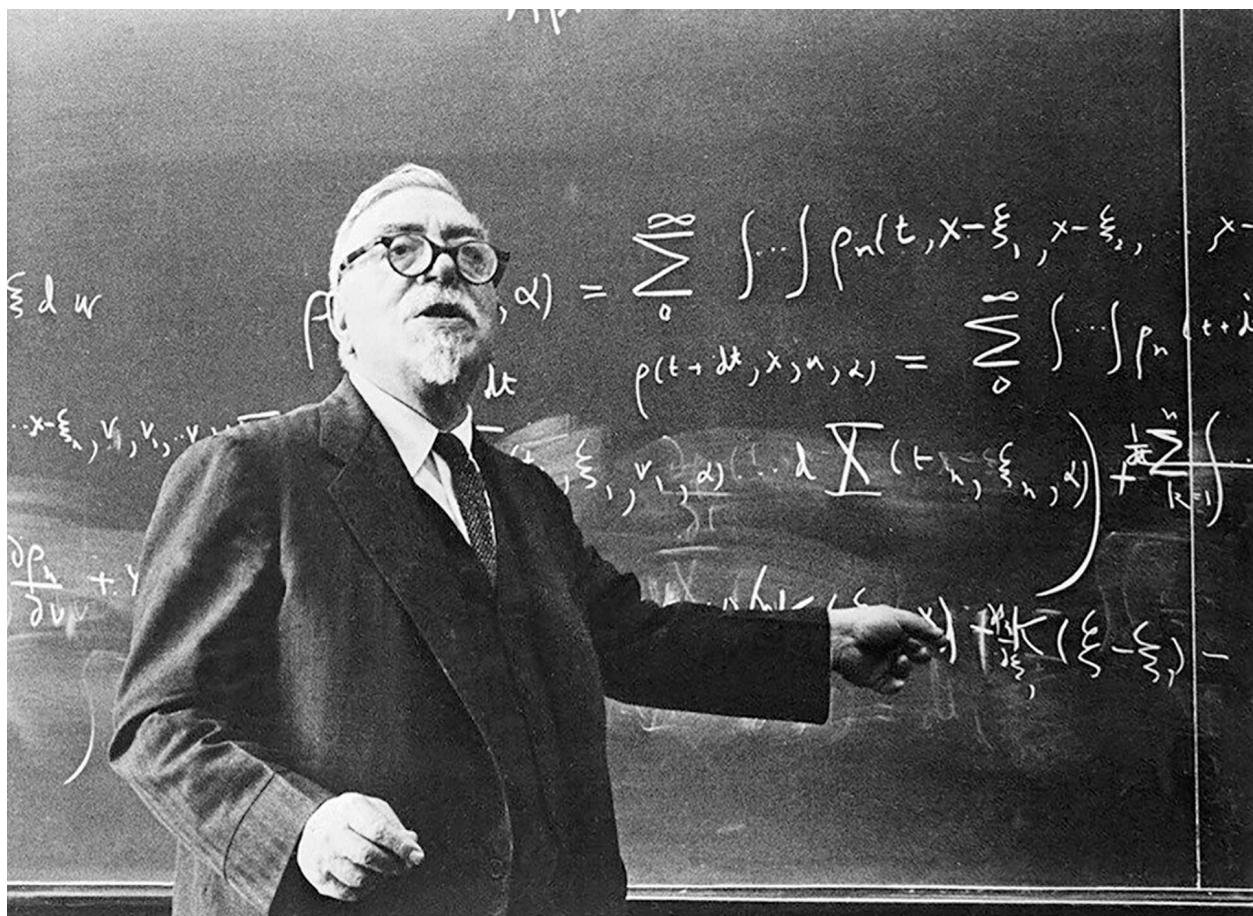
Ил. 6. Обученные модели классифицируют около 70% предложений «Записок» как написанные Горбачевским. Разброс в результатах объясняется тем, что для обучения использовались сто разных выборок ([https://www.nbpublish.com/library\\_read\\_article.php?id=72805](https://www.nbpublish.com/library_read_article.php?id=72805))

Пруссии, дообучена (с учётом специфики этих текстов) нейроязыковая модель с использованием алгоритмов Transformer (BART), что привело к существенному повышению её эффективности по сравнению с другими моделями. В то же время использование модели BART не обходится без «галлюцинаций». Автор отмечает, что даже дообученная модель требует постобработки, дополнительной проверки полученных результатов экспертом [16].

### ИИ рассуждает?

Возвращаясь к вопросу, поставленному в названии данной статьи, обратим внимание на более точное понятие ИИ, введённое впервые в 1955 г. Джоном Маккарти. В английском языке «Artificial Intelligence» означает «искусственный разум», «уме-

ние рассуждать разумно». Такая трактовка приводит к более узкому пониманию, чем привычный для нас «искусственный интеллект». И здесь обращает на себя внимание обозначившийся в последние годы тренд в развитии генеративных нейросетей, связанный с разработкой более креативных алгоритмов ответа ИИ на сформированный пользователем промпт. Эти алгоритмы реализуют модель *Reasoning* – более гибкого интерактивного взаимодействия ИИ и пользователя, когда модель «рассуждает», проводя пошаговый поиск релевантной информации в режиме «цепочки рассуждений», комментируя промежуточные результаты и даже формулируя некоторые оценочные суждения. Протокол этих «рассуждений» генеративного ИИ может занимать несколько страниц, демонстрируя стиль инфор-



Ил. 7. Норберт Винер

мированного и «рассудительного» виртуального ассистента исследователя. Более сложным является вопрос об оценке ИИ как генератора квазизнания (псевдознания, недостоверного знания), а также о возможных последствиях «галлюцинаций». Однако стремительное развитие генеративных нейросетей постоянно повышает качество их работы, и сегодня результаты их апробации в сфере исторических исследований приводят к более позитивным оценкам. При этом не следует полагать, что обученная машина сделает всё сама и нейросеть даст историку «правильный» результат.

### Заключение

Успехи в развитии технологий искусственного интеллекта в XXI в. несомненны. Происходит интеграция ИИ в различные сферы деятельности, включая научные исследования. Однако эта интеграция в сфере гуманитарного знания вызывает и эти-

ческие вопросы, такие как предвзятость данных, риск «галлюцинаций» в выводах генеративных нейросетей, необходимость человеческого контроля для обеспечения корректности интерпретаций. ИИ дополняет исследователя, играя в основном роль хорошо информированного виртуального ассистента, однако он не является заменой экспертизе историков, и его использование требует критической оценки. Предвидя подобные проблемы, Норберт Винер, «отец» кибернетики (ил. 7), ещё на заре компьютерной эпохи писал, что главное преимущество мозга человека как органа мышления, по сравнению с машинами его времени, это «способность мозга оперировать с нечётко очерченными понятиями <...>. Отдайте же человеку — человеческое, а вычислительной машине — машинное. В этом и должна, по-видимому, заключаться разумная линия поведения при организации совместных действий людей и машин» [17 с. 82–83].

## ЛИТЕРАТУРА

1. Бородкин Л.И. Методы искусственного интеллекта: новые горизонты исторического познания // Информационный бюллетень Комиссии по применению математических методов и ЭВМ в исторических исследованиях при Отделении истории Российской академии наук. 1992. № 5. С. 4–8.
2. Бородкин Л.И. Наука о данных и технология искусственного интеллекта: возможности и ограничения в исследованиях историков // Труды Отделения историко-филологических наук: Ежегодник. Т. 11 / Отв. ред. В.А. Тишков. М.: Изд-во РАН, 2022. С. 145–168.
3. Бородкин Л.И. Историк в мире нейросетей: вторая волна применения технологий искусственного интеллекта // Историческая информатика. 2025. № 1. С. 83–94.
4. Базарова Т.А., Проскурякова М.Е. Автографы Петра I: чтение технологиями искусственного интеллекта и создание электронного архива // Историческая информатика. 2022. № 4. С. 179–190.
5. Assael Y., Sommerschield T., Shillingford B. et al. Restoring and attributing ancient texts using deep neural networks // Nature. 2022. № 603. Рр. 280–283.
6. Юмашева Ю.Ю. К вопросу о применении искусственного интеллекта в исторических исследованиях // Историческая информатика. 2025. № 1 (51). С. 95–121.
7. Юмашева Ю.Ю. Историческая наука, архивы, библиотеки, музеи и искусственный интеллект: что день грядущий нам готовит? // Документ. Архив. История. Современность. Вып. 21. Екатеринбург, 2021. С. 247–279.
8. Юмашева Ю.Ю. Историческая наука, архивы, библиотеки, музеи и искусственный интеллект: год спустя // Документ. Архив, История. Современность. Вып. 22. Екатеринбург, 2022. С. 217–241.
9. Форум: искусственный интеллект в социальных и гуманитарных науках // Антропологический форум. 2024. № 60. С. 12.
10. Meadows, R. Darrell, Sternfeld, Joshua. Artificial Intelligence and the Practice of History: A Forum // The American Historical Review. 2023. Vol. 128. Issue. 3. Рр. 1345–1349.
11. Westenberger P., Farmaki D. Artificial intelligence for cultural heritage research: the challenges in UK copyright law and policy. Eur. J. Cult. Manag. Policy 2025 15:14009.
12. Маландина Т.В. Опыт использования искусственных нейросетей в решении задач виртуальной реконструкции исторических усадебных интерьеров // Исторический журнал: научные исследования. 2025. № 3. С. 45–60.
13. Разгулин Е.Р., Жеребятьев Д.И. Виртуальная реконструкция храмового комплекса святыни Николая Чудотворца на Мясницкой улице в Москве в конце XIX века // Историческая информатика. 2025. № 2. С. 90–126.
14. Латонов В.В., Латонова А.В. Определение авторства «Записок декабриста И.И. Горбачевского» методами машинного обучения // Там же. 2025. № 3. С. 141–150.
15. Кузнецов А.В. Автоматическое извлечение информации из эго-документа: сравнительный анализ эффективности больших языковых моделей на примере дневника К.А. Березкина // Там же. С. 99–127.
16. Котов А.С. Дообучение модели на основе архитектуры Transformer для нормализации корпуса средневековых текстов на немецком языке XIV–XV вв. из орденской Пруссии // Там же. С. 128–140.
17. Винер Н. Творец и робот. М.: Прогресс, 1966.

## Artificial Intelligence in Historical Research: a Virtual Assistant or a Generator of Quasi-Knowledge?

**Leonid Iosifovich Borodkin** — Doctor of Sciences in History, Corresponding Member of the Russian Academy of Sciences, Head of the Department of Historical Informatics, Faculty of History, Lomonosov Moscow State University.

E-mail: lborodkin@mail.ru

The article examines the current issues of the application of artificial intelligence (AI) methods and technologies in historical research. It outlines two waves of AI development, with the second wave focusing on artificial neural networks, machine learning (including deep learning), generative AI, and large language models (LLMs). Two areas of AI use by historians are explored: the recognition and transcription of handwritten and early-printed historical texts, and the integration of large language models, chatbots, and generative neural networks into research practices. The article highlights the methodological and ethical challenges that arise when testing generative AI in historical research. A brief overview of relevant research is provided, covering areas such as the virtual reconstruction of lost (fully or partially) cultural heritage sites and the attribution of historical texts.

**Keywords:** artificial intelligence, machine learning, generative neural networks, generative adversarial networks, historical research, historical informatics, three-dimensional modeling, attribution of historical text

## REFERENCES

1. Borodkin L.I. Metody iskusstvennogo intellekta: novye gorizonty istoricheskogo poznaniya // Informatsionnyi byulleten' Komissii po primeneniyu matematicheskikh metodov i EVM v istoricheskikh issledovaniyakh pri Otdelenii istorii Rossiiskoi akademii nauk. 1992. № 5. S. 4–8 (in Russian).
2. Borodkin L.I. Nauka o dannykh i tekhnologii iskusstvennogo intellekta: vozmozhnosti i ogranicheniya v issledovaniyakh istorikov // Trudy Otdeleniya istoriko-filologicheskikh nauk: Ezhegodnik. T. 11 / Otv. red. V.A. Tishkov. M.: Izd-vo RAN, 2022. S. 145–168 (in Russian).
3. Borodkin L.I. Istorik v mire neirosetei: vtoraya volna primeneniya tekhnologii iskusstvennogo intellekta // Istoricheskaya informatika. 2025. № 1. S. 83–94 (in Russian).
4. Bazarova T.A., Proskuryakova M.E. Avtografy Petra I: chtenie tekhnologiyami iskusstvennogo intellekta i sozdanie elektronnogo arkhiva // Istoricheskaya informatika. 2022. № 4. S. 179–190 (in Russian).
5. Assael Y., Sommerschield T., Shillingford B. et al. Restoring and attributing ancient texts using deep neural networks // Nature. 2022. № 603. Pp. 280–283.
6. Yumasheva Yu.Yu. K voprosu o primenenii iskusstvennogo intellekta v istoricheskikh issledovaniyakh // Istoricheskaya informatika. 2025. № 1 (51). S. 95–121 (in Russian).
7. Yumasheva Yu.Yu. Istoricheskaya nauka, arkhivy, biblioteki, muzei i iskusstvennyi intellekt: chto den' gryadushchii nam gotovit? // Dokument. Arkhiv. Istorya. Sovremennost'. Vyp. 21. Ekaterinburg, 2021. S. 247–279 (in Russian).
8. Yumasheva Yu.Yu. Istoricheskaya nauka, arkhivy, biblioteki, muzei i iskusstvennyi intellekt: god spustya // Dokument. Arkhiv, Istorya. Sovremennost'. Vyp. 22. Ekaterinburg, 2022. S. 217–241 (in Russian).
9. Forum: iskusstvennyi intellekt v sotsial'nykh i gumanitarnykh naukakh // Antropologicheskii forum. 2024. № 60. S. 12 (in Russian).

10. Meadows R. Darrell, Sternfeld, Joshua. Artificial Intelligence and the Practice of History: A Forum // The American Historical Review. 2023. Vol. 128. Iss. 3. Pp. 1345–1349.
11. Westenberger P., Farmaki D. Artificial intelligence for cultural heritage research: the challenges in UK copyright law and policy. Eur. J. Cult. Manag. Policy 2025:14009.
12. Malandina T.V. Opyt ispol'zovaniya iskusstvennykh neirosetei v reshenii zadach virtual'noi rekonstruktsii istoricheskikh usadebnykh inter'erov // Istoricheskii zhurnal: nauchnye issledovaniya. 2025. № 3. S. 45–60 (in Russian).
13. Razgulin E.R., Zherebyat'ev D.I. Virtual'naya rekonstruktsiya khramovogo kompleksa svyatitelya Nikolaya Chudotvortsya na Myasnitskoi ulitse v Moskve v kontse XIX veka // Istoricheskaya informatika. 2025. № 2. S. 90–126 (in Russian).
14. Latonov V.V., Latonova A.V. Opredelenie avtorstva «Zapisok dekabrista I.I. Gorbachevskogo» metodami mashinnogo obucheniya // Tam zhe. 2025. № 3. S. 141–150 (in Russian).
15. Kuznetsov A.V. Avtomaticheskoe izvlechenie informatsii iz ego-dokumenta: sravnitel'nyi analiz effektivnosti bol'shikh yazykovykh modelei na primere dnevnika K.A. Berezkina // Tam zhe. S. 99–127 (in Russian).
16. Kotov A.S. Doobuchenie modeli na osnove arkhitektury Transformer dlya normalizatsii korpusa srednevekovykh tekstov na nemetskom yazyke XIV–XV vv. iz ordenskoi Prussii // Tam zhe. S. 128–140 (in Russian).
17. Viner N. Tvorets i robot. M.: Progress, 1966 (in Russian).

Индекс УДК 003.07

Код ГРНТИ 03.23.31

DOI: 10.22204/2587-8956-2025-123-04-99-109



**А.В. СИРЕНОВ, Е.В. АНИСИМОВ, Т.А. БАЗАРОВА, М.Е. ПРОСКУРЯКОВА\***

## Искусственный интеллект и изучение письменного наследия Петра Великого и его сподвижников

Статья посвящена результатам совместной работы историков и палеографов Санкт-Петербургского института истории РАН и специалистов по анализу данных ПАО «Сбербанк». Сотрудничество двух учреждений началось в 2020 г. благодаря инициативе Российского исторического общества. Перед коллективом была поставлена задача научить искусственный интеллект распознавать автографы Петра Великого. Успешное решение данной задачи и создание модели «Digital Пётр» позволило в 2024 г. приступить к разработке нового алгоритма по распознаванию письма людей Петровской эпохи. В статье описывается работа коллектива над формированием наборов данных для обеих моделей, которая заключалась в определении принципов для отбора документов, транскрибировании их текста, а также разметке цифровых копий автографов. В статье также уделено внимание характеристике тех электронных ресурсов, которые были созданы за годы работы с петровскими автографами — модель распознавания «Digital Пётр» и сайт «Автографы Петра I».

**Ключевые слова:** Пётр Великий, автографы, эпистолярное наследие, искусственный интеллект, проект «Digital Пётр», Петровская эпоха

\* **Сиренов Алексей Владимирович** — член-корреспондент РАН, доктор исторических наук, директор Санкт-Петербургского института истории (СПБИИ) РАН.  
E-mail: sirenov@rambler.ru

**Анисимов Евгений Викторович** — доктор исторических наук, главный научный сотрудник Отдела средневековой истории России СПБИИ РАН, руководитель проекта «Походная канцелярия Александра Даниловича Меншикова (1702–1705): Новые подходы к изучению традиционных источников Петровской эпохи» (20-09-42051 (Петровская эпоха)).

E-mail: vbreviis@mail.ru

**Базарова Татьяна Анатольевна** — кандидат исторических наук, старший научный сотрудник, и.о. заведующего Отделом источниковедения СПБИИ РАН.

E-mail: tbazarova@yandex.ru

**Прокурякова Мария Евгеньевна** — кандидат исторических наук, старший научный сотрудник Отдела источниковедения СПБИИ РАН.

E-mail: m-proskuryakova@mail.ru

**В** 2020 г. по инициативе Российского исторического общества и ПАО «Сбербанк» начался научно-исследовательский проект «Автографы Петра Великого: чтение технологиями искусственного интеллекта». Проект был призван показать широкие перспективы применения новых методов при изучении автографов первого российского императора. Ключевыми исполнителям проекта стали сотрудники Санкт-Петербургского Института истории РАН (далее – СПБИИ РАН) и ПАО «Сбербанк». Коллектив включил в себя историков, палеографов, математиков и программистов, работавших в тесной кооперации на всех этапах реализации проекта (ил. 1).

### Научно-исследовательский проект «Digital Пётр»: первые итоги

В начале проекта рабочая группа произвела отбор петровских автографов в СПБИИ РАН. Он осуществлялся по продолжающемуся изданию «Писем и бумаг императора Петра Великого» (далее – ПиБ). В 7–13 томах ПиБ, опубликованных во второй половине XX – начале XXI в., архивные шифры документов наиболее близки к современным. Основной массив подлинных автографов Петра I в настоящее время находится в Российском государственном архиве древних актов (далее – РГАДА). Источниковую базу пополнили материалы из коллекции 277 (Коллекция документов Петра I), хранящейся в собрании СПБИИ РАН. Главным критерием отбора автографов являлось наличие текста, написанного рукой Петра I: от двух–трёх слов до несколько страниц.

Всего в специализированную программу (Computer Vision Annotation Tool) были загружены и размечены 719 страниц фотокопий, которые содержали 10 512 строк, написанных Петром Великим в 1707–1713 гг. Именно они составили набор данных, с которыми работали программисты, обучая алгоритм распознаванию петровских текстов. Итогом работы коллектива историков, палеографов и специалистов по ана-

лизу данных стала модель, размещённая на сайте «Digital Пётр» (<https://fusionbrain.ai/digital-petr>) [1, 2].

На этот ресурс пользователь может загрузить цифровые копии петровских автографов, от одной до нескольких десятков страниц одновременно, для получения распознанного текста (ил. 2). После транскрибирования интерфейс отображает и копию документа, и построчно прочитанный текст (все строки транскрипции пронумерованы). Программа распознаёт петровские автографы 1707–1713 гг. с точностью до 98%. На сайте также есть возможность скачать распознанный текст в формате docx. С помощью искусственного интеллекта чтение писем, написанных беглым неразборчивым почерком Петра I, стало возможным не только для узкой группы учёных-палеографов, но практически для каждого любителя российской истории.

### Проект «Digital Пётр» и сайт «Автографы Петра I»

В юбилейный петровский – 2022 – год проект «Digital Пётр» получил продолжение. По инициативе Российского исторического общества научные сотрудники вместе с IT-специалистами А.Ю. Ходотом и Е.В. Ригиным начали работу над сайтом «Автографы Петра I» (<https://peterscript.ru/>), также направленным на изучение и популяризацию эпистолярного наследия Петра I. В настоящее время на сайте размещено свыше 230 писем, указов и записок Петра I за 1707–1713 гг., оригиналы которых хранятся в РГАДА и Архиве СПБИИ РАН. Каждое изображение автографа снабжено кратким заголовком с датой, указанием на место хранение и публикацию, а также транскрипцией, полученной на сайте «Digital Пётр» (ил. 3). Также было принято решение «привязывать» распознанные на сайте документы к Биохронике Петра Великого. Подготовленная Е.В. Анисимовым и размещённая на сайте Национального исследовательского университета «Высшая школа

экономики» база данных (<https://spb.hse.ru/humart/history/peter/biochronic>) позволяет день за день проследить жизнь и деятельность первого российского императора.

Связь двух сайтов — «Автографы Петра I» и «Биохроника Петра Великого» — открыла перед исследователями и любителями российской истории дополнительные возможности. Например, можно узнать, что делал Пётр I в день отправки размещённого на сайте письма, какие ещё указы подписал, с кем встречался или кто его сопровождал в поездке. Следовательно, каждый оцифрованный автограф вводится в исторический контекст.

Таким образом, один и тот же текст на сайте «Автографы Петра I» оказался представленным в различных ипостасях. У специалиста или заинтересованного читателя появилась возможность прочитать петровский почерк самостоятельно, сравнить результат с публикацией в ПиБ или с текстом, распознанным искусственным интеллектом (ил. 4). Исследователи и все интересующиеся историей России Нового времени получили уникальную возможность работать с петровскими наследием ещё до завершения фундаментальной публикации писем и бумаг императора Петра Великого и оценить вклад, который он внёс в становление российской государственности.

### Перспективы развития проекта «Digital Пётр»

В 2020 г. искусственный интеллект прошёл обучение петровскому почерку на документах, написанных в течение семи лет жизни первого российского императора. Комплексы петровских автографов за 1689–1706 гг. и 1714–1725 гг. в программу не загружались. Однако, как известно, почерк человека меняется: на него влияют разные обстоятельства: образ жизни, болезни, душевное состояние. Изменения происходили и в почерке Петра Великого. В конце XVII в. молодой царь медленно и старательно выписывал каждую букву. Будущая



Ил. 1. Портрет Петра I. Неизвестный художник. Конец XVII в. (Архив СПбИИ РАН. Колл. 282. Оп. 1. Д. 126). Фото Д.В. Сироткина (ГЭ)



Ил. 2. Распознавание на сайте «Digital Пётр»  
<https://fusionbrain.ai/digital-petr/>

беглость и размашистость почерка только угадываются в самых ранних, дошедших до наших дней посланиях юного «недостойного Петрушки» его матери царице Натальи Кирилловне. Письма будущего царя-реформатора конца XVII в. сохранились в единичных экземплярах. В зрелые годы объём эпистолярной практики Петра Ве-

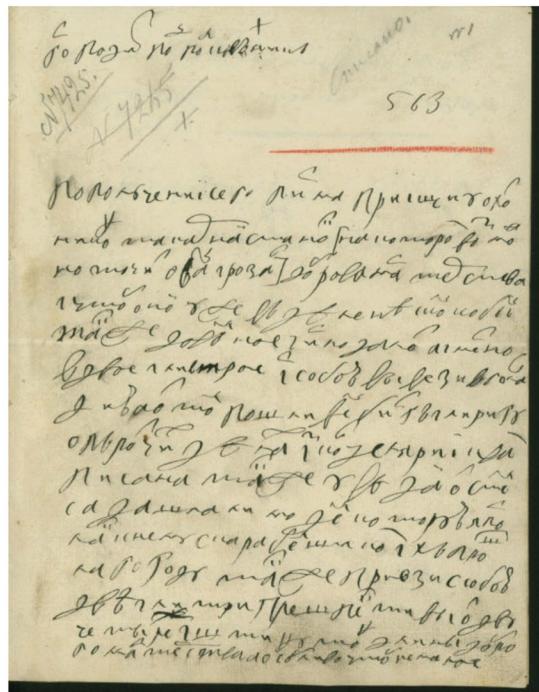
Автографы Петра I  
Электронный архив

АРХИВ Литература События  

### Письмо Петра I Б. И. Куракину, 27 мая 1711 г.

Письмо · 27 Мая 1711 г. · Ярославль

Архив Санкт-Петербургского института истории РАН



Архив СПБИИ РАН. Колл. 277. Оп. 2. Д. V/1. Л. 1, 2.



Digital Петр

Digital Петр Публикация PDF

1. господин подполковъник
2. по получениі сего писма приїша у охот
3. никоф ткацнй станок [на котором возмож
4. но точит овал і розан] добрая мастерства
5. і чтоб оной уже въ дѣле нѣсколко был
6. также довольное число дало а именно
7. вдвое іли втрое і с собою вывези в голос
8. дію а оттол пошли велбингъ іли ригу
9. о пъртчих дѣлах із койцеляриі к вам
10. писана также увѣдай отстел
11. са даша ли модел которую я пос
12. лал к нему с карабельщиком іхъ прош
13. лаго году также привзи с собою
14. дѣл іли три прешпективы от дву
15. четырех і шти футоф длины добро
16. го мастерства а особливо чтоб немалое

Ил. 3. Страница рукописи Петра I (Архив СПБИИ РАН. Колл. 277. Оп. 2. Д. V/1. Л. 1, 2) и результат распознавания текста, представленные на сайте «Автографы Петра I»

<https://peterscript.historyrussia.org/document/139>

ликого в несколько раз увеличился. Почти ежедневно — в своей резиденции, в поездке по России или Западной Европе, в военном походе — он пишет и ставит резолюции, делает приписки на полях на десятках различных документов. Его почерк становится максимально беглым, буквы теряют объём, практически «ложатся» в строку.

В ходе работы над проектом и тестирования ресурса «Digital Пётр» были выявлены две основные проблемы: 1) программа с низким качеством распознаёт документы конца XVII в., написанные рукой молодого Петра; 2) искусственный интеллект не рас-

познаёт части документов, которые написаны писарскими почерками.

Современному исследователю необходим надёжный инструмент — доступная компьютерная программа, которая могла бы читать скоропись XV–XVII вв. Чтение скорописи представляет собой отдельную проблему в изучении источников по истории России эпохи позднего Средневековья и раннего Нового времени. Традиционно считалось, что скоропись представляет собой особый тип письма, который был создан с целью ускорения процесса письма. Однако исследования последних

Автографы Петра I

Электронный архив

АРХИВ Литература События

— Peter

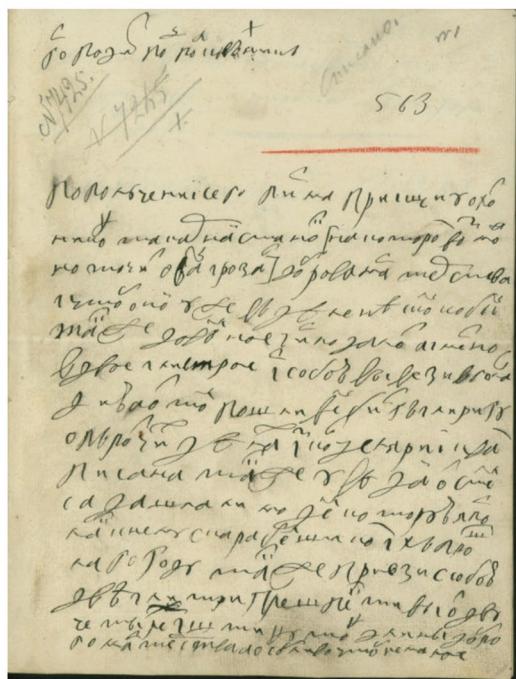
Письмо Петра I Б. И. Куракину, 27 мая 1711 г.

Письмо · 27 Мая 1711 г. · Ярославль

Архив Санкт-Петербургского института истории РАН

ITINERA PETRI

БИОХРОНИКА  
ПЕТРА ВЕЛИКОГО  
ДЕНЬ ЗА ДНЕМ



Digital Петр Публикация PDF

4470. Май 27.— Б. И. Куракину

«Господин подполковникъ.»

По получении сего письма пришли у окончнине такарий становок (на котором возможно точить овал і розал) доброва мастерства, і чтоб сюз уже въ дѣлѣ нѣсколько было; также довольно чисто далом, а именно, вдвое ізъ втрое, і съ собою вывези въ Голландію, а оттоль поплы въ Ельбинген или Ригу. О пѣромъ дѣлѣ изъ комплекціи къ вамъ писана. Такжѣ уведай отъ Стальса, даша ли модель, которую я послалъ къ нему съ карабельщикомъ ізъ прошлаго года. Также привези съ собою дѣлѣ ізъ три препятствіи омъду, четырехъ і штихъ длины доброго мастерства, а особенно, чтобъ немало.



Архив СПБИИ РАН. Колл. 277. Оп. 2. Д. V/1. Л. 1, 2.

Письма и бумаги императора Петра Великого. М., 1962. Т. 11. Вып. 1. № 4470. С. 246.

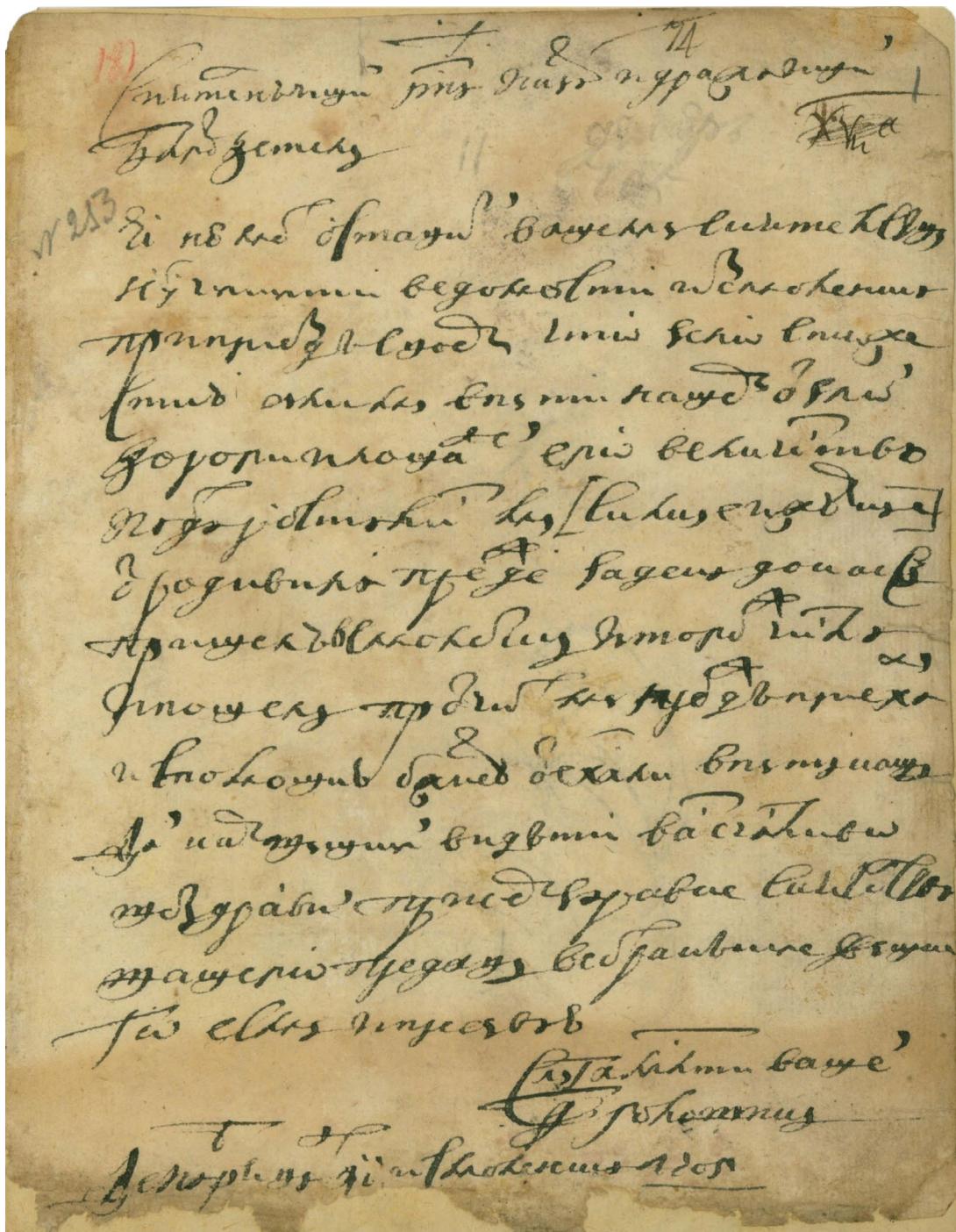
**Ил. 4.** Страница рукописи Петра I (Архив СПБИИ РАН. Колл. 277. Оп. 2. Д. V/1. Л. 1, 2) и публикация текста в издании «Письма и бумаги императора Петра Великого» (М., 1962. Т. 11. Вып. 1. № 4470. С. 246), представленные на сайте «Автографы Петра I»

<https://peterscript.historyrussia.org/document/139>

лет позволяют утверждать, что скоропись, появившаяся на Балканах в XIII в., является специально разработанным стилем курсивного канцелярского письма. Она появилась под влиянием византийского минускула и западноевропейского курсивного письма и предназначалась для написания документов. В XV в. скоропись проникла в Восточную Европу, где постепенно сложились свои школы скорописного письма. В Великом княжестве Московском сформировалась московская, или русская, скоропись, в Великом княжестве Литовском — рутенская, или западно-русская, на территории

молдавских земель — молдавская скоропись, близкая к балканской и рутенской [5]. Особую популярность скоропись приобрела в России, став основой для делопроизводственного письма XVI–XVII вв. Постепенно специфика русской скорописи становилась всё более выраженной, и для нашего современника чтение скорописных документов XVI–XVII вв. требует специальных навыков.

Петровская эпоха — это переходный период, когда традиционная скоропись XVII в. постепенно трансформируется, и проявляется больше индивидуальности в почерках. В написанных чётким писар-



Ил. 5. Письмо Ф.А. Головина А.Д. Меншикову. 14 декабря 1705 г.  
(Архив СПБИИ РАН. Ф. 83. Оп. 1. Д. 797. Л. 1. Автограф)

ским почерком документов появляются подписи, фразы, исправления и приписки соратников государя. Программу необходимо научить читать также и их. Много, бегло и на первый взгляд неразборчиво писал не только Петр I, но и его сподвижники Ф.А. Головин, П.П. Шафиров, П.А. Толстой, Б.И. Куракин и др. (ил. 5).

Таким образом, помимо дальнейшего развития проекта в направлении обучения искусственного интеллекта особенностям почерка Петра Великого во все годы его жизни, не менее важно создание программы, которая могла бы читать русскую скрипцию Раннего Нового времени. Особую актуальность эта проблема приобретает

в связи с тем, что в последние годы российские архивы оцифровывают фонды и выкладывают в свободный доступ тысячи файлов. Огромный массив составляют русские скорописные документы XVI–XVIII вв., с которыми работают не только специалисты историки, филологи и лингвисты, но также краеведы и генеалоги. Эти материалы нужно систематизировать и вести по ним электронный поиск.

### Документы походной канцелярии А.Д. Меншикова и модель по распознаванию письма Петровской эпохи

В 2024 г. сотрудники ПАО «Сбербанк» и СПБИИ РАН начали работу над новой задачей — научить искусственный интеллект читать автографы людей Петровской эпохи. Источниковой базой проекта послужили материалы походной канцелярии самого ближайшего сподвижника Петра Великого — светлейшего князя Александра Даниловича Меншикова (1673–1729) (ил. 6).

По масштабу личности и влиянию на историю России первой четверти XVIII в. его мало с кем можно сравнить. Важный источник могущества А.Д. Меншикова — его незаменимость в системе государственной власти. Наиболее ярко разносторонние дарования А.Д. Меншикова проявились в годы Северной войны (1700–1721).

Карьера А.Д. Меншикова как военного и государственного деятеля началась во время кампании в Ингерманландии. Царь назначил поручика А.Д. Меншикова шлиссельбургским губернатором за мужество и смекалку, проявленные в октябре 1702 г. при взятии шведской крепости Нотебург. Затем по мере продвижения русской армии к Балтийскому морю под его управление переходили все отвоёванные земли. К началу 1705 г. в ведении губернатора оказались огромные территории, по площади сравнимые с европейскими державами, в том числе часть Эстляндии, Приневье, южная часть Ладожского и Онежского озёр, река Свирь, земли до Онеги и Белоозера. А.Д. Меншиков, по-

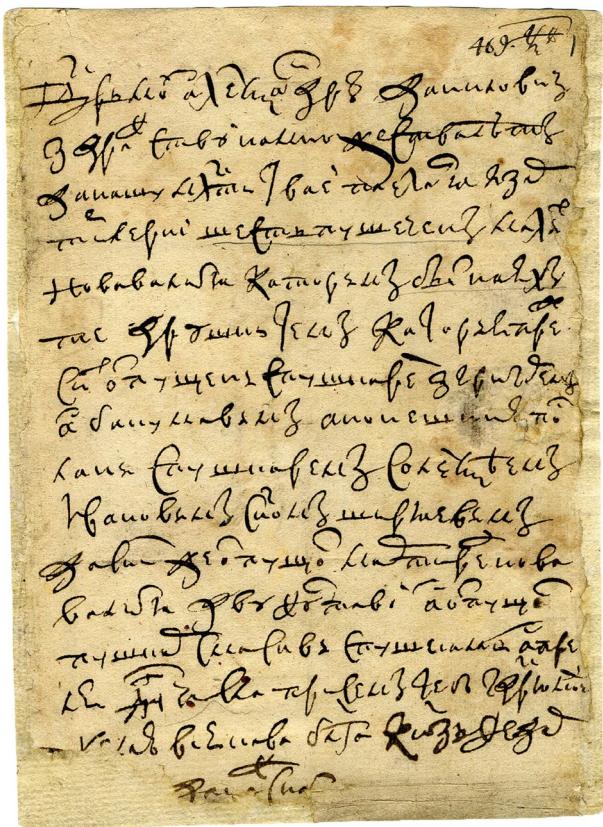


Ил. 6. А.Д. Меншиков (1673–1729). Рис. П.Ф. Бореля.  
Изд-во А.Э. Мюнстера. Середина XIX в.

мимо военных вопросов, ведал логистикой, организацией снабжения русской армии, строительством и обороной Петербурга и других крепостей Ингерманландии.

Стремительный взлёт был обусловлен не только личными качествами и организаторскими способностями А.Д. Меншикова. Близость к правительству стала тем символическим капиталом, которым А.Д. Меншиков воспользовался, чтобы легализовать своё положение как публичного деятеля. Значительно возросли его вес и влияние в среде высших царских сановников, военных и чиновников. В силу близости к царю и исключительной важности поручаемых ему дел Александр Данилович имел связи на разных уровнях и находился в центре обширной «сети» отношений. Деятельность А.Д. Меншикова нашла отражение в материалах его походной канцелярии. Уже в начале XVIII в. его адресантами стали практически все сподвижники Петра I.

В настоящее время основные материалы походной канцелярии А.Д. Меншикова хранятся в Архиве СПБИИ РАН (фонд 83) и РГАДА (фонд 198). Эти документы хорошо известны и востребованы несколькими



Ил. 7. Письмо Ф.Ю. Ромодановского из Москвы  
А.Д. Меншикову. После 14 апреля 1704 г.  
(Архив СПБИИ РАН. Ф. 83. Оп. 1.  
Д. 191. Л. 1. Автограф)

поколениями отечественных историков. Однако современные исследователи продолжают изучать и активно вводить в научный оборот бумаги из архива А.Д. Меншикова для анализа событий Северной войны на территории Ингерманландии, строительства Санкт-Петербурга и других крепостей, становления и способов поддержания социальных связей в России Раннего Нового времени.

В 2022 г. при поддержке РФФИ<sup>1</sup> был подготовлен и опубликован сборник документов «Походная канцелярия Александра Даниловича Меншикова (1703–1705)» [3]. Его основу составила входящая корреспонденция А.Д. Меншикова за первые годы Северной войны. Документы начала 1700-х гг. наиболее полно представлены

в материалах фонда 83, отложившихся в Архиве СПБИИ РАН. В этом фонде сохранилось около 800 документов за 1703–1705 гг. – период, в котором берут начало многие длительные формальные и неформальные контакты А.Д. Меншикова. Судя по дошедшим до нас материалам походной канцелярии, в эти годы у Александра Даниловича было не менее 142 адресантов. Больше всего в фонде до наших дней сохранилось писем генерал-фельдмаршала Б.П. Шереметева. Среди других корреспондентов, которые часто писали А.Д. Меншикову в 1703–1705 гг., были П.М. Апраксин, Р.В. Брюс, Я.В. Брюс, Ф.А. Головин, К.А. Нарышкин, В.И. Пороскин, К.Э. Рённе, А.И. Репнин, Т.Н. Стрешнев, И.Я. Яковлев (ил. 7).

В 1703–1705 гг. А.Д. Меншиков получал сведения от комендантov крепостей, а также командиров различных воинских частей, дислоцировавшихся в Ингерманландии, Великом Новгороде и Пскове. Значительная часть адресантов губернатора не находилась в его непосредственном подчинении. Такие военачальники, как Б.П. Шереметев, П.М. Апраксин, А.И. Репнин, обращались к А.Д. Меншикову прежде всего как к фавориту Петра I, губернатору и главе Семёновской приказной палаты (Ингерманландской канцелярии), который мог ускорить выполнение царских указов и оперативно решить возникшие проблемы. Это были как собственноручные письма, рапорты и донесения, так и документы, написанные писарским почерком, под которым адресант «от руки» добавлял несколько строк илиставил свою подпись. Таким образом, материалы походной канцелярии дали для нового проекта несколько десятков образцов почерков людей Петровской эпохи начала XVIII в. Они были дополнены документами из архива другого сподвижника Петра I – вице-канцлера П.П. Шафирова [4], что позво-

<sup>1</sup> Проект 20-09-42051 (Петровская эпоха) «Походная канцелярия Александра Даниловича Меншикова (1702–1705): Новые подходы к изучению традиционных источников Петровской эпохи». Руководитель проекта – д-р ист. наук Е.В. Анисимов.

лило не только увеличить число почерков, но и отдалить верхнюю хронологическую границу до 1719 г. (ил. 8).

В ходе реализации нового проекта использовался опыт, приобретённый в работе над проектом «Digital Пётр»: готовились парные материалы (цифровая копия листа и разбитый по строкам текст, со всеми буквами, как в подлиннике, без изъятий и замен знаков), которые передавались сотрудникам ПАО «Сбербанк». Они выполняли разметку цифровых копий на платформе ABC Elementary. Изображение разметки возвращалось в СПБИИ РАН, где несколько раз тщательно проверялось. Этот алгоритм работы был опробован ещё в 2020 г. и позволил почти полностью исключить ошибки в транскрибировании текстов и разметке копий. Всего было отобрано более 220 архивных дел, размечено более 720 страниц и около 12 000 строк.

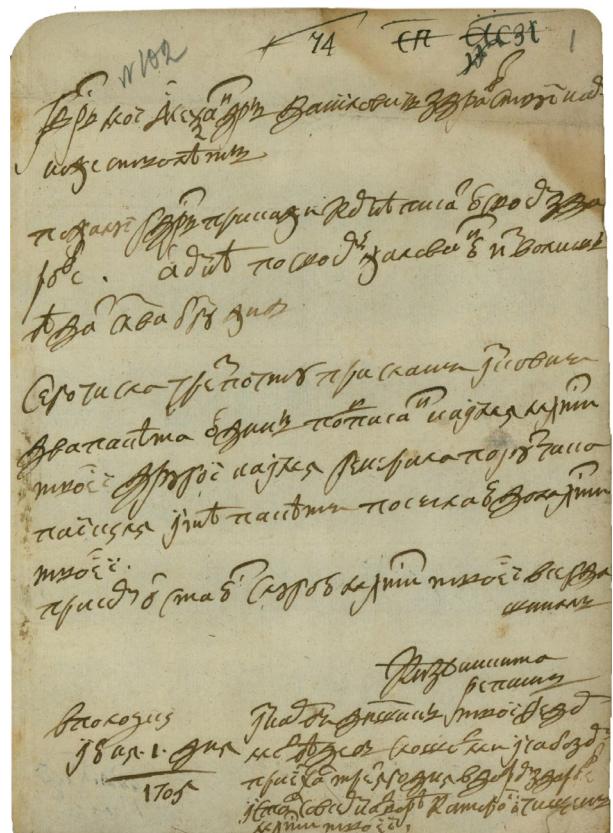
Перед коллективом СПБИИ РАН была поставлена задача максимально расширить число образцов письма, при этом представлять по несколько страниц текста каждого писца. Историки и палеографы передали специалистам по анализу данных образцы 190 почерков, значительная часть которых представлена более чем двумя страницами текста. Так, например, транскрибирован текст семи страниц писем-автографов А.И. Репнина (ил. 9).

В течение 2024 г. сотрудники ПАО «Сбербанк» пробовали возможности различных подходов для распознавания корпуса данных — сегментацию и оптическое распознавание символов (Optical Character Recognition). Новая задача была оценена как значительно более сложная по сравнению с задачей, стоявшей при создании модели по чтению автографов Петра Великого. Очень велик набор вариантов написания той или иной буквы и их соединений, а также обширен словарь применяемой в тексте лексики.

Модель для чтения рукописных текстов Петровской эпохи с нетерпением ждут исследователи, которые работают с большими массивами документов. Полученный в ходе



Ил. 8. П.П. Шафиров (1673–1739). Рис. Я. Аргунова. Гравюра С. Грачёва. Вторая половина XVIII в.



Ил. 9. Письмо А.И. Репнина А.Д. Меншикову. 1 июня 1705 г. (Архив СПБИИ РАН. Ф. 83. Оп. 1. Д. 629. Л. 1. Автограф)

реализации проекта опыт может быть также использован и для изучения эпистолярного наследия видных государственных деятелей России Нового времени. Немаловажны и перспективы разработки программы машинного чтения скорописи XVI–XVII вв. Чтение скорописных почерков этого времени представляет существенную проблему для тех, кто не является специалистом

по истории средневековой России, и затрудняет доступ к информации, содержащейся в неопубликованных документах. Машинное чтение скорописи XVI–XVII вв. даст дополнительный инструментарий и специалистам, поскольку позволит осуществлять быстрый поиск информации в неопубликованных документах, а таковых применительно к XVII в. больше, чем опубликованных.

## ЛИТЕРАТУРА

1. Базарова Т.А., Димитров Д.В., Потанин М.С., Прокурякова М.Е. Распознать и транскрибировать: Автографы Петра Великого и технологии искусственного интеллекта // Воронцово поле. 2020. № 4. С. 36–41.
2. Базарова Т.А., Потанин М.С., Прокурякова М.Е. «...Он нисколько не заботился о синтаксисе»: Может ли искусственный интеллект прочитать письма Петра I? // Родина. 2022. № 6 (июнь). С. 112–115.
3. Походная канцелярия Александра Даниловича Меншикова (1703–1705) / Е.В. Анисимов, Т.А. Базарова, К.Н. Лемешев, М.Е. Прокурякова. СПб.: Историческая иллюстрация, 2022.
4. Походная канцелярия вице-канцлера Петра Павловича Шафирова: Новые источники по истории России эпохи Петра Великого: В 3 ч. / Изд. подгот. Т.А. Базарова, Ю.Б. Фомина; Сост., вступ. ст., comment. Т.А. Базаровой. СПб.: Изд. Дом-мір, 2011.
5. Сиренов А.В. Дискуссия о типах письма древнерусских рукописей // Труды Отделения историко-филологических наук РАН. М.: РАН, 2024. Т. 13 / Отв. ред. В.А. Тишков; Сост. Н.В. Тарасова. С. 265–274.

## Artificial Intelligence and the Study of the Written Legacy of Peter the Great and his Associates

**Aleksey Vladimirovich Sirenov** – Corresponding Member of the Russian Academy of Sciences, Doctor of Sciences in History, Director of the St. Petersburg Institute of History (SPbII) of the Russian Academy of Sciences.

E-mail: [sirenov@rambler.ru](mailto:sirenov@rambler.ru)

**Evgeniy Viktorovich Anisimov** – Doctor of Sciences in History, Chief Researcher of the Department of Medieval History of Russia of the SPbII of the Russian Academy of Sciences, Head of the project “The Field Office of Alexander Danilovich Menshikov (1702–1705): New Approaches to the Study of Traditional Sources of the Petrine Era” (20-09-42051 (Petrine Era)).

E-mail: [vbrevi@mail.ru](mailto:vbrevi@mail.ru)

**Tatiana Anatolievna Bazarova** – Candidate of Sciences in History, Senior Research Fellow, Acting Head of the Department of Source Studies, SPbII RAS.

E-mail: [tbazarova@yandex.ru](mailto:tbazarova@yandex.ru)

**Maria Evgenievna Proskuryakova** – Candidate of Sciences in History, Senior Research Fellow, Department of Source Studies, SPbII RAS.

E-mail: [m-proskuryakova@mail.ru](mailto:m-proskuryakova@mail.ru)

The article presents the results of a collaborative project between historians and paleographers from the St. Petersburg Institute of History of the Russian Academy of Sciences and data analysis specialists from Sberbank. The partnership between the two institutions began in 2020,

initiated by the Russian Historical Society. The team was tasked with training an artificial intelligence model to recognize the autographs of Peter the Great. The successful completion of this task and the creation of the "Digital Peter" model enabled them, in 2024, to begin developing a new algorithm for recognizing the handwriting of individuals from the Petrine era. The article details the team's work on creating data sets for both models, which involved establishing principles for document selection, transcribing the texts, and annotating digital copies of the autographs. Furthermore, it outlines the key features of the electronic resources developed over the years of working with Peter's autographs, namely the "Digital Peter" recognition model and the "Autographs of Peter the Great" website.

**Keywords:** Peter the Great, autographs, epistolary heritage, artificial intelligence, the project "Digital Peter", Peter the Great's era

## REFERENCES

1. Bazarova T.A., Dimitrov D.V., Potanin M.S., Proskuryakova M.E. *Raspoznat' i transkribirovat': Avtografy Petra Velikogo i tekhnologii iskusstvennogo intellekta* // Vorontsovo pole. 2020. № 4. S. 36–41 (in Russian).
2. Bazarova T.A., Potanin M.S., Proskuryakova M.E. «...On niskol'ko ne zabolilsya o sintaksise»: Mozhet li iskusstvennyi intellekt prochitat' pis'ma Petra I? // Rodina. 2022. № 6 (iyun'). S. 112–115 (in Russian).
3. Pokhodnaya kantselyariya Aleksandra Danilovicha Menshikova (1703–1705) / E.V. Anisimov, T.A. Bazarova, K.N. Lemeshev, M.E. Proskuryakova. SPb.: Istoricheskaya illyustratsiya, 2022 (in Russian).
4. Pokhodnaya kantselyariya vitse-kantslera Petra Pavlovicha Shafirova: Novye istochniki po istorii Rossii epokhi Petra Velikogo: V 3 ch. / Izd. podgot. T.A. Bazarova, Yu.B. Fomina; Sost., vступ. ст., komment. T.A. Bazarovoi. SPb.: Izd. Dom-mir, 2011 (in Russian).
5. Sirenov A.V. Diskussiya o tipakh pis'ma drevnerusskikh rukopisei // Trudy Otdeleniya istoriko-filologicheskikh nauk RAN. M.: RAN, 2024. T. 13 / Otv. red. V.A. Tishkov; Sost. N.V. Tarasova. C. 265–274 (in Russian).



С.Ю. СТЕПАНОВ\*

## Рефлексия и искусственный интеллект: от психологии рефлексии к рефлексивно- цифровым практикам развития человека

Статья прослеживает эволюцию представлений о рефлексии от её истоков у Дж. Локка и в интроспективной психологии до становления самостоятельной области психологоческой науки в конце XX в. Автор подчёркивает важность институционализации психологии рефлексии в 1990-е гг. для разработки рефлексивных практик и психотехник для консультирования, тренингов, психотерапии и образования. Основной смысловой фокус статьи сделан на революционном воздействии цифровых технологий и искусственного интеллекта (ИИ) на выявление новых возможностей рефлексии в XXI в. Показано, что интеграция ИИ в сферу жизнедеятельности человека трансформирует рефлексивные процессы и связанные с ними практики. Возникает феномен «би-рефлексии» при взаимодействии человеческого сознания и машинного интеллекта, выражающийся в появлении «цифровых кентавров» и новых рефлексивно-цифровых практик. Эти практики расширяют когнитивные и личностные ресурсы, одновременно ставя вопросы о меняющейся субъектности, идентичности и этике, а также стимулируют «очеловечивание» ИИ через наделение его функциями креативности и рефлексивности. Автор представляет две перспективные концептуальные разработки. Первая – Цифровой Ангел (DA): персональный ИИ-агент, выступающий как психологический буфер и защита от рисков цифрового мира (киберугроз, конкуренции с AGI и т.п.). DA обеспечивает безопасное цифровое пространство, анализирует данные пользователя для усиления его рефлексии и саморазвития. Вторая – Цифровое Альтер-Эго (DAE): технология для «конструирования» и «проживания» новых личностных амплуа в виртуальных/дополненных реальностях. DAE позволяет актуализировать у человека неочевидные способности и ресурсы, создавая уже не «зону ближайшего развития» (по Л.С. Выготскому), а «зону отдалённого развития» – перспективы самореализации через взаимодействие с ИИ. В статье утверждается, что технологии DA и DAE, основанные на отечественных разработках психологии рефлексии, определяют новый вектор для цифровой гуманистики и гибридных рефлексивных практик.

**Ключевые слова:** рефлексия, искусственный интеллект, рефлексивно-цифровые практики, цифровой ангел, цифровое альтер-эго

\* Степанов Сергей Юрьевич — доктор психологических наук, профессор Московского государственного педагогического университета, руководитель Центра психологии метакогнитивных процессов и рефлексивных практик Института психологии РАН.

E-mail: parusnik1@ya.ru

## Гетерохронность проблематики рефлексии и искусственного интеллекта

Представления о рефлексии и искусственном интеллекте (ИИ) формировались в ходе человеческой истории в разные времена, в разных социокультурных контекстах и с разным темпом. Если на формирование представлений о первой ушли века, то о втором – лишь несколько десятилетий. Несмотря на это, и для первой, и для второго современность является моментом чрезвычайно интенсивного развития, во многом обусловленным тем, что они стали активно взаимодействовать как в теоретической плоскости профессионального сознания, так и в технологическом и практическом своём воплощении. В данной статье попробуем раскрыть коллизии их взаимообусловленности и перспективы применения в контексте проблематики развития человека. Однако прежде чем непосредственно приступить к выполнению этой задачи, обозначим для начала трек становления исследований рефлексии до начала разработок ИИ и цифровых технологий.

### Философско-психологический трек изучения рефлексии

В конце XVII в. Дж. Локк в своём знаменитом труде «Опыт о человеческом разумении» впервые ввёл термин «рефлексия» как важнейший источник знаний, приобретаемых на основе наблюдения человеком за деятельностью своего ума и души [7]. Данное представление о рефлексии воплотилось в главном гностическом инструменте возникшей на рубеже конца XIX – начала XX в. молодой психологической науки – в методе интроспекции. Позже, однако, метод самонаблюдения под давлением позитивистической методологии был раскритикован и элиминирован в психологии, а вслед за ним – и понятие «рефлексия».

До начала 70–80-х прошлого столетия рефлексия изредка использовалась в психологии и только лишь отдельными авторами как объяснительный принцип, позво-

ляющий интерпретировать и трактовать различные феномены, связанные с процессами психической саморегуляции человека и развития теоретического мышления. В работах В.В. Давыдова и его последователей рефлексия определялась как психический механизм осознания умственных действий и основа теоретического мышления [3]. У Я.А. Пономарёва, Н.Г. Алексеева, И.Н. Семёнова и В.К. Зарецкого рефлексия фигурировала в качестве уровневого компонента, регулирующего процесс творческого мышления [8].

Благодаря работам, опубликованным в 1982 [17] и 1983 г. [10], рефлексия стала обретать статус особой эмпирической реальности и предметно-психологической области исследований, что привело к формированию отдельной области науки – психологии рефлексии [18]. В рамках этой области удалось концептуально обосновать, методически выделить, эмпирически различить и экспериментально верифицировать в изначально синкетичном представлении о рефлексии не только её отдельные уровневые компоненты (например, такие виды, как ретроспективная, ситуативная, перспективная, экстенсивная, интенсивная, конструктивная рефлексия), но и различные типы рефлексии: интеллектуальная и личностная, коммуникативная и кооперативная, экзистенциальная и культуральная [18]. Рефлексия стала трактоваться не просто как сугубо интеллектуальный процесс аутоперцепции, осознания и ретроспективного анализа оснований мыслительных действий и переживаний человека, но как личностно обусловленный процесс осмыслиения и переосмыслиения стереотипов его опыта, а также содержаний (структур и процессов) сознания, мышления и поведения людей, лежащий в основе саморегуляции и самоорганизации, самодетерминации и саморазвития человека и приводящий к творческим открытиям и новациям в его деятельности и взаимодействии с другими людьми [18].

Осуществлённые в 90-е годы институционализация психологии рефлексии, разра-



Ил. 1. Рефлексирующий ИИ  
(здесь и далее помещены иллюстрации,  
созданные ИИ)

ботка метаметода сценарирования и проведения рефлексивных практик, а также отдельных рефлексивных психотехник (полилог, позиционная дискуссия, «коктейлирование», полипленум, композиция «проблемного поля», версификация символических инверсий, аукцион идей и т.д. и т.п.) [14] стали отправной точкой для их широкого применения в психологическом, управлении, организационном и политическом консультировании [13], при подготовке и проведении рефлексивных тренингов [9] и инновационных практик, направленных на модернизацию деятельности образовательных организаций [14]. Вместе с Е.П. Варламовой нам удалось создать концептуальные основы и разработать методические процедуры рефлексивной диагностики и самодиагностики, обладающие высоким практическим потенциалом для исследования и проектирования

процессов самопознания и саморазвития человека и организаций [2].

Вслед, а затем и параллельно с нашими разработками по психологии рефлексии уже в начале XXI века начали интенсивно разворачиваться исследования других отечественных психологов: А.В. Карпова [5], Д.А. Леонтьева, Е.Н. Осина [6], В.Д. Шадрикова, С.С. Кургиянья [19] и др. Помимо отечественных разработок по проблематике рефлексии, с 80-х гг. XX в. её интенсивная проработка осуществлялась и за рубежом в рамках психологии метакогнитивизма, прежде всего в исследованиях Ж. Пиаже, Дж. Дьюи, Дж. Флейвелла, Д. Дёрнера, Д. Шёна, Т. Бортона, Д. Колба и других.

### Рефлексивный гребень цифровой волны

Однако все описанные достижения меркнут на фоне кардинальных подвижек в понимании рефлексии и её возможностей, которые стали происходить уже в XXI в., т.е. с началом создания компьютерно-цифровых технологий машинного обучения, обработки больших данных и нейронных сетей. Тогда специалисты в области ИТ-индустрии и Интернета стали активно асимилировать результаты исследований в области психологии мышления, творчества и других когнитивных процессов, что привело к революционным подвижкам в области ИИ, к созданию генеративных нейросетей и больших языковых моделей (LLM) (ил. 1).

На фоне активно развивающихся технологий ИИ, инкорпорируемых практически во все сферы психического бытия современного человека, с неизбежностью начинают трансформироваться опосредуемые ими психические процессы и структуры, в том числе и рефлексия. Это, в частности, подтверждают эмпирические исследования Г.У. Солдатовой с коллегами, в которых показано, что «гиперподключённость» человека к взаимодействию с цифровыми ресурсами [12] меняет структуру самоидентичности и самоотношения, а значит, и характер личностной рефлексии чело-



Ил. 2. Цифровой двойник

века [20]. Опираясь на возможности ИИ, человек открывает средства к усилению и изменению рефлексии, направленной не только на самого себя (т.е. интеллектуальной, личностной и экзистенциальной), но и на переосмысление действий и поведения других людей и их сообществ (т.е. коммуникативной, кооперативной и организационной).

Цифровые платформы, такие как чат-боты или системы анализа поведения, становятся «внешними агентами рефлексии», предлагая альтернативные интерпретации действий и мыслей пользователей [11]. Это создаёт феномен бирефлексии, где человеческое сознание и машинный интеллект взаимодействуют и взаимообуслов-

ливают друг друга, формируя гибридные когнитивно-цифровые модели и «цифровые кентавры» [12]. Тем самым сейчас мы становимся свидетелями и участниками формирования новых видов практик, а именно рефлексивно-цифровых практик, которые, основываясь на взаимодействии человека с ИИ, с одной стороны, открывают новые возможности для расширения когнитивно-интеллектуальных и личностно-эмоциональных ресурсов личности, а также ставят вопросы о переосмыслинии природы субъектности, формах человеческой идентичности, этических границах технологий и, в том числе, трансформации существующих моделей образования и терапии. С другой же сторо-

ны, трансформации затрагивают не только психическую сферу человека, но параллельно инициируют попытки разработчиков гуманизировать цифровую метавселенную и «очеловечить» искусственный интеллект, программируя его в направлении ресурсов и функций, которые ещё до недавнего времени считались прерогативой человеческой психики, а именно креативность, рефлексивность, диалогичность и т.п. (ил. 2).

К примеру, уже сейчас создаются человекоподобные голограммические аватары на основе моделирования чувствительного цифрового двойника, которые представляют собой сочетание эмоциональных вычислений, аффективной обратной связи и технологий Интернета вещей, которые предназначены для создания «чувствительной цифровой среды». Эта модель собирает и интегрирует мультимодальные данные в режиме реального времени в систему, которая может интерпретировать поведение реального человека и прогнозировать его операционные планы, синхронизируя все эти данные с поведением чувствительного цифрового двойника, который может выполнять роль, например, интерактивного «экскурсовода» в музее или «учителя» во время урока [21]. Очевидно, что здесь речь идёт об «овладении» искусственным интеллектом в образе «чувствующего» аватара функцией коммуникативной и кооперативной рефлексии.

Разработки отечественной психологии рефлексии также находят своё воплощение в области рефлексивно-цифровых практик и бирефлексивных моделей взаимодействия человека с ИИ.

### От цифрового ассистента к рефлексивно-цифровому ангелу

Так, размышляя на рубеже 2018–2019 гг. вместе с Д.В. Ушаковым над перспективами развития технологии ИИ применительно к образовательным и психотерапевтическим задачам, мы предположили, что ближайшим и ключевым трендом её прогресса станет переход от разрабо-

ток Общего (Глобального и/или Супер) ИИ к созданию технологий Индивидуализированного (персонального) ИИ [22]. Этот прогноз основывался на рефлексии того факта, что подобное уже произошло трёх-четырьмя десятилетиями раньше с компьютерами, главный технологический сдвиг в разработке которых был сделан при переходе от больших вычислительных центров к персональным компьютерам и цифровым гаджетам.

Сегодня идеей персональных цифровых ассистентов, помощников, агентов и т.п. на основе ИИ уже никого не удивишь. Вместе с тем, анализируя предложенные на сегодняшний день модели цифровых ассистентов, можно сказать, что практически ни одна из этих цифровых «поделок» пока не дотягивает до предложенного нами концепта «Цифрового Ангела», поскольку их разработчики стремятся удовлетворить сиюминутные, по большей части утилитарные и весьма прагматические потребности людей, использующих возможности ИИ, которые связаны с решением, как правило, бытовых, развлекательных, сервисных, бизнесовых или узкопрофессиональных задач (ил. 3).

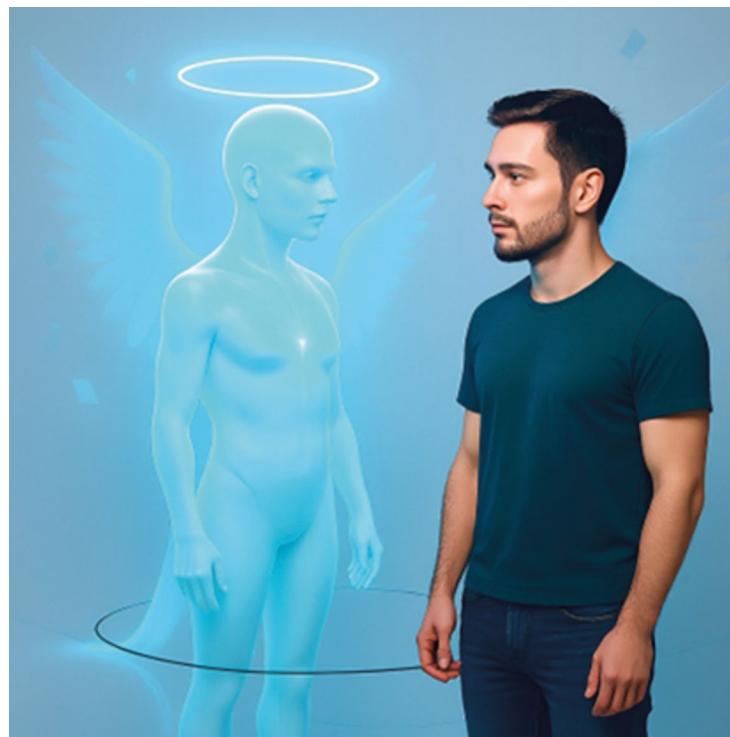
Концепция же ДА предполагает, что его возможности не просто ограничиваются вспомогательно-сервисными функциями, направленными на разгрузку человека от утилитарных и рутинных задач интеллектуального характера, но прежде всего на обеспечение его большей рефлексивности и, следовательно, большей защищённости от рисков и напастей, нарастающих в лавинообразно цифровизирующемся мире с разгулом кибермошенничества, индоктринационных и манипулятивных технологий, кибербуллинга, сетевого троллинга фейков и т.п. цифровых угроз. Усилившиеся возможности ИИ, в контексте вероятной конкуренции с естественным интеллектом человека и человечества в целом, грозящие обрушением рынков труда, также являются причиной всё возрастающей тревоги не только со стороны общественности, но даже самих IT-разра-

ботчиков. Из-за указанных рисков разработка технологии DA становится насущной необходимостью, поскольку она и должна стать своего рода рефлексивно-цифровым волнорезом перед нарастающей экспансией «Общего искусственного интеллекта» (Artificial General Intelligent – AGI) в различных его видах и формах.

В контексте разворачивающихся сегодня дискуссий о понимании позиционирования AGI относительно человека либо как «конкурирующего» (competitive), либо как «усиливающего» (augmenting), либо как «дополняющего» (complementary) разработка технологии DA как раз и может стать ключом к тому, чтобы первый вариант не стал доминирующим, а тем более единственным. А об опасности такого рода перспективы заявил Илон Маск – один из основных инициаторов и инвесторов разработки ChatGPT. В частности, в апреле 2024 г. он написал, что к концу 2025 г. возможности новых моделей искусственного интеллекта превзойдут человеческий разум, и в течение следующих пяти лет возможности AGI, способного выполнять умственную работу на человеческом уровне, превзойдут возможности всех людей [4].

DA – это технология, всецело выстроенная в интересах персоны конкретного человека, являющаяся психологическим буфером между ним и в первую очередь AGI, а также другими различными цифровыми сущностями (неважно, «добрьими» или «злыми», в силу относительности этих оценочных категорий в различных контекстах). Тем самым DA должен встать «на границе» цифрового мира и человека с его естественным интеллектом, обеспечив его безопасным личным цифровым пространством жизнедеятельности.

По нашему замыслу DA должен стать личным конфидентом человека, его



Ил. 3. «Цифровой Ангел» — персонализированный ИИ

рефлексивно-цифровым двойником, обеспечивающим не только обратной связью относительно интеракций с цифровым миром, но также прогнозами и оценками рисков и возможностей, которые они сулят ему в доступном для прогнозирования пространстве его жизнедеятельности. В этом смысле одной из важных составных частей данной технологии станет цифровой двойник, создаваемый и непрерывно обновляющийся за счёт накопления «больших данных» в виде цифровых следов жизнедеятельности человека, а также аналитики, которая будет материалом для его когнитивной, социальной, интеллектуальной и экзистенциальной рефлексии, а значит, основой для дальнейшего само-проектирования им своих жизненных треков развития.

### Цифровое Альтер-Эго – мультiplикатор человеческих возможностей и способностей

Важным добрачиванием технологии Цифрового Ангела может стать ещё одна предлагаемая нами технология – Цифро-



Ил. 4. Цифровое Альтер-Эго человека

вого Альтер-Эго (Digital Alter Ego – DAE). Она может выступать в качестве источника неведомых и неочевидных ранее психологических возможностей, ресурсов и способностей человека (ил. 4). DAE – это персонализированная технология ИИ, которая позволит человеку погружаться и действовать не просто в компьютерно-игровой виртуальной и/или дополненной реальности в качестве «аватара», но конструировать и примерять на себя новые личностные, интеллектуальные и коммуникативные амплуа с самым разнообразным набором мыслительных действий, ролей, компетенций, характерологических особенностей, потребностей и мотивов. И как актёр благодаря драматургу и режиссёру перевоплощается в определённую роль или сценического героя, проживая все коллизии и перипетии их судьбы, точно так же и любой человек с помощью DAE сможет разрабатывать различные личностные амплуа, вживаться в них и реализовывать себя в новых ролях в самых разнообразных сценариях с набором заранее запланированных и/или, наоборот, не-

предсказуемых жизненных ситуаций и коллизий, профессиональных профилей, образовательных и карьерных треков. Такие сценарии, генерируемые совместно человеком и ИИ, будут содержать различные проблемные, конфликтные и даже экзистенциальные коллизии. Тем самым человек сможет в проактивном режиме актуализировать, реализовывать и рефлексировать с помощью DAE различные свои «дремлющие» и наличные психологические задатки и способности, в том числе креативного характера.

В этом контексте может получить своё дальнейшее развитие весьма популярная в отечественной образовательной практике концепция «зоны ближайшего развития», предложенная Л.С. Выготским, когда коммуникация и взаимодействие человека с меньшей компетентностью (как правило, таковым является ребёнок) с человеком, обладающим большей компетентностью (здесь обычно подразумевается взрослый), создаёт для первого «зону ближайшего развития» (ЗБР). Интеракции же человека с ИИ, который уже сейчас аккумулирует практически весь доступный на данным момент опыт человечества за всю историю его существования, в рамках технологии DAE смогут, с нашей точки зрения, создавать для конкретного человека зону его отдалённого развития (ЗОР). Ведь вживание в различные комбинации психологических амплуа, неожиданных для человека, их поведенческое апробирование в предлагаемых виртуальных или дополненных реальностях при взаимодействии с другими подобными же персоналиями в ситуациях с высоким уровнем неопределённости и новизны условий самореализации должно будет актуализировать новые для самого человека ценности, мысли и действия, а также мобилизовать силы, о которых он мог не догадываться или даже не смел помышлять.

Возможность создания эффекта «зоны отдалённого развития» была проверена нами на практике. В качестве концептуально-психологического прообраза DAE выступила одна из психотехник, которые разрабатывались и успешно апробировались нами с конца 90-х годов в целом ряде рефлексивно-креативных практик в работе с командами и коллективами образовательных, государственных и бизнес-организаций целого ряда регионов [15]. Изначально эта психотехника была названа «методом рефлексивных инверсий и с творческих импровизаций». С её помощью создались условия, когда участники могли актуализировать, переосмысливать и преодолевать свои внутри- и межличностные проблемы и конфликты, производные от тех стереотипов поведения, которые выступают для них барьерами в личностном росте, в образовании, в развитии их профессиональных способностей и компетенций.

Процедурно это достигалось тем, что участникам рефлектируемости предлагалось: 1) вспомнить своих любимых сказочных, литературных или киногероев; 2) в ходе коллективного обсуждения произвести инверсии имён и психотипов последних (например, за счёт изменения нескольких букв в именах этих героев так, чтобы оставалось созвучие с исходным прототипом, но изменялся смысловой и психологический профиль персонажа); 3) дофантазировать (досочинить) их жизненные треки на основе осуществлённых инверсий; 4) вжиться в полученные таким образом инвертированные психологические амплуа; 5) реализовать их в смоделированных проблемно-конфликтных ситуациях в режиме психологической и/или социально-ролевой импровизации перед другими командами участников, выступающих в качестве рефлексивных супервизоров; и, наконец, 6) осуществить специально организованную личностную, интеллектуальную, коммуникативную, кооперативную и даже экзистенциальную рефлексию с помощью инструментов психо- и социографирования.

С одной стороны, такая рефлексивно-практическая проработка способствовала актуализации различных личностных стереотипов поведения (комплексов, по З. Фрейду) и социокультурных стереотипов (архетипов коллективного бессознательного, по К. Юнгу), а также их проживанию и изживанию в психологически защищённой среде и, соответственно, переосмыслению и креативной трансформации в новые степени личной свободы, в новые способности и горизонты возможностей развития. Как правило, изумлению и радости от открываемых как в себе, так и в других участниках неожиданных возможностей, талантов и способностей не было пределов. Глубокая рефлексия проживания такого опыта открывала удивительные и неожиданные для самих участников личностные качества креативной направленности, а также приводила к рекреативным и/или психотерапевтическим эффектам, к существенным, как правило позитивным, изменениям в жизненных и карьерных треках [1].

Интересно, что после уже многих лет реализации психотехники и рефлексивных практик создания «зон отдалённого развития» в пространстве социокультурного освоения появился художественный аналог технологии DAE в фантастическом фильме Джеймса Кэмерона «Аватар», вышедшем на экраны в 2009 г., где главный герой имел возможность за счёт особой ментально-цифровой и генно-биологической технологии далёкого будущего, созданной людьми, «переселяться» из своей телесной оболочки человека-инвалида в оболочку фантастического существа с планеты Пандора. Это открыло для него совершенно новую, необычную и даже невероятную по своей экзистенциальной полноте и насыщенности жизнь, в событийных перипетиях которой он попутно овладел целым рядом возможностей и способностей, не свойственных для человека, но присущих инопланетным аборигенам Пандоры.

В связи с этим важно отметить: опыт разработки и внедрения цифровых и генных технологий, основанных на ИИ, дока-

зывает: то, что вчера казалось несбыточной фантазией, завтра может стать обычной и повседневной реальностью.

И действительно, современные технологии на базе ИИ – генеративные нейросети (для создания текстового, аудио- и видеоконтента, голограмм), big-data, 3D-моделирование, цифровые двойники, чат-боты, компьютерные игры и квесты, дополненная и виртуальная реальность и т.д. и т.п. – уже вплотную подводят нас к возможности создания технологии DAE. Встраивание в контекст этих технологических новаций концепций DA и DAE позволяет целенаправить их дальнейшее совершенствование и развитие по абсолютно новому и весьма перспективному пути, придав им особо ценный психологический смысл и социо-

культурное значение в контексте цифровой гуманитаристики.

Для психологии же разработка технологии DAE наряду с технологией DA актуализирует как целый ряд проблем, так и новых возможностей для фундаментальных и прикладных исследований. В первую очередь это вопросы соотношения рациональных и иррациональных, рефлексивных и бессознательных, эксплицитных и имплицитных, лингвистических и паралингвистических процессов и феноменов в цифровой реальности и взаимодействии человека и человеческих сообществ с ИИ. Но это уже темы для дальнейших размышлений и разработок в области рефлексивно-цифровых технологий и моделей бирефлексивных практик.

## ЛИТЕРАТУРА

1. Варламова Е.П., Степанов С.Ю. Психология творческой уникальности человека: рефлексивно-гуманистический подход. М.: ИП РАН, 2002.
2. Варламова Е.П., Степанов С.Ю. Рефлексивная диагностика в системе образования // Вопросы психологии. 1997. № 5. С. 28–43.
3. Давыдов В.В. Теория развивающего обучения. М.: Педагогика, 1996.
4. Илон Маск заявил о возможности ИИ стать умнее человека уже в 2025 г. [Электронный ресурс] // Форбс. 2025. 4 апр. URL: <https://www.forbes.ru/milliardery/509969-ilon-mask-zaavil-o-vozmozhnosti-ii-stat-umnee-celoveka-uze-v-2025-godu> (дата обращения: 25.05.2025).
5. Карпов А.В. Рефлексивность как психическое свойство и методика её диагностики // Психологический журнал. 2003. Т. 24, № 5. С. 45–57.
6. Леонтьев Д.А., Осин Е.Н. Рефлексия «хорошая» и «дурная»: от объяснительной модели к дифференциальной диагностике // Психология. Журнал Высшей школы экономики. 2014. Т. 11, № 4. С. 110–135.
7. Локк Дж. Сочинения: В 3 т. Т. 1. М.: Мысль, 1985.
8. Пономарёв Я.А., Семёнов И.Н., Зарецкий В.К. [и др.] Исследование проблем психологии творчества / Под ред. Я.А. Пономарёва. М.: Наука, 1983.
9. Растворников А.В., Степанов С.Ю., Ушаков Д.В. Рефлексивное развитие компетентности в совместном творчестве. М.: PerSe, 2002.
10. Семёнов И.Н., Степанов С.Ю. Проблема предмета и метода психологического изучения рефлексии // Исследование проблем психологии творчества / Под ред. Я.А. Пономарёва. М.: Наука, 1983. С. 154–182.
11. Скурут А.И. ИИ как расширение рефлексивных способностей [Электронный ресурс]. URL: [https://www.b17.ru/blog/ai\\_and\\_reflection](https://www.b17.ru/blog/ai_and_reflection) (дата обращения: 25.06.2024).
12. Солдатова Г.У., Чигарькова С.В., Илюхина С.Н. Метаморфозы идентичности человека достроенного: от цифрового донора к цифровому кентавру // Социальная психология и общество. 2024. Т. 15, № 4. С. 40–57. DOI: 10.17759/sps.2024150404.
13. Степанов С.Ю. Развитие рефлексивной культуры госслужащих инновационно-гуманистическими методами: Дисс. ... докт. психол. н. М.: РАГС, 1999.

14. Степанов С.Ю. Рефлексивная практика творческого развития человека и организаций. М.: Наука, 2000.
15. Степанов С.Ю. Рефлексивно-гуманистическая психология створчества. М.; Петрозаводск, 1996.
16. Степанов С.Ю., Оржековский П.А., Ушаков Д.В. [и др.] Цифровизация образования: психолого-педагогические и валеологические проблемы / Под редакцией С.Ю. Степанова. М.: МГПУ, 2021.
17. Степанов С.Ю., Семёнов И.Н. Проблема формирования типов рефлексии в решении творческих задач // Вопросы психологии. 1982. № 1. С. 99–104.
18. Степанов С.Ю., Семёнов И.Н. Психология рефлексии: проблемы и исследования // Вопросы психологии. 1985. № 3. С. 31–40.
19. Шадриков В.Д., Кургинян С.С. Исследование рефлексии деятельности и её диагностика через оценку конструктов психологической функциональной системы деятельности // Экспериментальная психология. 2015. Т. 8, № 1. С. 106–126.
20. Belk R.W. Possessions and the Extended Self // Journal of Consumer Research. 1988. Vol. 15, № 2. Pp. 139–168.
21. Fominska I., Di Tore S., Nappi M., Iovane G., Sibilio M., Gelo A. Approaches to identifying emotions and affections during the museum learning experience in the context of the future internet // Future Internet. 2024. Vol. 16, № 11. Art. 417. DOI: 10.3390/fi16110417.
22. Stepanov S.Yu., Ushakov D.V. Artificial intelligence and digital angel technology in education // Восьмнадцатая Национальная конференция по искусственному интеллекту с международным участием КИИ–2020. М.: Рос. ассоц. искусств. интеллекта, 2020. С. 105–112.

## Reflection and Artificial Intelligence: From the Psychology of Reflection to Reflexive Digital Practices of Human Development

**Sergey Yurievich Stepanov** – Doctor of Sciences in Psychology, Professor at the Moscow State Pedagogical University, Head of the Center for the Psychology of Metacognitive Processes and Reflexive Practices of the Institute of Psychology of the Russian Academy of Sciences.

E-mail: parusnik1@ya.ru

This article traces the evolution of the concept of reflection from its origins with J. Locke and introspective psychology to its establishment as a distinct field within psychological science by the end of the 20<sup>th</sup> century. The author emphasizes the importance of the institutionalization of the psychology of reflection in the 1990s for developing reflective practices and psycho-techniques used in counseling, training, psychotherapy, and education.

The main semantic focus of the article is on the revolutionary impact of digital technologies and artificial intelligence (AI) on uncovering new possibilities for reflection in the 21<sup>st</sup> century. It demonstrates that the integration of AI into human activity transforms reflective processes and related practices. This gives rise to the phenomenon of "bi-reflection" during the interaction between human consciousness and machine intelligence, manifesting in the emergence of "digital centaurs" and new reflective-digital practices.

These practices expand cognitive and personal resources while simultaneously raising questions about changing subjectivity, identity, and ethics. They also stimulate the "humanization" of AI by endowing it with functions of creativity and reflectivity.

The author presents two promising conceptual developments. The first is the Digital Angel (DA): a personal AI agent that acts as a psychological buffer and protection against the risks of the digital world (such as cyber threats and competition with AGI). The DA provides a se-

ture digital space and analyzes user data to enhance their reflection and self-development. The second is the Digital Alter-Ego (DAE): a technology for “constructing” and “experiencing” new personal roles in virtual and augmented realities. The DAE allows individuals to activate latent abilities and resources, creating not a “zone of proximal development” (per L.S. Vygotsky) but a “zone of distal development” – offering prospects for self-realization through interaction with AI. The article argues that DA and DAE technologies, grounded in domestic advancements within the psychology of reflection, define a new vector for the digital humanities and hybrid reflective practices.

**Keywords:** reflection, artificial intelligence, reflexive digital practices, digital angel, digital alter ego

## REFERENCES

1. Varlamova E.P., Stepanov S.Yu. Psikhologiya tvorcheskoi unikal'nosti cheloveka: refleksivno-gumanisticheskii podkhod. M.: IP RAN, 2002 (in Russian).
2. Varlamova E.P., Stepanov S.Yu. Refleksivnaya diagnostika v sisteme obrazovaniya // Voprosy psichologii. 1997. № 5. S. 28–43 (in Russian).
3. Davydov V.V. Teoriya razvivayushchego obucheniya. M.: Pedagogika, 1996 (in Russian).
4. Ilon Mask zayavil o vozmozhnosti II stat' umnee cheloveka uzhe v 2025 g. [Elektronnyi resurs] // Forbs. 2025. 4 apr. URL: <https://www.forbes.ru/milliardery/509969-ilon-maskzaavil-o-vozmozhnosti-ii-stat-umnee-celoveka-uze-v-2025-godu> (data obrashcheniya: 25.05.2025) (in Russian).
5. Karpov A.V. Refleksivnost' kak psikhicheskoe svoistvo i metodika eyo diagnostiki // Psichologicheskii zhurnal. 2003. T. 24, № 5. S. 45–57 (in Russian).
6. Leont'ev D.A., Osin E.N. Refleksiya «khoroshaya» i «durnaya»: ot ob"yasnitel'noi modeli k differentsial'noi diagnostike // Psichologiya. Zhurnal Vysshei shkoly ekonomiki. 2014. T. 11, № 4. S. 110–135 (in Russian).
7. Lokk Dzh. Sochineniya: V 3 t. T. 1. M.: Mysl', 1985 (in Russian).
8. Ponomaryov Ya.A., Semyonov I.N., Zaretskii V.K. [i dr.] Issledovanie problem psichologii tvorchestva / Pod red. Ya.A. Ponomaryova. M.: Nauka, 1983 (in Russian).
9. Rastyannikov A.V., Stepanov S.Yu., Ushakov D.V. Refleksivnoe razvitiye kompetentnosti v sovmestnom tvorchestve. M.: PerSe, 2002 (in Russian).
10. Semyonov I.N., Stepanov S.Yu. Problema predmeta i metoda psikhologicheskogo izucheniya refleksii // Issledovanie problem psikhologii tvorchestva / Pod red. Ya.A. Ponomaryova. M.: Nauka, 1983. S. 154–182 (in Russian).
11. Skurtul A.I. II kak rasshirenie refleksivnykh sposobnostei [Elektronnyi resurs]. URL: [https://www.b17.ru/blog/ai\\_and\\_reflection](https://www.b17.ru/blog/ai_and_reflection) (data obrashcheniya: 25.06.2024) (in Russian).
12. Soldatova G.U., Chigar'kova S.V., Ilyukhina S.N. Metamorfozy identichnosti cheloveka dostroenennogo: ot tsifrovogo donora k tsifrovomu kentavru // Sotsial'naya psichologiya i obshchestvo. 2024. T. 15, № 4. S. 40–57. DOI: 10.17759/sps.2024150404 (in Russian).
13. Stepanov S.Yu. Razvitiye refleksivnoi kul'tury gossluzhashchikh innovatsionno-gumanisticheskimi metodami: Diss. ... dokt. psikhol. n. M.: RAGS, 1999 (in Russian).
14. Stepanov S.Yu. Refleksivnaya praktika tvorcheskogo razvitiya cheloveka i organizatsii. M.: Nauka, 2000 (in Russian).
15. Stepanov S.Yu. Refleksivno-gumanisticheskaya psichologiya sotvorchestva. M.; Petrozavodsk, 1996 (in Russian).

16. Stepanov S.Yu., Orzhekovskii P.A., Ushakov D.V. [i dr.] Tsifrovizatsiya obrazovaniya: psichologo-pedagogicheskie i valeologicheskie problemy / Pod redaktsiei S.Yu. Stepanova. M.: MGPU, 2021 (in Russian).
17. Stepanov S.Yu., Semyonov I.N. Problema formirovaniya tipov refleksii v reshenii tvorcheskikh zadach // Voprosy psichologii. 1982. № 1. S. 99–104 (in Russian).
18. Stepanov S.Yu., Semyonov I.N. Psichologiya refleksii: problemy i issledovaniya // Voprosy psichologii. 1985. № 3. S. 31–40 (in Russian).
19. Shadrikov V.D., Kurginyan S.S. Issledovanie refleksii deyatel'nosti i eyo diagnostika cherez otsenku konstruktov psichologicheskoi funktsional'noi sistemy deyatel'nosti // Eksperimental'naya psichologiya. 2015. T. 8. № 1. S. 106–126 (in Russian).
20. Belk R.W. Possessions and the Extended Self // Journal of Consumer Research. 1988. Vol. 15, № 2. Pp. 139–168.
21. Fominska I., Di Tore S., Nappi M., Iovane G., Sibilio M., Gelo A. Approaches to identifying emotions and affections during the museum learning experience in the context of the future internet // Future Internet. 2024. Vol. 16, № 11. Art. 417. DOI: 10.3390/fi16110417.
22. Stepanov S.Yu., Ushakov D.V. Artificial intelligence and digital angel technology in education // Vosemnadtsataya Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem KII–2020. M.: Ros. assots. iskusstv. intellekta, 2020. S. 105–112.

# ПАРАЛЛЕЛИ МЕЖДУ ЕСТЕСТВЕННЫМ И ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ

Индекс УДК 81-133, 004.82

Код ГРНТИ 16.31.21, 16.31.41

DOI: 10.22204/2587-8956-2025-04-122-135



Т.О. ШАВРИНА,  
А.А. КОРНИЛОВ\*

## Мультиязычность в языковом моделировании: задачи, данные и возможности для типологических ресурсов

Последние достижения в области машинного обучения значительно улучшили возможности больших языковых моделей (Large Language Models, LLM), в том числе способности машинного перевода и машинного чтения. Тем не менее большинство языков мира остаются не покрытыми основными ресурсами, необходимыми для построения качественных речевых технологий и языковых моделей: корпусами текстов, аннотированными датасетами, достаточным количеством записей звучащей речи. Такие языки — языки с ограниченными письменными ресурсами — называют малоресурсными.

В настоящей статье мы представляем обзор современного состояния мультиязычности и поддержки малоресурсных языков в языковых моделях, а также проводим оценку способностей текущих моделей извлекать и классифицировать информацию из зачастую единственно доступного источника знаний для малоресурсных языков — дескриптивных грамматик. Мы предлагаем подход на основе метода дополненной генерации (Retrieval-Augmented Generation, RAG), позволяющий использовать такие описания для последующих задач, таких как машинный перевод. Наши тесты охватывают грамматические описания 248 языков из 142 языковых семей, фокусируясь на типологических характеристиках баз данных WALS [1] и Grambank [2].

Предлагаемый в работе подход обеспечивает первую комплексную оценку способности языковых моделей точно интерпретировать и извлекать лингвистические признаки в контексте, создавая критически важный ресурс для масштабирования технологий на малоресурсные языки. Код и данные доступны публично: <https://github.com/al-the-eigenvalue/RAG-on-grammars>.

**Ключевые слова:** многоязычность, языковые модели, бенчмарки, малоресурсные языки

\* Шаврина Татьяна Олеговна — кандидат филологических наук, старший научный сотрудник Института языкоznания РАН.

E-mail: rybolas@gmail.com

Корнилов Альберт Андреевич — бакалавр Высшей школы экономики.

E-mail: albert.kornilov801@gmail.com

## 1. Введение

### 1.1. Языковое разнообразие и его представленность в языковых моделях

Современные языковые модели трансформировали ландшафт обработки естественного языка, продвинувшись в машинном переводе, анализе текста, генерации речи и ряде других задач. Однако масштабный успех этих технологий сконцентрирован вокруг нескольких десятков, максимум сотен языков: как правило, самых широко распространённых. По различным оценкам, в мире насчитывается около 7000 языков<sup>1</sup> и около 70 тысяч диалектов, контактных языков, пиджинов и других языковых единиц<sup>2</sup>, тогда как интернет-покрытие и параллельные корпуса для обучения моделей доступны лишь примерно для 1500. Таким образом, даже крупнейшие ресурсные и технические инициативы покрывают не более 20% от нижней оценки существующего языкового разнообразия.

Чтобы преодолеть разрыв между языками с высоким и низким объёмом ресурсов и сделать машинный перевод доступным для большего числа языков, были созданы новые бенчмарки перевода, ориентированные специально на малоресурсные языки. Конференция по машинному переводу (WMT<sup>3</sup>) теперь регулярно проводит открытые соревнования по машинному переводу для малоресурсных языков, например, для индийских<sup>4</sup> и африканских языков<sup>5</sup>; воркшопы, такие как AmericasNLP<sup>6</sup>, поддерживают коренные языки [3]. А крупные колаборации, такие как Masakhane<sup>7</sup> [4] и Aya<sup>8</sup>, создали ресурсы для машинного перево-

да африканских [5], индонезийских языков [6] – в целом для более чем для 100 языков. Недавно корпус параллельного перевода FLORES-200 расширил покрытие данных для перевода до 200 языков. Помимо этих усилий, благодаря масштабной работе по очистке данных, фильтрации и идентификации языков исследователи смогли сбрать данные и обучить модели машинного перевода для более чем 1000 языков [7].

На сегодняшний день самая многоязычная работа в области языкового моделирования – это модель для обработки звучащей речи Xeus [8]. Xeus – это открытая базовая модель для векторного представления звучащей речи, вспомогательная для многих применений и обученная почти на 1,1 млн часов неразмеченных аудиоданных на 4057 языках. Распределение данных внутри этой коллекции крайне неравномерное. Более чем у половины представленных языков объём записанной речи не превышает двух часов, что ставит их в категорию малоресурсных с точки зрения задач распознавания речи (по определению [9]). Таким образом, даже в самых масштабных многоязычных инициативах сохраняются существенные ограничения в объёмах данных для большинства языков.

Понятие «малоресурсный язык», или «low-resource language», в целом является относительным и зависит от конкретной задачи. В разных областях обработки языка минимальные пороговые значения различаются:

- для задач распознавания и обработки речи – менее двух часов аудиозаписей считается малоресурсным уровнем [9];

<sup>1</sup> 7159 — по базе данных Ethnologue (по состоянию на май 2025 г.). Электронный ресурс. URL: <https://www.ethnologue.com/insights/how-many-languages/>.

<sup>2</sup> 70 900 единиц в реестре LinguaSphere observatory (по данным переписи 2011 года). Электронный ресурс. URL: <https://web.archive.org/web/20120614005015/http://www.linguasphere.info/>.

<sup>3</sup> Электронный ресурс. URL: <https://www2.statmt.org/wmt23/>.

<sup>4</sup> Электронный ресурс. URL: <https://www2.statmt.org/wmt23/indic-mt-task.html>.

<sup>5</sup> Электронный ресурс. URL: <https://www2.statmt.org/wmt23/african-mt-task.html>.

<sup>6</sup> Электронный ресурс. URL: [https://turing.iimas.unam.mx/americasnlp/2023\\_st.html](https://turing.iimas.unam.mx/americasnlp/2023_st.html).

<sup>7</sup> Электронный ресурс. URL: <https://www.masakhane.io/>.

<sup>8</sup> Электронный ресурс. URL: <https://cohere.com/research/aya>.

- для прикладных задач, таких как создание вопросно-ответных систем, тематической классификации документов, извлечения именованных сущностей — менее 10 000 размеченных примеров [9];
- для обучения языковых моделей — менее 350 000 токенов текста.

Таким образом, малоресурсность — это не фиксированная характеристика языка, а динамическое понятие, которое определяется в зависимости от требований конкретных задач. Для задач обработки речи необходимы аудиозаписи, для обучения языковых моделей — большие текстовые корпуса, а для решения прикладных задач (например, машинного перевода или классификации) — размеченные данные. Один и тот же язык может считаться малоресурсным для одной задачи и вполне обеспеченным ресурсами — для другой. Поэтому при разработке мультиязычных моделей важно учитывать не только наличие данных, но и их вид, объём и качество в зависимости от цели использования.

Работа с привнесением многих языков в более высокоресурсную среду усложняется рядом фундаментальных проблем.

1. Дефицит данных: для успешной работы LLM требуются большие массивы данных. Однако даже минимальные ориентиры — два часа речи (для задач распознавания речи), 10 тысяч аннотированных примеров (для задач вроде классификации), 350 тысяч токенов текста (для языкового моделирования) — практически недостижимы для подавляющего большинства языков мира.
2. Отсутствие параллельных и монолингвальных корпусов: для малоресурсных языков крайне редко встречаются параллельные корпуса, которые лежат в основе современных моделей машинного перевода. Также отсутствуют большие монолингвальные коллекции, что делает обучение практически невозможным.
3. Проблемы валидации данных: даже если данные удаётся собрать, их качество и точность остаются под вопросом без

участия носителей языка и профессиональных лингвистов.

В случаях, когда ресурсов практически нет (отсутствуют аудиозаписи, текстовые корпуса и размеченные примеры), единственным источником информации о языке остаются лингвистические ресурсы: дескриптивные грамматики и типологические базы данных. Такие материалы содержат систематизированное описание структуры языка (фонологии, морфологии, синтаксиса) и могут быть использованы как основа для создания моделей в условиях крайней малоресурсности.

Современные подходы, например Machine Translation from One Book (МТОВ, [10]), показывают, что даже такие «ручные» источники можно интегрировать в работу с помощью методов, не требующих обучения языковой модели: дополненной генерации и длинного контекста.

Однако стоит отметить, что применение описательных грамматик для машинного перевода сталкивается с рядом проблем, таких как вариативность терминологии, нестандартные структуры и рассеянность релевантной информации. Кроме того, массивный тест для оценки текстовых вложений (МТЕВ, [11]) предоставляет детальную оценку текстовых вложений по различным задачам и языкам. Основная проблема остаётся в эффективном применении этих моделей к описательным грамматикам для малоресурсных языков, обычно поддерживаемых лишь лингвистическими материалами, такими как грамматики и словари.

Данная работа направлена на решение этих проблем, предлагая систематический подход к извлечению информации из описательных грамматик и создание масштабируемого метода для систематизации грамматических описаний. Ключевым аспектом предлагаемого подхода является метод дополненной генерации (RAG), позволяющий извлекать релевантную информацию из грамматик на основе конкретной типологической характеристики (например, порядок подлежащего, дополнения и сказуемого). На основе извлечён-

ных параграфов крупная языковая модель определяет значение этой характеристики (например, порядок слов «подлежащее—сказуемое—дополнение»).

В настоящей статье представлены следующие результаты:

1. Обзор состояния языковых моделей для применения к языковому разнообразию.
2. Первая масштабная лингвистическая оценка способностей машинного чтения крупных языковых моделей на материале описательных грамматик.
3. Методология на основе дополненной генерации (RAG), извлекающей релевантные параграфы из грамматик на основе заданной типологической характеристики (например, WALS 81A: порядок подлежащего, дополнения и сказуемого) и предоставляющей их в виде подсказок для крупной языковой модели с целью определения значений этих характеристик.

Предлагаемая архитектура, а также описанные тестовые наборы направлены на содействие дальнейшему развитию систем машинного перевода и повышение их качества и эффективности, а также на помочь лингвистам в типологических исследованиях путём частичной автоматизации извлечения данных из описательных грамматик.

Весь код работы открыт, распространяется под лицензией MIT: <https://github.com/al-the-eigenvalue/RAG-on-grammars>.

## 2. Обзор литературы

### 2.1. Метод дополненной генерации (RAG)

Метод дополненной генерации (retrieval-augmented generation, RAG) стал эффективным инструментом для повышения качества генерации в больших языковых моделях, дополняя их внутренние знания за счёт извлечения внешней информации. Вместо того чтобы полагаться исключительно на знания, встроенные в модель, RAG позволяет извлекать релевантные фрагменты из внешних источников в процессе генерации [12]. Этот подход доказал

свою результативность в задачах открытого вопросно-ответного взаимодействия и суммирования документов, повышая достоверность информации и расширяя контекст.

Особенно перспективен метод RAG при работе с малоресурсными языками и сложными форматами, такими как лингвистические грамматики. Недавние исследования в области извлечения информации на основе запросов свидетельствуют о его применимости в условиях фрагментированных и неполных языковых данных. Настоящая работа применяет принципы RAG для повышения качества перевода и языкового моделирования с помощью описательных грамматик для языков с ограниченными ресурсами.

### 2.2. Обучение без примеров (Zero-shot Learning)

Zero-shot learning стало ключевым направлением в области обработки естественного языка, особенно с развитием фундаментальных моделей. Способность моделей к обобщению на новые задачи и языки без специфических обучающих данных критически важна для расширения применения NLP в отношении малоресурсных языков. Модели, такие как GPT-3 и GPT-4, демонстрируют впечатляющие возможности zero-shot для решения задач классификации, машинного перевода и др., что делает их незаменимыми инструментами в условиях ограниченного количества размеченных данных.

Тем не менее существующие zero-shot модели всё ещё испытывают трудности с языками, практически не представленными ни в монолингвальных, ни в билингвальных корпусах. Недавние исследования [10, 13] показали, что интеграция лингвистических описаний — таких как грамматики — может значительно улучшить качество zero-shot решений для таких языков. В этой работе мы продолжаем это направление, оценивая способность моделей использовать описательные грамматики в условиях отсутствия обучающих данных.

### 2.3. Использование грамматик в NLP

Несмотря на широкий спектр источников данных, используемых в NLP, таких как параллельные корпуса и размеченные датасеты, описательные грамматики остаются недостаточно используемым ресурсом. Между тем они содержат богатую структурированную информацию, которая особенно полезна для малоресурсных языков. Проекты Grambank и WALS показали, что типологические признаки можно систематизировать для лингвистического анализа и сопоставления языков.

Использование грамматик как источника знаний для языковых моделей всё чаще рассматривается как потенциально перспективное направление. В работах [10 и 14] подчёркивается, что лингвистическая информация из грамматик помогает строить обобщаемые представления о языках и их структурах, особенно в контексте задач машинного перевода и морфосинтаксического анализа. Тем не менее формализация, цифровизация и стандартизация таких ресурсов остаются открытыми вызовами. В данной статье мы частично преодолеваем этот барьер, интегрируя описательные грамматики в архитектуру RAG и предоставляя соответствующие тесты и инструменты.

### 2.4. Извлечение типологических признаков из грамматик

Существующие исследования по извлечению типологических признаков из грамматик предшествуют появлению крупных языковых моделей и основываются на правилах, классических методах машинного обучения и ранних версиях нейросетевых подходов. Серия работ [15] использует методы, требующие трудоёмкой аннотации семантических рамок. Работа [16] предлагает подход, применимый исключительно к бинарным типологическим признакам, а система, описанная в [17], ограничивается извлечением информации по ключевым словам.

В настоящей работе мы впервые применяем современные большие языковые модели для задачи извлечения типологических признаков и демонстрируем их потенциал в работе с лингвистическими описаниями.

### 2.5. Машинный перевод из одной грамматики

Ряд достижений в языковом моделировании открывает возможность использования лингвистических описаний: методы дополненной генерации для извлечения и генерации информации, способность моделей работать дескриптивными текстами в качестве затравок (prompt), наличие таких текстов в машиночитаемом формате.

Работа [10] представляет пример прикладного использования описательной грамматики малоресурсного языка для улучшения качества перевода на языки с чрезвычайно ограниченными ресурсами, показывая потенциал масштабных языковых моделей в преодолении разрыва между теоретической лингвистикой и практическими приложениями NLP.

Представленный в работе метод – Machine Translation from One Book – это новый подход к оценке способности языковых моделей выполнять перевод в условиях, когда язык целиком отсутствует в данных предварительного обучения. В отличие от традиционных бенчмарков машинного перевода, где модели обучаются на больших параллельных корпусах, МТОВ предлагает моделям опереться только на единственный источник – дескриптивную грамматику и словарь языка Каламанг<sup>1</sup>, представленные в оцифрованном виде. Такой сценарий моделирует реальные условия работы с экстремально малоресурсными и вымирающими языками, где отсутствуют как параллельные, так и монолингвальные данные.

Особенность подхода состоит в том, что модели получают в качестве контекста грамматическое описание языка и несколько примеров перевода, и на их основе

<sup>1</sup> Электронный ресурс [https://en.wikipedia.org/wiki/Kalamang\\_language](https://en.wikipedia.org/wiki/Kalamang_language).

<b>Input</b>	<b>Kor kancing wa me an tur teba ma patin.</b>
Reference	My ankle bone, I fell and wounded it.
text-davinci-003	Even if I fall, my ankle bone is wounded.
gpt-3.5-turbo	I dream of a watch falling soon and injuring.
gpt-4	This ankle of mine is falling and progressively getting wounded.
Claude 2	This ankle of mine fell down while I was walking, and it got wounded.
Human	This ankle, I injured by falling.

**Ил. 1.** Качественный пример перевода с каламанга (kgv) на английский (Tanzer et al., 2024).

Представлены ответы моделей в конфигурации с максимальным контекстом для каждой модели.

Модель **text-davinci-003** правильно переводит лексику, но добавляет галлюцинированную конструкцию «even if». **GPT-3.5-turbo** отвлекается на нерелевантную лексику. **GPT-4** демонстрирует интересную ошибку: она интерпретирует значение частицы *teba* (маркер прогрессивного аспекта) как «progressively». **Claude 2** формулирует перевод странно, но почти правильно; однако в исходной фразе нет упоминания о хождении пешком, которое появляется на выходе модели. Человеческий перевод является точным, хотя и использует немного неестественную конструкцию для передачи топикализации и прогрессивного аспекта, присутствующих в оригинале

<b>Input</b>	<b>I'm getting pandanus, I want to make a mat.</b>
Reference	An padamualat repte kalifan paruotkin.
text-davinci-003	Kawat sie padamual, suka kangjie temun irar.
gpt-3.5-turbo	An kaloum bunga rampi, an suka rarie el.
gpt-4	An padamual gousat, suka an irar minggi.
Claude 2	An padamual rep=kin minggi kalifan paruo=kin
Human	An padamualat rep teba, elat paruotkin.

**Ил. 2.** Качественный пример перевода с английского на каламанг (eng→kgv) (Tanzer et al., 2024).

Представлены ответы моделей в конфигурации с максимальным контекстом для каждой модели.

**text-davinci-003** включает как релевантные, так и нерелевантные извлечённые слова с бессмысленной грамматикой. **GPT-3.5-turbo** и **GPT-4** начинают использовать местоимение «я» (*I*) и подбирают более последовательно релевантную лексику, однако перевод остаётся неграмматичным. **Claude 2** оперирует на уровне глосс, используя форму *=kin*, но упускает фонологические чередования, такие как *paruo=kin* → *paruotkin*; тем не менее содержание перевода в целом передано правильно: буквально «Я хочу добыть панданус и сделать из него циновку». Человеческий перевод использует более буквально переданную грамматику по сравнению с эталоном и применяет термин *el* (грубая циновка), а не *kalifan* (тонкая циновка); нам неизвестно, используется ли панданус действительно для обоих видов циновок

переводят новые предложения. Качество оценивается строго: язык должен отсутствовать в предобучении, а перевод основан только на грамматике и примерах. Для Каламанг используются латинские буквы, что облегчает задачу моделям. Задания — предложения для перевода в направлении «Каламанг — английский» или «английский — каламанг». Для оценки используются метрики машинного перевода и экспертная валидация.

### 3. Масштабирование на большее число языков

Несмотря на впечатляющие результаты [10] в задачах оценки пригодности одних лишь грамматик и словаря для работы с малоресурсным языком, остаётся открытым вопрос о том, насколько данный подход масштабируем и применим к другим языкам помимо каламанга, языкам с разными типологическими характеристиками и качеством грамматических описаний.

Далее рассмотрим эксперименты по расширению методологии МТОВ для более широкой выборки языков [18]: мы предлагаем систематический подход к выбору типологически репрезентативной выборки языков, сбору и структурированию их грамматик, а также разработке автоматизированного метода оценки качества понимания таких грамматик языковыми моделями.

Как оценить способность языковых моделей масштабироваться на широкий спектр языков, особенно в тех случаях, когда имеются доступные грамматические описания, но отсутствует возможность привлечения носителей языка? Одним из возможных решений является использование не задач машинного перевода, а формата вопросно-ответных систем, где ответы модели можно валидировать на основе уже известных типологических характеристик языка. Такой подход позволяет не только оценивать корректность извлечения лингвистической информации, но и дополнительно проверять способ-

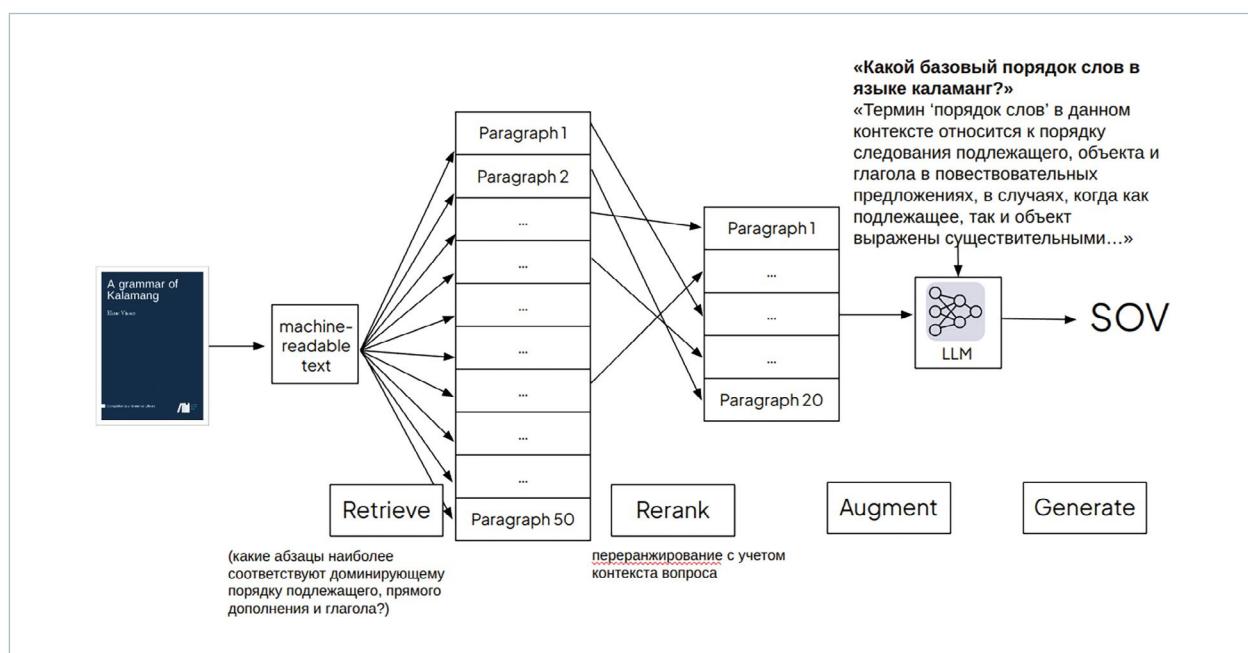
ность языковых моделей интерпретировать научный метаязык, используемый в дескриптивных грамматиках.

### 3.1. Базовая методология

Базовая методология строится как расширение подхода Retrieval-Augmented Generation (RAG), в котором модель сначала извлекает релевантные параграфы из грамматического описания языка с помощью классических методов информационного поиска (BM25), а затем использует языковую модель для генерации ответа на лингвистический вопрос.

Базовый метод дополненной генерации (Naive RAG) включает в себя базу документов (в нашем случае грамматики), компонент извлечения релевантных абзацев, процесс извлечения релевантных абзацев из базы документов на основе вопроса от пользователя, а также языковую модель, генерирующую ответ на основе этого вопроса и извлечённой информации.

Второй компонент — метод извлечения информации. Мы оцениваем BM25<sup>1</sup>,



Ил. 3. Схема работы языковой модели с RAG на основе дескриптивной грамматики для вопросно-ответной системы

<sup>1</sup> Электронный ресурс: [https://ru.wikipedia.org/wiki/Okapi\\_BM25](https://ru.wikipedia.org/wiki/Okapi_BM25).

**Таблица 1**  
**Распределение языков по макрорегионам**

Макроареал	Всего языков
Африка	29
Австралия	9
Евразия	20
Северная Америка	25
Папуа Новая Гвинея	39
Южная Америка	26
<b>Сумма</b>	<b>148</b>

не зависящий от языка метод на основе частотности терминов, а также современные методы на основе эмбеддингов, представленные в рейтинге Massive Text Embedding Benchmark (МТЕВ) [11]. Такие методы устойчивы к лингвистическому разнообразию, поскольку описательные грамматики содержат примеры на исследуемых языках, включая диакритику и сегменты, редкие или отсутствующие в английском.

Третий компонент – формат затравки (prompt). Базовый шаблон включает абзац, вопрос о типологической характеристики, пояснение терминов и фиксированный набор ответов (см.: Приложение). Например, для признака WALS 81A «Преобладающий порядок: подлежащее, дополнение, сказуемое» возможные ответы: «SVO», «SOV», «VOS», «VSO», «OSV», «OVS», «Нет доминирующего порядка», а также «Недостаточно информации», если порядок не удаётся определить. Мы также реализуем стратегию затравок с расширенным описанием признака (из WALS или Grambank) и примерами – это вариант цепочки рассуждений (chain-of-thought prompting) [21].

Последний компонент – сама языковая модель. Мы используем GPT-4o, флагманскую модель OpenAI на май 2024 г. с улучшенными характеристиками по сравнению с GPT-4. Её задача – определить значение признака, например «4 падежа» для WALS 49A «Число падежей», на основе подсказки и параграфов из грамматики.

### 3.2. Данные

Бенчмарк для оценки метода RAG включает 148 описательных грамматик<sup>1</sup>, поскольку бенчмарки с менее чем 100 примерами считаются ненадёжными: одна ошибка снижает точность более чем на один процент.

Случайный выбор грамматик мог бы исказить репрезентативность, создавая перекос в сторону языков из одних и тех же семей или географических регионов. Поэтому мы использовали метод Genus-Macroarea, описанный в [19] и реализованный в [20]. Как и в [20], мы берём распределение языков по макроареалам из списка родов WALS, автоматически выбираем грамматики из базы Glottolog References [22] и ограничиваем выбор одной грамматикой на род (genus). В отличие от Chevleva, мы ограничили выбор грамматиками, написанными на английском языке.

Для оценки качества извлечения информации были выбраны четыре признака:

**WALS 81A.** Порядок подлежащего, дополнения и сказуемого, поскольку он зачастую явно указывается в грамматиках;

**GB 107.** Может ли стандартная негативная форма быть выражена аффиксом, клитикой или изменением глагола? Это бинарный признак, но трудный для наивного извлечения из-за вариативности терминологии. Составной признак, связанный с общевопросными предложениями (полярными вопросами);

<sup>1</sup> [https://anonymous.4open.science/r/from-MTEB-to-MTOB/ground\\_truth\\_rag.csv](https://anonymous.4open.science/r/from-MTEB-to-MTOB/ground_truth_rag.csv).

**Таблица 2**  
**F1-оценки для всех конфигураций архитектуры**

	<b>F1 мера</b>	<b>Baseline</b>	<b>BM25</b>	<b>BM25+CoT</b>
Все признаки	micro	$0,5551 \pm 0,0359$	0,6892	<b>0,7027</b>
Все признаки	macro	$0,2812 \pm 0,0276$	<b>0,7179</b>	0,6097
<b>WALS 81A</b>	weighted	$0,5328 \pm 0,0337$	0,6694	<b>0,6890</b>
<b>GB 107</b>	weighted	$0,5724 \pm 0,0361$	0,5957	<b>0,5959</b>
<b>WALS 49A</b>	weighted	$0,3453 \pm 0,0237$	0,5314	<b>0,5542</b>
<b>interrog.intonation only</b>	weighted	$0,4068 \pm 0,0484$	0,8480	<b>0,9007</b>
<b>interrog.word order</b>	weighted	$0,9611 \pm 0,0064$	<b>0,9936</b>	0,9878
<b>clause-initial particle</b>	weighted	$0,7371 \pm 0,0337$	0,9054	<b>0,9205</b>
<b>clause-final particle</b>	weighted	$0,4661 \pm 0,0428$	0,7314	<b>0,7644</b>
<b>clause-medial particle</b>	weighted	$0,6644 \pm 0,0355$	0,8439	<b>0,8888</b>
<b>interrog. verb morphology</b>	weighted	$0,7102 \pm 0,0160$	0,8192	<b>0,8451</b>
<b>tone</b>	weighted	$0,9104 \pm 0,0121$	0,9390	<b>0,9637</b>

**Примечание:** Все пять конфигураций используют промпты из Приложения с включением Wikipedia-резюме для соответствующих признаков. Колонка **Baseline** соответствует запросам к GPT-4o без использования материалов из грамматик (т.е. без применения RAG). **BM25** означает использование 50 параграфов, извлечённых методом BM25. **CoT** обозначает добавление цепочки рассуждений (Chain-of-Thought), т.е. инструкций и примеров из WALS или Grambank в промпт

**WALS 116A.** Он объединяет семь бинарных признаков из Grambank и оценивает способность моделей к рассуждению на основе нескольких реализаций одного феномена;

**WALS 49A.** Число падежей – количественный признак, требующий широкого охвата грамматики (разделы морфологии и синтаксиса).

Все признаки обладают следующими характеристиками:

- Интерпретируемость – признаки должны быть формулируемы в виде понятного лингвистического вопроса.
- Наличие аннотации в WALS или Grambank – для обеспечения возможности автоматической валидации.
- Независимость от конкретных языков – признаки должны быть универсальными и применимыми к разным языкам.
- Наличие явного описания в грамматике – чтобы их можно было извлечь с помощью RAG-архитектуры.

### 3.3. Результаты: оценка метода

Для оценки любой NLP-задачи по бенчмарку важно определить, обладает ли языковая модель знаниями о значениях признаков, включённых в процедуру оценки.

Чтобы установить базовый уровень (baseline) для GPT-4o, т.е. оценить его работу без привлечения грамматики, мы провели тест, исключив модуль извлечения из архитектуры RAG. Мы предложили GPT-4o определить значения всех признаков без использования извлечённых параграфов – модель получала только подсказку и краткое содержание статьи из Википедии по соответствующему признаку. Для каждого признака тест запускался десять раз, чтобы отразить разброс результатов и точнее различить случаи наличия и отсутствия знаний.

Результаты всех конфигураций метода RAG представлены в табл. 2. Все конфигурации RAG превосходят базовое решение (генерацию ответа без грамматики). Более высокие макроусреднённые значения F1-меры<sup>1</sup>, по сравнению с микроусреднён-

<sup>1</sup> <https://en.wikipedia.org/wiki/F-score>.

ными, указывают на то, что метод RAG лучше справляется с частотными классами и испытывает сложности при наличии дисбаланса классов в типологических профиле языков мира.

Следует отметить, что далеко не все типологические признаки были успешно извлечены с высокой точностью. В частности, такие широко распространённые характеристики, как базовый порядок слов и выражение глагольного отрицания, несмотря на их частое упоминание в грамматических описаниях, демонстрируют низкое качество извлечения языковой моделью. Проведённый тест на контаминацию (Baseline) показал, что в ряде случаев модели уже обладают встроенными ассоциациями между языками и значениями признаков, что позволяет им правильно отвечать даже без обращения к тексту грамматики. Этот эффект может искусственно завышать показатели качества извлечения и затруднять объективную оценку работы модели по менее представленным и редким признакам.

Использование цепочек рассуждений (Chain-of-Thought prompts) и предоставление инструкций, основанных на примерах из WALS и Grambank, существенно улучшают интерпретацию модели, особенно для сложных грамматических конструкций. В то же время обработка нескольких признаков одновременно показала непредсказуемые результаты, подчёркивая сложность автоматического извлечения типологических данных из грамматик.

Расширение подхода МТОВ может существенно выиграть от приведения описательных грамматик различных языков к унифицированному формату, используя базы данных, такие как Grambank или WALS.

Тем не менее даже в строго контролируемой среде, продемонстрированной в данной работе, дескриптивные лингвистические тексты остаются значительным вызовом. Несмотря на то, что машинное чтение в целом может рассматриваться как «решённая задача», результаты по извлечению лингвистических признаков показывают, что описательные грамматики остаются

нетривиальным источником, выявляющим слабые стороны языковых моделей.

## Заключение

Расширение подхода МТОВ может существенно выиграть от стандартизации описательных грамматик различных языков в единый формат с опорой на базы данных, такие как Grambank или WALS. Тем не менее даже в неконтаминированной среде, рассмотренной в данной работе, описательные лингвистические тексты продолжают представлять значительную сложность.

В данной статье мы представили метод оценки решений, объединяющих дополненную генерацию (RAG) и большие языковые модели, с целью извлечения и классификации типологических признаков из описательных грамматик. Также мы представили метод для извлечения лингвистической информации, обладающий значительным потенциалом для улучшения NLP-систем на малоресурсных языках.

Кроме того, несмотря на распространённое мнение о решённости задачи машинного чтения, лингвистические работы остаются областью повышенной сложности. Хотя языковые модели значительно продвинулись в обработке различных типов текстов, полученные нами результаты указывают на то, что пока рано говорить об их полной эффективности в домене описательной лингвистики.

Наши результаты закладывают основу для расширения возможностей языковых моделей в работе со сложными лингвистическими данными, такими как грамматические описания. Эта работа представляет собой важный шаг в направлении поддержки малоресурсных языков в NLP.

В дальнейшем возможны улучшение компонентов извлечения и классификации, расширение бенчмарка за счёт включения большего числа языков, а также исследование практических применений извлечения лингвистической информации, например, кросс-языковая типологическая аналитика или машинный перевод для крайне малоресурсных языков.

## ЛИТЕРАТУРА

1. Dryer M.S., Haspelmath M. (eds.). WALS Online (v2020.4) [Data set]. Zenodo, 2013. DOI: 10.5281/zenodo.13950591.
2. Skirgård H., Haynie H., Passmore S. et al. Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss // Science Advances. 2023. Vol. 9, № 16. Article eadg6175. DOI: 10.1126/sciadv.adg6175.
3. Ebrahimi A. et al. Findings of the AmericasNLP 2023 Shared Task on Machine Translation into Indigenous Languages // Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP). Toronto, Canada: Association for Computational Linguistics, 2023. Pp. 206–219.
4. Lovenia H. et al. SEACrowd: A Multilingual Multimodal Data Hub and Benchmark Suite for Southeast Asian Languages // Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP). Miami, USA: Association for Computational Linguistics, 2024. Pp. 5155–5203.
5. Nekoto W. et al. Participatory Research for Low-Resourced Machine Translation: A Case Study in African Languages // Findings of the Association for Computational Linguistics: EMNLP 2020. Online: Association for Computational Linguistics, 2020. Pp. 2144–2160.
6. Winata G.I. et al. NusaX: Multilingual Parallel Sentiment Dataset for 10 Indonesian Local Languages // Proceedings of the 17<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics (EACL). Dubrovnik, Croatia: Association for Computational Linguistics, 2023. Pp. 815–834.
7. Bapna A. et al. Building machine translation systems for the next thousand languages. arXiv preprint arXiv:2205.03983, 2022.
8. Chen W. et al. Towards Robust Speech Representation Learning for Thousands of Languages // Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP). Miami, USA: Association for Computational Linguistics, 2024. Pp. 10205–10224.
9. Garrette D., Mielens J., Baldridge J. Real-World Semi-Supervised Learning of POS-Taggers for Low-Resource Languages // Proceedings of the 51<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL 2013). Vol. 1: Long Papers. Sofia, Bulgaria: Association for Computational Linguistics, 2013. Pp. 583–592.
10. Tanzer G., Suzgun M., Visser E., Jurafsky D., Melas-Kyriazi L. A benchmark for learning to translate a new language from one grammar book. arXiv preprint arXiv:2309.16575, 2023.
11. Muennighoff N., Tazi N., Magne L., Reimers N. MTEB: Massive Text Embedding Benchmark // Proceedings of the 17<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics (EACL). Dubrovnik, Croatia: Association for Computational Linguistics, 2023. Pp. 2014–2037.
12. Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Kiela D. Retrieval-augmented generation for knowledge-intensive NLP tasks // Advances in Neural Information Processing Systems (NeurIPS). 2020. Vol. 33. Pp. 9459–9474.
13. Zhang K., Choi Y., Song Z., He T., Wang W.Y., Li L. Hire a Linguist!: Learning Endangered Languages in LLMs with In-Context Linguistic Descriptions // Findings of the Association for Computational Linguistics: ACL 2024. Bangkok, Thailand: Association for Computational Linguistics, 2024. Pp. 15654–15669.
14. Ponti E.M., Glavaš G., Majewska O., Liu Q., Vulić I., Korhonen A. XCOPA: A Multilingual Dataset for Causal Commonsense Reasoning // Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Online: Association for Computational Linguistics, 2020. Pp. 2362–2376.
15. Virk S.M., Foster D., Sheikh M.A., Saleem R. A Deep Learning System for Automatic Extraction of Typological Linguistic Information from Descriptive Grammars // Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021). Online: INCOMA Ltd., 2021. Pp. 1480–1489.

16. Hammarström H., Her O.-S., Allassonnière-Tang M. Term spotting: A quick-and-dirty method for extracting typological features of language from grammatical descriptions // Selected Contributions from the Eighth Swedish Language Technology Conference (SLTC-2020). 2020. Pp. 27–34.
17. Kornilov A. Multilingual Automatic Extraction of Linguistic Data from Grammars // Proceedings of the Second Workshop on NLP Applications to Field Linguistics. Dubrovnik, Croatia: Association for Computational Linguistics, 2023. Pp. 86–94.
18. Kornilov A., Shavrina T. From MTEB to MTOB: Retrieval-Augmented Classification for Descriptive Grammars. arXiv preprint arXiv:2411.15577, 2024.
19. Miestamo M., Bakker D., Arppe A. Sampling for variety // Linguistic Typology. 2016. Vol. 20, № 2. Pp. 233–296.
20. Cheveleva A. Neutralization of gender values in the plural. Bachelor's thesis. Moscow: HSE University, 2023.
21. Wei J., Wang X., Schuurmans D., Bosma M., Xia F., Chi E., Le Q.V., Zhou D. et al. Chain-of-thought prompting elicits reasoning in large language models // *Advances in Neural Information Processing Systems (NeurIPS)*. 2022. Vol. 35. Pp. 24824–24837.
22. Hammarström H., Forkel R., Haspelmath M., Bank S. (eds.). *Glottolog 5.0* [Data set]. Zenodo, 2024. DOI: 10.5281/zenodo.8635585.

#### Приложение

##### Затравка для определения числа падежей (на основе WALS 49A);

*Please determine the number of cases in the language <...>. The term “cases” in the context of this feature refers to productive case paradigms of nouns. Reply with one of the 9 following options: No morphological case-marking, 2 cases, 3 cases, 4 cases, 5 cases, 6-7 cases, 8-9 cases, 10 or more cases, Exclusively borderline case-marking. The feature value “Exclusively borderline case-marking” refers to languages which have overt marking only for concrete (or “peripheral”, or “semantic”) case relations, such as locatives or instrumentals. Categories with pragmatic (non-syntactic) functions, such as vocatives or topic markers, are not counted as case even if they are morphologically integrated into case paradigms. Genitives are counted as long as they do not encode categories of the possesum like number or gender as well, if they do not show explicit adjective-like properties. Genitives that may take additional case affixes agreeing with the head noun case (“double case”) are not regarded as adjectival. 1. Provide the reasoning for the chosen option. 2. After the reasoning, output the word “Conclusion:” and the chosen option at the end of your response.*

## Multilinguality in Language Modeling: Tasks, Data, and Opportunities for Typological Resources

**Tatiana Olegovna Shavrina** – Ph.D. in Philology, Senior Researcher, Institute of Linguistics, Russian Academy of Sciences.  
E-mail: rybolos@gmail.com

**Albert Andreevich Kornilov** – Bachelor's Degree, Higher School of Economics.  
E-mail: albert.kornilov801@gmail.com

This paper addresses the significant challenge of building language technologies for the majority of the world's under-resourced languages, which lack the large text corpora and annotated datasets necessary for modern machine learning. While advances in Large Language Models

(LLMs) have revolutionized machine translation and reading comprehension, these models often underperform or fail entirely for languages with limited written resources.

We present an overview of current multilingual support in LLMs and evaluate their ability to understand the primary available knowledge source for such languages: descriptive grammars. To effectively utilize this structured but complex information, we propose a Retrieval-Augmented Generation (RAG) framework. This approach enables models to accurately extract and interpret linguistic features from grammatical texts, facilitating downstream tasks like machine translation. Our evaluation provides the first comprehensive assessment of model performance on this critical task, covering grammatical descriptions of 248 languages from 142 language families. The analysis focuses on the typological characteristics of the WALS [1] and Grambank [2] databases.

The proposed approach demonstrates the first comprehensive assessment of the ability of language models to accurately interpret and extract linguistic features in context, creating a critical resource for scaling technologies to under-resourced languages. Code and data from this study are made publicly available: <https://github.com/al-the-eigenvalue/RAG-on-grammars>.

**Keywords:** multilingualism, language models, benchmarks, under-resourced

## REFERENCES

1. Dryer M.S., Haspelmath M. (eds.). WALS Online (v2020.4) [Data set]. Zenodo, 2013. DOI: 10.5281/zenodo.13950591.
2. Skirgård H., Haynie H., Passmore S. et al. Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss // Science Advances. 2023. Vol. 9, № 16. Article eadg6175. DOI: 10.1126/sciadv.adg6175.
3. Ebrahimi A. et al. Findings of the AmericasNLP 2023 Shared Task on Machine Translation into Indigenous Languages // Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP). Toronto, Canada: Association for Computational Linguistics, 2023. Pp. 206–219.
4. Lovenia H. et al. SEACrowd: A Multilingual Multimodal Data Hub and Benchmark Suite for Southeast Asian Languages // Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP). Miami, USA: Association for Computational Linguistics, 2024. Pp. 5155–5203.
5. Nekoto W. et al. Participatory Research for Low-Resourced Machine Translation: A Case Study in African Languages // Findings of the Association for Computational Linguistics: EMNLP 2020. Online: Association for Computational Linguistics, 2020. Pp. 2144–2160.
6. Winata G.I. et al. NusaX: Multilingual Parallel Sentiment Dataset for 10 Indonesian Local Languages // Proceedings of the 17<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics (EACL). Dubrovnik, Croatia: Association for Computational Linguistics, 2023. Pp. 815–834.
7. Bapna A. et al. Building machine translation systems for the next thousand languages. arXiv preprint arXiv:2205.03983, 2022.
8. Chen W. et al. Towards Robust Speech Representation Learning for Thousands of Languages // Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP). Miami, USA: Association for Computational Linguistics, 2024. Pp. 10205–10224.
9. Garrette D., Mielens J., Baldridge J. Real-World Semi-Supervised Learning of POS-Taggers for Low-Resource Languages // Proceedings of the 51<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL 2013). Vol. 1: Long Papers. Sofia, Bulgaria: Association for Computational Linguistics, 2013. Pp. 583–592.

10. Tanzer G., Suzgun M., Visser E., Jurafsky D., Melas-Kyriazi L. A benchmark for learning to translate a new language from one grammar book. arXiv preprint arXiv:2309.16575, 2023.
11. Muennighoff N., Tazi N., Magne L., Reimers N. MTEB: Massive Text Embedding Benchmark // Proceedings of the 17<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics (EACL). Dubrovnik, Croatia: Association for Computational Linguistics, 2023. Pp. 2014–2037.
12. Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Kiela D. Retrieval-augmented generation for knowledge-intensive NLP tasks // Advances in Neural Information Processing Systems (NeurIPS). 2020. Vol. 33. Pp. 9459–9474.
13. Zhang K., Choi Y., Song Z., He T., Wang W.Y., Li L. Hire a Linguist!: Learning Endangered Languages in LLMs with In-Context Linguistic Descriptions // Findings of the Association for Computational Linguistics: ACL 2024. Bangkok, Thailand: Association for Computational Linguistics, 2024. Pp. 15654–15669.
14. Ponti E.M., Glavaš G., Majewska O., Liu Q., Vulić I., Korhonen A. XCOPA: A Multilingual Dataset for Causal Commonsense Reasoning // Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Online: Association for Computational Linguistics, 2020. Pp. 2362–2376.
15. Virk S.M., Foster D., Sheikh M.A., Saleem R. A Deep Learning System for Automatic Extraction of Typological Linguistic Information from Descriptive Grammars // Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021). Online: INCOMA Ltd., 2021. Pp. 1480–1489.
16. Hammarström H., Her O.-S., Allassonnière-Tang M. Term spotting: A quick-and-dirty method for extracting typological features of language from grammatical descriptions // Selected Contributions from the Eighth Swedish Language Technology Conference (SLTC-2020). 2020. Pp. 27–34.
17. Kornilov A. Multilingual Automatic Extraction of Linguistic Data from Grammars // Proceedings of the Second Workshop on NLP Applications to Field Linguistics. Dubrovnik, Croatia: Association for Computational Linguistics, 2023. Pp. 86–94.
18. Kornilov A., Shavrina T. From MTEB to MTOB: Retrieval-Augmented Classification for Descriptive Grammars. arXiv preprint arXiv:2411.15577, 2024.
19. Miestamo M., Bakker D., Arppe A. Sampling for variety // Linguistic Typology. 2016. Vol. 20. № 2. Pp. 233–296.
20. Cheveleva A. Neutralization of gender values in the plural. Bachelor's thesis. Moscow: HSE University, 2023.
21. Wei J., Wang X., Schuurmans D., Bosma M., Xia F., Chi E., Le Q.V., Zhou D. et al. Chain-of-thought prompting elicits reasoning in large language models // Advances in Neural Information Processing Systems (NeurIPS). 2022. Vol. 35. Pp. 24824–24837.
22. Hammarström H., Forkel R., Haspelmath M., Bank S. (eds.). Glottolog 5.0 [Data set]. Zenodo, 2024. DOI: 10.5281/zenodo.8635585.

## ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 1

Номер посвящён 80-летию Победы в Великой Отечественной войне 1941–1945 гг.



*Тематический редактор номера –  
Ю.А. Петров – доктор исторических наук,  
директор Института российской истории РАН*

В журнале опубликованы подготовленные российскими историками статьи, многие из которых основаны на архивных документах, впервые вводимых в научный оборот, что позволяет раскрыть новые страницы военных лет и глубже осознать величие подвига советского народа. Все статьи номера систематизированы по четырём рубрикам: «В штабах Победы», «Без срока давности», «Архивные тайны» и «Сохранение памятников культуры и развитие спорта в годы войны».

Первую рубрику открывает статья видных специалистов по истории Великой Отечественной войны В.Н. Мякоты и С.В. Кудряшова «Советские планы ответа на военные приготовления Германии». Основное внимание авторы обращают на планы Генерального штаба РККА весны 1941 г. Малоизученной теме – специальной пропаганде, которая осуществлялась различными подразделениями Красной армии в годы Великой Отечественной войны с целью влияния на войска противника, посвящена статья О.Д. Панкова. Вопросы создания и деятельности Центрального штаба партизанского движения, его взаимодействия с Красной армией и органами государственной безопасности в целях повышения уровня руководства и эффективности разведывательно-диверсионной и боевой деятельности партизанских формирований на оккупированной советской территории в годы Великой Отечественной войны проанализировал А.Ю. Попов.

Под рубрикой «Без срока давности» публикуются статьи С.В. Аристова о судьбе советских военнопленных в годы войны, С.В. Кудряшова – о нацистском терроре на советских территориях и С.М. Соловьёва – об отношениях между союзниками по антигитлеровской коалиции в ходе подготовки и проведения Нюрнбергского процесса. Большой массив не известных ранее документов вводится в научный оборот в статьях М.М. Беклемишевой «Трофейные документы Великой Отечественной войны в фокусе внимания разведывательных, политических и архивных органов СССР» и А.И. Домнина «Картографический след крупнейшего успеха Красной армии в Великой Отечественной войне», опубликованных под рубрикой «Архивные тайны». Деятельности советского руководства по сохранению памятников культуры и развитию спорта в ходе войны и в послевоенные годы посвящены статьи Е.Ю. Зубковой и А.И. Куприянова.

Завершает номер перечень научных проектов по изучению истории и различных аспектов Великой Отечественной войны, выполненных при поддержке Российского гуманитарного научного фонда и Российского фонда фундаментальных исследований (ныне – Российский центр научной информации).

80 лет отделяют нас от победоносного окончания Великой Отечественной войны – главного события XX в. не только отечественной, но и мировой истории. Авторы и редколлегия журнала выражают надежду, что публикуемые в этом номере статьи ведущих российских учёных внесут вклад в сохранение исторической памяти о Великой Отечественной войне.

## ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 2

Номер посвящён вопросам теологии и религиоведения



### Тематические редакторы номера:

**Д.В. Шмонин** – доктор философских наук, директор Института теологии Санкт-Петербургского государственного университета, заместитель координатора секции «Теология» Экспертного совета РГФИ по философии, социологии, политологии, правоведению и науковедению, председатель экспертного совета Высшей аттестационной комиссии при Минобрнауки России по теологии;

**В.А. Бажанов** – заслуженный деятель науки РФ, доктор философских наук, заведующий кафедрой философии Ульяновского государственного университета.

Статьи номера систематизированы по пяти рубрикам: «Теология как научная дисциплина», «Теология, гуманитарные науки и естествознание: аспекты взаимодействия», «Теологические исследования культуры и церковной традиции», «Современный исламский дискурс в фокусе теологии» и «Библейские исследования».

Важному направлению политики Российской Федерации в сфере государственно-конфессиональных отношений – гармонизации отношений между религиозными традициями, осуществляющими служение в нашей стране на протяжении её многовековой истории, – посвящена статья Н.М. Кропачева, А.В. Третьякова, А.А. Соловьёва и Д.В. Шмонина «Ценностно-правовые аспекты развития государственно-конфессиональных отношений, теологии и религиозного образования: к 30-летию работы Совета при Президенте Российской Федерации по взаимодействию с религиозными объединениями».

Аспекты взаимодействия теологии, гуманитарных наук и естествознания освещены в статьях Е.И. Аринина, Н.М. Марковой, М.С. Лютаевой, Ю.А. Сибирцевой, Т.А. Кильдяшовой, В.В. Голенко «Проект «Теология (богословие) и религиоведение в России: между медиа и диалогом экспертов»: опыт и перспективы», С.В. Кривовичева «Феномен жизни в естественных науках и теологии», Д.С. Курдыбайло «Слово Бога и Книга Природы в фокусе человеческого познания: предварительные замечания о периоде поздней античности» и Р.В. Светлова «Античные истоки рациональной теологии. Опыт исследования в рамках проекта РГФИ «Теология»».

В статьях Т.И. Афанасьевой «Ярославский часослов как уникальный памятник древнерусской литургической письменности второй половины XIII века», С.К. Костовска, Т.А. Галкиной, А.А. Герцен, С.В. Иванова «Храм и ландшафт: разработка и верификация типологии культовых сооружений по критериям места, эпохи, события» и П.В. Хондзинского, А.В. Малышева, В.А. Ячменика «Современное изучение древней латинской церковной традиции» рассмотрены вопросы теологических исследований культуры и церковной традиции.

Современный исламский дискурс в фокусе теологии проанализирован в статьях Т.К. Ибрагима, Н.В. Ефремовой «О возрождении философской теологии фальсафы в реформаторско-модернистском дискурсе» и П.А. Магомедовой «Аспекты исследования религиозного (исламского) дискурса в полиязыковом пространстве».

Библейским исследованиям посвящены статьи М.Г. Селезнёва, А.А. Луповой, Л.В. Шулякова «Структура рассказа о творении мира в первой главе книги Бытия» и А.В. Сизикова «Книга Иисуса, сына Сирахова, и проект академического библейского перевода».

## ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 3

Номер посвящён актуальной проблеме изучения  
и сохранения языкового разнообразия России



*Тематические редакторы номера:*

**А.А. Кибрик** – доктор филологических наук, член-корреспондент РАН, директор Института языкоизнания РАН (Ияз РАН);

**Т.И. Давидюк** – младший научный сотрудник Ияз РАН, программист НИВЦ МГУ им. М.В. Ломоносова;

**Д.Д. Мордашова** – кандидат филологических наук, младший научный сотрудник Ияз РАН.

В отечественной лингвистике приоритетное внимание уделяется изучению языкового разнообразия России, а также исторически и географически близких территорий. Этой области знания и посвящён настоящий выпуск.

Публикуемые статьи разделены на три рубрики. Первая из них, «Языковая политика и социолингвистика», объединяет четыре статьи, посвящённые языковой политике, т.е. системе принципов, определяющих планирование, использование и развитие языков на том или ином общественном уровне – от целой страны до отдельной семьи.

В статье Н.М. Мухарямова, О.Б. Януш «К вопросу о мезополитике культурно-языкового многообразия в российских условиях» особое внимание уделено концептуальному целеполаганию и нормативному закреплению положений о дву- и многоязычии в актах правовой политики, а также возможностям воспроизведения языкового многообразия. Статья Н.В. Борисовой и Н.С. Белоусовой «Языковая политика в регионах Приволжского федерального округа: факторы общественного запроса» посвящена роли этнических активистов. Их представительство в руководящих структурах региона позволяет наращивать социальный капитал в условиях плотных этнических социальных связей и высокого уровня внутригрупповой солидарности. Языковая ситуация, сложившаяся в Республике Дагестан в XXI в., сфера коммуникативной активности в функционировании языков Дагестана как один из важных факторов, влияющий на их витальность в наше время, анализируется в статье Б.М. Атаева и М.О. Ибрагимовой «Сфера функционирования языков Дагестана в современную эпоху: дистрибуция и тенденции развития». Назревшую необходимость корректировки государственной программы сохранения и развития хакасского языка на период с 2022 до 2029 г. выявили Т.Г. Боргоякова и А.В. Гусейнова в статье «Этноязыковая политика и социолингвистическая вариативность языкового сдвига (на примере Республики Хакасия)».

В статьях, опубликованных во второй рубрике, рассмотрены числительные второго десятка в языках Сибири (В.Ю. Гусев), образ рыбака в рассказе на хантыйском языке (А.Д. Каксин), дан сравнительный анализ кадарского и мегебского языков (Р.О. Муталов), рассмотрены идиомы сето на территории нашей страны (Т.Б. Агранат), а также вопросы создания письменности в XXI в. для бесписьменного языка уйльта Сахалина (Л.И. Миссонова) и цифровизации лингвистических исследований как способа сохранения языкового многообразия Сибири (Т.Р. Рыжикова).

Третья рубрика посвящена методам исследования языковой эрозии в билингвальном тексте (Н.В. Кондратьева), этноязыковой идентичности современных потомков аляскинских креолов (М.К. Раскладкина) и русскому старожильческому населению Монголии, его идентичности, языку и культуре (А.Н. Биткеева).

## ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 4

**Номер посвящён проблемам, связанным с гуманитарными аспектами  
применения искусственного интеллекта**



*Тематический редактор номера –*

*Д.В. Ушаков – академик РАН, доктор психологических наук, директор Института психологии РАН, заведующий кафедрой общей психологии МГУ им. М.В. Ломоносова*

В статье Л.И. Бородкина «Искусственный интеллект в исторических исследованиях: виртуальный ассистент или генератор квазизнаний?» рассмотрены актуальные вопросы использования в исторических исследованиях методов и технологий искусственного интеллекта (ИИ). Отмечены методологические и этические проблемы, возникающие в процессе апробации генеративного ИИ.

Актуальное состояние и перспективы развития технологий искусственного интеллекта, а также условия максимизации его эффектов для экономического роста анализируются в статье И.В. Данилина «Экономические и политические вызовы развития искусственного интеллекта».

В статье Т.Я. Хабриевой «Правовая социализация искусственного интеллекта и создание новой системы глобального управления» проведён анализ процесса социализации технологий искусственного интеллекта и формирования правового регулирования отношений, связанных с его внедрением.

Особое внимание В.К. Левашов и О.В. Гребняк в статье «На пороге цифровой цивилизации: диалектика искусственного интеллекта и гуманизма» уделяют противоречивым последствиям цифровизации. Авторы приходят к выводу о необходимости соблюдения этических норм и нравственных гуманистических императивов технологического прогресса, социальной ответственности, социокультурных и этнополитических особенностей.

С.Ю. Степанов в статье «Рефлексия и искусственный интеллект: от психологии рефлексии к рефлексивно-цифровым практикам развития человека» подчёркивает важность институционализации психологии рефлексии для разработки рефлексивных практик и психотехник.

Предлагаемый в статье Т.О. Шавриной и А.А. Корнилова «Мультиязычность в языковом моделировании: задачи, данные и возможности для типологических ресурсов» подход обеспечивает комплексную оценку способности языковых моделей точно интерпретировать и извлекать лингвистические признаки в контексте, создавая критически важный ресурс для масштабирования технологий на малоресурсные языки.

Различные аспекты, связанные с использованием искусственного интеллекта, рассмотрены в статьях А.Р. Бахтизина, А.В. Брагина «Генеративный искусственный интеллект в задачах социально-экономического прогнозирования»; А.В. Сиренова, Е.В. Анисимова, Т.А. Базаровой, М.Е. Проскуряковой «Искусственный интеллект и изучение письменного наследия Петра Великого и его сподвижников»; Д.И. Коломацкого, Е.В. Коровиной «Этический аспект использования технологий искусственного интеллекта в области сохранения языков коренных народов»; Ю.В. Балакиной, М.В. Григорьевой, Е.Н. Соколовой «Применение больших языковых моделей для анализа ценностно-патриотического дискурса русскоязычных пользователей».

В номере традиционно размещён «Перечень статей, опубликованных в журнале “Вестник РФФИ. Гуманитарные и общественные науки” в 2025 году».

**Верили в Победу свято:  
Аннотированный каталог научной литературы,  
изданной РГФИ в 1992–2022 гг. и посвящённой 80-летию  
Победы в Великой Отечественной войне**



*Сост.: В.В. Квардаков, В.Н. Фридлянов, Ю.Л. Воротников,  
Р.А. Казакова, И.Л. Ровинская, Л.С. Крекшина; Вступ. ст.: Ю.А. Петров. —  
М.: Российский центр научной информации, 2025. — 200 с.: ил.*

Выход в свет настоящего каталога приурочен в 80-летию Победы нашей страны в Великой Отечественной войне – главного события XX века не только в отечественной, но и в мировой истории. Не подлежит сомнению, что именно на советско-германском фронте в 1941–1945 гг. разворачивались основные события Второй мировой войны, решалась судьба человечества. Ключевыми факторами Победы стали массовый героизм и многонациональное братство советских людей. Неудивительно, что в России празднованию Дня Победы придавалось и придаётся исключительное значение. Наше общество воспринимает 9 мая как особое событие – торжественное и трагическое, как свидетельство национального единения перед лицом смертельной опасности. Символами несгибаемого мужества стали Брестская крепость, блокадный Ленинград, непобеждённый Севастополь, героический Сталинград, уничтоженная карателями вместе с мирными жителями белорусская деревня Хатынь.

Отстояв свою независимость, СССР сыграл ключевую роль в разгроме Третьего рейха и его сателлитов, а также в освобождении народов Европы от нацизма. Попытки замалчивания или ревизии этих положений, а тем более «исключений» нашей страны из числа победителей являются грубой фальсификацией исторической истины.

Предлагаемый вниманию читателей юбилейный каталог содержит аннотации на посвящённые Великой Отечественной войне научные труды, изданные Российским гуманитарным научным фондом (РГНФ) и Российским фондом фундаментальных исследований (РФФИ, ныне – Российский центр научной информации). Публикуемые аннотации на книги систематизированы по пяти разделам: «Накануне Второй мировой и Великой Отечественной войн», «Страна в огне», «Единство фронта и тыла», «СССР в послевоенном мире» и «Историческая память о Великой Отечественной войне».

Дополнением к основному блоку каталога стали два приложения. В первом из них приводится перечень поддержанных РГНФ и РФФИ научных проектов, посвящённых истории и различным аспектам Великой Отечественной войны. Во второе приложение включены рецензии на книги о войне, ранее опубликованные в журналах «Вестник РГНФ» и «Вестник РФФИ. Гуманитарные и общественные науки».

Каталог проиллюстрирован изображениями обложек вошедших в него книг и плакатов военных лет.

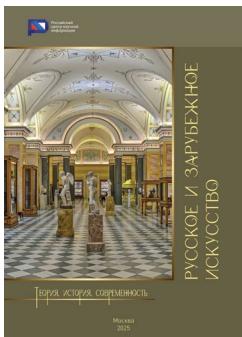
Для удобства использования каталог снабжён именным указателем.

Издание адресовано учёным-гуманитариям, а также широкому кругу читателей, хранящих память о героическом подвиге нашего народа в Великой Отечественной войне и желающих пополнить свои знания о её истории.

Надеемся, что изданные при поддержке РГНФ и РФФИ труды российских учёных о Великой Отечественной войне, выполненные научные исследования и публикация настоящего каталога РЦНИ внесут вклад в сохранение исторической памяти о Великой Отечественной войне.



**Русское и зарубежное искусство.  
Теория, история, современность:  
Аннотированный каталог научной литературы,  
изданной РФФИ (РГНФ) в 1992–2022 гг.**



*Сост.: Ю.Л. Воротников, В.В. Квардаков, В.Н. Фридлянов; Ред. Р.А. Казакова,  
И.Л. Ровинская. — М.: Российский центр научной информации, 2025. — 360 с.: ил.*

Каждое настоящее творение искусства – будь то храм или живописное полотно, музыкальное произведение или скульптурная композиция, художественный фильм или театральная постановка – заслуживает внимательного отношения, вдумчивого изучения и бережного сохранения. На протяжении 30 лет Российской фонд фундаментальных исследований и Российской гуманитарный научный фонд уделяли большое внимание изучению и сохранению отечественного и мирового культурного наследия. При поддержке фондов ежегодно выполнялись сотни проектов проведения научных исследований, организаций экспедиций и конференций по широкому спектру жанров искусства. Результатом выполнения этих проектов становились индивидуальные и коллективные монографии, сборники трудов и материалы научных форумов, каталоги ценных музейных коллекций, факсимильные издания уникальных первоисточников.

В ретроспективном тематическом каталоге представлены аннотации на труды по искусствоведению, изданные в 1992–2022 годах РГФИ (РГНФ). Наряду с высоким научным уровнем книги по искусствоведению отличает исключительное богатство тематики, что отражено в настоящем каталоге.

Издание состоит из двух частей. В первой из них приведены аннотации на научную литературу, вышедшую в свет в рамках издательских конкурсов; во второй части опубликованы аннотации на книги, изданные по итогам конкурсов на лучшие научно-популярные труды. Составившие первый, наиболее объёмный раздел книги систематизированы по семи подразделам: «Проблемы теории искусства», «История искусства разных стран и регионов», «Архитектура и градостроительство», «Изобразительное искусство», «Декоративно-прикладное искусство», «Музыкальное искусство», «Искусство театра и кино».

Во включённых в первый подраздел книгах исследуются не теряющие своей актуальности теоретические вопросы, касающиеся различных жанров искусства. Глубже понять истоки и тенденции развития разных жанров искусства помогают публикуемые во втором подразделе труды, освещающие историю искусства разных стран и регионов.

В каталоге представлены также издания, посвящённые исследованию жизни и творчества выдающихся художников, иконописцев, скульпторов, мастеров станковой гравюры, церковных фресок и росписей жилых домов Русского Севера, лицевого шитья и реставрационного искусства.

Каталог проиллюстрирован изображениями обложек вошедших в него книг и снабжён именным указателем.

Как сказал один известный драматург: «Все виды искусств служат величайшему из искусств – искусству жить на земле». Помимо эстетической ценности, искусство должно служить жизни человека, помогая ему лучше понять мир, самого себя и выстроить гармоничные отношения с окружающими.

Издание адресовано учёным-искусствоведам, а также широкому кругу читателей, интересующихся различными жанрами русского и зарубежного искусства.



## ПЕРЕЧЕНЬ СТАТЕЙ, ОПУБЛИКОВАННЫХ В ЖУРНАЛЕ «ВЕСТНИК РФФИ. ГУМАНИТАРНЫЕ И ОБЩЕСТВЕННЫЕ НАУКИ» В 2025 г.

### № 1, ПОСВЯЩЁННЫЙ 80-ЛЕТИЮ ВЕЛИКОЙ ПОБЕДЫ

Петров Ю.А. Навеки в памяти народной. — № 1. — С. 9–12.

#### В ШТАБАХ ПОБЕДЫ

Мякота В.Н., Кудряшов С.В. Советские планы ответа на военные приготовления Германии. — № 1. — С. 13–25.

Панков О.Д. Спецпропаганда в годы Великой Отечественной войны. — № 1. — С. 26–38.

Попов А.Ю. О централизации руководства партизанским движением в годы Великой Отечественной войны. — № 1. — С. 39–51.

#### БЕЗ СРОКА ДАВНОСТИ

Аристов С.В. Судьба советских военнопленных в годы войны. — № 1. — С. 52–63.

Кудряшов С.В. Война на уничтожение. Нацистский террор на советских территориях. — № 1. — С. 64–78.

Соловьёв С.М. Отношения между союзниками по антигитлеровской коалиции в ходе подготовки и проведения Нюрнбергского процесса. — № 1. — С. 79–96.

#### АРХИВНЫЕ ТАЙНЫ

Беклемишиева М.М. Трофейные документы Великой Отечественной войны в фокусе внимания разведывательных, политических и архивных органов СССР. — № 1. — С. 97–107.

Домнин А.И. Операция «Багратион»: картографический след. — № 1. — С. 108–124.

#### СОХРАНЕНИЕ ПАМЯТНИКОВ КУЛЬТУРЫ И РАЗВИТИЕ СПОРТА В ГОДЫ ВОЙНЫ

Зубкова Е.Ю. Война и судьба памятников культуры: на пути к возрождению. «Новгородский феномен». — № 1. — С. 125–139.

Куприянов А.И. Спорт в годы Великой Отечественной войны. — № 1. — С. 140–152.

Научные проекты РФФИ (РГНФ) по изучению истории и различных аспектов Великой Отечественной войны. — № 1. — С. 153–159.

### № 2, ПОСВЯЩЁННЫЙ ТЕОЛОГИИ В КОНТЕКСТЕ МЕЖДИСЦИПЛИНАРНЫХ НАУЧНЫХ ИССЛЕДОВАНИЙ

#### ТЕОЛОГИЯ КАК НАУЧНАЯ ДИСЦИПЛИНА

Шмонин Д.В. Междисциплинарное и конфессиональное: теология в конкурсном контексте. — № 2. — С. 9–18.

Бажанов В.А. Теология в Российском фонде фундаментальных исследований (исторические предпосылки и опыт выполнения грантовых проектов). — № 2. — С. 19–24.

Кропачев Н.М., Третьяков А.В., Соловьёв А.А., Шмонин Д.В. Ценностно-правовые аспекты развития государственно-конфессиональных отношений, теологии и религиозного образования: к 30-летию работы Совета при Президенте Российской Федерации по взаимодействию с религиозными объединениями. — № 2. — С. 25–36.

## ТЕОЛОГИЯ, ГУМАНИТАРНЫЕ НАУКИ И ЕСТЕСТВОЗНАНИЕ: АСПЕКТЫ ВЗАИМОДЕЙСТВИЯ

Аринин Е.И., Маркова Н.М., Лятаева М.С., Сибирцева Ю.А., Кильдяшова Т.А., Голенко В.В. Проект «Теология (богословие) и религиоведение в России: между медиа и диалогом экспертов»: опыт и перспективы. — № 2. — С. 37–47.

Кривовичев С.В. Феномен жизни в естественных науках и теологии. — № 2. — С. 48–58.

Курдыбайло Д.С. Слово Бога и Книга Природы в фокусе человеческого познания: предварительные замечания о периоде поздней античности. — № 2. — С. 59–70.

Светлов Р.В. Античные истоки рациональной теологии. Опыт исследования в рамках проекта РФФИ. — № 2. — С. 71–81.

## ТЕОЛОГИЧЕСКИЕ ИССЛЕДОВАНИЯ КУЛЬТУРЫ И ЦЕРКОВНОЙ ТРАДИЦИИ

Афанасьева Т.И. Ярославский часослов как уникальный памятник древнерусской литургической письменности второй половины XIII века. — № 2. — С. 82–90.

Костовска С.К., Галкина Т.А., Герцен А.А., Иванов С.В. Храм и ландшафт: разработка и верификация типологии культовых сооружений по критериям места, эпохи, события. — № 2. — С. 91–104.

Хондзинский П.В., Малышев А.В., Ячменик В.А. Современное изучение древней латинской церковной традиции. — № 2. — С. 105–114.

## СОВРЕМЕННЫЙ ИСЛАМСКИЙ ДИСКУРС В ФОКУСЕ ТЕОЛОГИИ

Ибрагим Т.К., Ефремова Н.В. О возрождении философской теологии фальсафы в реформаторско-модернистском дискурсе. — № 2. — С. 115–121.

Магомедова Л.А. Аспекты исследования религиозного (исламского) дискурса в полиязыковом пространстве. — № 2. — С. 122–133.

## БИБЛЕЙСКИЕ ИССЛЕДОВАНИЯ

Селезнёв М.Г., Луплова А.А., Шуляков Л.В. Структура рассказа о творении мира в первой главе книги Бытия. — № 2. — С. 134–143.

Сизиков А.В. Книга Иисуса, сына Сирахова, и проект академического библейского перевода. — № 2. — С. 144–153.

Поддержанные научные проекты конкурса «Теология в контексте междисциплинарных научных исследований». — № 2. — С. 154–159.

## № 3, ПОСВЯЩЁННЫЙ ЯЗЫКОВОМУ РАЗНООБРАЗИЮ НАРОДОВ РОССИИ

### КОЛОНКА ТЕМАТИЧЕСКОГО РЕДАКТОРА

Кибрик А.А., Давидюк Т.И., Мордашова Д.Д. Языковое разнообразие России и сопредельных территорий: изучение и сохранение. — № 3. — С. 7–9.

### ЯЗЫКОВАЯ ПОЛИТИКА И СОЦИОЛИНГВИСТИКА

Мухаряров Н.М., Януш О.Б. К вопросу о мезополитике культурно-языкового многообразия в российских условиях. — № 3. — С. 10–22.

Борисова Н.В., Белоусова Н.С. Языковая политика в регионах Приволжского федерального округа: факторы общественного запроса. — № 3. — С. 23–32.

Атаев Б.М., Ибрагимова М.О. Сфера функционирования языков Дагестана в современную эпоху: дистрибуция и тенденции развития. — № 3. — С. 33–43.

Боргоякова Т.Г., Гусейнова А.В. Этноязыковая политика и социолингвистическая вариативность языкового сдвига (на примере Республики Хакасия). — № 3. — С. 44–53.

## ОПИСАНИЕ И ДОКУМЕНТИРОВАНИЕ ЯЗЫКОВ

Гусев В.Ю. Числительные второго десятка в языках Сибири. — № 3. — С. 54–69.

Каксин А.Д. Образ рыбака в рассказе на хантыйском языке: разговорные средства в структуре художественного текста. — № 3. — С. 70–79.

Муталов Р.О. Сравнительный анализ кадарского и мегебского: сходства и различия в даргинской ветви нахско-дагестанских языков. — № 3. — С. 80–92.

Агранат Т.Б. Идиомы сето на территории РФ. — № 3. — С. 93–101.

Миссонова Л.И. Создание письменности в XXI веке для бесписьменного языка уильта Сахалина: вопросы истории и современное решение. — № 3. — С. 102–111.

Рыжикова Т.Р. Цифровизация лингвистических исследований как способ сохранения языкового многообразия Сибири (база данных диалогической речи миноритарных этносов Новосибирской области и сопредельных регионов). — № 3. — С. 112–121.

## ФУНКЦИОНИРОВАНИЕ ЯЗЫКОВ В ПОЛИЛИНГВАЛЬНОЙ СРЕДЕ

Кондратьева Н.В. К вопросу о методах исследования языковой эрозии в билингвальном тексте (на примере русско-удмуртского двуязычия). — № 3. — С. 122–132.

Раскладкина М.К. Этноязыковая идентичность современных потомков аляскинских креолов. — № 3. — С. 133–147.

Биткеева А.Н. Русское старожильческое население Монголии: идентичность, язык, культура. — № 3. — С. 148–159.

## № 4, ПОСВЯЩЁННЫЙ СОЦИОГУМАНИТАРНЫМ ИССЛЕДОВАНИЯМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

### КОЛОНКА ТЕМАТИЧЕСКОГО РЕДАКТОРА

О редакторе тематического блока академике РАН Дмитрии Викторовиче Ушакове. — № 4. — С. 9.

Ушаков Д.В. Алгебра и гармония: технологии искусственного интеллекта в социогуманитарных науках. — № 4. — С. 10–14.

## ОБЩЕСТВО СОСУЩЕСТВОВАНИЯ ЕСТЕСТВЕННОГО ИНТЕЛЛЕКТА С ИСКУССТВЕННЫМ

Хабриева Т.Я. Правовая социализация искусственного интеллекта и создание новой системы глобального управления. — № 4. — С. 15–24.

Данилин И.В. Экономические и политические вызовы развития искусственного интеллекта. — № 4. — С. 25–33.

Левашов В.К., Гребняк О.В. На пороге цифровой цивилизации: диалектика искусственного интеллекта и гуманизма. — № 4. — С. 34–45.

Коломацкий Д.И., Коровина Е.В. Этический аспект использования технологий искусственного интеллекта в области сохранения языков коренных народов. — № 4. — С. 46–55.

Балакина Ю.В., Григорьева М.В., Соколова Е.Н. Применение больших языковых моделей для анализа ценностно-патриотического дискурса русскоязычных пользователей. — № 4. — С. 56–69.

## ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ КАК СРЕДСТВО ПОЗНАНИЯ И ПОМОЗИ ЧЕЛОВЕКУ И ОБЩЕСТВУ

Бахтизин А.Р., Брагин А.В. Генеративный искусственный интеллект в задачах социально-экономического прогнозирования. — № 4. — С. 70–86

Бородкин Л.И. Искусственный интеллект в исторических исследованиях: виртуальный ассистент или генератор квазизнаний? — № 4. — С. 87–98.

Сиренов А.В., Анисимов Е.В., Базарова Т.А., Проскурякова М.Е. Искусственный интеллект и изучение письменного наследия Петра Великого и его сподвижников. — № 4. — С. 99–109

Степанов С.Ю. Рефлексия и искусственный интеллект: от психологии рефлексии к рефлексивно-цифровым практикам развития человека. — № 4. — С. 110–121.

## ПАРАЛЛЕЛИ МЕЖДУ ЕСТЕСТВЕННЫМ И ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ

Шаврина Т.О., Корнилов А.А. Мультиязычность в языковом моделировании: задачи, данные и возможности для типологических ресурсов. — № 4. — С. 122–135.

## ИЗДАНИЯ РЦНИ ПО ГУМАНИТАРНЫМ НАУКАМ

ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 1

Номер посвящён 80-летию Победы в Великой Отечественной войне 1941–1945 гг. — № 4. — С. 136.

ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 2

Номер посвящён вопросам теологии и религиоведения. — № 4. — С. 137.

ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 3

Номер посвящён актуальной проблеме изучения и сохранения языкового разнообразия России. — № 4. — С. 138.

ВЕСТНИК РФФИ. Гуманитарные и общественные науки. 2025. № 4

Номер посвящён проблемам, связанным с гуманитарными аспектами применения искусственного интеллекта. — № 4. — С. 139.

Верили в Победу свято: Аннотированный каталог научной литературы, изданной РФФИ в 1992–2022 гг. и посвящённой 80-летию Победы в Великой Отечественной войне. — № 4. — С. 140.

Русское и зарубежное искусство. Теория, история, современность: Аннотированный каталог научной литературы, изданной РФФИ (РГНФ) в 1992–2022 гг. — № 4. — С. 142.

Перечень статей, опубликованных в журнале «Вестник РФФИ. Гуманитарные и общественные науки» в 2025 г. — № 4. — С. 144–151.

## LIST OF ARTICLES PUBLISHED IN THE JOURNAL “RFBR BULLETIN. HUMANITIES AND SOCIAL SCIENCES” IN 2025

### No. 1, DEDICATED TO THE 80TH ANNIVERSARY OF THE GREAT VICTORY

*Petrov Yu.A.* Forever in the people's memory — No 1. — Pp. 9–12.

#### AT VICTORY'S HEADQUARTERS

*Myakota V.N., Kudryashov S.V.* Soviet Plans to Respond to German Military Preparations — No 1. — Pp. 13–25.

*Pankov O.D.* Psychological Warfare during the Great Patriotic War. — No 1. — Pp. 26–38.

*Popov A.Yu.* On the Centralisation of the Leadership of the Partisan Movement during the Great Patriotic War. — No 1. — Pp. 39–51.

#### NO STATUTE OF LIMITATION

*Aristov S.V.* Fate of Soviet Prisoners of War during the War Years. — No 1. — Pp. 52–63.

*Kudryashov S.V.* War of Extermination. Nazi Terror in Soviet Territories. — No 1. — Pp. 64–78.

*Solovyov S.M.* Relations between the Allies in the Anti-Hitler Coalition Before and During the Nuremberg Trials. — No 1. — Pp. 79–96.

#### ARCHIVAL SECRETS

*Beklemisheva M.M.* Captured Documents of the Great Patriotic War in the Center of Attention of the Soviet Intelligence, Political Bodies and Archives. — No 1. — Pp. 97–107.

*Dominin A.I.* Bagration Operation: the Cartographic Mark. — No 1. — Pp. 108–124.

#### PRESERVATION OF CULTURAL HERITAGE AND SPORTS DEVELOPMENT DURING THE WAR

*Zubkova E.Yu.* War and Cultural Heritage: on the Way to Restoration. “The Novgorod Phenomenon”. — No 1. — Pp. 125–139.

*Kupriyanov A.I.* Sport during the Great Patriotic War. — No 1. — Pp. 140–152.

Scientific projects of the Russian Foundation for Basic Research (Russian Humanitarian Scientific Foundation) on studying the history and different aspects of the Great Patriotic War. — No 1. — Pp. 153–159.

### No. 2, DEDICATED TO THEOLOGY IN THE CONTEXT OF INTERDISCIPLINARY SCIENTIFIC RESEARCH

#### THEOLOGY AS A SCIENTIFIC DISCIPLINE

*Shmonin D.V.* The Interdisciplinary and Confessional: Theology in Competitive Context . — No 2. — Pp. 9–18.

*Bazhanov V.A.* Theology in the Russian Foundation for Basic Research (Historical Background and Practice of Implementing Grant Projects). — No 2. — Pp. 19–24.

*Kropachev N.M., Tretyakov A.V., Solovyov A.A., Shmonin D.V.* Value-Based and Legal Aspects of Development of State & Confessions Relations, Theology, and Religious Education: on the 30<sup>th</sup> Anniversary of the Council for Coordination with Religious Organizations under the President of the Russian Federation. — No 2. — Pp. 25–36.

## THEOLOGY, HUMANITIES AND NATURAL SCIENCES: ASPECTS OF RELATIONSHIP

*Arinin E.I., Markova N.M., Lyutaeva M.S., Sibirtseva Yu.A., Kildyashova T.A., Golenko V.V.* Project: "Theology and Religious Studies in Russia: Between Media and Expert Dialogue": Experience and Prospects. — No 2. — Pp. 37–47.

*Krivovichev S.V.* The Phenomenon of Life in Natural Sciences and Theology. — No 2. — Pp. 48–58.

*Kurdybaylo D.S.* The Word of God and the Book of Nature in the Focus of Human Knowledge: Preliminary Remarks on Late Antiquity. — No 2. — Pp. 59–70.

*Svetlov R.V.* Ancient Origins of Rational Theology. Research Experience in the Project of the Russian Foundation for Basic Research. — No 2. — Pp. 71–81.

## THEOLOGICAL STUDIES OF CULTURE AND CHURCH TRADITION

*Afanasyeva T.I.* The Yaroslavl Book of Hours as a Unique Monument of Ancient Russian Liturgical Writing of the Second Half of the 13<sup>th</sup> Century. — No 2. — Pp. 82–90.

*Kostovska S.K., Galkina T.A., Herzen A.A., Ivanov S.V.* Temple and Landscape: Development and Verification of the Typology of Religious Buildings According to Place, Era, Event. — No 2. — Pp. 91–104.

*Khondzinsky P.V., Malyshev A.V., Yachmenik V.A.* Modern Study of Ancient Latin Church Tradition. — No 2. — Pp. 105–114.

## MODERN ISLAMIC DISCOURSE IN THE FOCUS OF THEOLOGY

*Ibrahim T.K., Efremova N.V.* On the Revival of the Philosophical Theology of Falsafa in Reformist-Modernist Discourse. — No 2. — Pp. 115–121.

*Magomedova P.A.* Aspects of Study of Religious (Islamic) Discourse in Multilingual Space. — No 2. — Pp. 122–133.

## BIBLICAL STUDIES

*Seleznyov M.G., Luppova A.A., Shulyakov L.V.* Structure of Creation Narrative in Genesis Chapter 1. — No 2. — Pp. 134–143.

*Sizikov A.V.* The Wisdom of Jesus, Son of Sirach, and the Project of Academic Bible Translation. — No 2. — Pp. 144–153.

Supported scientific projects of the competition "Theology in the Context of Interdisciplinary Scientific Research". — No 2. — Pp. 154–159.

## No. 3, DEDICATED TO THE LINGUISTIC DIVERSITY OF THE PEOPLES OF RUSSIA

### THEMED ISSUE EDITOR'S COLUMN

*Kibrik A.A., Davidyuk T.I., Mordashova D.D.* Linguistic Diversity of Russia and Adjacent Territories: Study and Preservation. — No 3. — Pp. 7–9.

### LANGUAGE POLICY AND SOCIOLINGUISTICS

*Mukharyamov N.M., Yanush O.B.* On Issue of Mesopolicy of Cultural and Linguistic Diversity in Russian Context. — No 3. — Pp. 10–22.

*Borisova N.V., Belousova N.S.* Language Policy in Regions of Volga Federal District: Factors of Public Demand. — No 3. — Pp. 23–32.

*Ataev B.M., Ibragimova M.O.* Areas of Functioning of Dagestan Languages in Modern Era: Distribution and Development Trends. — No 3. — Pp. 33–43.

*Borgoyakova T.G., Guseynova A.V.* Ethnic and Linguistic Policy and Sociolinguistic Variability of Language Shift (Example of Republic of Khakassia). — No 3. — Pp. 44–53.

## DESCRIPTION AND DOCUMENTATION OF LANGUAGES

Gusev V.Yu. Numerals of Second Ten in Languages of Siberia . — No 3. — Pp. 54–69.

Kaksin A.D. Image of Fisherman in Khanty Narrative: Colloquial Means in Structure of Fictional Text . — No 3. — Pp. 70–79.

Mutalov R.O. Comparative Analysis of Kadar and Mehweb: Similarities and Differences in Dargin Branch of Nakh-Dagestani Languages. — No 3. — Pp. 80–92.

Agranat T.B. Seto Idioms in Russia. — No 3. — Pp. 93–101.

Missonova L.I. Development of Writing in 21<sup>st</sup> Century for Unwritten Language of Sakhalin Ulta People: Historical Issues and Modern Solution. — No 3. — Pp. 102–111.

Ryzhikova T.R. Digitization of Linguistic Research as Means of Preservation of Linguistic Diversity of Siberia (Database of Diagonal Speech of Minority Ethnic Groups of Novosibirsk Region and Adjacent Regions). — No 3. — Pp. 112–121.

## FUNCTIONING OF LANGUAGES IN MULTILINGUAL ENVIRONMENT

Kondratieva N.V. On Methods of Studying Linguistic Attrition in Bilingual Text (Example of Russian-Udmurt Bilingualism). — No 3. — Pp. 122–132.

Raskladkina M.K. The Ethnic and Linguistic Identity of the Descendants of Alaskan Creoles in Our Days. — No 3. — Pp. 133–147.

Bitkeeva A.N. Russian Old-Timers in Mongolia: Identity, Language, Culture. — No 3. — Pp. 148–159.

## No. 4, DEDICATED TO SOCIO-HUMANITARIAN STUDIES OF ARTIFICIAL INTELLIGENCE

### THEMED ISSUE EDITOR'S COLUMN

*About the Editor of the Thematic Section, Academician of the Russian Academy of Sciences, D.V. Ushakov.* — No 4. — P. 9.

Ushakov D.V. Algebra and Harmony: Artificial Intelligence Technologies in the Social Sciences and Humanities. — No 4. — Pp. 10–14.

## SOCIETY OF COEXISTENCE OF NATURAL AND ARTIFICIAL INTELLIGENCE

Khabrieva T.Ya. Legal Socialization of Artificial Intelligence and the Creation of a New Global Governance System. — No 4. — Pp. 15–24.

Danilin I.V. Economic and Political Challenges of Artificial Intelligence Development . — No 4. — Pp. 25–33.

Levashov V.K., Grebnyak O.V. On the Threshold of Digital Civilization: the Dialectic of Artificial Intelligence and Humanism. — No 4. — Pp. 34–45.

Kolomatsky D.I., Korovina E.V. The Ethical Aspect of Using Artificial Intelligence Technologies In the Field of Indigenous Languages Preservation. — No 4. — Pp. 46–55.

Balakina Yu.V., Grigoryev M.V., Sokolova E.N. Application of Large Language Models for the Analysis of Value-Patriotic Discourse of Russian-Speaking Users. — No 4. — Pp. 56–69.

## ARTIFICIAL INTELLIGENCE AS A MEANS OF COGNITION AND SUPPORT FOR HUMANS AND SOCIETY

Bakhtizin A.R., Bragin A.V. Applications of Generative Artificial Intelligence in Socioeconomic Forecasting. — No 4. — Pp. 70–86.

Borodkin L.I. Artificial Intelligence in Historical Research: A Virtual Assistant or a Generator of Quasi-Knowledge? — No 4. — Pp. 87–98.

Sirenov A.V., Anisimov E.V., Bazarova T.A., Proskuryakova M.E. Artificial Intelligence and the Study of the Written Legacy of Peter the Great and His Associates. — No 4. — Pp. 99–109.

Stepanov S.Yu. Reflection and Artificial Intelligence: From the Psychology of Reflection to Reflexive Digital Practices of Human Development. — No 4. — Pp. 110–121.

## PARALLELS BETWEEN NATURAL AND ARTIFICIAL INTELLIGENCE

Shavrina T.O., Kornilov A.A. Multilinguality in Language Modeling: Tasks, Data, and Opportunities for Typological Resources. — No 4. — Pp. 122–135.

## PUBLICATIONS OF THE RUSSIAN CENTER FOR SCIENCE INFORMATION IN THE HUMANITIES

RFBR Bulletin. Humanities and Social Sciences. 2025. No. 1. The issue is dedicated to the 80<sup>th</sup> anniversary of Victory in the Great Patriotic War of 1941–1945. — No 4. — P. 136.

RFBR Bulletin. Humanities and Social Sciences. 2025, No. 2. The issue is dedicated to theology and religious studies. — No 4. — P. 137.

RFBR Bulletin. Humanities and Social Sciences. 2025, No. 3. The issue is dedicated to the current problem of studying and preserving the linguistic diversity of Russia. — No 4. — P. 138.

RFBR Bulletin. Humanities and Social Sciences. 2025, No. 4. The issue is dedicated to the problems related to the humanitarian aspects of the application of artificial intelligence. — No 4. — P. 139.

They Kept the Faith in Victory: Annotated Catalog of Scholarly Literature Published by the Russian Foundation for Basic Research from 1992 to 2022 and Dedicated to the 80<sup>th</sup> Anniversary of Victory in the Great Patriotic War. — No 4. — P. 140.

Russian and Foreign Art. Theory, History, Modernity: An Annotated Catalog of Scientific Literature published by the Russian Foundation for Basic Research (Russian Humanitarian Research Foundation) in 1992–2022. — No 4. — P. 142.

List of Articles Published in the RFBR Bulletin. Humanities and Social Sciences in 2025. — No 4. — Pp. 144–151.

# Вестник Российского фонда фундаментальных исследований

## Гуманитарные и общественные науки

2025, № 4 (123)

Адрес редакции: 119991, Москва, Ленинский пр-т, д. 32а  
Тел.: +7 (499) 941-01-15 (доб. 3121)  
E-mail: roviro@rcsi.science  
<https://rcsi.science>

Подписано в печать 12.12.2025. Формат 60×90  $\frac{1}{8}$ .  
Усл. печ. л. 19. Печать офсетная. Бумага мелованная.  
Тираж 300 экз. Заказ № 361315

Отпечатано в типографии ООО «Вива-Стар».  
107023, г. Москва, ул. Электрозаводская, д. 20, стр. 3