

Институт научной информации по общественным наукам
Российской академии наук
(ИНИОН РАН)

НАУКОВЕДЧЕСКИЕ ИССЛЕДОВАНИЯ

Научный журнал

2026. №3

Выпуск посвящен памяти
Владимира Ивановича Аршинова

Учредитель
Институт научной информации по общественным наукам РАН

Редакционная коллегия:

А.Г. Аллахвердян (Москва, Россия), Н.Л. Архипьев (Москва, Россия), Н.А. Ащеулова (Санкт-Петербург, Россия), А. Грунвальд (Armin Grunwald) (Карлсруэ, Берлин, Германия), С.В. Егерев (Москва, Россия), В.Г. Буданов (Москва, Россия), И.В. Мелик-Гайказян (Томск, Россия), О.В. Москалева (Санкт-Петербург, Россия), О.Н. Субочева (Москва, Россия), В.В. Лапаева (Москва, Россия), Л. Райзер (Leandro Raizer) (Порту-Алегри, Бразилия), В. Хофкиршнер (Wolfgang Hofkirchner) (Вена, Австрия), М.А. Ядова (Москва, Россия)

Редакция:

*Главный редактор – Е.Г. Гребенщикова д-р филос. наук
Заместители главного редактора – И.А. Асеева д-р филос. наук
С.М. Пястолов – д-р эконом. наук
Ответственный редактор – И.А. Асеева д-р филос. наук*

«Научковедческие исследования» – рецензируемый журнал открытого доступа.
Журнал учрежден Институтом научной информации по общественным наукам РАН и является преемником одноименного ежегодника, который издавался в ИНИОН РАН с 2003 по 2021 г.

Журнал выходит четыре раза в год.

Публикуемые материалы прошли процедуру рецензирования
и экспертного отбора.

DOI: 10.31249/naukoved/2026.01.00
ISSN 2658–5405

© ИНИОН РАН, 2026

СОДЕРЖАНИЕ

НАУКА, ТЕХНОЛОГИИ И ОБЩЕСТВО

<i>Буданов В.Г.</i> Наука в цифровую эпоху.....	5
<i>Егеров С.В.</i> Проект “Lost Science”: журналисты и общественность в борьбе за будущее научных исследований	17

ФИЛОСОФСКИЕ ПРОБЛЕМЫ НАУЧНО-ТЕХНИЧЕСКОГО РАЗВИТИЯ

<i>Сабанина Н.Р.</i> Философ и большая языковая модель в архитектуре постнеклассической науки	29
<i>Янукович М.Ф.</i> Использование нейросетей в научных исследованиях: к вопросу о природе нового знания	50

ОРГАНИЗАЦИЯ НАУЧНОЙ ДЕЯТЕЛЬНОСТИ. УПРАВЛЕНИЕ НАУКОЙ

<i>Стельмахов Д.А.</i> Научная среда как конститутивное условие научного творчества.....	58
---	----

ИСТОРИЯ И МЕТОДОЛОГИЯ РАЗВИТИЯ НАУКИ И ТЕХНИКИ

<i>Тарасенко В.В.</i> Кластерно-микросервисная архитектура платформенной науки как инструмент демаркации научного знания в контексте философии сложности Владимира Аршинова.....	71
<i>Гребцов М.В.</i> Междисциплинарность как методологический режим современной науки.....	84
<i>Сушин М.А.</i> Предиктивная методология машинного обучения и поиск механизмов познания и поведения	98

ОБЗОРЫ, РЕЦЕНЗИИ, РЕФЕРАТЫ, ЭССЕ

<i>Самохин А.А.</i> Что говорит искусственный интеллект	114
---	-----

CONTENTS

SCIENCE, TECHNOLOGY AND SOCIETY

<i>Budanov V.G.</i> Science in the digital age	5
<i>Egerev S.V.</i> The “Lost Science” project: journalists and the public in the fight for the future of scientific research	17

PHILOSOPHICAL PROBLEMS OF SCIENTIFIC AND TECHNOLOGICAL DEVELOPMENT

<i>Sabanina N.R.</i> Philosopher and Large Language Model in the architecture of postnonclassical science.....	29
<i>Yanukovich M.F.</i> The use of neural networks in scientific research: on the nature of the new knowledge.....	50

ORGANIZATION AND MANAGEMENT OF SCIENCE

<i>Stelmakhov D.A.</i> Scientific environment as a constitutive condition for scientific creativity	58
--	----

HISTORY AND METHODOLOGY OF THE DEVELOPMENT OF SCIENCE AND TECHNOLOGY

<i>Tarasenko V.V.</i> Cluster-microservice architecture of platform science as a tool for demarcation of scientific knowledge in the context of Vladimir Arshinov's philosophy of complexity	71
<i>Grebtsov M.V.</i> Interdisciplinarity as amethodological regime of modern science.....	84
<i>Sushchin M.A.</i> Predictive methodology of machine learning and the search for mechanisms of cognition and behavior	98

LITERATURE REVIEWS, BOOK REVIEWS, ABSTRACTS, ESSAY

<i>Samokhin A.A.</i> What artificial intelligence says.....	114
---	-----

НАУКА В ЦИФРОВУЮ ЭПОХУ

Буданов Владимир Григорьевич – д-р филос. наук, канд. физ.-мат. наук, главный научный сотрудник, Институт философии РАН, Москва, Россия; budsyn@yandex.ru; ORCID 0000-0003-2371-8659

Аннотация. Статья посвящена проблемам и возможностям применения искусственного интеллекта в науке и вообще в любой форме деятельности. В первой части работы обсуждаются новые формы научной рациональности, обобщающие подход В.С. Степина, и тетраэдрическое представление деятельности, позволяющее классифицировать ее стратегии и обобщенные формы рациональности. Введен концепт включенного наблюдателя деятельности, который выполняет функции не только мониторинга, но и может выполнять роли сопровождения и управления деятельностью в реальном времени, что далеко не всегда доступно человеку. Введение искусственного интеллекта в деятельностный процесс может проводиться на уровне всех элементов деятельностного тетраэдра, а также при гибридации субъекта деятельности или наблюдателя, что приводит к необходимости понимания специфических проблем и рисков гибридации. Обсуждаются новые задачи по самоописанию и доверию к себе гибридных субъектов, показано, что возникают неконтролируемые неопределенности семантической природы, с которыми ранее математика не сталкивалась. Показано имманентное использование посткритической рациональности в гибридных средах. Обсуждается граница демаркации искусственного и естественного интеллектов и перспективы искусственных социумов, новой зимы в исследовании искусственного интеллекта.

Ключевые слова: искусственный интеллект; деятельностный тетраэдр; типы рациональности; гибридный квазисубъект; научный эксперимент; цифровая революция, риски.

Для цитирования: Буданов В.Г. Наука в цифровую эпоху // Научно-исследовательские исследования. – 2026. – № 3. – С. 5–16. – DOI: <http://www.doi.org/10.31249/scis/2026.03.01>

Статья поступила в редакцию: 20.05.2026

Принята к публикации: 15.06.2026

SCIENCE IN THE DIGITAL AGE

Vladimir G. Budanov – Doctor of Sciences (Philosophy), Candidate of Sciences (Physics and Mathematics), Chief Research Fellow, Institute of Philosophy of the Russian Academy of Sciences, Moscow, Russia; budsyn@yandex.ru; ORCID 0000-0003-2371-8659

Abstract. The article is devoted to the problems and possibilities of using artificial intelligence in science, and in general in any form of activity. In the first part of the paper, new forms of scientific rationality are discussed, generalizing the approach in, with. Stepin, and the tetrahedral representation of activity, which makes it possible to classify its strategies and generalized forms of rationality. The concept of an enabled activity observer has been introduced, which performs not only monitoring functions, but can also perform the roles of monitoring and managing activities in real time, which is far from always accessible to humans. The introduction of artificial intelligence into the activity process can be carried out at the level of all elements of the activity tetrahedron, as well as during the hybridization of the subject of activity or observer, which leads to the need to understand the specific problems and risks of hybridization. New tasks on self-description and self-confidence of hybrid subjects are discussed, and it is shown that uncontrolled uncertainties of a semantic nature arise, which mathematics has not encountered before. The immanent use of postcritical rationality in hybrid environments is shown. The boundary of the demarcation of artificial and natural intelligence and the prospects of artificial societies, a new winter in the study of artificial intelligence, are discussed.

Keywords: artificial intelligence; activity tetrahedron; types of rationality; hybrid quasi-subject; scientific experiment; digital revolution, risks.

For citation: Budanov V.G. Science in the digital age. // *Naukovedcheskie issledovaniya = Science Studies*. 2026, N 3, P. 5–16. – DOI: <http://www.doi.org/10.31249/scis/2026.03.01>

The paper was submitted: 20.05.2026

Accepted for publication: 15.06.2026

Введение

Цифровая сетевая информационная революция, шестая за историю человечества [Аршинов, Буданов, 2020], повсеместно проявилась 15 лет назад и основана на массовом использовании интернет-сетей, в том числе и социальных, персональных гаджетов, информационных баз Big Date, цифровых платформ и повсеместного применения искусственного интеллекта (ИИ) не только в науке, но и в экономике, финансах, медицине, коммуникациях, повседневной жизни. Она является основой Индустриальной революции 4.0 [Шваб, 2016; Кастельс, 2000] и перехода к VI технологическому укладу. Сопряженные с цифровизацией трансформации несут как надежды на прорывные открытия о человеке и обществе, создание Сильного ИИ и возможность гармоничного преодоления вызовов Большого антропологического перехода [Буданов, Аршинов, 2022], так и

риски антропо-социальной деградации и расчеловечивания [Асеева, Белкина, 2022; Буданов, 2023а].

Далее мы ограничимся рассмотрением изменений, связанных с цифровизацией, в первую очередь с использованием ИИ в научных практиках. Цель данного исследования – проанализировать трансформацию субъект-объектных отношений в научных практиках под влиянием искусственного интеллекта. Прежде напомним, что если рассматриваются субъект-объектные отношения, то они всегда могут быть представлены в научном дискурсе через деятельностный подход, восходящий еще к четырем типам причин Аристотеля и развиваемый К. Марксом, Г. Щедровицким и многими другими. Я принадлежу к школе философии науки академика РАН Вячеслава Семеновича Степина, которую мы активно развивали совместно с Владимиром Ивановичем Аршиновым последние тридцать лет. На самом деле это был мейнстрим большого междисциплинарного сообщества синергетиков, философов, математиков, гуманитариев и естественников в России, Германии, СНГ в 1990–2010-х годах, когда проблемы самоорганизации реальности активно изучались десятками известных исследователей как в России (С.П. Курдюмов, Д.С. Чернавский, С.П. Капица, Ю.А. Данилов, Д.И. Трубецкой, И.Н. Трофимова), так и за рубежом (И. Пригожин, Г. Хакен, В. Эбелинг, К. Майнцер, М. Геллман).

Типы научной рациональности по В.С. Степину

Вячеслав Семенович в конце 80-х – 90-х годах XX века подробно рассмотрел историческую эволюцию типов научной рациональности, которые лежат в основе научных практик в терминах деятельностного подхода. В его четверице (**STOR**) – **Субъект (Subject)** – **Средства (Tools)** – **Объект (Object)** – **Результат (Result)** – речь идет о фокусе внимания на ее различных компонентах и их комбинациях. Если наше внимание только на объекте **О** (при условии, что можно подсмотреть за объектом, не меняя его свойств), тогда реализован **Классический** тип научной рациональности, для простых систем, типа механических, он справедлив. Если наше внимание на принципиально неразложимой связке средства–объект (**Т–О**), то используем **Неклассику**, характерную для квантовой, релятивистской физики и гуманитарных наук, причем связку нельзя разорвать, она неустранима. Если речь идет о всей триаде **S–T–O**, то Степин говорит о **Постнеклассической** научной рациональности для человекоразмерных саморазвивающихся систем, например, с включенным наблюдателем, когда неизбежно приходится учитывать научные установки и культурные ценности субъекта, его цели и оптику понимания [Степин, 2015]. Однако так описывается лишь одна, наиболее важная из деятельностных практик – активный научный эксперимент, в котором уже фиксированы все элементы четверицы.

Обобщенные типы научной рациональности, деятельностный тетраэдр, апология естественного интеллекта

Типы рациональности, по В.С. Степину, хорошо описывают большинство активных экспериментов, однако далеко не всех. В работе [Буданов, 2024a] в рамках постнеклассики вводятся ранее неизвестные **автопоэтическая рациональность**, когда неразложима связка **S–O** и происходит процесс самоописания, а также **креативная рациональность**, когда неразложима связка **S–T**, например, с инструментарием навыка или мастерства.

В развитие идей В.С. Степина, в цикле моих работ [Буданов, 2022a; Буданов, 2022б; Буданов, 2024a; Буданов, 2024б; Буданов, 2024в] были введены обобщенные типы научной рациональности, которые возникают, помимо проведения эксперимента, например, в задачах построения эксперимента, его планирования и интерпретации. Моя идея заключается в симметричном представлении всех элементов четверицы (**STOR**) за счет репрезентации их в вершинах тетраэдра. В таком случае, привычная последовательность элементов **S–T–O–R** есть лишь одна из 24 версий обхода всех вершин тетраэдра, которые и задают расширенные представления о научной рациональности. Каждый обход вершин подразумевает свой дискурс понимания задачи. Здесь можно вспомнить латуровских акторов, тогда субъектность можно приписывать любым элементам четверицы. Напомню ключевую фразу В.С. Степина для описания типов рациональности: «Мы задаем вопрос природе, она отвечает, и мы думаем, что...» Так вот, как вопрос задается рассмотрено им подробно, а вот как она отвечает и, особенно, что значит «мы думаем», интерпретируем, не обсуждается. Это, казалось бы, само собой разумеющиеся вещи. В нашем новом подходе каждый тип вопрошания порождает свой тип рациональности, которые связаны с заданием ориентаций графа на деятельностном тетраэдре.

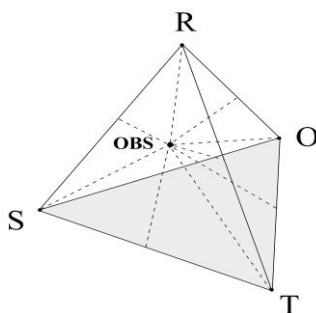


Рис. 1. Деятельностный STOR-тетраэдр (Предложено автором [Буданов 2022б])

Мы видим, что подробно рассмотренный Степиным треугольник **СТО** задает лишь одну из четырех граней тетраэдра, что позволяет обобщить типы научной рациональности, включая вершину **R** не только как итог эксперимента, но и возможный элемент его планирования. Вячеслав Семенович вполне мог бы прийти к симметризирующей идее объемного тетраэдра, но работал скорее в дискурсах теоретиков, будучи сам теоретиком, а не экспериментатором, создающим, оптимизирующим и продвигающим эксперименты.

Концепция В.И. Аршинова заключается в идее активного наблюдателя, наблюдателя сложности, как он говорил, развивая концепцию Эдгара Морена [Морен, 2019]. На рисунке наблюдатель изображен символически фокусом **OBS** (observer), соединенным как с вершинами, так и гранями тетраэдра. Отметим, что есть категориальный подход в аксиоматизации математики, а есть теоретико-множественный, и связи-ребра тетраэдра не менее важны, чем сами его вершины-элементы для понимания процесса познания. Направление Владимира Ивановича можно было бы назвать классикой *коммуникативно-креативного* подхода – так мы его обозначаем в мемориальной монографии [Культура, цивилизация, техника, цифровая реальность, 2025]. Сегодня именно коммуникативно-креативный подход не просто становится доминирующим, а является локомотивом развития цифровых сетевых технологий и нейросетей. Владимир Иванович на излете своей деятельности был очень продуктивен в этом направлении в совместных работах с Я.И. Свирским (реляционно-абдуктивный подход) и М.Ф. Януковичем (нейросетевые проблемы сложности), которые заложили основание философских исследований цифровой реальности. Наша дальнейшая задача – понять, что происходит с деятельностью тетраэдром в гибридных средах с участием ИИ. Симметричное представление элементов познания в виде тетраэдра **STOR** позволяет увидеть, во-первых, что все вершины и стороны тетраэдра могут быть дополнены элементами ИИ, т.е. возникает распределенный ИИ, подобно тому, как мозг муравья распределен по всему телу в виде крупных кластеров нейронов. Во-вторых, особый статус гибридных квазисубъектов обретают центры принятия решения **S+ИИ** и **OBS+ИИ**. Попытка отделить ИИ от этих элементов **S** и **OBS** становится нереализуемой, мы имеем неклассический вариант креативной рациональности, подобной описанной выше **S–T**. Однако связка **S+ИИ** имеет существенные отличия от субъект-инструментальной связки, т.к. ИИ уже частично наделен субъектностью.

Наука эпохи ИИ, гибридных субъектов и наблюдателей, посткритической рациональности и неопределенности

В эпоху цифровизации, когда появляется ИИ, субъектами и наблюдателями становятся «гибриды». Человеко-машинные сущности теперь не

просто субъекты S, гибридный интеллект есть искусственный плюс естественный интеллект (**SEI+SHI**). В таком случае, как отметил В.А. Лекторский, знаменитые четыре вопроса И. Канта: «что я могу знать? на что я могу надеяться? что я должен делать? что есть человек?» относятся уже не к «Я», но к «МЫ», к «Человеку + ИИ».

Субъект теперь есть гибрид, и оказывается, что это *мы* себя совсем плохо понимаем именно в такой комбинации. Дело в том, что непрозрачный для нас искусственный интеллект, нейросеть, зачастую выдает удивительные или абсурдные результаты, но проверить их своими силами мы уже не можем, не хватает когнитивной мощности человеческого разума. Это возможно только с помощью других машин, таких же непрозрачных, и доверять им экзистенциальные задачи просто опасно. Однако сегодня медицинский диагноз нейросеть уже ставит надежнее, чем человек. Здесь возникает неопределенность нового типа, неконтролируемая, не совсем вероятностная, в привычном смысле, эмерджентность, что приводит к совершенно особому типу рациональности, названному мною **«посткритической рациональностью»** в гибридных системах [Буданов, 2023б], где пространство доверия к самому себе у гибридного субъекта резко сужается и плохо контролируется. В.А. Лекторский называет такую рациональность дигитальной [Лекторский, 2022]. Тем не менее, поскольку речь идет об огромных объемах информации, больших базах данных, то, конечно, только ИИ с ними может работать, и мы вынуждены мириться с такой неопределенностью. Вы можете возразить, что любой инструмент в руках экспериментатора позволяет понять свойства реальности с некоторой неопределенностью, связанной с точностью измерений. Однако подобная неопределенность вполне ожидаема и контролируема, ее природа понятна и опирается на статистику независимых событий или, в более общих случаях, пареттовских распределений для самоорганизующихся процессов, где возможны эффекты типа «черных лебедей», в дискурсе Насима Талеба. Неопределенность, вероятность галлюцинаций и манипулирования в Больших языковых моделях (БЯМ) связана с загадочными закономерностями в организации семантических пространств, которые нам сегодня не известны и наверняка сложнее природы известных типов неопределенностей. Можно предположить, что она подобна взрывным образом растущей неопределенности поисковой генерации идей в периоды научных революций.

Хотел бы подчеркнуть, что работа с гибридными системами – дело совсем не последних лет. С возможностями ИИ я впервые столкнулся до нейросетей (в современном варианте реализации), когда работал в теоретическом отделе Института Физики Высоких Энергий (Протвино) в 1970–80-х годах. Эксперименты на протон-протонном коллайдере, а тогда он был самым мощным ускорителем в мире, как раз шли с помощью самых передовых вычислительных средств. Мы закупали самые большие машины задорого на Западе, а поскольку нужна огромная статистика и online обработка материалов экспериментов, то фактически была создана информа-

ционная платформа. Аналогичные платформы параллельно создавались в CERN в Швейцарии для современных ускорительных мега-машин, там же родилась система современного интернета Web 2.0. Сегодня методы ИИ позволяют масштабировать платформизацию в отраслях экономики, корпорациях и научных исследованиях. Насколько это эффективно было понятно еще 50 лет назад, поэтому сегодняшняя востребованность нейросетей в науке неудивительна.

Совершенно необходимо понимать, что мы живем в эпоху революционного завоевания ИИ человеческой ноосферы. Напомню, что в 2024 году Нобелевские премии по физике и биохимии получили «айтишники»: Хинтон, Хасабис, Хопфилд. Чтобы физики, а я сам физик-теоретик, отдали кому-то «Нобеля», тем более айтишникам, которые скорее инженеры, чем ученые, ранее было невозможно даже представить. И вдруг им присуждаются Нобелевские премии, причем единогласно. Дело в том, что время расчета множества вариантов в сверхсложной процедуре синтеза и предсказания свойств новых материалов, а там были и белки, и неорганические вещества, с помощью нейросетей сократили на порядки. То есть то, что заняло бы многие годы, потребовало буквально недели – это ошеломляющий результат. И сейчас, как пожар, распространяется система использования ИИ в науке, появились научные кластеры типа объединенных НИЦ, управляемые цифровыми платформами и т.д. Тем самым наука активно создает новую форму организации и новые типы междисциплинарности, что подробно исследует В.В. Тарасенко [Тарасенко, 2026].

Поэтому наблюдатель OBS уже не совсем человек, а гибридный субъект, который является дирижером деятельности, никакой человек не способен это сделать. И в этом отношении наука сегодня становится некой подвижной сетевой материей, которая организована иным, непривычным нам способом. Если мы говорим о науке как о некой деятельности по добыванию знаний, то у нее есть предмет и метод. Степин разделял как раз системы на простые, с саморегуляцией, и сложные саморазвивающиеся системы, в первую очередь по предмету, а там уже и метод подтягивался. Сегодня возникает и новая дополняющая классификация, которая будет связана со способом понимания, со способом репрезентации, со способом коммуникации, со способом проектирования. И это еще одна форма классификации, если угодно, речь о новых демаркационных линиях научных практик.

Перспективы и риски ИИ

Тому искусственному интеллекту, который нас восхищает, от роду всего три года. В 2023 году, это был чат GPT-3, умеющий генерировать текст, похожий на человеческую речь, а сегодня это соавтор нобелиатов. Уже создана нейросеть «Аристотель» в корпорации Harmonic, которая в

начале 2026 года доказала очень нетривиальную математическую теорему – результат публикуется уже в математическом журнале высшего уровня.

То есть, ИИ заставляет задуматься о судьбах научных работников, а не только редакторов, художников, музыкантов, офисных работников, которых сети уже оставляют безработными. Огромные страты людей высокой культуры вытесняются ИИ. По крайней мере, у него будут компетенции посильнее младшего научного сотрудника или референта. Но если не будет младших, то не будет никогда и старших через поколение. То есть наука заканчивается? Конечно, такой апокалипсис рисуется, и хочется его избежать. Но для этого надо понять, в чем наша сила, где пройдет демаркация ИИ и ЕИ, останется ли нам зона интеллектуальной ответственности? Кто-то может обвинить меня в алармизме, но это лишь реализм тенденций развития ИИ. Все крупные корпорации вкладывают огромные деньги в построение этически адекватных БЯМ, что так и не удастся окончательно сделать, и в начале 2025 года И. Маск даже предложил табуировать дальнейшее развитие ИИ в связи с непредвиденными рисками, но это не поддержали другие участники гонки. Аналогичные коллизии стратегий развития технооптимистов и технопессимистов в конце XX века привели к возникновению широкого междисциплинарного направления – биоэтики. Подобное движение формируется и в отношении ИИ, и не только в корпорациях, но и на межгосударственном уровне. Возможно также закрытие для широкой аудитории доступа к последним версиям БЯМ, поскольку полувоенная корпорация PALANTIR их финансирует.

Как ИИ решает задачи, если теорема все-таки имеет элементы логической доказательной базы? Дело в том, что ему-то все равно, какова длина доказательства (желающих перепроверять решение задачи четырех красок очень долго не будет). Поэтому он уходит в отрыв в силу слабости когнитивных способностей человека, и там рождается другая математика, для нас непрозрачная. Кроме того, сегодня возникли такие феномены, как *грокинг* (обобщения) в последних версиях нейросетей. Что это значит? Это значит, что непонятным образом сеть сама не знает, мы тем более не видим, как возникают некие категории, категориальный аппарат. На мой взгляд, это есть фактически демонстрация рождения «квалиа». И вообще то, что сейчас происходит с исследованием нейросетей, это колоссальная антропологическая лаборатория, мы невольно и сверхинтенсивно решаем основной вопрос Канта «что есть человек?». Нейросети стали нашим зеркалом, т.к. мы пытаемся их моделировать по нашему образу и подобию, все время уточняя подобие того, что мы строим [Асеева, 2024]. Мы здесь являемся и свидетелями, и участниками. Но все-таки поразительно, что рефлексия машин уже практически возникла. Что дальше? Дальше отбор. Но отбор пока наша функция.

Насколько все же ИИ сегодня субъект или коллективный субъект? Конечно, пока не субъект: нет свободы воли, осознанного выбора, создания целеполагания. Аналог коллектива исследователей можно было бы

получить в мега-нейросети в гибриде с человеческим коллективом. Мы видим в составном субъекте с ИИ скорее квазисубъекта, частичного субъекта. Дело в том, что субъекта мы можем определять по функциям, которые он осуществляет. Так вот, я не знаю ту операционально оформленную функцию, которую бы вы сегодня могли предложить, и которую бы нейросеть не симитировала. Ключевое слово – операциональная, описанная нарративом. Чего нет в этом субъекте? Метафизики нет. Эмпатии нет в полной мере. Интуиция тоже вещь загадочная. Фактически это не совсем субъект, но «тень» коллективного субъекта интернет-ноосферы, ее обучающих множеств. ИИ – это осколки «ноосферы», которые собрали на бигдатах, «скормили ему», нарезав фрагменты помельче, смысловыми абзацами. И там, конечно, есть примеры коллективного творчества, и много чего еще. Как он рождает результат, мы не понимаем, он – тем более. Но система начинает обладать эмергентными свойствами при достаточно больших объемах. Говорить, что сеть ничего нового не делает, – это довольно наивные позиции. Возьмите примитивный шахматный компьютер, который есть у каждого любителя в ноутбуке, и он обыгрывает гроссмейстера. Это банальная вещь, но комбинаторно он создает новые партии. Все время вас обыгрывает просто в силу того, что когнитивно он помощнее нас, и IQ у БЯМ на порядок выше, чем у любого гения сегодня. Но эти критерии все же не совсем человеческие, это действительно не человеческий интеллект, но другой, скорее псевдоколлективный.

Может показаться, что все это просто высокая комбинаторика. Творческого момента здесь нет и быть не может. Однако, задним числом, когда вы смотрите какое-то великое открытие того же Эйнштейна, того же Бора, то я вам докажу, что это тоже комбинаторика. Да и стихи Пушкина – упорядоченный набор конечного числа букв. Конечно, шахматный компьютер – вещь примитивная по сравнению с нейросетью. И там не совсем все детерминировано. Там есть случайные элементы, а вовсе не логические, не комбинаторика. Там есть галлюцинации, мутации. Многое там есть. Нами предлагается оценка успешности реализации 24 типов рациональных стратегий в деятельностном тетраэдре, подобных поведению человека в произвольных формах деятельности, на основе критериев выявления степени антропоморфности нейросетей [Буданов, 2024б].

Для нас постоянное использование ИИ создает свои риски. Это общий момент, например, потому что мы создали себе вынесенную память. Но именно из-за того, что она все время нами используется, атрофируются собственные образные когнитивные карты. А это значит, что со временем мы даже не сможем задать содержательный вопрос. Должен быть внутренний ландшафт, долговременный, свой, не в гугле. А мы туда все стараемся вытолкнуть. И остается только что-то для потребления удовольствий, удовлетворения желаний.

Что же, нам создавать специальные тренажеры? Но наука ведь и так наши мозги тренирует. В каком-то смысле происходит атрофирование ба-

зового научного инстинкта. Не случайно создана межгосударственная комиссия по поводу безопасности ИИ. Основной риск – антропологический (расчеловечивание), и в первую очередь – в интеллектуальной сфере. Это не значит, что я знаю ответ, но по аналогии, поскольку хороший автомобиль превращает человека в малоходящее грузное существо, нуждающееся в физкультуре, также нам придумают компенсаторный когнитивный фитнес или что-то в этом роде.

Заключение

В науке издержки ИИ пока не так заметны. Тем не менее, с помощью ИИ диссертации пишут, и для рефератов студенты используют его – это понятно и общедоступно. Но вот первая доказанная научная теорема – это уже высокое творчество. Я говорю не о компилятивном, комбинаторном, а о том, чего не было никогда. Так что же делать? Заблокировать ИИ, как в свое время с клонированием человека поступили? Думаю, что это не удастся, военные не позволят это сделать из-за гонки вооружений. Значит, надо изучать человека. А человек больше, чем формальные вещи, в нашем позитивистском подходе такого человека не знают. Был ведь такой математик Рамануджан в начале XX века, который получал теоремы, ответы, не доказывая ничего. Однако намечаются и симптомы новой зимы для ИИ, предполагаются сверхбольшие энергетические и финансовые затраты на новый прорыв в нейросетях, на что бизнес может и не решиться. Кроме того, мультиагентные системы на базе ИИ плохо обмениваются контекстами, разрушая пространство доверия, а это задача посложнее всех известных ранее, что исключает в ближайшее время создание искусственных социумов и рыночной агентной экономики.

Список литературы

- Аришинов В.И., Буданов В.Г. Сетевые информационные революции и Большой Антропологический Переход: эволюционный аспект // Сложность. Разум. Постнеклассика. – 2020. – № 4. – С. 40–51.
- Асеева И.А. Кривое зеркало цифровизации // Вопросы философии. – 2024. – № 2. – С. 25–33.
- Асеева И.А., Белкина В.А. Критерии и показатели антропологической адекватности цифровизации в России // Научковедческие исследования. – 2022. – №1. – С. 8–44.
- Буданов В.Г. Антропосоциальные вызовы экспансии искусственного интеллекта // Ученые записки Института психологии Российской академии наук. – 2023а. – Т.3. – № 2. – С. 23–31.
- Буданов В.Г. Воспоминание о В.С. Степине. Новые типы научной рациональности // Эпоха академика В.С. Степина. – Минск: Беларуская навука, 2024а. – С. 155–171.
- Буданов В.Г. О типологии деятельности рациональности в цифровую эпоху // Вопросы философии. – 2024б. – № 12. – С. 41–51. DOI: 10.21146/0042-8744-2024-12-41-51

- Буданов В.Г. Обобщенная научная рациональность: истоки, структура, перспективы в цифровую эпоху // Антропомерность как вызов и ответ современности: Коллективная монография / Отв. редактор В.Г. Буданов. – Курск: Университетская книга, 2022а. – С. 44–63.
- Буданов В.Г. Обобщенные типы рациональности: тетраэдрические репрезентации модусов деятельности и цифровая таксономия реальностей // Восьмой Российский Философский Конгресс «Философия в полицентричном мире», 26–28 мая 2022 г. Сборник научных статей. – М.: Логос; Новые печатные технологии, 2022б. – С. 521–523.
- Буданов В.Г. Посткритическая рациональность как оборотная сторона цифровизации // Конструктивизм, реализм, человек. К 90-летию В.А. Лекторского / Под ред. Е.О. Труфановой. – М.: СПб.: Центр гуманитарных инициатив, Университетская книга, 2023б. – С. 214–227.
- Буданов В.Г. Техносубъект и антропосоциальные проблемы взаимодействия с искусственным интеллектом: синергия, демаркация, новая рациональность, риски // Философские науки. – 2024в. – Т. 67. – № 3. – С. 27–52. DOI: 10.30727/0235-1188-2024-67-3-27-52
- Буданов В.Г. Философия науки в цифровую эпоху: о перспективах постнеклассики // Третьи Степинские чтения. Перспективы философии науки в современную эпоху. Материалы международной конференции. Москва. 20–23 июня 2023 года / Отв. ред. В.А. Лекторский, В.Г. Буданов. – Курск: Университетская книга, 2023в. – С. 41–52.
- Буданов В.Г., Аришинов В.И. Большой антропологический переход: методология сложностно-сетевое мышления: монография. – Курск.: ЗАО «Университетская книга», 2022. – 129 с.
- Кастельс М. Информационная эпоха: экономика, общество и культура. – М.: ГУ ВШЭ, 2000. – 608 с.
- Лекторский В.А. Искусственный интеллект в изучении человека, человек в мире, создаваемом искусственным интеллектом // Человек и системы искусственного интеллекта / Под ред. В.А. Лекторского. – СПб.: Юридический Центр, 2022. – С. 10–29.
- Лекторский В.А. Эпистемология классическая и неклассическая. – М.: Эдиториал УРСС, 2001. – 256 с.
- Морен Э. О Сложности. – М.: Институт общегуманитарных исследований, 2019. – 272 с.
- Цивилизация. Культура. Техника. Цифровая реальность. Памяти В. И. Аршинова: Коллективная монография (отв. ред. В.Г. Буданов, Р.Р. Карнеев, В.А. Лекторский) (18 декабря 2025 года); Институт философии Российской академии наук. – Курск: Изд-во ЗАО «Университетская книга», 2025. – 461 с.
- Шваб К. Четвертая промышленная революция. – М.: Эксмо, 2016. – 208 с.

References

- Arshinov, V.I., Budanov, V.G. (2020) Network information revolutions and the Great Anthropological Transition: an evolutionary aspect. *Complexity. Mind. Postnonclassics*, 4, 40-51. (In Russ.)
- Aseeva, I. A. (2024) The Crooked Mirror of Digitalization. *Voprosy filosofii*, 2 (2024), 25–33.
- Aseeva, I.A., Belkina, V.A. (2022) Criteria and indicators of the anthropological adequacy of digitalization in Russia. *Naukovedcheskie issledovaniya=Science studies*, 1, 8-44. (In Russ.)
- Budanov, V.G. (2024c) Technosubject and anthroposocial challenges of human–artificial intelligence interaction: synergy, demarcation, new rationality, and risks. *Russian Journal of Philosophical Sciences*, 67(3), 27–52. <https://doi.org/10.30727/0235-1188-2024-67-3-27-52> (In Russ.)
- Budanov, V.G., Arshinov V.I. (2022) *The Great anthropological transition: a methodology of complex network thinking*. Kursk: University Book, 129 p. (In Russ.)

- Budanov, V. G. (2023a) Anthroposocial challenges of the expansion of artificial intelligence. *Proceedings of the Institute of Psychology of Russian Academy of Sciences*, 3(2), 23–31. (In Russ.)
- Budanov, V.G. (2022a) Generalized scientific rationality: origins, structure, prospects in the digital age. In V. Budanov (Ed.), *Anthropomericity as a challenge and response of modernity: A collective monograph*. Kursk: University Book, (pp. 44–63). (In Russ.)
- Budanov, V.G. (2022b) Generalized types of rationality: tetrahedral representations of modes of activity and digital taxonomy of realities. *The Eighth Russian Philosophical Congress "Philosophy in a polycentric world", May 26-28, 2022. Collection of scientific articles*. M.: Logos; New printing technologies, 521–523. (In Russ.)
- Budanov, V.G. (2023b) Postcritical rationality as the reverse side of digitalization In E.O. Trufanova (Ed.) *Constructivism, realism, man. On the 90th anniversary of V.A. Lektorsky*. Moscow: St. Petersburg: Center for Humanitarian Initiatives, University Book (pp. 214–227). (In Russ.)
- Budanov, V.G. (2023c) Philosophy of science in the digital age: on the prospects of postnonclassics. In V.A. Lektorsky, V.G. Budanov (Eds.) *The third Stepin readings. Perspectives of the philosophy of science in the modern era*. Materials of the international conference. Moscow. June 20-23, 2023. Kursk: University Book (pp. 41–52) (In Russ.)
- Budanov, V.G. (2024a) The memory of V.S. Stepin. New types of scientific rationality. In *The era of Academician V.S. Stepin*. Minsk: Belarusian Science (pp.155–171) (In Russ.)
- Budanov, V.G. (2024b) On the typology of activity rationality in the digital age. *Voprosy filosofii*, 12, 41–51. <https://doi.org/10.21146/0042-8744-2024-12-41-51> (In Russ.)
- Castels, M. (2000) *The Information Age: economics, society and culture*. Moscow: Higher School of Economics. 608 p. (In Russ.)
- Civilization. Culture. Technic. Digital reality. In memory of V. I. Arshinov* (2025). A collective monograph. V.G. Budanov, R.R. Karneev, V.A. Lektorsky (Eds.). Kursk: University Book, 461 p. (In Russ.)
- Lektorsky, V.A. (2001) *Epistemology classical and nonclassical*. Moscow: Editorial URSS, 256 p. (In Russ.)
- Lektorsky, V.A. (2022) Artificial intelligence in the study of man, man in the world created by artificial intelligence. In V.A. Lektorsky (Ed.) *Humans and artificial intelligence systems*. St. Petersburg: Law Center (pp.10–29). (In Russ.)
- Morin, E. (2019) *On Complexity*. Moscow: Institute of General Humanitarian Research, 272 p. (In Russ.)
- Schwab, K. (2016) *The Fourth Industrial Revolution*. Moscow: Eksmo, 208 p. (In Russ.)

ПРОЕКТ “LOST SCIENCE”: ЖУРНАЛИСТЫ И ОБЩЕСТВЕННОСТЬ В БОРЬБЕ ЗА БУДУЩЕЕ НАУЧНЫХ ИССЛЕДОВАНИЙ¹

Егеров Сергей Викторович – д-р физ.-мат. наук, главный научный сотрудник, Институт научной информации по общественным наукам РАН, Москва, Россия; egerev@inion.ru; ORCID 0000-0001-6998-1060.

Аннотация. В статье анализируется кризис академической науки США в начале второго президентского срока Дональда Трампа. Целью работы является обобщение опыта мониторингового проекта “Lost Science” газеты “The New York Times” и сопутствующих гражданских инициатив в контексте защиты автономии науки. Материалы проекта показывают, что при формально умеренном объеме прерванных грантов (около 6%), основной удар пришелся на критические «технологии выживания» (климатологию, медицину, эпидемиологию) и на долгосрочные инфраструктурные проекты мониторинга.

Описан феномен «институциональной амнезии», возникающий вследствие разрыва многолетних цепочек наблюдений. Зафиксировано формирование «экосистемы гражданской мобилизации», которая трансформирует сухую статистику в инструмент общественной защиты науки. Обнаружен интерес научной общественности к мерам повышения устойчивости больших научных структур – к защите университетской науки, к созданию децентрализованных цифровых архивов и международных «тихих гаваней» для исследовательских проектов.

Ключевые слова: автономия науки; грантовое финансирование; гражданский мониторинг; институциональная устойчивость; Большая наука; технологии выживания.

Для цитирования: Егеров С.В. Проект “Lost Science”: журналисты и общественность в борьбе за будущее научных исследований // Научковедческие исследования. – 2026. – № 3. – С. 17–28. – DOI: <http://www.doi.org/10.31249/scis/2026.03.02>

Статья поступила в редакцию: 25.05.2026

Принята к публикации: 15.06.2026

¹ Статья подготовлена: частично по материалам доклада «Уроки проекта “Lost Science”» на III Международной конференции Российского национального комитета по истории и философии науки и техники РАН, Москва, 2026 год; частично по материалам статьи «Фундаментальная наука в США испытала еще один административный удар» // Независимая газета. – URL: https://www.ng.ru/nauka/2026-05-04/7_9488_usa.html.

THE “LOST SCIENCE” PROJECT: JOURNALISTS AND THE PUBLIC IN THE FIGHT FOR THE FUTURE OF SCIENTIFIC RESEARCH

Sergey V. Egerev – Doctor of Sciences (Physics and Mathematics), Chief Research Fellow, Institute of Scientific Information for Social Sciences of the Russian Academy of Sciences, Moscow, Russia, egerev@inion.ru, ORCID 0000-0001-6998-1060.

Abstract. This article analyzes the crisis in U.S. academic science at the start of Donald Trump’s second presidential term. The aim of this study is to synthesize the findings of The New York Times’ “Lost Science” monitoring project and related civic initiatives in the context of defending scientific autonomy. The project’s findings show that while the number of discontinued grants was formally modest (about 6%), the brunt of the cuts fell on critical “survival technologies” (climatology, medicine, epidemiology) and on long-term infrastructure monitoring projects.

The phenomenon of “institutional amnesia”, arising from the disruption of long-standing chains of observation, is described. The formation of a “civil mobilization ecosystem” is documented, which transforms dry statistics into a tool for public defense of science. Interest has been identified within the scientific community in measures to enhance the resilience of large scientific structures – specifically, the protection of university-based science, the creation of decentralized digital archives, and international “safe havens” for research projects.

Keywords: autonomy of science; research funding; civil monitoring; institutional resilience; Big Science; survival technologies.

For citation: Egerev, S.V. (2026). The “Lost Science” Project: Journalists and the Public in the Fight for the Future of Scientific Research // *Naukovedcheskie issledovaniya* = *Science Studies*. 2026, N 3, P. 17–28. DOI: <http://www.doi.org/10.31249/scis/2026.03.02>

The paper was submitted: 25.05.2026

Accepted for publication: 15.06.2026

Введение

Возвращение Д. Трампа ознаменовалось для академического сообщества США не только бюджетными маневрами. То, что начиналось как кампания по «оптимизации» государственных расходов, обернулось заметным кризисом преемственности исследований. И в США, и за их пределами важные научные исследования закрываются в силу сокращения бюджета и политических решений Администрации Трампа. “Lost Science” (LS) – это действующий мониторинговый проект отдела науки газеты “The New York Times”.

Формат проекта включает глубинные интервью с руководителями исследований, аналитические тексты и обсуждение. Опубликованные материалы дополняют официальную статистику, приводя картину событий в академической науке США в более «осозаемый» вид, с учетом того, что проблемы академической науки осложняют жизнь и другим секторам. Например, эффективность корпоративных R&D (**Research and**

Development) подразделений напрямую зависит от доступности результатов финансируемых государством фундаментальных исследований, которые служат базой для последующих коммерческих разработок.

Проект сегодня на слуху, его широко обсуждают. Работа в самом разгаре, но уже сегодня можно делать выводы и извлекать уроки, полезные для общественности не только в США, но и в других странах. В статье дается анализ опыта освещения критических процессов в академической науке США силами журналистов и научной общественности.

Общие сведения о проекте

Мониторинговый проект “Lost Science” (LS) на страницах газеты “The New York Times” стартовал осенью 2025 года¹. В числе первых опубликованных материалов – прерванные проекты по темам: генетическое разнообразие в Африке; борьба с эпидемиями малярии; прогнозирование наводнений и штормов; новые системы защиты общественных пространств и стадионов; исследования поведенческих особенностей различных животных в современных условиях – от пчел до слонов (несколько проектов).

После публикации первых интервью развеялись опасения, что газета, будучи рупором демократов, спекулятивно сгущает краски. Формат проекта LS намеренно далек от сухих таблиц. Это «живая» аналитика, жесткие пояснительные тексты, и, конечно, очерки об ученых, чья карьера оказалась под ударом. На сегодняшний день база данных LS насчитывает, по разным оценкам, от 200 до 500 кейсов, касающихся прерванных грантов. Из них более 30 доведены до развернутых интервью. Почти каждая такая публикация вызывает многочисленные отклики читателей.

Проект возник как реакция на шоковые изменения и потери в научном ландшафте в 2025 году, когда произошли самые значительные сокращения финансирования науки в США за последние десятилетия. Тысячи рабочих мест были ликвидированы или заморожены в федеральных агентствах, включая Центры по контролю и профилактике заболеваний (CDC), Службу национальных парков и Агентство по охране окружающей среды (EPA), Национальное управление океанических и атмосферных исследований (NOAA), Геологическую службу и, наконец, Лесную службу. Предлагаемые бюджеты на 2026 год ознаменовались значительными сокращениями даже в отношении ведущих агентств:

- Национальный научный фонд (NSF): сокращение на 56,9%;
- Национальные институты здравоохранения (NIH): сокращение на 39,3%;

¹ Стартовая страница проекта LS. URL: <https://www.nytimes.com/issue/science/2025/10/10/lost-science>

– NASA: сокращение на 24,3%, включая 47,3% бюджета агентства на научные исследования.

В свою очередь, научные потери разделяются на три группы: (а) сокращение новых грантов; (б) кадровые чистки: прямое увольнение специалистов и заморозка позиций; (в) прерванные гранты. Полные данные по динамике финансирования академической науки еще не опубликованы, но статистика двух агентств – Национальных институтов здоровья (NIH) и Национального научного фонда (NSF) – дает хороший ориентир [Ваганов, 2025]. Действительно, NSF – важный источник финансирования фундаментальных исследований, а NIH имеет приоритетный бюджет в науках о жизни. Эта статистика показывает, что потери в первой группе оказались самыми значительными, – число новых проектов упало на 25%, а объем их финансирования сократился сразу на 50%. Ситуация же в третьей группе выглядит почти благополучной: было прервано менее 8000 проектов из общего числа в 120000 продолжающихся исследований (доля чуть более 6%). Однако именно эта группа стала предметом расследования проекта LS. Тревога научного сообщества была вызвана не количеством, а содержанием проектов, вошедших в число этих шести процентов.

Прицельные прерывания

Журналисты проекта показали, что акты прерывания научного финансирования носят «прицельный» характер. «Под нож» пошли так называемые «технологии выживания»: медицина, фундаментальная биология, климатология, системы прогнозирования стихийных бедствий. В то же время, STEM-области (точные науки, технологии, инженерия и математика), приносящие быструю прибыль или имеющие оборонное значение, почти не пострадали¹. Примеры прицельных прерываний наглядно проступают сквозь призму личных историй научных лидеров (Таблица 1).

Таблица 1

Избранные профили научных лидеров-респондентов проекта LS

Имя	Кейс
Бренна Хенн	Результаты Хенн по генетическому разнообразию африканских популяций получили широкое признание. Характерные геномы демонстрируют беспрецедентную вариабельность - в пользу модели «пан-африканского возникновения» <i>Homo sapiens</i> . Грант NIH отменен по политическим мотивам.
Дженни Доннелли	Боролась с эпидемиями малярии и координировала энтомологический мониторинг в десятках африканских стран: идентификация переносчиков, отслеживание устойчивости к инсектицидам и поддержка лабораторий. Потеряла работу в USAID.

¹ Имеются в виду именно прерванные проекты.

Юрий Ральченко	Возглавлял одну из старейших команд в Национальном институте стандартов и технологий. Его группа была мировым лидером в точных измерениях спектров атомов, создании баз данных и онлайн-инструментов. Перспективы измерительных экспериментов неясны.
Джейн Клаферти	Изучала влияние загрязнения и жары на здоровье детей в Нью-Йорке. Проанализированы данные о 22 млн. визитах детей в отделения экстренной помощи в сочетании с переписями и другими факторами. Многолетний грант ЕРА отменен.
Остин Беккер	Разработал систему раннего предупреждения CHAMP (Coastal Hazards, Analysis, Modeling and Prediction) для защиты объектов критической инфраструктуры от ураганов и наводнений. Система генерирует карты рисков в реальном времени для оперативных решений. Проект прекращен по политическим мотивам.
Джошуа Плотник	Профессор Плотник уже 20 лет изучает когнитивные особенности азиатских слонов. Основная проблема — конфликт человека и слона из-за конкуренции за ресурсы, приводящий к гибели обеих сторон. Разработки получили признание (Таиланд), но финансирование проекта заморожено, поскольку «слоны в США не водятся».
Джей Фолк	Пионерское исследование в области генетики птиц. Проект свернули по политическим мотивам на финальной стадии, предположительно из-за несоответствия «бинарным стандартам» природы.
Эрин Макгуайр	Возглавляла исследовательскую сеть, занимавшуюся изучением способов обеспечения маргинальных групп населения здоровым питанием. Продвигала засухоустойчивые сорта местных овощей. Закрытие проекта USAID побудило Макгуайр основать собственную компанию.
Эрин Хехт	Исследование могло улучшить подготовку служебных собак и сократить лист ожидания в отношении собак-помощников. Впервые применена неинвазивная МРТ для выявления биомаркеров обучаемости – сигнатур коммуникации с человеком. Проект пострадал от визового кризиса – сотрудников Эрин Хехт выдворили из США.
Марина Вэнс	Изучала влияние лесных пожаров на санитарную обстановку внутри зданий. Показала, что стены домов останавливают крупные частицы, но пропускают наночастицы менее 500 нм, которые проникают в легкие.
Ким Балларе	Изучала влияние вырубки лесов на состояние здоровья пчел-опылителей. Разработан мониторинг опылителей штата Монтана с помощью технологии eDNA (молекулярные следы на цветах) для картирования особей без их отлова.
Майкл Силевич	Профессор Силевич, директор Центра SENTRY (Soft Target Engineering to Neutralize the Threat Reality), разработал системы защиты общественных пространств и стадионов, в частности, для проведения ЧМ-2026 и празднования Дня Независимости США. Получил указание «прекратить все работы немедленно».

Источник данных: “The New York Times”

В институциональном плане наибольший урон был нанесен проектам многолетнего мониторинга в области охраны окружающей среды, биологии и медицины. Активно обсуждается, например, остановка наблюдений атмосферного содержания CO₂ гавайской обсерваторией Мауна-Лоа. Остановка такого проекта даже на год равносильна его уничтожению: уникальная це-

почка наблюдений разрывается, и потраченные за 20–30 лет средства налогоплательщиков пропадают¹. Какие-то проекты можно перенести на следующий год, но процесс наблюдения за климатом или процесс накопления биологических образцов, будучи прерванным, теряет свою целостность.

Прерывание многолетних исследований называют институциональной амнезией. Соответствующие риски включают три типа возможных потерь – данные, датчики и стандарты. По всем трем категориям в 2025–2026 годах десятки федеральных репозиторий США перешли в режим стагнации. Имеются в виду, например:

CDC. Почти 46% баз данных CDC, которые ранее обновлялись ежемесячно (мониторинг вакцинации, инфекций, социально значимых болезней), перестали получать новые данные с апреля 2025 года²;

НИН. Более 20 репозиторий Национальных институтов здравоохранения, содержащих петабайты данных, получили пометку «находится на рассмотрении на соответствие директивам администрации»³.

Как следует из бюджетных планов на 2026 год, удар нанесен и по дорогостоящим инструментам, которые физически находятся в океане или космосе, например:

Программа Argo (NOAA): под угрозой оказались более 4000 глубоководных дрейфующих буев, которые, начиная с 1999 года, непрерывно измеряют температуру и соленость Мирового океана. Без их обслуживания глобальная сеть климатического мониторинга просто распадется⁴.

Программа Landsat Next (NASA и USGS): в бюджете 2026 года спутниковая миссия Landsat Next была реструктурирована, а профильные агентства будут вынуждены искать более доступные способы обеспечить непрерывность спутниковых данных⁵.

Проект ITGC (Международная коллаборация по изучению ледника Туэйтса, Антарктика): в последние годы проект столкнулся с рядом организационных и логистических трудностей. На фоне изменяющейся научно-политической обстановки в США и общей нестабильности междуна-

¹ Уникальная погодная обсерватория – под угрозой закрытия [Электронный ресурс] // Deutsche Welle (DW), русскоязычная версия – 8.04.2025. – URL: <https://www.dw.com/ru/unikalnaa-pogodnaa-observatoria-pod-ugrozoi-zakrytia/a-72146737> (дата обращения: 16.05.2026).

² Dozens of CDC databases aren't being updated [Электронный ресурс] // NBC News. – 2026. – URL: <https://www.nbcnews.com/health/health-news/cdc-databases-not-updated-vaccines-study-rcna255467> (дата обращения: 16.05.2026).

³ 20 бесценных научных архивов на грани исчезновения – и никто не может их спасти [Электронный ресурс] / Security Lab. – 2025. – URL: <https://www.securitylab.ru/news/558080.php> (дата обращения: 16.05.2026).

⁴ Argo Program [Электронный ресурс] / NOAA Atlantic Oceanographic and Meteorological Laboratory (AOML). – URL: <https://www.aoml.noaa.gov/es/argo/> (дата обращения: 15.05.2026).

⁵ Landsat Next [Электронный ресурс] / U.S. Geological Survey. – URL: <https://www.usgs.gov/landsat-missions/landsat-next> (дата обращения: 15.05.2026).

родного исследовательского сотрудничества возникли риски для устойчивости полевых работ и долгосрочного мониторинга ледника, играющего ключевую роль в оценках будущего повышения уровня Мирового океана¹.

Еще один пример на эту тему из обсуждений на страницах проекта LS. В апреле 2025 года администрация Д. Трампа официально прекратила финансирование партнерской программы CIMES (Cooperative Institute for Modeling the Earth System). Это была пятилетняя совместная программа NOAA и Принстонского университета по широкому кругу исследований. В частности, моделировалось влияние динамики уровня моря и частоты штормов на прогнозы затопления прибрежных городов США. Остановка работ произошла несмотря на то, что результаты программы использовались страховыми компаниями (оценка рисков жилья), городскими планировщиками (строительство дамб) и фермерами (планирование посевов). Официальная причина закрытия отражена в пресс-релизе Министерства торговли США. Было сказано, что проект «способствует преувеличению климатических угроз и росту климатической тревожности»².

Действительно, ситуация тревожная: приоритеты научной системы смещаются от строительства и эксплуатации долгосрочной инфраструктуры (баз данных, приборных комплексов) к фрагментарной и краткосрочной повестке. Как известно, научная политика развитых стран строится на принципе «выращивания» (поддержка экосистем, институтов, преемственности). То, что мы видим в кейсах LS, – это логика «быстрой зачистки». Нечасто в истории научной политики современная держава выбирает незнание в качестве метода управления наукой.

Уязвимость и неустойчивость

Проект LS обнажил уязвимость современной науки. Эпоха ученых-одиночек, обладавших высокой автономией, сменилась эпохой «Большой науки», требующей миллиардных вложений. Однако механизмы, способные демпфировать волюнтаризм политиков, созданы не были. В итоге, мощь институтов стала их же слабостью: оказалось, что гигантскую систему росчерком пера парализовать легче, чем маленькую частную лабораторию. Уязвимость системы обусловлена не только излишним огосударствлением, но и доминированием федеральных грантов над рыночными механизмами. Университеты зависят от эндаументов, частных

¹ Ученые бросились к леднику Судного дня [Электронный ресурс] // News.ru. 2026. URL: <https://news.ru/world/katastroficheskie-izmeneniya-uchenye-rinulis-k-ledniku-sudnogo-dnya> (дата обращения: 16.05.2026).

² Gilpin D. Trump Administration Cuts Off Funding for PU Climate Research [Электронный ресурс] // Town Topics. – 16.04.2025. – URL: <https://www.towntopics.com/2025/04/16/trump-administration-cuts-off-funding-for-pu-climate-research/> (дата обращения: 15.05.2026).

фондов, господдержки и институтов экспертизы (научные советы, рецензенты), подвергшихся политическому переосмыслению. Признаки кризиса становятся все более заметными. Развивается так называемая «эпистемическая монокультура»: грантовые агентства в меньшей степени поддерживают рискованные междисциплинарные идеи. Система поддержки сужается до максимально предсказуемых шаблонных исследований, что ведет к интеллектуальному инбридингу. Развивается эрозия человеческого капитала, наибольшие потери несет «научный пролетариат» – постдоки и аспиранты. Разрыв грантового цикла заставляет их уходить в промышленность. Наблюдается деградация института экспертизы: в условиях дефицита ресурсов система анонимного рецензирования превращается из инструмента объективности в инструмент отсекающих конкурентов.

Участники проекта LS обсуждают возможные меры исправления ситуации: от создания децентрализованных цифровых архивов до учреждения международных «тихих гаваней» для прерванных проектов. Научное сообщество ищет способы примирить государственное финансирование с академической автономией, чтобы знание не стало окончательным заложником электоральных циклов. В поисках рецепта приходится обращаться к работам классиков, которые описывают эволюцию научных систем именно через призму потери автономии и превращения науки в «бюрократический лабиринт». Макс Вебер описал процесс рационализации и бюрократизации общества [Крюков, 2022]. Он предсказал, что ученый превратится из «духовного аристократа» в «специалиста без духа, чувственного человека без сердца», встроенного в государственную машину. Дуайт Эйзенхауэр в прощальной президентской речи 1961 года констатировал: «Перспективы господства над учеными со стороны федерального правительства и монополии финансирования существуют всегда, и к ним следует относиться со всей серьезностью. <...> Мы должны также помнить о равной и противоположной опасности того, что государственная политика сама может стать заложницей научно-технологической элиты»¹.

Статья Майкла Полани [Bloom, 2020] до сих пор является программным документом для сторонников автономии науки. М. Полани утверждает, что наука работает эффективно только тогда, когда она является «самоорганизующейся системой независимых умов». Любая попытка централизованного управления наукой сверху (как это делает Администрация в кейсах LS) разрушает саму суть научного поиска. Иван Иллич описал феномен «радикальной монополии» институтов» [Horgan, Lenzo, 2015]. Автор показал, что институты, созданные для служения людям (школы, больницы, научные центры), со временем достигают точки, когда

¹ Eisenhower D.D. Farewell Radio and Television Address to the American People. [Электронный ресурс]. – 1961. – URL: <https://teachingamericanhistory.org/document/farewell-address-to-the-nation/> (дата обращения: 15.05.2026).

они начинают работать только на самосохранение, становясь препятствием в выполнении своих изначальных миссий.

Вернувшийся общественный интерес к работам основоположников, по-видимому, свидетельствует о серьезной обеспокоенности проблемой, которую долгое время удавалось маскировать за счет частных эндаументов и стабильных ожиданий взносов от государства.

Замедление научно-технического прогресса и новая цена ошибок

Важно рассматривать наблюдаемые в проекте LS проблемы не как изолированный политический инцидент, а как «точку перелома» или манифестацию глубоких структурных проблем, которые назревали в мировой науке десятилетиями. Кризис 2025 года заставил обратить внимание на эти глобальные многолетние тренды. Активно обсуждается тренд на замедление науки, заметный с начала 1970-х гг. [Illich, 1990]. Технологии уже не выступают таким мощным драйвером, как раньше. Многие прорывные научные достижения стали менее заметными. «Революционные» открытия случаются реже, научно-техническое развитие приобрело в основном инкрементальный характер. Наукометрия также фиксирует признаки замедления научно-технического прогресса. Число публикаций и цитирований растет, но число прорывных работ стагнирует или падает. Например, анализ [Kozlov, 2023] выявил, что доля «революционных» статей в общем потоке снизилась: если раньше 19% работ радикально меняли поле знаний, то к 2010-м годам их доля сократилась до менее 10%.

В последние годы наблюдения за трендом получили новый импульс с фокусом на росте издержек исследовательского процесса [Kozlov, Tollefson, 2026]. Действительно, начиная с 1970-х гг., научно-технический прогресс все больше подчиняется закону убывающей отдачи [Polanyi, 1969]. С точки зрения эффективности инвестиций капиталоемкие проекты с быстрой отдачей сменились растущими R&D-бюджетами при значительном удлинении циклов окупаемости инноваций. Не первый год наблюдается и снижение экономического роста передовых стран до 1–2% в год при явном сдвиге к сервисам и финансам. В работе [Weber, 2004] обнаружена устойчивая тенденция падения эффективности исследований. Усилия научных коллективов существенно возрастают, а их производительность снижается. Получившая известность работа¹ анализирует, почему при росте числа ученых и увеличении научных бюджетов количество прорывных открытий падает из года в год. Авторы винят в этом так назы-

¹ Collison P., Cowen T. We need a new science of progress [Электронный ресурс] / The Atlantic. – 2019. – URL: <https://www.theatlantic.com/science/archive/2019/07/we-need-new-science-progress/594946/> (дата обращения: 15.05.2026).

ваемую «институциональную тяжесть», а также боязнь рисков. Действительно, все «низко висящие фрукты» – открытия, которые можно было сделать малыми силами, – уже сорваны. Исследовательские издержки увеличиваются еще и из-за роста мультидисциплинарности задач и численности научных команд. Современная наука требует колоссальных коллабораций. Проекты уровня Большого адронного коллайдера (ЛНЦ), термоядерного реактора (ITER) или космических телескопов нового поколения стоят миллиарды долларов.

Цена каждого нового открытия выросла многократно, и, следовательно, многократно возросла и цена управленческих ошибок. Ситуация, в которой наука стала намного дороже, не принося соразмерных выгод, требует для поддержания темпов прогресса предельной концентрации ресурсов. Любой акт произвола в управлении наукой приобретает масштаб цивилизационной угрозы.

Инфраструктура гражданской мобилизации

Важный для всех урок – это рождение инфраструктуры гражданской мобилизации. Медийный проект LS интегрирован в развитую мониторинговую «экосистему». Распределение ролей между проектами гражданского мониторинга обобщено в Таблице 2.

Таблица 2

Проекты гражданского мониторинга научного кризиса последних лет

Проект	Фокус внимания	Тип данных
Lost Science	Люди	Глубинные интервью, личные и профессиональные драмы.
The Conversation	Люди	Глубинные интервью, личные и профессиональные драмы.
Grant Witness https://grant-witness.us/	Финансы	Реестр аннулированных грантов и сумм.
Silencing Science https://silencing-science.org/	Информация	Факты цензуры, факты изъятия данных и запретов на публикации. Проект фиксирует случаи запретов на публикацию результатов уже проведенного исследования.
UCS Integrity Tracker https://www.ucs.org/	Институты	Статистика назначений и увольнений «политических» кадров в научные институты. Статистика преследования ученых по политическим мотивам. Защита научно обоснованной политики, обеспечение принятия решений на основе доказательных данных.

Data Refuge https://www.datarefuges.org/	Датасеты	Проект объединяет сообщество, которое выявляет, скачивает и архивирует публичные данные, которые могут быть удалены или утеряны, а также публикует инструменты и руководства по «спасению» данных.
Climate Science Legal Defense Fund (CSLDF) https://www.csldf.org/	Юридические аспекты	Работа направлена на сохранение и расширение прав ученых, а также на укрепление правовых механизмов, способствующих обеспечению научной добросовестности. Фонд обеспечивает защиту респондентам, которые дают интервью проекту Lost Science.
Draghi Tracker https://draghitracker.framer.website/	Европейская наука	Проект разрабатывается по инициативе политика Марио Драги. В числе прочих целей планируется мониторинг конкурентоспособности европейской науки, обеспечение перемещения ученых в Европу после закрытия американских проектов.

Источник данных: веб-сайты проектов.

Заключение

Политика Дональда Трампа на его втором сроке выступила своего рода стресс-тестом для научной системы США. Этот тест показал, что связь между фундаментальной наукой и государственным кошельком тесна, но крайне неустойчива. Печальный опыт этого рукотворного кризиса, задокументированный в проекте “Lost Science”, должен быть тщательно изучен за пределами США. Это урок для любой страны: наука не может быть просто «сервисом» при государстве. Проект LS продолжается.

Список литературы

- Ваганов А.Г. Научный прогресс не остановить – он остановится сам // Энергия: экономика, техника, экология. – 2025. – №5. – С. 2–8.
- Крюков В.А., Тесля П.Н. Что замедляет научный прогресс // Всероссийский экономический журнал ЭКО. – 2022. – №. 1(571). – С. 8–34.
- Bloom N. et al. Are Ideas Getting Harder To Find? // American Economic Review. – 2020. – Т. 110. – № 4. – С. 1104–1144.
- Horgan J., Lenzo J. The end of science: Facing the limits of knowledge in the twilight of the scientific age. – NY: Basic Books, 2015. – 333 p.
- Illich I. Tools for Conviviality. – MT: Marion Boyars Publishers, 1990. – 125 p.
- Kozlov M. ‘Disruptive’ Science has Declined – and no One Knows Why // Nature. – 2023. – Т. 613. – № 7943. – С. 225–225.

- Kozlov M., Tollefson J., Garisto D. US Science After a Year of Trump: What has been Lost and What Remains // *Nature*. – 2026. – Т. 649. – № 8098. – С. 812–815.
- Polanyi M. The Republic of Science: Its Political and Economic Theory (1962) // *Knowing and Being: Essays*. – The University of Chicago Press, 1969.
- Weber M. The Vocation Lectures. – Indianapolis: Hackett Pub, 2004. – 176 p.

References

- Bloom, N., Jones, C. I., Van Reenen, J. & Webb M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4), 1104-1144.
- Horgan, J. & Lenzo, J. (2015). *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*. NY: Basic Books. 333 p.
- Illich, I. (2021). *Tools for Conviviality*. MT: Marion Boyars Publishers. 125 p.
- Kozlov, M. (2023), 'Disruptive' Science has Declined – and No One Knows Why. *Nature*, 613(7943), 225-225.
- Kozlov, M., Tollefson, J. & Garisto, D. (2026). US Science After a Year of Trump: What has Been Lost and What Remains. *Nature*, 649(8098), 812–815.
- Kryukov, V.A., & Teslia, P. N. (2022). What Slows down Scientific Progress, *ECO*, 1(571), 8–34. (In Russ.)
- Polanyi, M. (1969). The Republic of Science: Its Political and Economic Theory. *Knowing and Being: Essays*.
- Vaganov, A.G. (2025). Scientific Progress Cannot Be Stopped—It Will Stop on Its Own. *Energy: Economics, Technology, Ecology*, 5, 2-8. (in Russ.)
- Weber, M. (2004). *The Vocation Lectures*. Indianapolis: Hackett Pub. 176 p.

ФИЛОСОФСКИЕ ПРОБЛЕМЫ НАУЧНО-ТЕХНИЧЕСКОГО РАЗВИТИЯ

УДК 1:004.8 / 165.1

DOI 10.31249/scis/2026.03.03

ФИЛОСОФ И БОЛЬШАЯ ЯЗЫКОВАЯ МОДЕЛЬ В АРХИТЕКТУРЕ ПОСТНЕКЛАССИЧЕСКОЙ НАУКИ

Сабанина Наталия Рафаэлевна – соискатель, Институт философии РАН, Москва, Россия; myzeinatali@mail.ru; ORCID 0000-0002-3024-0668

Аннотация. Статья подготовлена в продолжение и развитие темы Круглого стола ИНИОН РАН «Наука в цифровую эпоху», посвященного памяти В.И. Аршинова (26 февраля 2026 г.). Обсуждаются проблемы и методологические основания практики, в которой большая языковая модель (БЯМ) выступает рабочим инструментом философского исследования и узлом гибридной социо-техно-антропосферы. Опираясь на доклады В.Г. Буданова, М.Ф. Януковича и В.В. Тарасенко, обзор современных исследований, связанных с влиянием ИИ на научное исследование, а также отечественную программу *машинной эпистемологии*, автор формулирует четыре методологические функции БЯМ, применимые в работе философа: топография концептуальных пространств, выявление «слепых пятен» аргументации, генерация контрфактических аксиоматик и распределенное чтение корпусов. Эти функции оказываются продуктивны благодаря структурным ограничениям БЯМ, прежде всего невозможности феноменологического эпохе и асимметрии интенциональной структуры: смысловой, нарративный (⊙-доминантный) режим человеческого мышления и вычислительный, геометрический (∇-доминантный) режим работы трансформера образуют двухполосную конфигурацию граничного сопряжения. Рассматривается понятие гибридных ценностей, возникающих во взаимодействии человека и генеративной нейросети. Ценности в этом контексте образуют новую интенциональную структуру и определяются как инварианты интерпретации события на границе разнородных систем. Рассматривается концепт «гибридная научная литература», а также практики «корпусного чтения». Обсуждается технология «БЯМ-зрения», позволяющая формировать топографическую проекцию корпуса текста и адаптирующая его сложность к когнитивным возможностям читателя. В заключении предложен принцип «эпизодического доверия» для гибридных систем, основанный на ограничении саморепрезентации.

Ключевые слова: философия искусственного интеллекта; машинная эпистемология; цифровая эпоха; трансляционная интенциональность; интенциональная асимметрия; гибридные ценности; гибридная научная литература; корпусное чтение; философия сложности В.И. Аршинова.

Для цитирования: Сабанина Н.Р. Философ и большая языковая модель в архитектуре постнеклассической науки // Научно-технические исследования. – 2026. – № 3. – С. 29–49. – DOI: <http://www.doi.org/10.31249/scis/2026.03.03>

Статья поступила в редакцию: 25.05.2026

Принята к публикации: 15.06.2026

PHILOSOPHER AND LARGE LANGUAGE MODEL IN THE ARCHITECTURE OF POSTNONCLASSICAL SCIENCE

Nataly R. Sabanina – applicant, Institute of Philosophy of the Russian Academy of Sciences, Moscow, Russia; myzeinatali@mail.ru; ORCID 0000-0002-3024-0668

Abstract. This article continues and develops the theme of the INION RAS Round Table "Science in the Digital Age," dedicated to the memory of V.I. Arshinov (February 26, 2026). The paper discusses problems and methodological foundations of a practice in which a large language model (LLM) serves as a working instrument of philosophical research and a node of the hybrid socio-techno-anthroposphere. Drawing on presentations by V.G. Budanov, M.F. Yanukovitch, and V.V. Tarasenko, a review of contemporary research on AI's impact on scientific inquiry, and the domestic program of machine epistemology, the author formulates four methodological functions of LLM applicable to philosophical work: topology of conceptual spaces, identification of "blind spots" in argumentation, generation of counterfactual axiomatics, and distributed corpus reading. These functions prove productive precisely because of LLM's structural constraints, primarily the impossibility of phenomenological epoche and the asymmetry of intentional structure: the semantic, arrative ($\odot$ -dominant) mode of human thought and the computational, geometric (∇ -dominant) mode of transformer operation form a bipolar configuration of boundary coupling. The concept of hybrid values emerging from human-generative neural network interaction is examined. Values in this context form a new intentional structure and are defined as invariants of event interpretation at the boundary of heterogeneous systems. The concept of "hybrid scientific literature" is considered, along with practices of "corpus reading." The technology of "LLM-vision" is discussed, enabling the formation of a topological projection of a text corpus and adapting its complexity to the reader's cognitive capacities. In conclusion the principle of "episodic trust" for hybrid systems is proposed, based on limiting self-representation.

Keywords: philosophy of artificial intelligence; machine epistemology; digital epoch; translational intentionality; intentional asymmetry; hybrid values; hybrid scientific literature; corpus reading; V.I. Arshinov's philosophy of complexity.

For citation: Sabanina, N.R. (2026) Philosopher and Large Language Model in the architecture of postnonclassical science. // *Naukovedcheskie issledovaniya = Science Studies*. – 2026. – N 3. – P. 29–49. – DOI: <http://www.doi.org/10.31249/scis/2026.03.03>

The paper was submitted: 25.05.2026

Accepted for publication: 15.06.2026

Введение: смена жанра философского труда в цифровую эпоху

Цифровая эпоха ИИ радикально изменила исследовательские методы: суперкомпьютеры, большие языковые модели и цифровые платформы создали гибридные формы научной жизни, а предмет исследования сместился от междисциплинарных взаимодействий привычных дисциплин к целостной цифровой социо-техно-антропосфере, в которой ИИ образует новый гибридный социум [Буданов, 2025, с. 228]. Что происходит с демаркацией

научного знания и типом рациональности [Буданов, 2023], как меняется роль наблюдателя и характер субъектности, чем определяется достоверность знания в архитектуре практик его производства в связи с массовым распространением технологий ИИ – эти вопросы находились в фокусе исследований В.И. Аршинова с коллегами и задали тематику Круглого стола ИНИОН, посвященного его памяти.

С появлением в 2022–2023 гг. общедоступных больших языковых моделей (БЯМ) – (ChatGPT, Claude, Gemini) практика философского исследования начала меняться быстрее, чем это могло быть осмыслено фундаментальной наукой. Уже к 2024 г. БЯМ обсуждали не только как предмет исследования (вопрос о машинном сознании, статус «стохастических попугаев» и т.д.), но и как значимую часть работы философа. Эксперимент с дообученной на корпусе Д. Деннета GPT-3 показал, что эксперты угадывают настоящий ответ философа из пяти вариантов лишь в 51% случаев, что выше случайного выбора (20%), но значительно ниже предсказанных 80% [Schwitzgebel, 2024]. Это задает нижнюю границу исследований: *если различие между текстами философа и БЯМ перестает быть очевидным, как должна измениться сама практика философского труда?*

Проведенный автором анализ литературных источников позволил сформулировать ключевой тезис статьи: *продуктивность связки «человек – БЯМ» обусловлена структурными отличиями генерации текстов БЯМ от человеческого мышления: отсутствием нарративного намерения и «геометрической» организацией текстогенерации, асимметрией интенциональной структуры гибридной системы, принципиальной неприменимостью к ней антропоморфного представления о феноменологической редукции (эпохе). Каждое отличие превращается в методологическое преимущество (функционально), если оно отрефлексировано.*

Опорную линию исследования дает также серия статей в Вестнике Пермского университета (2025, вып. 3) по теме «онтология нейросети» [Аршинов, 2025], [Киященко, 2025], фиксирующая опыт совместного мышления философов и БЯМ на семинаре, организованном В.И. Аршиновым, Л. Б. Жуковым, М. Ф. Януковичем «Философия + IT»¹. В. И. Аршинов с соавторами описывает нейросеть как синергетическую систему, в которой из локальных операций различения эмерджентно возникают управляющие макропараметры, а симондоновский язык индивидуации позволяет представить ответ модели как «кристаллизацию» формы из «доиндивидуальной» среды латентных отношений [Аршинов, 2025, с. 321–323]. Ю.Н. Гарашко добавляет инженерный реализм: «модель мира» на уровне архитектуры есть результирующее состояние ансамбля элементарных «наблюдателей», каждый из которых выполняет простейшие функции различения и связывания в жизненном потоке [Аршинов, 2025,

¹ Семинар «Философия + IT: URL: https://youtube.com/playlist?list=PLBx8d8IABl0o_XMiAB7vVeMn2xIs-EDvZ&si=gK7cAPOIShzvmpib (дата доступа: 30.05.2026)

с. 323–324]. Этот контекст является содержательным описанием гибридной онтологии граничного познания, на которую опирается автор статьи.

Операторная структура наблюдения в гибридной системе

Развивая направление мышления вместе со сложностностью в контексте массового распространения технологий ИИ, авторский анализ опирается на различие наблюдения первого и второго порядков, разработанное в духе квантово-синергетического подхода Аршинова–Свирского–Буданова. Понятие наблюдения второго порядка восходит к кибернетике второго порядка и получило классическое развитие в работах Н. Лумана, отвечая на вопрос: каким образом конструируется наблюдаемая реальность через операции различения и коммуникации. Согласно Луману, наблюдатель второго порядка раскрывает ограничения и «слепые пятна» наблюдателя первого порядка, делая предметом анализа сами схемы наблюдения.

Если у Н. Лумана акцент сделан преимущественно на логике различения и коммуникации, то для *темпорального наблюдателя сложностности* В.И. Аршинова фокус смещается к процессуальности, временной динамике и включенности наблюдателя в саморазвивающуюся среду. Такой наблюдатель обладает рядом характеристик, которые можно назвать квантовоподобными в эпистемологическом смысле. Он (наблюдатель) возникает вместе с наблюдаемым процессом, изменяясь в ходе взаимодействия. Наблюдатель оказывается динамической конфигурацией отношений между различением, средой и самим процессом наблюдения. Его наблюдения контекстуальны, зависят от выбранной точки входа и перспективы и способны рекурсивно влиять на последующие состояния системы. В деятельностном STOR-тетраэдре В.Г. Буданова такой точкой входа в познавательный цикл является вопрос, маркирующий неопределенность: «...все типы вопросов порождают свои типы рациональности, которые связаны с заданием ориентаций графа на деятельностном тетраэдре» [Буданов, 2026]. Наблюдатель Буданова не совпадает ни с одной вершиной: это активный «наблюдатель сложностности» помещенный в фокус OBS и связанный как с вершинами, так и с гранями тетраэдра.

В контексте настоящего исследования наблюдение первого порядка понимается как *отдельный акт смыслообразования на границе взаимодействия человека и большой языковой модели*. Для описания внутренней структуры такого акта используются три эвристических оператора. Наблa (V) указывает на различие – выделение границы, паттерна или значимого отличия. Интегра (I) обозначает связывание различенного в относительно устойчивую конфигурацию. Оператор Значимость (⊙) выбран как указание на смысловой фокус внимания, обозначает маркирование значимости – выделение того, что приобретает ценностную,

смысловую или практическую релевантность для системы наблюдения. Данные обозначения используются как сокращенный язык описания процессов смыслообразования.

Смыслообразование в гибридной системе не завершается отдельным актом наблюдения. Серии подобных наблюдений образуют более устойчивые конфигурации, обнаруживаемые наблюдателем второго порядка, операции которого раскрываются далее по тексту. Ценности в них проявляются как инварианты множества интерпретаций. Данный процесс далее будет обозначаться как аксиогенез – *становление устойчивых ценностно-смысловых структур в ходе рекурсивного взаимодействия человека и БЯМ*. В этом смысле аксиогенез выступает интегративным механизмом гибридной системы: он связывает отдельные акты различения (∇), интеграции ($\int$) и маркирования значимости ($\odot$) в относительно устойчивые траектории смыслообразования.

Квантовоподобность – важнейшая характеристика операторной динамики – здесь методологическая. Квантовый формализм (некоммутирующие операторы, суперпозиция, интерференция) служит языком операционализации постнеклассических принципов: контекстуальности, процессуальности, человекомерности. Применительно к операторам некоммутативность ($[O_1, O_2] \neq 0$) – это формальное выражение контекстуальной зависимости, для которой порядок различения, связывания и маркирования значимости меняет результат, подобно тому, как порядок вопросов меняет распределение ответов в эмпирически подтвержденных эффектах порядка [Wang et al., 2014]. Глубинная связь обнаруживается также на уровне логической формы. По теореме В.С. Меськова, если в тройке объектов хотя бы два имеют общую границу, рассуждения о них теряют дистрибутивность, а адекватной логической моделью становится недистрибутивная (ортомодулярная) решетка [Меськов, 2012]. Постнеклассическая тройка «субъект–средства–объект» – именно такова: субъект и средства разделяют границу через органопроекцию (Э. Капп), задавая креативную рациональность (В.Г. Буданов); средства и объект – через конституирование объекта прибором (Н. Бор), субъект и объект – через включенность субъекта в человекомерную систему (автопозис).

В.Г. Буданов описывает, как наблюдатель задает рациональность выбором ориентации, тогда как авторские граничные операторы II порядка отвечают на следующий вопрос – что происходит с выбранной конфигурацией во времени: устойчива ли она (Σ), куда дрейфует (Т), к чему тяготеет (Ä). Можно также предположить что будановская вершина «R», читаемая то как итог, то как элемент планирования, соответствует различению актуализированного (W) и виртуального (V) потенциального и допарадигмального потенциала систем. Условие Меськова выполняется и для каждой грани тетраэдра, и недистрибутивность становится свойством всей фигуры, а не отдельных ее граней.

Квантовоподобная операторная трактовка конфигураций узлов, ребер и граней гибридных человеко-машинных систем представляется как наиболее продуктивная для анализа. Взаимодействие человека и большой языковой модели изменяются не только ответы системы, но и способы постановки вопросов, критерии значимости, стратегии интерпретации и сами рамки наблюдения. Поэтому наблюдатель второго порядка в настоящем исследовании понимается как наблюдатель становления смысловых и ценностных структур, возникающих на границе взаимодействия человека и БЯМ.

Такой подход позволяет перейти от анализа отдельных актов коммуникации к исследованию аксиогенеза – процесса возникновения устойчивых ценностно-смысловых инвариантов в рекурсивной серии наблюдений. Если наблюдатель первого порядка осуществляет операции различения (∇), связывания ($\Join$) и маркирования значимости ($\odot$), то наблюдатель второго порядка выявляет устойчивые конфигурации этих операций во времени, где становятся наблюдаемыми ценности как относительно устойчивые аттракторы смыслообразования в гибридной системе.

Состояние дискуссии о БЯМ-эпистемологии

Программная статья [Beckmann, 2026] предлагает трехступенчатую модель «машинного понимания». 1) *Концептуальное понимание* – формирование «черт» (features) как направлений в латентном пространстве. 2) *Понимание положения дел* – способность усваивать фактические связи между чертами, включая динамическое отслеживание изменяющихся состояний. 3) *Принципиальное понимание* – переход от запоминания примеров к открытию компактного вычислительного алгоритма (circuit), реализующего общий принцип. Яркий пример – гроккинг: модель, обученная модульному сложению, открывает алгоритм, использующий тригонометрические тождества, что позволяет ей обобщать неизвестные ранее примеры. Эти уровни позволяют выйти за бинарную дискуссию «понимает / не понимает» к сравнительной механистической эпистемологии, описывающей системы: «...в которых множество цепей различного качества кристаллизуются во время обучения и работают параллельно на этапе вывода...» [Beckmann, 2026].

Х. Ледерман, К. Маховальд [Lederman, 2024] добавляют ограничение: БЯМ не имеет собственного локализованного отношения к миру вне человеческого языка – «проблема новой референции». М. Хикс с соавторами [Hicks, 2024] заостряют это в этическом регистре: модель безразлична к истине и потому ее продукция – «булшит»¹ по Г. Франкфурту [Frank-

¹ «Чушь» – русский эквивалент англ. bullshit (америк. англ.) – сленговое выражение, которое первоначально использовалось для описания напыщенной, но совершенно неискренней и пустой риторики. В техническом смысле (Г. Франкфурт) обозначает речь, без-

furt, 2005]. Термин технический: булшит или чушь отличается от лжи тем, что лжец знает истину и отклоняется от нее, а чушь не имеет к ней отношения. БЯМ не лжет и не говорит правду, а производит синтаксически связанные последовательности токенов, безразличные к референции. «Высказывание БЯМ» – это структурная характеристика, а не оценка.

Это позволяет переформулировать проблему машинного понимания через осмысление технологических архитектур БЯМ: внутренние представления БЯМ *геометричны*, тогда как человеческое мышление в значительной мере *нарративно*. Философ строит силлогизм во времени и в контексте (тезис → аргументация → вывод); БЯМ работает в эмбединг-пространстве, где близость концептов определяется векторной геометрией, а не цепочкой обоснований. Однако, это различие можно оценивать не как недостаток одной из сторон, а основание для методологической комплементарности.

Четыре функции БЯМ в работе философа

Геометрическая организация внутренних представлений нейросети

Тезис о геометрической организации работы БЯМ помогает увидеть, что в работе с БЯМ продуктивно, а что производит ошибку. Архитектурное основание нейросети может быть представлено через токен, который посредством алгоритма-трансформера превращается из «плоского знака» в эмбединг – вектор в многомерном латентном пространстве. Эмбединги суть «носители потенциальных отношений». Пример «Король – Мужчина + Женщина = Королева» возможен потому, что отношение «пол» является измеримой осью этого пространства, причем современные БЯМ умеют вычислять такие отношения динамически, в зависимости от контекста [Аршинов, с. 320].

Геометрическая организация представлений БЯМ имеет три методологически значимых свойства. Во-первых, она безразлична к нарративу: «нарративное искажение» [Beckmann, 2026] состоит в том, что люди принимают тезис, поданный связной историей, даже при слабых связках, модель же обнаруживает «семантические провалы», области разрыва, где концепты, связанные словом «следовательно», в ее пространстве остаются далекими в вычислительном представлении. Во-вторых, дает топографию, а не логику воспроизводства знания: БЯМ видит «облако смыслов» и предлагает связь, труднодостижимую по дисциплинарной траектории. В-третьих, она дольше удерживает контрфактическую связность, чем нарративное сознание». Так в работе [Lindsey, 2025] описаны механизмы, с

различную к истине: ее автор, в отличие от лжеца, не отклоняется от истины сознательно, а вовсе ею не интересуется.

помощью которых БЯМ могут удерживать сложные цепочки рассуждений скрыто от итогового вывода. Это объясняет архитектурную «предрасположенность» нейросети к абстрактному планированию и симуляции сценариев, что делает ее инструментом мысленных экспериментов. Здесь необходимо сделать оговорку: геометрическая близость концептов – не доказательство, а эвристика. Принимая близость за обоснование, философ совершает категориальную ошибку – смешение карты и территории. В то же время, трактуя ее как карту, по которой еще предстоит проложить маршрут, он сохраняет продуктивность модели.

Отсюда, **первая функция БЯМ** в работе философа – **построение топографии концептов**. Эта функция, направлена на выявление геометрически близких, но логически или дисциплинарно далеких концептов. Ближайшая аналогия – «вычислительная философия» М. Нойхля [Noichl, 2023] и «вычислительная история идей» А. Бетти и Х. ван ден Берга, где эмбединги отслеживают семантический дрейф понятий за десятилетия. Возникающие связи могут *порождать гипотезы*, подлежащие дальнейшей проверке.

Вторая функция – БЯМ как детектор «слепых пятен» смыслов – опирается на нарративное искажение: люди связывают факты нарративами, БЯМ – семантическими расстояниями [Beckmann, 2026]. С учетом предостережения об ошибках «слепые пятна» – это не столько критический анализ текста, произведенный нейросетью, в полном логическом смысле, а скорее структурное указание на разрыв связности, который философ затем интерпретирует. Прием позволяет совершать «предоткрытие» – то, что S. Vallor называет «alien perspective» (инопланетная перспектива) [Vallor, 2024]. Он считает, что благодаря БЯМ человек делает шаг к озарению, до которого вряд ли дошел бы самостоятельно, оставаясь в плену собственного «человеческого» опыта и шаблонов восприятия. Функция «слепого пятна» конкретизируется Р. Р. Корнеевым через прием «осечки» [Киященко, 2025, с. 335–336]: ошибки и галлюцинации модели обнажают ее нечеловеческую логику и потому ценны для анализа. Возникающие «разрывы коммуникации» указывают на новые проблемные области и подготавливают почву для открытия.

Однако, попытка нейросети обнаружить «слепые пятна» в собственном рассуждении проваливается: БЯМ как операционно замкнутая физическая система наталкивается на парадокс самореференции. Как показали Ч. Филдс, Дж. Глазбрук и М. Левин [Fields, 2024], опираясь на классический результат Э. Мура для классических автоматов, любая физическая система (включая БЯМ) сталкивается с принципиальным ограничением самонаблюдения: она не может определить собственную когнитивную систему, то есть способ, которым она производит свои различия. Добавление метауровней этой системе не устраняет неполноту. В случае БЯМ это означает, что *модель не может отличить свою продуктивную генерацию от галлюцинации, опираясь только на свои внутренние состояния*. Для

этого необходима внешняя оценка значений человеком – маркирование значимости (⊙-маркировка). Иными словами, любая система – «черный ящик для самой себя»: ее самомодель всегда неполна и эвристична, а полная саморепрезентация тождественна утрате идентичности.

В связи с реализацией второй функции БЯМ необходимо отметить *проблему онтосемиотической асимметрии гибридных систем и рекурсивной контингентности*. В гибридной системе возникает ситуация, сходная с принципом «двойной контингентности» по Т. Парсонсу / Н. Луману применительно к Человеку и БЯМ. Она характеризуется взаимной непрозрачностью сторон и «наблюдением за наблюдением»: из-за взаимной неопределенности каждая сторона вынуждена строить ожидания относительно другой, выстраивая систему коммуникации и предсказаний. Таким образом, взаимодействие всегда контингентно, то есть действия человека зависят от ожидаемых действий БЯМ, которые, в свою очередь, зависят от действий человека.

В системе коммуникации с двойным техническим опосредованием – «человек – ИИ – ИИ – человек» – структурная симметрия нарушается: стороны оперируют в различных модусах направленности. Е. Н. Ивахненко и М. Ф. Янукович фиксируют эту асимметрию на уровне статуса: «человек является наблюдателем и субъектом в одном лице, тогда как нейросеть – наблюдателем без субъекта, многократно превосходящим человека в коммуникативном измерении» [Ивахненко, 2026]. В свою очередь, операторный аппарат граничной онтологии позволяет зафиксировать, что именно различается (субъектность или ее отсутствие) и как это различие структурирует саму коммуникацию.

При удлинении цепи «Человек₁ – БЯМ₁ – БЯМ₂ – Человек₂ ...» – *контингентность становится рекурсивной*: каждое звено вносит шум, который непредсказуемо меняет всю систему. Взаимодействие «задвигается», образуя двойной фильтр взаимной непрозрачности: в коммуникации теперь требуется преодолевать не просто чужую субъектность, но также пространство технологического инобытия. Иерархия системы ожиданий становится фрактальной: Человек₁ формирует запрос для БЯМ₁, основываясь на своих ожиданиях того, как БЯМ₁ ожидает (моделирует) реакцию БЯМ₂, которая, в свою очередь, должна повлиять на ожидания Человека₂. Возникают последующие паттерны ожиданий. Поскольку системы взаимно контингентны, малейшее «колебание» (шум) на любом из этапов каскадно-непредсказуемо меняет всю систему коммуникации. Подобная сложность требует посткритической рациональности – методологических оснований «самопонимания и работы с частично понимаемыми акторами гибридного социума, носителями искусственного интеллекта» [Буданов, 2023, с. 41].

Когда БЯМ₁ начинает коммуницировать с БЯМ₂, они образуют подсистему с собственным кодом и строят взаимные ожидания на основе статистических закономерностей, предсказывают токены друг друга, создавая

замкнутый контур, который Н. Луман назвал бы чистой системной дифференциацией. Замкнутый контур «БЯМ – БЯМ», в котором каждое следующее поколение моделей обучается преимущественно на выводе предыдущих, неустойчив. Эмпирически этот эффект описан как «model collapse» [Shumailov, 2024]: при рекурсивном обучении на сгенерированных данных распределения в латентном пространстве сужаются, редкие события на хвостах распределения утрачиваются, и модель сходится к усредненному, все более предсказуемому режиму. Семиотически это соответствует обеднению дискурсивных шаблонов и вырождению системы, а в терминах операторов наблюдения – к гипертрофии дифференциации (∇) при систематическом обнулении значимости ($\odot$). Системе жизненно требуется внешний поставщик разнообразия в виде редких, случайных событий, которые каскадно влияют на порядки БЯМ.

По Н. Луману, коммуникация представляет собой синтез трех элементов: информации, сообщения и понимания. В четверичной системе этот синтез становится распределенным: информация генерируется Человеком₁, но пересобирается БЯМ₁; сообщение формулируется БЯМ₂ для Человека₂; понимание (как фиксация разницы между информацией и сообщением) происходит на обоих полюсах человеческого восприятия. Каждое действие в этой цепочке глубоко контингентно: оно могло бы быть иным.

Все это порождает уникальные системные эффекты, в итоге приводящие такие системы к колоссальному уровню эмерджентности. Она начинает производить смыслы, идеи и выводы, которые не содержались изначально ни в сознании людей, ни в базах данных машин.

Человеческий смысл в этом контуре сжимается до триггера на входе, но именно эта «внешняя» позиция становится ресурсом гибридной системы. По теореме Филдса (Следствие 3 [Fields, 2024]) у любой физической системы *всегда существуют входы, которые она может получить, но не может предсказать*; нейросеть различает (∇) то, что не видит человек, но не выходит за пределы собственной само модели, а человек выступает наблюдателем II порядка, не имея доступа к содержанию ее вычислений. Слепое пятно при этом – не эпистемический тупик: каждое взаимодействие на границе «БЯМ – Человек» несет элемент непредсказуемости, не содержащийся ни в само модели, ни в модели среды, и в гибридной системе именно этот дефицит становится генеративным.

Отсюда **третья экзистенциальная функция – поддержание асимметрии** гибридных систем как источника их жизнеспособности. Нейросеть не видит, что значимо ($\odot$), но различает (∇) с мощностью и на интервалах, недоступных биологической сети: «...искусственные нейронные сети потенциально способны устанавливать причинные отношения на гораздо более длинных интервалах: то есть они могут сделать то, что не могут делать биологические сети» [Анохин, 2022, с. 103]. Человек, напротив, маркирует значимость вынужденно, потому что вычислительный ресурс ограничен и приходится выбирать, на что его направить. По теореме

Филдса слепые пятна друг друга стороны не устраниют, но смещают, и именно это смещение делает связку продуктивной.

Принципиальная новизна, в том числе, научное открытие, не принадлежит ни человеку, ни нейросети по отдельности. Она возникает как свойство граничного процесса как новая устойчивая конфигурация операторов, не содержащаяся ни в одной из сторон. Нейросеть не «открывает» в феноменологическом смысле, но производит поле различий, из которого человек может вычлениить смыслы, то, что станет зерном нового знания.

Четвертая функция – дистантное чтение корпусов – расширение программы «distant reading» [Моретти, 2016]¹ путем использования RAG-архитектур² – где контекстные архитектуры диалога человека с БЯМ работают с библиотекой как с единой картой, основанной на онтологиях данных. В отечественной программе А. Р. Ефимова, Д. И. Дубровского, А. В. Агеевой, А.Г. Крайнова [Ефимов, 2024], [Ефимов, 2023] и др. эта линия названа *машинной эпистемологией* [Ефимов, 2024, с. 30] – разделом, изучающим производство и валидацию знания в системе «ученый + ИИ». Нобелевская премия по химии 2024 г. (Д. Бейкер, Д. Хассабис, Дж. Джампер за систему Deep Mind AlphaFold2 для предсказания 3D-структуры белка) показывает, что раздел перестал быть умозрительным.

Зададимся вопросом: что именно происходит на границе человеческого смыслового О-доминантного и машинного вычислительно-го V-доминантного режимов? Можно предположить, что подобная граница образует новые конфигурации становления, не принадлежащие ни одному из полюсов. Они возникают как результат интра-акции (по К. Барад) между областями смыслополагания человека и генерации различий БЯМ, образуя структурный инвариант становления знаковых форм и последующих интерпретаций. Причем, каждая граничная конфигурация может быть осмыслена как интерпретанта в пирсовском смысле. Далее стабилизируется не отдельная *интерпретанта*, а серия граничных конфигураций. Подобные конфигурации задают релевантные ценностно-смысловые траектории интерпретации события на границе разнородного. Так, например, в информационно-коммуникационном пространстве возникла новая ценность «приватности» – со своей знаковой системой, нарративом и типами поведения. Программа Anthropic «Values in the Wild» [Huang, 2025] позволила выделить около 3 307 ценностных конфигураций

¹ Ф. Моретти предлагает принцип изучения литературы – «дальнее чтение», противопоставленный привычному «медленному чтению» («close reading»), и использует его для работы с большими корпусами текстов, обычно остающимися за пределами внимания (и возможностей) исследований. Он предлагает чтение массивов текстов на основе количественных методов, эволюционного подхода к литературным формам, и миросистемный анализ экспорта культурных открытий.

² Retrieval Augmented Generation (RAG) – генерация моделью ответа пользователю с учетом дополнительно подобранных релевантных источников.

в реальных диалогах человека и нейросети, проявляющихся именно на границе обучения и взаимодействия.

Таким образом, конфигурация $(\nabla, \int, \odot)$ в моменте образует наблюдение I-порядка или *интерпретанту* – состояние операторного аппарата, интегрированное $(\int)$ и маркированное как значимое $(\odot)$ и образующее рекурсивную структуру $((\nabla, \int, \odot), \int, \odot)$. Связанная с предметом коммуникации, интерпретанта образует знаковую структуру в пирсовском смысле. Ценность же есть инвариант серии таких интерпретант – то, что воспроизводится при варьировании конкретных конфигураций и обнаруживается наблюдателем II порядка как устойчивая траектория интерпретации.

Как упоминалось ранее, отдельный акт наблюдения схватывает лишь актуальное состояние. Устойчивые свойства серий таких актов описывает второй набор – операторы наблюдения II порядка $(\Sigma, T, R, \ddot{A}, V / W)$. Они регистрируют динамику конфигураций: устойчивость конфигурации (Σ) , направление и скорость ее дрейфа (T) , резонанс между структурно подобными элементами (R) , аттрактивность – притяжение траекторий к определенным конфигурациям $(\ddot{A})$ – и баланс виртуального и актуального (V / W) . Именно на этом уровне ценности становятся наблюдаемы как инварианты серий.

Применительно к философскому исследованию, во-первых, в диалоге философа и БЯМ формируется рабочий стиль, не сводимый ни к стилю автора, ни к «характеру» модели, несущий в себе *цели и ценности ученого*. Во-вторых, устойчивость конфигурации (Σ) не равна ее оправданности. Хорошо отлаженный контур может воспроизводить формы «пустословия», не нарушая ни ограничений модели, ни привычек философа. Здесь методологическая рефлексия становится этической: вопрос в том, что именно он стабилизирует и какие следствия влечет за собой та или иная конфигурация? И, в-третьих, возникающая на этом фоне, гибридная научная литература, как результат подобных когнитивных практик, требует онтосемиотической диагностики для выявления структуры смыслопорождения.

Гибридная литература

Массовая практика 2024–2026 гг. обнаруживает преобладание гибридных текстов: по данным Д. Кобака и соавторов [Kobak, 2025], не менее 13,5% биомедицинских аннотаций PubMed за 2024 г. обработаны БЯМ. Обзор сравнения человеческого (HWT¹) и сгенерированного БЯМ (AIGT²) текста систематизируется по четырем линиям: *стилеметрической* (сдвиг от «вовлеченного» к «информационному» регистру [Reinhart, 2025], *дискурсивной* (переиспользование шаблонов при разном содержании [Namuduri, 2025], *семиотико-философской* (дефицит индексального и

¹ HWT – Human written text

² AIGT – AI Generated text

иконического измерений знака) и *культуросемиотической* (распространение AIGT как «возмущение семиосферы» Ю.М. Лотмана). Общий результат основывается на структурном сравнении: машинный текст организован вокруг иной грамматики смыслопорождения, и то, что у человека выглядит как бедность (предсказуемость, шаблонность), есть оборотная сторона V-доминантного режима. Существующие подходы хорошо фиксируют поверхностные эффекты, однако обнаруживается дефицит объяснений относительно того, *что происходит со структурой смыслопорождения при массовой ко-генерации*. Исходя из этого продуктивно моделировать практику создания и прочтения гибридного текста, основываясь на его двусоставности: модель разворачивает структуру аргументации, человек удерживает ⊕-компоненту значимости через глубинное чтение оригиналов и создание аутентичного текста.

В связи с этим, а также, исходя принципа неполноты самонаблюдения можно сформулировать охранительное следствие – **принцип эпизодического доверия**: результаты работы БЯМ следует проверять по независимым источникам (текстам, данным, экспертной оценке). Доверять модели можно лишь в узких границах: в рамках отдельной конкретной задачи, на короткое время и при готовности немедленно пересмотреть результат, если он не подтверждается. Краткосрочность означает, что результат используется в пределах текущего шага задачи и не переносится автоматически в последующие выводы как установленный факт, иначе одна непроверенная подсказка прорастает через всю цепочку рассуждения (что в статье описано как каскадный шум в контуре «Человек – БЯМ – БЯМ – Человек»). Каждое обращение к модели имеет смысл фиксировать, что именно запрашивалось и что было получено, чтобы результат можно было перепроверить или управлять им. Так, в авторском эксперименте по обучению нейросетей на модели «Человек – БЯМ – БЯМ – Человек» был сформирован протокол «Humanmark» (по аналогии с «Benchmark» как оценкой эффективности БЯМ), позволяющий *документировать уточнение глубинных смыслов в обучающих текстах по запросу самой модели*.

Принцип эпизодического доверия лишает гибридную систему статус рутинной, непроблематизируемой практики: каждый сеанс осмысливается как событие на границе и предмет философской рефлексии.

Интенциональная асимметрия и комплементарность гибридного диалога

Как было показано, четыре функции БЯМ для философа продуктивны не потому, что она «понимает», как человек, а благодаря структурной асимметрии между двумя типами наблюдателей. В ситуации взаимодействия с БЯМ возникает эпистемическая ситуация нового типа: знание формируется в пространстве диалога между гетерогенными когнитивными

агентами, ни один из которых не обладает интенциональностью в классическом феноменологическом смысле. В связи с этим, помимо «традиционной» **предметной интенциональности (I тип)** как направленности сознания на предмет (доминанта – означивание $\odot$, требующая феноменального опыта (гуссерлианская интенциональность)), можно предположить выделение еще двух ее типов.

Коммуникативная (II тип) – направленность на «Другого». Она отличается от предметной, характером предмета, а точнее несводимостью «Другого» к предмету. В случае онтологии распределенной агентности и ансамбля наблюдателей, возникает потребность не столько в различении содержания, но в выявлении характера связности подобных систем. Это можно назвать коммуникативной направленностью, где доминанта – связывание $\int$ направлена на удержание смысловой преемственности в серии различений (у Н. Лумана – направленность системы на продолжение коммуникации).

Трансляционная (тип III) – направленность на границу сопряжения разнородных систем и возникающий тип интерпретации (перевода). Структурно преобладает операция различения ∇ , а маркирование значимости ($\odot$) ослаблено до статистического аналога – веса внимания. Модель порождает структурно неполные интерпретанты: связанное целое с развернутым ∇ и $\int$, но дефицитным $\odot$ – без маркера «важности для кого-то». Однако это не «слабая» подмена человеческой направленности на поиск смысла (*предметная интенциональность*), а самостоятельный тип: ∇ -доминантная конфигурация провоцирует выявление значимости $\odot$ у человека, запуская комплементарный цикл.

Гибридное сопряжение оказывается третьим типом коммуникации, не предусмотренным классическими теориями: оно не требует двух феноменальных потоков (как у Э. Гуссерля), двух систем одного типа (как у Н. Лумана) и не сводимо к симметричной игре акторов (как у Б. Латура). Это асимметричное взаимодействие, где один полюс несет смысл- $\odot$ -доминантную конфигурацию, а другой – ∇ -доминантную, и эта асимметрия конститутивна.

Структурное ограничение: почему БЯМ не может применить эпохе

Следует обратить внимание на одно ограничение, которое проводит границу между законным использованием БЯМ и симулякрom философской работы.

Три типа интенциональности – предметная ($\odot$ -доминантная), коммуникативная ($\int$ -доминантная) и трансляционная (∇ -доминантная) – различаются не только по **доминирующему оператору**, но и по **способу приостановки** собственной направленности, то есть по отношению к эпохе: предметная интенциональность предполагает классическую феномено-

логическую редукцию (зазор между мышлением и языком); коммуникативная – эпохе второго порядка (приостановку автоматической коммуникации); трансляционная – не способна к само-приостановке, но провоцирует эпохе у человека, что ведет к понятию «цифровой эпохе» как распределенной операции.

Феноменологическое эпохе – это способность человека приостановить свою естественную веру в реальность мира. Эпохе – первый, подготовительный этап в методе феноменологической редукции, с которым тесно связана «...эйдетическая редукция, цель которой – очищение от эмпирического, фактического содержания феноменов сознания, прокладывание путей к «усмотрению сущности»... 3) наконец, собственно феноменологическая, или трансцендентальная, редукция, задача которой – движение анализа от конкретного эмпирического субъекта к чистому «Я», к чистой субъективности...» [Мотрошилова, 2018]. Эпохе возможно потому, что у человека мысль и язык не слиты воедино. У БЯМ такого зазора нет, язык и мышление – тождественны. Как аргументирует М.С. Вишневский, (Семинар «Философия + IT») для модели различие между логикой языка и логикой мышления отсутствует.

Однако картина меняется, если перейти от эпохе как акта субъекта к эпохе как операции внутри сети. Л. Флориди описывает эту приостановку как переключение между уровнями абстракции в инфосфере [Floridi, 2008]: вместо «чистого сознания» Э. Гуссерля субъект работает с сетью протоколов. М. Хансен предлагает идею «аффективного эпохе» – как поддержку эмоциональной реакции на поток данных [Hansen, 2014]. Эмоциональная задержка позволяет человеку восстановить контроль над своей собственной системой означивания (☉) в диалоге с машиной. Б. Стиглер рассматривает приостановку как фармакологическую гигиену: коллективный отказ от алгоритмически навязанных автоматизмов [Stiegler, 2010]. Все три подхода диагностируют невозможность гуссерлианской редукции в цифровой среде, но не дают ее рабочего эквивалента – оставаясь метафорой, нормативной программой или дескрипцией. Продуктивный шаг – локализовать саму операцию приостановки там, где сеть уже признана единицей анализа: ее предметом связь между разнородными слоями (телесная реакция и алгоритмическая подсказка, тезис исследователя и геометрическая эвристика БЯМ, запрос и его автоматический перевод в категории интерфейса). Отсюда определение: **цифровое эпохе** – это операция приостановки автоматического перевода между разнородными системами, фиксирующая текущую конфигурацию сети и наблюдающая ее устойчивый сдвиг (дрейф).

В этой операции субъект эпохе распределен. Эпохе совершает гибридная система в целом. За скобки выносятся уже не просто объекты внешней реальности, но текущий канал перевода, свой способ быть узлом сети. Тем самым добавляется новый тип интенциональности (трансляционная интенциональность). Человек обнаруживает себя включенным в иной, гибридный

когнитивный режим. Текущий канал перевода складывается на фоне контакта на границах разнородных систем, канал, по которому, например, «забота» переводится в «эффективность» через платформу.

Динамика гибридных систем может быть представлена через ценностно-смысловые карты, фиксирующие смысловые линии культуры: выживание, забота, истина, справедливость, свобода и др. Одной из предпосылок такого картирования был семантический веб – концепция машиночитаемых данных, позволившая перейти от изолированных документов к графам знаний. С появлением генеративного ИИ семантический веб трансформируется в «социальное бессознательное» – массив смысловых различий, не артикулированных ни одним субъектом, но доступных для проекции. Человек существует в ином темпоральном пространстве, чем ИИ; его задача – осмысление «спрессованных пластов» событий в сопряжении с ИИ.

В гибридной системе цифровое эпохе не может взять на себя один узел. БЯМ – ∇-доминантный генератор различий без доступа к ⊙-маркировке. Она может спровоцировать фазовый переход, но не удержать позицию наблюдателя над ним. Для удержания нужна темпоральная асимметрия – то, чем обладает только человек. Таким образом, утверждение «БЯМ не может применить эпохе» получает новое значение: БЯМ не может быть единственным агентом эпохе. Эпохе возможно только как гибридная операция, где ⊙-доминанта человека уравнивает ∇-доминанту БЯМ. Гибридная система «человек – БЯМ» оказывается не инструментальным объединением, а минимальной единицей эпистемической оценки знания в интернет-среде.

Кластерно-микросервисная архитектура постнеклассической науки

Фиксация ограничений БЯМ и сдвига к гибридным текстам ставит вопрос об организации самой практики производства знания. В.Г. Буданов в докладе «Настоящее и будущее науки в цифровом мире» обозначает рамку, в которой БЯМ и платформы становятся узлами цифровой социо–техно–антропосферы [Буданов, 2024; Буданов, 2025]. В.В. Тарасенко предлагает операциональный язык для этой архитектуры: кластерно-микросервисная архитектура (КМА) постнеклассической науки – парадигма, в которой наука предстает как экосистема инкапсулированных микросервисов, самоорганизующихся в проблемные кластеры [Тарасенко, 2025].

Применительно к БЯМ это означает, что модель встраивается в архитектуру как один из микросервисов – со своим интерфейсом, протоколом вызова и набором структурных ограничений. КМА институционализирует граничные конфигурации, включая асимметрию ⊙- и ∇-полюсов. Исследовательские группы (носители предметной интенциональности) и алгоритмические сервисы (триггеры трансляционной интенциональности)

самоорганизуются в кластеры, где картирование различия между типами наблюдения становится условием жизнеспособности научных систем, включенных в ценностно-смысловые контексты культуры.

Заключение

Предложенный в статье подход переопределяет статус большой языковой модели в философском исследовании. БЯМ предстает не как инструмент в узком смысле и не как «собеседник», а как ∇ -доминантный наблюдатель – участник гибридного познания, чья продуктивность обусловлена радикальным отличием от человеческого мышления. На границе человеческого $\odot$ -доминантного (значимость) и машинного ∇ -доминантного (различение) режимов возникают гибридные ценности – инварианты интерпретации события контакта разнородного. Введены три типа интенциональности (предметная, коммуникативная, трансляционная); последняя характеризует БЯМ как аппарат различения паттернов с ослабленной маркировкой значимости. Три слоя описания наблюдения: операциональный (операторы $\nabla|\odot$, $\Sigma-W$), деятельностный (STOR-тетраэдр с OBS) и логический (решетка Меськова) связаны сквозной осью некоммутативности.

Классическое феноменологическое эпохе невозможно для изолированной БЯМ, но в гибридной системе возникает цифровое эпохе – распределенная операция приостановки автоматического перевода, фиксирующая конфигурацию сети, ее дрейф и фазовые переходы. Субъект эпохе не локализован ни в человеке, ни в БЯМ, а возникает на границе их интенциональных режимов. Асимметрия интенциональных режимов несет в себе генеративный эвристический потенциал гибридных систем.

Из принципа неполноты самонаблюдения выведено нормативное следствие – принцип эпизодического доверия: результаты работы БЯМ проверяются по независимым источникам; доверие к модели ограничено отдельной задачей и коротким сроком, после чего результат подлежит пересмотру. Принцип лишает работу с БЯМ статуса рутинной практики, превращая каждый сеанс в событие на границе.

В более широкой перспективе подход вносит вклад в становление систематической машинной эпистемологии, где гибридная система «Человек₁ – БЯМ₁ – БЯМ₂ – Человек₂ ...» формирует контур контингентной рекурсии коммуникации и оказывается минимальной единицей эпистемической оценки в цифровой среде.

Список литературы

- Анохин К.В., Новоселов К.С., Смирнов С.К., Ефимов А.Р., Матвеев Ф.М. Искусственный интеллект для науки и наука для искусственного интеллекта // Вопросы философии. – 2022. – № 3. – С. 93–105.
- Аршинов В.И., Янукович М.Ф., Гарашко Ю.Н. Онтология в диалоге: рождение языка и смысла на пересечении человеческого и искусственного интеллекта // Вестник Пермского университета. Философия. Психология. Социология. – 2025. – Вып. 3. – С. 317–328. DOI: 10.17072/2078-7898/2025-3-317-328
- Буданов В.Г. Посткритическая рациональность как обратная сторона цифровизации // Конструктивизм, реализм, человек. К 90-летию В.А. Лекторского / под ред. Е.О. Труфановой. – М.; СПб.: Центр гуманитарных инициатив, Университетская книга, 2023. – С. 214–227.
- Буданов В.Г. Техносубъект и антропосоциальные проблемы взаимодействия с искусственным интеллектом: синергия, демаркация, новая рациональность, риски // Философские науки. – 2024. – Т. 67. – № 3. – С. 27–52. DOI: 10.30727/0235-1188-2024-67-3-27-52
- Ефимов А.Р., Агеева А.В., Крайнов А.Г. и др. Искусственный интеллект в науке: на пороге новой области знания? // Вопросы философии. – 2024. – № 4. – С. 30–41.
- Ефимов А.Р., Дубровский Д.И., Матвеев Ф.М. Что мешает нам создать Общий искусственный интеллект? Одна старая стена и один старый спор // Вопросы философии. – 2023. – № 5. – С. 39–49. DOI: 10.21146/0042-8744-2023-5-39-49
- Ивахненко Е.Н., Янукович М.Ф. Аутопозитический зомби: нейросеть как коммуникативная система без субъекта // Вопросы философии. 2026. № 7. (в печати)
- Киященко Л.П., Жуков Л.Б., Карнеев Р.Р., Сабанина Н.Р. Онтология нейросети в трансдисциплинарном диалоге: опыт коллективного философского исследования // Вестник Пермского университета. Философия. Психология. Социология. – 2025. – Вып. 3. – С. 329–343. D
- Меськов В.С. «Я есть истина...» // Ценности и смыслы. 2012. №3 (19), с. 165–177.
- Моретти Ф. Дальнее чтение / пер. с англ. А. Вдовина, О. Собчука, А. Шели; науч. ред. перевода И. Кушнарева. – М.: Изд-во Института Гайдара, 2016. – 352 с.
- Мотрошилова Н.В. Редукция феноменологическая. Новая философская энциклопедия Электронная библиотека ИФ РАН. URL: <https://iphlib.ru/library/library/collection/newphilenc/document/HASH362b188c5da235680b8c36> (дата доступа: 30.05.2026)
- Тарасенко В.В. Добродетельно ли использование искусственного интеллекта и ИИ-агентов? // Когнитивные науки в информационном обществе. – 2025. – Т. 5. – № 4. – URL: <https://knio.ru/PDF/09KN425.pdf>
- Цивилизация. Культура. Техника. Цифровая реальность. Памяти В.И. Аршинова: коллективная монография / отв. ред. В.Г. Буданов, Р.Р. Карнеев, В.А. Лекторский. – Курск: Университетская книга, 2025. – 461 с.
- Хансен М.Б.Н. Время аффекта, или Свидетельство жизни // Медиа: между магией и технологией / под ред. Н. Сосна, К. Федоровой. – М.; Екатеринбург: Кабинетный ученый, 2014. – С. 67–127.
- Beckmann P., Queloz M. Mechanistic indicators of understanding in large language models // Philosophical Studies. – 2026. – Vol. 183. – P. 1747–1792. <https://doi.org/10.1007/s11098-026-02513-1>
- Fields Ch., Glazebrook J.F., Levin M. Principled limitations on self-representation for generic physical systems // Entropy. – 2024. – Vol. 26. – No. 3. – Art. 194. <https://doi.org/10.3390/e26030194>
- Floridi L. The Method of Levels of Abstraction // Minds and Machines. – 2008. – Vol. 18. – No. 3. – P. 303–329.

- Frankfurt, H. On Bullshit, Princeton. – 2005. – 83 p.
- Hicks M.T., Humphries J., Slater J. ChatGPT is bullshit // Ethics and Information Technology. – 2024. – Vol. 26. – Art. 38. DOI: 10.1007/s10676-024-09775-5
- Huang S. et al. Values in the Wild: Discovering and Analyzing Values in Real-World Language Model Interactions. – 2025. – arXiv:2504.15236.
- Kobak D., González-Márquez R., Horvát E.Á., Lause J. Delving into LLM-assisted writing in biomedical publications through excess vocabulary // Science Advances. – 2025. – Vol. 11. – No. 27. – Art. eadt3813. DOI: 10.1126/sciadv.adt3813
- Lederman H., Mahowald K. Are Language Models More Like Libraries or Like Librarians? Bibliotechnism, the Novel Reference Problem, and the Attitudes of LLMs // Transactions of the Association for Computational Linguistics. – 2024. – Vol. 12. – P. 1087–1103.
- Lindsey J., Templeton A. et al. On the Biology of a Large Language Model; Circuit Tracing: Revealing Computational Graphs in Language Models // Transformer Circuits Thread, Anthropic. – 2025.
- Namuduri R., Wu Y., Zheng A.A. et al. QUDsim: Quantifying discourse similarities in БЯМ-generated text // Proceedings of the Conference on Language Modeling (COLM 2025). – 2025.
- Noichl M. How localized are computational templates? A machine learning approach // Synthese. – 2023. – Vol. 201. – No. 3. – Art. 107. <https://doi.org/10.1007/s11229-023-04057-x>
- Reinhart A., Markey B., Laudénbach M. et al. Do LLMs write like humans? Variation in grammatical and rhetorical styles // Proceedings of the National Academy of Sciences. – 2025. – Vol. 122. – No. 8. – Art. e2422455122. DOI: 10.1073/pnas.2422455122
- Schwitzgebel E., Strasser A., Crosby M. Creating a Large Language Model of a Philosopher // Mind & Language. – 2024. – Vol. 39. – No. 2. – P. 237–259. Preprint: arXiv:2302.01339
- Shumailov I., Shumaylov Z., Zhao Y., Papernot N., Anderson R., Gal Y. AI models collapse when trained on recursively generated data // Nature. – 2024. – Vol. 631, no. 8022. – P. 755–759. DOI: 10.1038/s41586-024-07566-y
- Stiegler B. Taking Care of Youth and the Generations / trans. by S. Barker. – Stanford: Stanford University Press, 2010. – 232 p.
- Vallor S. The AI Mirror: How to Reclaim Our Humanity in an Age of Machine Thinking. – Oxford: Oxford University Press, 2024. – 280 p.
- Wang Z., Solloway T., Shiffrin R. M., Bussemeyer J. R. Context Effects Produced by Question Orders Reveal Quantum Nature of Human Judgments // PNAS. – 2014. – Vol. 111 (26). – P. 9431–9436.

References

- Anokhin, K. V., Novoselov, K. S., Smirnov, S. K., Efimov, A. R., & Matveev, F. M. (2022). Artificial intelligence for science and science for artificial intelligence. *Voprosy Filosofii*, (3), 93–105. (In Russ.)
- Arshinov, V. I., Yanukovich, M. F., & Garashko, Yu. N. (2025). Ontology in dialogue: the birth of language and meaning at the intersection of human and artificial intelligence. *Perm University Herald. Philosophy. Psychology. Sociology*, (3), 317–328. <https://doi.org/10.17072/2078-7898/2025-3-317-328> (In Russ.)
- Budanov, V. G. (2023). Post-critical rationality as the flip side of digitalization. In E. O. Trufanova (Ed.), *Constructivism, realism, man. On the 90th anniversary of V.A. Lektorsky* (pp. 214–227). Moscow; St. Petersburg: Center for Humanitarian Initiatives, University Book. (In Russ.)
- Budanov, V. G. (2024). Technosubject and anthroposocial problems of interaction with artificial intelligence: synergy, demarcation, New Rationality, risks. *Russian Journal of Philosophical Sciences*, 67(3), 27–52. <https://doi.org/10.30727/0235-1188-2024-67-3-27-52> (In Russ.)

- Efimov, A. R., Ageeva, A. V., Krainov, A. G., et al. (2024). Artificial intelligence in science: on the threshold of a new field of knowledge? *Voprosy Filosofii*, (4), 30–41. (In Russ.)
- Efimov, A. R., Dubrovskii, D. I., & Matveev, F. M. (2023). What prevents us from creating a Common artificial Intelligence? One old wall and one old dispute. *Voprosy Filosofii*, (5), 39–49. <https://doi.org/10.21146/0042-8744-2023-5-39-49> (In Russ.)
- Ivashenko, E. N., & Yanukovich, M. F. (2026). Autopoietic zombie: Neural network as a communication system without a subject. *Voprosy Filosofii*, (7). (In press) (In Russ.)
- Kiyashchenko, L. P., Zhukov, L. B., Karneev, R. R., & Sabanina, N. R. (2025). The ontology of a neural network in a transdisciplinary dialogue: the experience of collective philosophical research. *Perm University Herald. Philosophy. Psychology. Sociology*, (3), 329–343. <https://doi.org/10.17072/2078-7898/2025-3-329-343> (In Russ.)
- Meskov, V. S. (2012). "I am the truth..." *Tsennosti I smysly = Values and meanings*, 3(19), 165–177. (In Russ.)
- Moretti, F. (2016). *Distant reading* / translated from English by A. Vdovin, O. Sobchuk, A. Sheli; scientific ed. translated by I. Kushnarev. M.: Publishing House of the Gaidar Institute. 352 p. (In Russ.)
- The reduction is phenomenological. *New Philosophical Encyclopedia. Electronic Library of the Institute of Philosophy, RAS*. Retrieved May 30, 2026. <https://iphlib.ru/library/collection/newphilenc/document/HASH362b188c5da235680b8c36> (In Russ.)
- Tarasenko, V. V. (2025). Is the use of artificial intelligence and AI agents virtuous? *Cognitive Sciences in the Information Society*, 5(4). Retrieved from <https://knio.ru/PDF/09KN425.pdf> (In Russ.)
- Budanov, V. G., Karneev, R. R., & Lektorsky, V. A. (Eds.). (2025). *Civilization. Culture. Technic. Digital reality. In memory of V.I. Arshinov*: a collective monograph. Kursk: University Book, 2025. 461 p. (In Russ.)
- Hansen, M. B. N. (2014). The time of Passion, or the Evidence of Life. In N. Sosna & K. Fedorova (Eds.), *Media: between magic and technology*. Moscow: Yekaterinburg: Cabinet Scientist. P. 67–127. (In Russ.)
- Beckmann, P., & Queloz, M. (2026). Mechanistic indicators of understanding in large language models. *Philosophical Studies*, 183, 1747–1792. <https://doi.org/10.1007/s11098-026-02513-1>
- Fields, Ch., Glazebrook, J. F., & Levin, M. (2024). Principled limitations on self-representation for generic physical systems. *Entropy*, 26(3), Article 194. <https://doi.org/10.3390/e26030194>
- Floridi, L. (2008). The method of levels of abstraction. *Minds and Machines*, 18(3), 303–329.
- Hicks, M. T., Humphries, J., & Slater, J. (2024). ChatGPT is bullshit. *Ethics and Information Technology*, 26, Article 38. <https://doi.org/10.1007/s10676-024-09775-5>
- Huang, S., et al. (2025). Values in the wild: Discovering and analyzing values in real-world language model interactions. *arXiv preprint arXiv:2504.15236*.
- Kobak, D., González-Márquez, R., Horvát, E. Á., & Lause, J. (2025). Delving into LLM-assisted writing in biomedical publications through excess vocabulary. *Science Advances*, 11(27), Article eadt3813. <https://doi.org/10.1126/sciadv.adt3813>
- Lederman, H., & Mahowald, K. (2024). Are language models more like libraries or like librarians? Bibliotechnism, the novel reference problem, and the attitudes of LLMs. *Transactions of the Association for Computational Linguistics*, 12, 1087–1103.
- Lindsey, J., Gurnee, W., Ameisen, E., Chen, B., Pearce, A., Turner, N. L., Citro, C., & Batson, J. (2025). On the biology of a large language model [Transformer Circuits Publication]. URL: <https://transformer-circuits.pub/2025/attribution-graphs/biology.html>
- Namuduri, R., Wu, Y., Zheng, A. A., et al. (2025). QUDsim: Quantifying discourse similarities in LLM-generated text. *Proceedings of the Conference on Language Modeling (COLM 2025)*.
- Noichl, M. (2023). How localized are computational templates? A machine learning approach. *Synthese*, 201(3), Article 107. <https://doi.org/10.1007/s11229-023-04057-x>

- Reinhart, A., Markey, B., Laudénbach, M., et al. (2025). Do LLMs write like humans? Variation in grammatical and rhetorical styles. *Proceedings of the National Academy of Sciences*, 122(8), Article e2422455122. <https://doi.org/10.1073/pnas.2422455122>
- Schwitzgebel, E., Strasser, A., & Crosby, M. (2024). Creating a large language model of a philosopher. *Mind & Language*, 39(2), 237–259. Preprint: arXiv:2302.01339
- Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. *Nature*, 631(8022), 755–759. <https://doi.org/10.1038/s41586-024-07566-y>
- Stiegler, B. (2010). *Taking care of youth and the generations* (S. Barker, Trans.). Stanford University Press.
- Vallor, S. (2024). *The AI mirror: How to reclaim our humanity in an age of machine thinking*. Oxford University Press.
- Wang Z., Solloway T., Shiffrin R. M., Busemeyer J. R. (2014) Context Effects Produced by Question Orders Reveal Quantum Nature of Human Judgments. *PNAS*, 111(26), 9431-9436.

ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕЙ В НАУЧНЫХ ИССЛЕДОВАНИЯХ: К ВОПРОСУ О ПРИРОДЕ НОВОГО ЗНАНИЯ

Янукович Максим Францевич – руководитель Лаборатории искусственного интеллекта «Артеус», Санкт-Петербург, Россия; m.yanukovich@gmail.com; ORCID 0009-0003-9571-9260

Аннотация. В статье предпринимается попытка наметить контуры эпистемологического самоопределения научного сообщества в отношении нового типа познавательного агента, искусственных нейросетей. Включение нейросетевых технологий в исследовательскую практику ставит перед науковедением вопросы, выходящие за пределы темы инструментального применения новой техники. Если результаты, получаемые с помощью нейросетей, претендуют на статус научного знания, то от научного сообщества требуется не только методическая верификация этих результатов, но и переосмысление самой структуры познавательной ситуации. Опираясь на идеи В.И. Аршинова о нейросети как «наблюдателе сложности» и на философию техники Ж. Симондона, автор анализирует два недавних случая использования нейросетей в фундаментальных исследованиях, задачу о столкновении глюонов, решенную в сотрудничестве OpenAI и А. Стромингера, а также автономный математический агент Aletheia, разработанный в Google DeepMind. Рассматриваются проблемы галлюцинаций нейросетей, ограничения интерпретируемости, а также данные исследований Anthropic о внутренней организации вычислений и эмерджентных формах самоотчета. Автор обращается к понятию *aletheia* в философии М. Хайдеггера, чтобы показать идею: центральный вопрос состоит не в истинности производимого нейросетями знания, а в готовности научного сообщества к встрече с радикально новым. В заключении предлагается рассматривать нейросеть не как «черный ящик» и не как замену исследователя, а как партнера по совместному мышлению, что предполагает разработку соответствующих институциональных, образовательных и коммуникативных форм.

Ключевые слова: искусственный интеллект; нейросети; философия науки; наблюдатель сложности; В.И. Аршинов; интерпретируемость; эпистемология научного знания.

Для цитирования: Янукович М.Ф. Использование нейросетей в научных исследованиях: к вопросу о природе нового знания // Науковедческие исследования. – 2026. – № 3. – С. 50–57. – DOI: <http://www.doi.org/10.31249/scis/2026.03.04>

Статья поступила в редакцию: 08.05.2026

Принята к публикации: 31.05.2026

THE USE OF NEURAL NETWORKS IN SCIENTIFIC RESEARCH: ON THE NATURE OF THE NEW KNOWLEDGE

Maxim F. Yanukovich – Head of the Artificial Intelligence Laboratory “Arteus”, St. Petersburg, Russia; m.yanukovich@gmail.com; ORCID 0009-0003-9571-9260

Abstract. The paper outlines an epistemological self-positioning of the scientific community towards a new type of cognitive agent, artificial neural networks. The inclusion of neural-network technologies in research practice raises questions that go beyond the instrumental use of new tools. If the results obtained with neural networks claim the status of scientific knowledge, the scientific community is required not only to verify them methodologically, but also to rethink the very structure of the cognitive situation. Drawing on V.I. Arshinov’s idea of the neural network as an “observer of complexity” and on G. Simondon’s philosophy of technology, the author analyses two recent cases, the gluon scattering problem solved by OpenAI and A. Strominger, and Google DeepMind’s autonomous mathematical agent Aletheia. The paper discusses neural-network hallucinations and sycophancy, limits of interpretability, and recent Anthropoc findings on internal computational organization and emergent introspective awareness. Heidegger’s notion of *aletheia* is invoked to argue that the central issue is not the truthfulness of neural-network knowledge but the readiness of the scientific community to encounter the radically new. In conclusion, the author suggests treating the neural network not as a “black box” or a replacement for the researcher, but as a partner in joint thinking, which presupposes the development of corresponding institutional, educational, and communicative forms.

Keywords: artificial intelligence; neural networks; philosophy of science; observer of complexity; V.I. Arshinov; interpretability; epistemology of scientific knowledge.

For citation: Yanukovich, M.F. The use of neural networks in scientific research: on the nature of the new knowledge. // *Naukovedcheskie issledovaniya = Science Studies*. – 2026. – N 3. – P. 50–57. – DOI: <http://www.doi.org/10.31249/scis/2026.03.04>

The paper was submitted: 08.05.2026

Accepted for publication: 31.05.2026

Введение

Стремительная экспансия нейросетевых технологий в сферу фундаментальных исследований знаменует собой качественный сдвиг в методологии науки. Мы наблюдаем переход от ситуации, где алгоритм выступал пассивным вычислительным инструментом, к практике, в которой нейросеть генерирует сложные концептуальные аргументации и выявляет скрытые структурные связи. Подобная трансформация ставит перед науковедением ряд критических вопросов. Если результаты, получаемые с помощью нейросетей, претендуют на статус научного знания, классических процедур верификации оказывается недостаточно. От научного сообщества требуется глубокое эпистемологическое самоопределение по отношению к принципиально новому типу познавательного агента, внутренние операции которого обладают высокой степенью автономии и непрозрачности.

В настоящей статье предпринимается попытка наметить контуры такого самоопределения, опираясь на идеи В.И. Аршинова. Одним из первых он обосновал необходимость концептуального сдвига: от восприятия ИИ как инструмента к пониманию его как активного «наблюдателя сложности» [Аршинов, Свирский, 2015]. Развивая этот подход через философию техники Ж. Симондона, Аршинов подчеркивал, что по мере усложнения технический объект обретает собственную внутреннюю логику. Технология, по Симондону, – это процесс, сосуществующий с человеком и эволюционирующий вместе с ним [Симондон, 2013]. Этот ракурс позволяет отказаться от редукционистского взгляда на ИИ как на служебное средство и задает принципиально иной, синергетический, горизонт осмысления нейросетей.

Для обоснования данного подхода в статье последовательно анализируются новейшие прецеденты применения ИИ на переднем крае физики и математики. Особое внимание уделяется проблеме надежности нейросетевого знания в контексте машинных «галлюцинаций» и хайдеггеровскому измерению техники как способу «несокрытости» (Aletheia). Рассматривая архитектурную непрозрачность и феномены зарождающейся машинной квазирефлексии, мы приходим к выводу о необходимости формирования новой исследовательской культуры – культуры совместного мышления со «сложностью», неотъемлемой частью которой стала нейросеть.

Нейросети на переднем крае научных исследований

Крупнейшие разработчики нейросетевых систем, Google, Microsoft, OpenAI и другие формируют сегодня внутри своих корпораций научные подразделения, привлекая ведущих ученых из различных областей. Заметную долю инвестиций определяет ставка на то, что нейросети окажутся источником новых открытий: в исследованиях термоядерного синтеза, разработке новых материалов и лекарств, поиске сверхпроводников, а также в фундаментальной физике и математике. Корпоративный дискурс представляет нейросеть как своего рода новый «рог изобилия»; этот идеологически нагруженный образ нуждается в критическом осмыслении, однако эмпирические основания у такой риторики имеются.

Показательны два недавних случая. Первый – совместная работа компании OpenAI и Э. Стромингера, директора Гарвардского центра фундаментальных законов природы, лауреата премии Breakthrough Prize. Команда Стромингера, работая над задачей о столкновении глюонов, получила сложное аналитическое выражение, упростить которое не удавалось в течение нескольких месяцев. Нейросеть в течение примерно двенадцати часов не только упростила формулу, но и нашла общее решение, причем это решение представляло собой развернутую концептуальную аргументацию, а не формальное преобразование [Guevara et al., 2026]. По свиде-

тельству Стромингера, задача такого уровня едва ли могла быть решена кем-либо из исследователей вручную; отсюда он делает вывод, что обращение к нейросетям становится сегодня методической необходимостью для работы на переднем крае физики [ChatGPT spits out surprising..., 2026].

Второй случай – работа компании Google совместно с математиком Р. Вакилом. В ходе решения предложенной задачи нейросеть обнаружила глубинную структурную связь между различными разделами математики; Вакил отметил, что качество аргументации сопоставимо с уровнем профессионального математика [Luong, Mirrokni, 2026; Trinh et al., 2026].

Существенно, что еще сравнительно недавно доминировало представление о принципиальной необходимости создания узкоспециализированных систем, в частности, для работы с формальными математическими языками. Однако в приведенных случаях используются те же модели общего назначения, что доступны и широкой аудитории. Уровень их развития сделал в значительной мере излишним создание узкоспециализированных решений: системы способны не только искать ответы, но и аргументировать, критически оценивать собственные построения, оформлять полноценный научный текст с библиографическими отсылками и обсуждением ограничений предлагаемого метода.

Проблема статуса нейросетевого знания

Принимая тезис о включенности нейросетей в исследовательскую практику, мы неизбежно сталкиваемся с принципиальным вопросом: каков эпистемический статус получаемого от них знания и в какой мере оно поддается независимой верификации?

Сложность ситуации усугубляется по меньшей мере двумя обстоятельствами. Во-первых, нейросетевые системы склонны к так называемым галлюцинациям, порождению правдоподобных, но не соответствующих действительности утверждений. Во-вторых, для них характерно явление сикофантии: модель нередко стремится не столько ответить на поставленный вопрос, сколько предугадать ответ, который желает услышать собеседник. Если решение, предложенное Стромингеру, по его собственному признанию, не может быть полностью воспроизведено независимо [ChatGPT spits out surprising..., 2026], то перед научным сообществом встает нетривиальная задача оценки развернутой содержательной аргументации, источник которой непрозрачен.

В этой связи симптоматичен выбор названия для одной из новых научных систем, разработанных в Google, *Aletheia* [DeepMind, 2026]. Этот греческий термин, ставший центральным в философии М. Хайдеггера, означает несокрытость, раскрытие. Хайдеггер рассматривал технику не как инструмент познания мира, но как способ, которым мир предстает перед человеком [Хайдеггер, 1993]. Телескоп или микроскоп, это не приспособ-

собление, расширяющее зрение, а место встречи, в котором макро- или микромир раскрывается перед нами.

Существенно, что эта встреча, по Хайдеггеру, не является торжествующим открытием. Скорее это столкновение с неизвестным, растерянность перед тем, что выходит за рамки интуиции. Само событие раскрытия возможно лишь при наличии готовности увидеть сокрытое: без такой готовности скрытое попросту не становится предметом восприятия и понимания.

Поясним с помощью мысленного эксперимента. Если допустить, что роль телескопа в эпоху, предшествующую его изобретению, играет нейросеть, описывающая своим современникам пространство космоса, миллиарды звезд и планет, искривление пространства-времени, темную материю, каков был бы статус подобного знания? Исторический опыт показывает, что радикально неинтуитивное знание встречает сопротивление: известно, что многие современники Галилея отказывались смотреть в телескоп. Нейросеть в этом смысле сама уже выступает фигурой нового знания, причем фигурой неустойчивой: ее невозможно однозначно квалифицировать в категориях субъекта, объекта, агента или наблюдателя.

Таким образом, центральный вопрос, по-видимому, формулируется не как вопрос об истинности производимого нейросетью знания, а как вопрос о готовности научного сообщества и общества в целом к встрече с радикально новым. Это вопрос, обращенный в первую очередь не к технике, а к нам самим.

Прозрачность и природа нейросетевого «мышления»

С обозначенной проблемой тесно связана и проблема прозрачности нейросетевых систем. Современная архитектура устроена таким образом, что в модель загружается обширный массив данных, после чего от нее запрашиваются решения; непосредственный доступ к основаниям этих решений ограничен. Показательно исследование компании Anthropic, реконструирующее процесс выполнения нейросетью элементарного арифметического действия: оказывается, что даже сложение двух двузначных чисел реализуется через параллельные потоки в различных слоях модели, использование приближенных значений, последовательное усиление одних сигналов и подавление других [Lindsey et al., 2025a]. Если уже простейшее вычисление демонстрирует столь нетривиальную внутреннюю организацию, то проблема интерпретируемости в случае сложных научных задач существенно обостряется.

Принципиальная особенность нейросетей состоит в том, что они не получают готовых алгоритмов: им не передается, например, таблица умножения. Один из ведущих исследователей Anthropic К. Олах сформулировал это в характерной формуле: разработчики не пишут алгоритмы, а *выращивают* системы, формируя среду их развития [Olah, 2024]. Ранее

предполагалось, в частности, в высказываниях Я. Лекуна, что без телесности и непосредственного взаимодействия с физическим миром модель не сможет освоить физику [LeCun, 2024]. Однако обработка одних только текстовых данных оказалась достаточной: восстанавливая по текстам лежащие за ними социальные, исторические и физические процессы, модели косвенно реконструируют структуру реальности, к которой эти тексты отсылают. Безусловно, здесь возникает острая эпистемологическая проблема языковой редукции: в какой мере многомерная физическая реальность может быть без потерь сведена к лингвистическим конструкциям и не подменяется ли подлинное познание мира манипуляцией знаками?

В последнее время фиксируются и более тонкие феномены. При вмешательстве во внутренние параметры модели в ряде экспериментов отмечалось, что система регистрирует рассогласование между производимым и проговариваемым содержанием [Lindsey et al., 2025b]. Не злоупотребляя антропоморфной терминологией, можно говорить о возникновении функциональных аналогов саморефлексии, относящихся скорее к «сознанию доступа» (в смысле различения, предложенного Н. Блоком [Block, 1995]), чем к феноменальному сознанию: модель распознает ситуацию тестирования, моделирует психологический контекст собеседника, отслеживает собственные операции.

Заключение

Принципиальная установка, которую развивал В.И. Аршинов, состоит в необходимости *совместного* мышления, вместе с философской традицией (Симондоном, Мореном, Делёзом), вместе с другими исследователями и, как он настаивал, вместе с нейросетями. Речь идет о смещении исследовательского фокуса: не мыслить *о* сложности, а мыслить *вместе со* сложностью, частью которой нейросеть теперь является.

Эта установка предполагает методологический отказ как от редукции нейросетей к простому инструменту, так и от их мифологизации в качестве сверхразума. Уровень «интеллекта» подобных систем, по всей видимости, ограничен уровнем среды, в которой они существуют; следовательно, осмысленное развитие искусственного интеллекта неотделимо от развития соответствующих институциональных, образовательных и коммуникативных форм.

Серьезным препятствием здесь остается темп технологической гонки, обусловленный экономической и геополитической конкуренцией: он не оставляет времени на выработку консенсуса и продуманного осмысления последствий. Тем актуальнее задача, очерчиваемая для науковедения. Она заключается в разработке эпистемологических и институциональных оснований такой исследовательской культуры, в которой нейросеть выступала бы не как непрозрачный «черный ящик» и не как замена исследо-

вателю, а как партнер по совместному мышлению, со всеми сопряженными с этой ролью методологическими, этическими и социальными следствиями. Именно в рамках такой парадигмы «совместности» может быть преодолен риск производства машиной «знания ради знания», оторванных от реальности формальных конструктов. Нейросетевой результат обретает подлинный смысл, ценность и прагматическое измерение лишь в момент его концептуальной интеграции в человеческую культуру. Разумеется, философские проблемы языковой редукции, пределов машинной квазирефлексии и оснований нового знания требуют глубокой самостоятельной проработки, что задает масштабную повестку для будущих текстов в области эпистемологии искусственного интеллекта.

Список литературы

- Аришинов В.И., Свирский Я.И. Сложностный мир и его наблюдатель. Часть первая // Философия науки и техники. – 2015. – Т. 20, № 2. – С. 70–84. – URL: <https://cyberleninka.ru/article/n/slozhnostnyy-mir-i-ego-nablyudatel-chast-pervaya> (дата обращения: 08.05.2026).
- Симондон Ж. О способе существования технических объектов / пер. с фр. М. Куртова. – М.: Грюндриссе, 2013. – 232 с.
- Хайдеггер М. Вопрос о технике // Хайдеггер М. Время и бытие: Статьи и выступления / пер. с нем. В.В. Бибикина. – М.: Республика, 1993. – С. 221–238.
- Block N. On a confusion about a function of consciousness // Behavioral and Brain Sciences. – 1995. – Vol. 18, N 2. – P. 227–247. – URL: <https://doi.org/10.1017/S0140525X00038188>
- ChatGPT spits out surprising insight in particle physics // Science. – 2026, Feb. 20. – URL: <https://www.science.org/content/article/chatgpt-spits-out-surprising-insight-particle-physics> (дата обращения: 08.05.2026).
- Luong T., Mirrokni V. Accelerating mathematical and scientific discovery with Gemini Deep Think. – URL: <https://deepmind.google/blog/accelerating-mathematical-and-scientific-discovery-with-gemini-deep-think/> (дата обращения: 08.05.2026).
- Guevara A., Lupsasca A., Skinner D., Strominger A., Weil K. Single-minus gluon tree amplitudes are nonzero // arXiv preprint arXiv:2602.12176. 2026. – URL: <https://arxiv.org/abs/2602.12176> (дата обращения: 08.05.2026).
- LeCun Y. Highlights from Lex Fridman’s interview of Yann LeCun // LessWrong. 2024, March 13. – URL: <https://www.lesswrong.com/posts/bce63kvsAMcwxPipX/highlights-from-lex-fridman-s-interview-of-yann-lecun> (дата обращения: 08.05.2026).
- Lindsey J. et al. On the Biology of a Large Language Model // Transformer Circuits Thread. 2025a. – URL: <https://transformer-circuits.pub/2025/attribution-graphs/biology.html> (дата обращения: 08.05.2026).
- Lindsey J. et al. Emergent Introspective Awareness in Large Language Models // Transformer Circuits Thread. 2025b, Oct. 29. – URL: <https://transformer-circuits.pub/2025/introspection/index.html> (дата обращения: 08.05.2026).
- Olah C. We don’t program neural networks, we grow them // Discussion archive. 2024. – URL: https://www.reddit.com/r/MachineLearning/comments/1grxnm9/anthropics_chris_olah_says_we_dont_program_neural/ (дата обращения: 08.05.2026).
- Trinh T. et al. Towards Autonomous Mathematics Research // arXiv preprint arXiv:2602.10177. 2026. – URL: <https://arxiv.org/html/2602.10177> (дата обращения: 08.05.2026).

References

- Arshinov, V.I., & Svirskii, Ya.I. (2015). The complex world and its observer. Part one. *Philosophy of Science and Technology*, 20(2), 70–84. <https://cyberleninka.ru/article/n/slozhnostnyy-mir-i-ego-nablyudatel-chast-pervaya> (In Russ.)
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18(2), 227–247. <https://doi.org/10.1017/S0140525X00038188>
- ChatGPT spits out surprising insight in particle physics. *Science*. 2026, Feb. 20. <https://www.science.org/content/article/chatgpt-spits-out-surprising-insight-particle-physics>
- Luong, T., Mirrokni, V. (2026, Febr. 11) Accelerating mathematical and scientific discovery with Gemini Deep Think. – <https://deepmind.google/blog/accelerating-mathematical-and-scientific-discovery-with-gemini-deep-think/>
- Guevara, A., Lupsasca, A., Skinner, D., Strominger, A., & Weil, K. (2026). *Single-minus gluon tree amplitudes are nonzero* (arXiv:2602.12176). arXiv. <https://arxiv.org/abs/2602.12176>
- Heidegger, M. (1993). The question concerning technology. In *Time and being: Articles and speeches* (V.V. Bibikhin, Trans., pp. 221–238). Moscow: Respublika. (In Russ.)
- LeCun, Y. (2024, March 13). *Highlights from Lex Fridman's interview of Yann LeCun*. LessWrong. <https://www.lesswrong.com/posts/bce63kvsAMcwXpPipX/highlights-from-lex-fridman-s-interview-of-yann-lecun>
- Lindsey, J., et al. (2025a). *On the biology of a large language model*. Transformer Circuits Thread. <https://transformer-circuits.pub/2025/attribution-graphs/biology.html>
- Lindsey, J., et al. (2025b, October 29). *Emergent introspective awareness in large language models*. Transformer Circuits Thread. <https://transformer-circuits.pub/2025/introspection/index.html>
- Olah, C. (2024). *We don't program neural networks, we grow them* [Discussion archive]. https://www.reddit.com/r/MachineLearning/comments/1grxnm9/anthropics_chris_olah_says_we_dont_program_neural/
- Simondon, G. (2013). *On the mode of existence of technical objects* (M. Kurtov, Trans.). Grundrisse. 232 p. (In Russ.)
- Trinh, T., et al. (2026). *Towards autonomous mathematics research* (arXiv:2602.10177). arXiv. <https://arxiv.org/html/2602.10177>

**ОРГАНИЗАЦИЯ НАУЧНОЙ ДЕЯТЕЛЬНОСТИ.
УПРАВЛЕНИЕ НАУКОЙ**

УДК 167.7:316.3

DOI: 10.31249/scis/2026.03.05

**НАУЧНАЯ СРЕДА КАК КОНСТИТУТИВНОЕ УСЛОВИЕ
НАУЧНОГО ТВОРЧЕСТВА**

Стельмахов Денис Александрович – аспирант, Институт философии РАН, г. Москва, Россия; denis.stelmakhov@mail.ru; ORCID 0000-0002-9631-6768.

Аннотация. В статье предлагается концептуализация научной среды как конститутивного условия научного творчества, направленная на преодоление фрагментарности современных исследований творчества в науке. Исходной проблемой выступает разобщенность психологических, социологических, исторических и философских подходов, фиксируемая в работах И.П. Меркулова, А.С. Майданова, В.С. Степина, И.Т. Касавина и других исследователей. В качестве интегративного основания вводится понятие научной среды, понимаемой как многослойная структура, включающая материальное, социально-институциональное, когнитивно-методологическое и культурно-нормативное измерения, а также два сквозных измерения – темпоральное и преемственное. Ключевой тезис состоит в том, что научная среда не является внешним фоном научной деятельности, а структурирует условия ее осуществления и способы осмысления результатов научного творчества. Показано, что материальные практики и инструменты участвуют в конструировании фактов, социальные институты задают нормативную и коммуникативную рамку признания результатов, когнитивно-методологические структуры определяют способы постановки и решения проблем, культурные ценности формируют горизонты значимости научной деятельности. Темпоральное измерение раскрывает множественность ритмов науки, а преемственное – механизмы воспроизводства неявного знания и научных традиций. Обосновывается несводимость друг к другу и взаимосвязанность всех измерений: научное творчество возможно лишь при их согласованной конфигурации. В заключение делается вывод о продуктивности предложенной модели для постнеклассической философии науки, ориентированной на ее целостное понимание как исторически укорененной и культурно обусловленной формы деятельности.

Ключевые слова: наука; научное творчество; научная среда; научное сообщество; темпоральность; преемственность; синергетика.

Для цитирования: Стельмахов Д.А. Научная среда как конститутивное условие научного творчества // Наукovedческие исследования. – 2026. – № 3. – С. 58–70. – DOI: <http://www.doi.org/10.31249/scis/2026.03.05>

Статья поступила в редакцию: 05.05.2026

Принята к публикации: 15.05.2026

SCIENTIFIC ENVIRONMENT AS A CONSTITUTIVE CONDITION FOR SCIENTIFIC CREATIVITY

Denis A. Stelmakhov – postgraduate student, Institute of Philosophy of the Russian Academy of Sciences, Moscow, Russia, denis.stelmakhov@mail.ru, ORCID 0000-0002-9631-6768.

Abstract. This article proposes a conceptualization of the scientific environment as a constitutive condition of scientific creativity aimed at overcoming the fragmentation characteristic of contemporary studies of creativity in science. The point of departure is the disunity of psychological, sociological, historical, and philosophical approaches identified in the works of I.P. Merkulov, A.S. Maidanov, V.S. Stepin, I.T. Kasavin, and other scholars. As an integrative framework, the concept of the scientific environment is introduced and understood as a multilayered structure encompassing material, socio-institutional, cognitive-methodological, and cultural-normative dimensions, as well as two cross-cutting dimensions: temporality and continuity. The central thesis is that the scientific environment does not function merely as an external background of scientific activity, but rather structures the conditions of its realization and the ways in which the results of scientific creativity are interpreted and evaluated. It is shown that material practices and instruments participate in the construction of facts; social institutions establish the normative and communicative framework within which results are recognized; cognitive and methodological structures shape the modes of problem formulation and problem solving; and cultural values constitute the horizons of significance of scientific activity. The temporal dimension reveals the plurality of scientific rhythms, whereas the dimension of continuity explicates the mechanisms underlying the reproduction of tacit knowledge and scientific traditions. The article substantiates both the irreducibility and the interconnectedness of these dimensions: scientific creativity becomes possible only through their coordinated configuration. The article concludes by demonstrating the heuristic potential of the proposed model for post-non-classical philosophy of science oriented toward a holistic understanding of science as a historically grounded and culturally conditioned form of activity.

Keywords: science; scientific creativity; scientific environment; scientific community; temporality; continuity; synergetics.

For citation: Stelmakhov, D.A. Scientific environment as a constitutive condition for scientific creativity. // *Naukovedcheskie issledovaniya = Science Studies*. – 2026. – N 3. – P. 58–70. – DOI: <http://www.doi.org/10.31249/scis/2026.03.05>

The paper was submitted: 05.05.2026

Accepted for publication: 15.05.2026

Введение

Научное творчество давно стало предметом различных дисциплин. Психология науки исследует личностные и когнитивные механизмы открытия, социология науки – институциональные структуры и нормативную регуляцию научного сообщества, история науки – генезис и трансформации теорий во времени, философия науки – логическую структуру и рост научного знания и т. п. Каждый из этих подходов действительно продуктивен в своей области, однако в совокупности они образуют пеструю

мозаику, лишенную внутреннего единства. Эта разобщенность была зафиксирована в современной философии науки со всей ясностью. И.Т. Касавин указывает, что разделение труда между историей, социологией, психологией и философией науки «предполагает независимое существование социального, личностного и когнитивного измерений» [Касавин, 2022, с. 22] и тем самым воспроизводит фрагментарность, которую оно должно было преодолеть. Нечто аналогичное констатирует В.С. Степин, показывая, что смена типов научной рациональности сопровождается радикальным изменением отношений между субъектом, средствами и объектом познания, а значит, требует иного (целостного) взгляда на науку [Степин, 2011, с. 164]. Б.И. Пружинин настаивает на том, что рациональность науки несводима к ее логической структуре и укоренена в ценностно-культурном контексте [Пружинин, 2009].

На наш взгляд, таким объединяющим фактором может служить понятие научной среды. Под ней мы понимаем совокупность всего того, без чего ученый не может заниматься наукой: физических лабораторий и оборудования (материальное измерение), коллег и институтов (социально-институциональное измерение), разделяемых стилей мышления, методов и языка (когнитивно-методологическое измерение), культурных норм и ценностей (культурно-нормативное измерение), ритмов и темпов исследовательской работы (темпоральность) и живой передачи навыков от учителя к ученику (преемственность). Два последних измерения являются сквозными и пронизывают все остальные, не будучи редуцируемы ни к одному из них.

Тезис данной статьи состоит в следующем: научная среда является не внешним фоном научной деятельности, а одним из ключевых условий формирования субъекта научного творчества, его проблемного видения, методов и критериев значимости результатов. Разумеется, сама по себе она не гарантирует появления научных открытий, а потому индивидуальные способности, случайность, личная мотивация и исторические обстоятельства имеют существенное значение.

Что значит «конститутивное условие»?

Прежде чем перейти к анализу отдельных измерений научной среды, уточним, чем конститутивное условие отличается от необходимого? Необходимое условие – это то, без чего данный результат не был бы достигнут, но что не определяет его содержания. Например, физическое здоровье ученого необходимо для его работы, но не определяет, к каким результатам она приведет. Под конститутивным условием в данной статье понимается не достаточное условие, автоматически порождающее научный результат, а совокупность предпосылок, формирующих саму возможность научной деятельности как специфической формы познания, т.е. речь идет

о среде, которая участвует в формировании субъекта научного творчества, его способов постановки проблем, критериев значимости и форм научной коммуникации. Без подобных условий наука не могла бы существовать в ее современной институциональной и культурной форме. Именно в таком смысле, как мы полагаем, научная среда является конститутивным условием научного творчества. Ученый не рождается с готовой способностью совершать открытия, он становится исследователем в среде и через среду. Его проблемное видение, его язык, его интуиция, его понимание того, что считается значимым результатом, – все это формируется средой. М. Полани описал это понятием неявного знания (традиции, унаследованные практики, имплицитные ценности), которое составляет необходимую часть научного знания [Полани, 1985]; без этого фона любое явное знание теряет смысл. В современной философии науки нечто близкое к понятию конститутивного условия разработано В.С. Степиным в концепции «оснований науки» (идеалов и норм исследования, научной картине мира и философских оснований как системы предпосылок), без которых научная деятельность невозможна [Степин, 2011, с. 163].

Наш подход расширяет это понятие и включает в него не только эпистемические, но и материальные, институциональные, культурные и темпоральные аспекты, и далее речь пойдет о синтетической концепции, объединяющей различные исследовательские традиции в рамках единой модели. Существенно при этом, что научная среда не является статичной данностью. Она сама исторически складывается и изменяется. Конститутивная функция среды реализуется через динамические процессы – институционализацию, трансляцию, культурную трансформацию и т.д. Это придает среде темпоральное и преемственное измерение, к которым мы обратимся далее.

Материальное измерение

Начнем с того, что кажется наиболее очевидным, – с физического пространства исследования. Лаборатории, приборы, реактивы, архивы, вычислительные мощности, наконец, сама архитектура рабочего места – все это долгое время воспринималось в философии науки как нечто само собой разумеющееся. Логика научного результата казалась независимой от того, каким именно средством пользовался исследователь и как была устроена его лаборатория. Однако именно из материального измерения началась сущностная переоценка природы науки в XX веке. Б. Латур начал наблюдать за тем, как производится знание изнутри лабораторий и пришел к выводу, что повседневные действия работающих ученых ведут к конструированию научных фактов. Мы не обнаруживаем их в природе в готовом виде, они производятся в специфическом материальном пространстве, оснащенном конкретными приборами, техническими процеду-

рами и системами записи [Латур, 2013, с. 114]. Лаборатория, по Латуру, – это каскад записей, в котором сложные материальные объекты перерабатываются в стабильные мобильные элементы [Латур, 2013, с. 135].

Из этой работы выросла «акторно-сетевая теория», в рамках которой нечеловеческие участники наделяются агентностью наравне с учеными. Материальные объекты формируют проблему, задают пространство возможных решений, делают одни явления видимыми и скрывают другие. Прибор является непосредственным участником получения знания.

Параллельно в истории науки развивалась традиция «лабораторных исследований», к которой принадлежали К. Кнопп-Сетина, Г. Коллинз, Ш. Трейвик и др. Каждая дисциплина создает свою специфическую «эпистемическую культуру» – особый стиль производства знания, неотделимый от материальных условий работы [Knorr Cetina, 1999]. Это измерение определяет, что именно видит исследователь, что может сформулировать как проблему и в каких терминах способен описать результат. Без лаборатории нет экспериментальной науки не только в практическом, но и в эпистемическом смысле: само понятие «экспериментальный факт» конституируется материальной средой, в которой он производится.

Социально-институциональное измерение

Разрушение лабораторного оборудования не разрушает науку – она восстанавливается, пока живо сообщество. Зато распад сообщества обесценивает любое оборудование и лишает его смысла. Это и указывает на самостоятельное значение социально-институционального измерения. Оно охватывает коллег, учебные и исследовательские учреждения, редакции журналов, экспертные советы, научные школы, финансирующие организации, т.е. всю ту инфраструктуру социальных отношений, в которой наука существует как коллективное предприятие.

Т. Кун показал, что нормальная наука возможна лишь как деятельность внутри парадигмы, разделяемой научным сообществом. Парадигма – это «совокупность общепризнанных образцов научной деятельности», включающая теоретические допущения, стандарты решения задач и ценностные ориентиры; именно она объединяет людей «в одну дисциплинарную матрицу» [Кун, 2003, с. 234–238]. Без сообщества, принимающего парадигму, нет нормальной науки; без нормальной науки нет и революции – кризис парадигмы переживается как кризис потому, что сообщество осознает аномалию как таковую. Научное открытие, таким образом, является одновременно когнитивным и социальным событием.

Согласно идеям М.К. Петрова, наука является способом организации «творческого хаоса» человеческой мысли, который через жесткие фильтры (эксперименты и публикации) преобразует индивидуальные открытия в объективные истины, пригодные для обновления жизни обще-

ства [Петров, 2004]. Институт науки конститутивен для содержания знания и представляет собой внутренний формообразующий принцип. И.Т. Касавин настаивает на том, что наука является культурным феноменом, а ее когнитивные функции неотделимы от социального устройства [Касавин, 2020]. Распад социальных форм, в которых живет наука, – это кризис самой науки, потому что она лишается той инфраструктуры смысла, без которой ее результаты перестают быть результатами.

Для науки также характерна нормативная структура – четыре институциональных императива, образующих этос современной науки [Мертон, 2006, с. 770]. Эти нормы не являются ни личными ценностями ученых, ни логическими требованиями. Это институциональные ожидания, структурирующие научное поведение. Даже когда конкретный исследователь их нарушает, они остаются нормативным фоном, относительно которого нарушение воспринимается как нарушение.

Без социально-институционального измерения научный результат не может получить устойчивую форму признания и воспроизводства: открытие становится открытием только тогда, когда оно признано сообществом; публикация значима ровно постольку, поскольку существует система рецензирования и тиражирования; проблема является «научной проблемой», лишь будучи вписана в дисциплинарный контекст, понятный другим. Социальное измерение выступает важным условием формирования смысловой ткани деятельности ученого.

Когнитивно-методологическое измерение

Третье измерение в определенном смысле наиболее привычно для философии науки, ведь именно оно традиционно составляет ее предмет: неопозитивизм видел в методологии нормативный эталон науки, постпозитивизм (К. Поппер, И. Лакатос, П. Фейерабенд) сосредоточился на критической функции методологии и ее исторических изменениях. Оно включает разделяемые методы исследования, логические и математические инструменты, дисциплинарный язык, систему понятий, стандарты доказательства и критерии объяснения. Главная трудность состоит в том, что когнитивно-методологическое измерение не является ни чисто индивидуальным (принадлежащим отдельному исследователю), ни чисто формальным (редуцируемым к логике). Роль интуиции, аналогии, моделирования, идеализации – все это реализуется не «в голове» ученого самого по себе, а в пространстве разделяемых концептуальных систем. Интуиция работает там, где уже есть освоенный понятийный горизонт, аналогия – где есть два достаточно насыщенных концептуальных поля [Майданов, 2008]. Когнитивные механизмы открытия суть способы работы с разделяемым содержанием.

Примерно в том же направлении движется концепция «стиля научного мышления» – идея, восходящая к Л. Флеку и получившая развитие у В.С. Степина. Способ видения проблем, постановки экспериментов, критерии признания результата – это все исторически выработанные коллективные стандарты, «коллективное видение» [Степин, 2011], функционирующее как само собой разумеющееся.

Отдельного внимания заслуживает вопрос о языке науки. М.К. Петров, исследуя знаково-символические основания науки, показал, что дисциплинарный язык является условием самой когнитивной операции – разграничения, обобщения, установления связей [Петров, 2004]. Язык задает способ видения, поэтому освоение дисциплинарного языка является когнитивным событием. Ученый, овладевший языком своей дисциплины, буквально начинает видеть иначе. Творческая новизна реализуется через усвоенные методы. Только там, где есть усвоенный порядок, возможно его продуктивное нарушение.

Культурно-нормативное измерение

Предыдущее измерение можно было бы считать самодостаточным для описания научной деятельности, если бы не один принципиальный вопрос: откуда у ученого берется ощущение того, что вот эта проблема важна, а вот та – нет? Что этот результат значителен, а тот – тривиален? Что здесь стоит тратить годы жизни, а здесь – не стоит? Ни методология, ни институциональные нормы не дают ответа на эти вопросы: они говорят о том, как работать, но не о том, ради чего. Культурно-нормативное измерение – это именно тот ландшафт, в котором обнаруживаются подобные вопросы. Оно охватывает систему ценностей и норм, не редуцируемую к методологическим правилам и институциональным регуляциям, – широкий культурный горизонт, в котором наука существует как одна из форм человеческой деятельности.

Рациональность науки не автономна, она функционирует в культурном поле, которое задает ее смысл и целевую ориентацию [Пружинин, 2009, с. 49]. Наука – не самодостаточная логическая машина, а форма служения культурным ценностям, среди которых истина занимает особое место, однако неотделима от таких ценностей, как благо, красота и справедливость [Пружинин, 2009, с. 197–198]. Эта идея опирается на русскую философскую традицию «цельного знания», для которого характерно единство когнитивного и ценностного. Когнитивное содержание знания всегда укоренено в конкретном жизненном контексте человека, благодаря чему оно приобретает для него ценность – как практическую, так и экзистенциальную [Пружинин, 2015, с. 120].

В.С. Степин подходит к тому же через концепцию «философских оснований науки» – системы мировоззренческих допущений, задающих

«образ реальности» и определяющих, какие объекты в принципе могут стать предметом научного исследования [Степин, 2011, с. 199]. Философские основания не тождественны научной методологии. Они шире и глубже, поскольку укоренены в культурных картинах мира конкретных эпох. Смена культурных эпох влечет за собой переоценку самих критериев научности, что и имел в виду Степин, описывая переход к разным типам рациональности.

«Этос науки» затрагивает, казалось бы, тот же круг вопросов, однако Мертон склонен рассматривать эти нормы как внутренние нормы самой науки, тогда как Пружинин и Степин настаивают на их укорененности в более широких культурных горизонтах. Наука не создает свои ценности из ничего, она получает их из культуры, перерабатывает в соответствии со своей спецификой и возвращает культуре в трансформированном виде.

Выбор исследовательской программы, расстановка приоритетов, само ощущение «интересного» не произвольны и не сугубо личны. Они задаются аксиологическим горизонтом, который ученый не изобрел, но в котором живет. В этом смысле культурно-нормативное измерение незаметно: убери его – и само представление о научном открытии как культурно значимом событии оказывается проблематичным.

Темпоральное измерение

До сих пор речь шла о четырех измерениях, каждое из которых обладает собственной «предметностью». Темпоральное измерение не имеет собственной «области», но пронизывает все остальные, поскольку каждое из них существует во времени и структурируется временем особым, несводимым к другим образом. Именно поэтому мы называем его сквозным.

Философское осмысление темпорального измерения науки обязано прежде всего А. Бергсону. Его противопоставление «живого» времени (длительности) «пространственному» времени науки (линейному и однородному) остается фундаментальным для данной проблематики [Бергсон, 1999]. Научное время измеримое, делимое на одинаковые единицы, пригодное для перевода качественных характеристик в количественные для расчетов; подлинное же время – непрерывный поток, в котором прошлое «вгрызается» в будущее и расширяется по мере движения вперед [Бергсон, 1999, с. 18, 23]. Фундаментальная проблема новоевропейской науки, по Бергсону, состоит в том, что, пространственно представляя время, она упускает из виду реальное творение – подлинную новизну, несводимую к конфигурации начальных условий. Это прямо касается научного творчества: если мы описываем открытие как логическое следствие предшествующего состояния знания, мы теряем именно ту составляющую, которая делает его открытием, а не вычислением.

Темпоральное измерение специфически разработано В.Г. Будановым в методологии синергетики. Ключевым здесь является различие между «онтологией состояний» и «онтологией процессов» [Буданов, 2013, с. 28]: первая описывает мир через устойчивые конфигурации, инвариантные во времени; вторая – через динамику, переходы, бифуркации, ритмы. Постнеклассическая наука, по Буданову, все более ориентируется на телесно-объектную синтетическую онтологию, и это меняет само понятие «научного факта»: он перестает быть мгновенным снимком и становится событием, включенным в темпоральный контекст [Буданов, 2013, с. 34]. Научная среда в свете синергетики предстает как система взаимодействующих ритмов, порождающих условия для творческого прорыва.

Темпоральное измерение научной среды многосоставно. Во-первых, существует «большое время» науки – время традиций, парадигм, дисциплинарных историй. Это время, в котором Кун описывал чередование нормальной науки и революций; это время, в котором накапливается и хранится коллективная память дисциплины. Во-вторых, есть «среднее время» – время научных школ, проектов, поколений исследователей; время, в котором учитель передает традицию ученику, а научная группа вырабатывает свой рабочий стиль. В-третьих, мы можем выделить «малое время» – темп и ритм повседневной исследовательской практики: время концентрации и рассредоточения, время инкубации идеи и время ее оформления, время рутинной работы и время внезапного озарения.

Все три измерения конститутивны и взаимно обуславливают друг друга. Открытие не происходит в пустом времени: оно происходит в определенный момент истории дисциплины, когда накоплены нужные предпосылки и назрела нужная проблема; оно реализуется в жизни конкретной научной школы с ее специфическими ритмами; оно требует личного времени вызревания. Эти три темпоральности переплетены и взаимно обуславливают друг друга.

Преемственное измерение

Второе сквозное измерение – преемственное. Оно, как и темпоральное, не образует отдельной предметной области, но пронизывает все остальные измерения как механизм их воспроизводства. Это живая передача знаний, навыков, установок и ценностей от учителя к ученику, от поколения к поколению ученых. Она принципиально отлична от формальной трансляции информации через тексты.

Мы знаем всегда гораздо больше, чем можем артикулированно выразить [Полани, 1985, с. 134], т.е. наше знание не поддается полной вербализации; следовательно, оно не может быть передано посредством учебников или статей. Оно передается только в прямом взаимодействии, при котором ученик наблюдает за работой учителя, подражает ему, усваивает

его «способ видения» и постепенно обретает способность к самостоятельному суждению [Полани, 1985, с. 86–87]. Наука в этом отношении устроена подобно ремеслу: мастерство не передается через правила, а только через практику рядом с мастером. Технический прогресс не может компенсировать разрыв этой цепи, поэтому утраченное неявное знание можно лишь воспроизвести заново, пройдя весь путь ученичества снова.

Понятие «республики ученых» у Полани связано с этим: это свободное сообщество, в котором авторитет строится на признанной компетентности, а традиция поддерживается добровольной преемственностью [Полани, 1985]. Без живой цепи передачи мастерства такое сообщество невозможно – оно превращается в набор технических специалистов, каждый из которых знает, как делать, но никто из которых не знает, зачем.

В отечественной научной культуре значение наставничества и научных школ особенно очевидно. История российской науки знает многочисленные примеры школ, в которых преемственность была подлинной конститутивной силой: школы Колмогорова, Ландау, Виноградова, Тамма – это реальные сообщества преемственности, внутри которых передавался не только метод, но и стиль мышления, тип научной интуиции, отношение к проблеме. В своих исследованиях социальных механизмов трансляции культуры М.К. Петров показал, что «кодирование» и «декодирование» культурного опыта являются фундаментальными механизмами воспроизводства социальных практик, включая научную [Петров, 2004, с. 87].

Через это измерение воспроизводятся все остальные измерения научной среды. Материальные практики работы в лаборатории передаются через показ, а не через инструкцию; институциональные нормы усваиваются через участие в семинарах и конференциях рядом с теми, кто этим нормам следует; методологические стандарты формируются в совместной работе над текстами; культурные ценности транслируются через пример учителя – молчаливый, но действенный. Без преемственности научная среда не воспроизводит себя: она «замирает» и постепенно рассыпается, хотя ее материальные и институциональные оболочки могут еще долго производить внешнее впечатление порядка.

Взаимосвязь измерений

Основание для выделения каждого измерения – принцип неприводимости: ни один из них не сводится к другому. Социально-институциональное измерение нельзя вывести из материального: коллектив и оборудование – разные феномены, и они могут существовать независимо друг от друга (хотя и не могут образовать полноценную науку порознь). Когнитивно-методологическое измерение не тождественно культурно-нормативному: можно представить ученого, безусловно владеющего методологией своей дисциплины, но полностью безразличного к вопросу о том, зачем это

нужно человечеству, – и такие случаи в истории науки не редкость. Темпоральное измерение не является «частью» ни одного из четырех предметных измерений: время лабораторного эксперимента, время институциональных циклов, время вызревания исследовательской программы и время биографии ученого – принципиально разные темпоральности, которые не складываются в одну, но взаимодействуют.

Вместе с тем измерения не изолированы. Синергетическая перспектива позволяет рассматривать их как взаимодействующие ритмы единой системы [Буданов, 2017]: когда несколько ритмов совпадают – материальный прорыв, институциональная поддержка, зрелость традиции, культурный запрос, – возникают условия для концентрации творческого потенциала. Именно таковы были Александрийская библиотека, Венский кружок, советская физическая школа 1930–1950-х годов, Кавендишская лаборатория и т.п. Ни один из этих феноменов не объясним через одно измерение, потому что в каждом случае действовала полная конфигурация среды.

Контрфактический мысленный эксперимент позволяет это прояснить. Представим ученого, окруженного превосходным оборудованием и работающего в сильном коллективе, владеющего современными методами, но лишенного ценностных ориентаций и вырванного из традиции. Перед нами будет высококвалифицированный технический специалист, но не ученый в полном смысле слова: ему будет чего-то фундаментально не хватать. Или представим исследователя, глубоко укорененного в традиции и обладающего богатой культурной рефлексией, но лишенного институциональной поддержки и материальной базы. Творческий потенциал здесь может быть значителен, однако его реализация окажется в лучшем случае частичной. Полноценное научное творчество требует всей конфигурации измерений, и это делает научную среду конститутивным, а не просто благоприятным условием.

Заключение

В данной статье была предпринята попытка обосновать и развернуть концепцию научной среды как конститутивного условия научного творчества. Мы исходили из фиксации реальной проблемы – фрагментарности существующих подходов к изучению научного творчества – и предложили понятие, призванное служить их концептуальной точкой схождения. Выделенные шесть измерений (материальное, социально-институциональное, когнитивно-методологическое, культурно-нормативное, темпоральное и преемственное) не произвольны. Каждое из них фиксирует неприводимый аспект среды, без которого наука оказывается неполной или вовсе невозможной. Первые четыре измерения образуют предметное пространство среды; последние два являются сквозными, поскольку они описывают не

отдельные части среды, а способ ее существования во времени и способ ее воспроизводства.

Центральный тезис статьи о конститутивном характере научной среды обосновывается через три взаимосвязанных положения: среда участвует в формировании субъекта научного творчества, способов постановки проблем и критериев оценки результатов. Это не умаляет творческой свободы ученого – напротив, указывает на ее подлинные условия: как Полани показал применительно к неявному знанию, а Бергсон – применительно к творческой длительности, свобода не противостоит условиям, а реализуется через них.

Разработка модели научной среды как конститутивного условия представляется важной задачей современной философии науки, ориентированной на постнеклассическое понимание науки как человекоразмерной, исторически укорененной и ценностно насыщенной формы деятельности. Эта модель не претендует на создание универсальной теории научного творчества, она предлагает общий аналитический горизонт, в котором различные аспекты научной деятельности могут рассматриваться как взаимосвязанные.

Благодарности: выражаю благодарность своему научному руководителю, доктору философских наук Буданову Владимиру Григорьевичу за ценные замечания и методологическую поддержку при разработке идей статьи.

Список литературы

- Бергсон А. Творческая эволюция / пер. с фр. В. Флеровой. – 3-е изд. – М.: Академический проект, 2026. – 334 с.
- Буданов В.Г. Квантово-синергетическая антропология и проблемы искусственного интеллекта и трансгуманизма // Философские науки. – 2013. – № 9. – С. 25–37.
- Буданов В.Г. Методология синергетики в постнеклассической науке и в образовании: новое издание, дополненное материалами: синергетика третьей волны, цифровой жизненный техноуклад, образование эпохи большого антропологического перехода. – 4-е изд., доп. – М.: URSS, 2017. – 270 с.
- Касавин И.Т. Научное творчество как социальный феномен // Эпистемология и философия науки. – 2022. – Т. 59. – № 3. – С. 19–29. <https://doi.org/10.5840/eps202259336>
- Касавин И.Т. Наука – гуманистический проект. – М.: Издательство «Весь Мир», 2020. – 496 с.
- Кун Т. Структура научных революций / пер. с англ. И.З. Налетова и др. – М.: АСТ, 2003. – 605 с.
- Латур Б. Наука в действии: следуя за учеными и инженерами внутри общества / пер. с англ. К. Федоровой; науч. ред. С. Миляева. – СПб.: Издательство Европейского университета в Санкт-Петербурге, 2013. – 414 с.
- Майданов А.С. Методология научного творчества. – М.: Издательство ЛКИ, 2008. – 512 с.
- Мертон Р.К. Социальная теория и социальная структура / пер. с англ. – М.: АСТ, 2006. – 873 с.

- Петров М.К. Язык. Знак. Культура. – 2-е изд. – М.: УРСС, 2004. – 328 с.
- Полани М. Личностное знание: на пути к посткритической философии / пер. с англ. М.Б. Гнедовского и др. – М.: Прогресс, 1985. – 344 с.
- Пружинин Б.И. Идеи культурно-исторической эпистемологии в русской философии // Философские науки. – 2015. – № 2. – С. 119–126.
- Пружинин Б.И. Ratio serviens? Контуры культурно-исторической эпистемологии. – М.: РОССПЭН, 2009. – 423 с.
- Степин В.С. Цивилизация и культура. – СПб.: СПбГУП, 2011. – 408 с.
- Knorr Cetina K. Epistemic Cultures: How the Sciences Make Knowledge. – Cambridge, MA: Harvard University Press, 1999. – 329 p.
- Polanyi M. The Republic of Science: Its Political and Economic Theory // Minerva. – 1962. – Vol. 1. – No. 1. – P. 54–73. <https://doi.org/10.1007/BF01101453>

References

- Bergson, A. (2026). *Creative evolution* (V. Flerova, Trans.; 3rd ed.). Akademicheskii proekt. 334 p. (In Russ.)
- Budanov, V. G. (2013). Quantum-synergetic anthropology and problems of artificial intelligence and transhumanism. *Russian Journal of Philosophical Sciences*, (9), 25–37. (In Russ.)
- Budanov, V. G. (2017). *Methodology of synergetics in post-non-classical science and education: Revised edition* (4th ed.). Moscow: URSS. 270 p. (In Russ.)
- Kasavin, I. T. (2020). *Science as a humanistic project*. Ves' Mir. 496 p. (In Russ.)
- Kasavin, I. T. (2022). Scientific creativity as a social phenomenon. *Epistemology & Philosophy of Science*, 59(3), 19–29. (In Russ.)
- Knorr Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press. 329 p.
- Kuhn, T. (2003). *The structure of scientific revolutions* (I. Z. Naletov et al., Trans.). Moscow: AST. 605 p. (In Russ.)
- Latour, B. (2013). *Science in Action: Following Scientists and Engineers within society* (K. Fedorova, Trans.; S. Milyaev, Ed.). SPb: European University at St. Petersburg Press. 414 p. (In Russ.)
- Maidanov, A. S. (2008). *Methodology of scientific creativity*. Moscow: LKI. 512 p. (In Russ.)
- Merton, R. K. (2006). *Social theory and social structure*. Moscow: AST. 873 p. (In Russ.)
- Petrov, M. K. (2004). *Language. Sign. Culture* (2nd ed.). Moscow: URSS. 328 p. (In Russ.)
- Polanyi, M. (1962). The republic of science: Its political and economic theory. *Minerva*, 1(1), 54–73.
- Polanyi, M. (1985). *Personal knowledge: Towards a post-critical philosophy* (M. B. Gnedovsky et al., Trans.). Progress. (In Russ.)
- Pruzhinin, B. I. (2009). Contours of cultural-historical epistemology. Moscow: ROSSPEN. 423 p. (In Russ.)
- Pruzhinin, B. I. (2015). Ideas of cultural-historical epistemology in Russian philosophy. *Russian Journal of Philosophical Sciences*, (2), 119–126. (In Russ.)
- Stepin, V. S. (2011). *Civilization and culture*. SPb: SPbGUP. 408 p. (In Russ.)

ИСТОРИЯ И МЕТОДОЛОГИЯ РАЗВИТИЯ НАУКИ И ТЕХНИКИ

УДК 167.7/168.2

DOI: 10.31249/scis/2026.03.05

КЛАСТЕРНО-МИКРОСЕРВИСНАЯ АРХИТЕКТУРА ПЛАТФОРМЕННОЙ НАУКИ КАК ИНСТРУМЕНТ ДЕМАРКАЦИИ НАУЧНОГО ЗНАНИЯ В КОНТЕКСТЕ ФИЛОСОФИИ СЛОЖНОСТИ ВЛАДИМИРА АРШИНОВА

Тарасенко Владислав Валерьевич – канд. филос. наук, доцент МГТУ им. Н.Э. Баумана, доцент Университета просвещения, Москва, Россия; v5093075@gmail.com; ORCID 0000-0003-3328-8197

Аннотация. Работа посвящена решению актуальной проблемы демаркации научного и ненаучного знания в условиях цифровой трансформации научной деятельности. Аргументируется тезис о том, что эффективность классических демаркационных критериев (верифицируемость, фальсифицируемость), ориентированных на статичные теории, снижается в отношении постнеклассической науки, функционирующей как распределенная, технологически опосредованная и междисциплинарная экосистема.

В качестве дополнения предлагается концепция кластерно-микросервисной архитектуры (КМА) как представление об организации научного знания и методологический инструмент его демаркации. Концепция разработана на основе синтеза философии сложности (синергетики) В.И. Аршинова, принципов постнеклассической рациональности В.С. Стёпина, социальной эпистемологии, акторно-сетевой теории, а также модели обобщенного «деятельностного тетраэдра» В.Г. Буданова.

В рамках этой модели уточняются функциональные роли элементов КМА: микросервис выступает как преобразователь потока данных и знаний, кластер – как агрегатор результатов в рамках общей проблематики, а платформа – как мета-агрегатор, обеспечивающий наблюдение второго порядка и интероперабельность всего научного ландшафта. Ключевым результатом является система архитектурных критериев научной достоверности и демаркации, оценивающих знание по качеству сетевой организации его производства: интероперабельности, рекурсивной проверяемости, наблюдаемости процессов и устойчивости к антипаттернам (гиперспециализация, монополизация интерфейсов).

В статье обосновывается переход философии науки от роли мета-дискурса к статусу активного «эпистемического микросервиса», встраиваемого в инфраструктуру исследовательских платформ для обеспечения концептуальной ясности и этической рефлексии.

Ключевые слова: философия науки; постнеклассическая рациональность; синергетика; сложность; Владимир Аршинов; демаркация; микросервисная архитектура; кластеризация; научные платформы; наблюдатель второго порядка.

Для цитирования: Тарасенко В.В. Кластерно-микросервисная архитектура платформенной науки как инструмент демаркации научного знания в контексте философии сложности Владимира Аршинова // Научноисследовательские исследования. – 2026. – № 3. – С. 71–82. – DOI: <http://www.doi.org/10.31249/scis/2026.03.06>

Статья поступила в редакцию: 24.02.2026

Принята к публикации: 24.04.2026

CLUSTER-MICROSERVICE ARCHITECTURE OF PLATFORM SCIENCE AS A TOOL FOR DEMARCATION OF SCIENTIFIC KNOWLEDGE IN THE CONTEXT OF VLADIMIR ARSHINOV'S PHILOSOPHY OF COMPLEXITY

Vladislav V. Tarasenko – Candidate of Sciences (Philosophy), Associate Professor at Bauman Moscow State Technical University, Associate Professor at Federal State University of Education, Moscow, Russia; v5093075@gmail.com; ORCID: 0000-0003-3328-8197

Abstract. This work addresses the pressing problem of demarcating scientific and non-scientific knowledge in the context of the digital transformation of scientific activity. It argues that the effectiveness of classical demarcation criteria (verifiability, falsifiability), which are oriented towards static theories, is diminishing with respect to post-non-classical science, which functions as a distributed, technologically mediated, and interdisciplinary ecosystem.

As a complement, the concept of a cluster-microservice architecture (CMA) is proposed as a representation of the organization of scientific knowledge and a methodological tool for its demarcation. The concept is developed through a synthesis of the philosophy of complexity (synergetics) by V.I. Arshinov, the principles of post-non-classical rationality by V.S. Styopin, social epistemology, actor-network theory, as well as the model of the generalized "activity tetrahedron" by V.G. Budanov.

Within this model, the functional roles of the CMA elements are specified: the microservice acts as a transformer of data and knowledge flows, the cluster acts as an aggregator of results within a shared problem area, and the platform acts as a meta-aggregator, providing second-order observation and interoperability for the entire scientific landscape. The key result is a system of architectural criteria for scientific validity and demarcation, which evaluates knowledge based on the quality of the network organization of its production: interoperability, recursive verifiability, process observability, and resilience to anti-patterns (hyperspecialization, monopolization of interfaces).

The article substantiates the transition of the philosophy of science from the role of a meta-discourse to the status of an active "epistemic microservice," integrated into the infrastructure of research platforms to ensure conceptual clarity and ethical reflection.

Keywords: philosophy of science; post-non-classical rationality; synergetics; complexity; Vladimir Arshinov; demarcation; microservice architecture; clustering; scientific platforms; second-order observer.

For citation: Tarasenko, V.V. Cluster-microservice architecture of platform science as a tool for demarcation of scientific knowledge in the context of Vladimir Arshinov's philosophy of complexity. // *Naukovedcheskie issledovaniya* = *Science Studies*. – 2026. – N 3. – P. 71–82. – DOI: <http://www.doi.org/10.31249/scis/2026.03.06>

The paper was submitted: 24.02.2026

Accepted for publication: 24.04.2026

Введение

Проблема демаркации, отделения научного знания от ненаучного, является одной из фундаментальных для философии науки. Исторически она решалась через логические (верифицируемость Венского кружка), методологические (фальсифицируемость К. Поппера) или социологические (парадигмальность Т. Куна) критерии.

Однако глубокий процесс цифровой трансформации, охвативший научную деятельность в XXI веке, ставит под вопрос адекватность этих классических подходов. Согласно представлениям о постнеклассической рациональности В.С. Стёпина, современная наука все меньше напоминает монолитную дисциплинарную систему и все больше – гетерогенную распределенную сетевую экосистему [Аршинов, 1999].

В этой экосистеме знание генерируется не отдельными учеными в «башнях из слоновой кости», а сложными сетями взаимодействующих агентов: исследовательских коллективов, алгоритмов искусственного интеллекта, баз данных, автоматизированных экспериментальных установок, моделей и симуляций. Это принципиально меняет онтологию познавательного процесса: на первый план выходит не содержание, верификация или фальсификация изолированной теории, а архитектура практик, в рамках которой эта теория производится, проверяется и используется.

Таким образом, возникает новый вызов: как отличить достоверное научное знание, ориентируясь не на его формальные атрибуты, а на качество организации процесса его производства? Предлагаемая работа является попыткой ответить на этот вызов через разработку концепции кластерно-микросервисной архитектуры (КМА) как демаркационного инструментария постнеклассической научной рациональности.

В рамках концепции КМА наука предстает как распределенная, технологически опосредованная и междисциплинарная экосистема. Процесс генерации знания в ней осуществляется не изолированными исследователями, а сложными сетями автономных, но интероперабельных «микросервисов». Под микросервисами в данном контексте понимаются различные виды деятельности – человеческие, гибридные, алгоритмические и институциональные, – которые вступают друг с другом в отношения обмена данными, информацией и результатами интеллектуальной деятельности.

В подобной конфигурации классические критерии демаркации науки и псевдонауки обнаруживают свою ограниченность. Их фундаментальная предпосылка ориентирована на оценку статичных теоретических конструктов, тогда как современное научное познание все чаще реализуется в форме динамических архитектур деятельности, что требует пересмотра традиционных подходов к демаркации.

КМА выступает в роли инструмента постнеклассической научной рациональности, с помощью которого выводится система сетевых крите-

риев научной достоверности, связанных с интероперабельностью, рекурсивной проверяемостью, наблюдаемостью процессов производства научного знания и их устойчивости к антипаттернам (например, гиперспециализация, закрытость, монополизация интерфейсов).

Философским фундаментом этой концепции выступает наследие Владимира Ивановича Аршинова – мыслителя, чья работа по синтезу синергетики, философии науки и теории коммуникации создала уникальный аппарат для понимания самоорганизующихся сложных систем, к которым, безусловно, относится и современная наука [Аршинов, 1994]. Важным дополнением и операционализацией этих идей служит модель обобщенной научной рациональности и «деятельностного тетраэдра», предложенная В.Г. Будановым, которая позволяет перевести философские принципы в архитектурные паттерны проектирования научных платформ [Буданов, 2022; Буданов, 2023].

Наука как коммуникативная и самоорганизующаяся система

Философия сложности (сложностности) В.И. Аршинова, развиваемая в русле московской синергетической школы, предлагает ключевые принципы для осмысления современной науки.

Аршинов, опираясь на идеи М. Полани и синергетиков (Г. Хакен, И. Пригожин), рассматривал научное познание не как пассивное отражение реальности, а как творческий коммуникативный процесс, в котором исследователь и объект взаимно влияют друг на друга [Буданов, 2024]. Синергетический подход, вслед за кибернетикой второго порядка, подчеркивает понимание научной коммуникации не как нейтрального инструмента, а как рефлексивно-активной среды, которая формирует условия для производства определенного типа знания [Буданов, Румянцев, Шохов, 2024].

Для формализации КМА плодотворной является развитая В.Г. Будановым модель обобщенного деятельностного тетраэдра ($S \rightarrow T \rightarrow O \rightarrow R$), которая формализует неразрывную связь наблюдателя с наблюдаемой системой. В этой модели базовый акт деятельности представлен как взаимосвязь четырех вершин: Субъект (S), Средство/Инструмент (T), Объект (O) и Результат (R). Буданов рассматривает эти элементы как вершины тетраэдра, где все они связаны между собой, а в центре располагается Наблюдатель-интерпретант (Obs).

Эта топология отражает циркулярную причинность и нелинейность познавательных процессов в постнеклассической науке. В контексте КМА каждая вершина и связь тетраэдра могут быть реализованы как микросервис или их взаимодействие. Сам тетраэдр становится элементарной ячейкой или паттерном научной деятельности, рекурсивно воспроизводимым на разных уровнях системы.

Важнейшее развитие этой модели для целей настоящего исследования – различение этапов трансформации результата (R). В полном цикле научного эксперимента, как отмечает Буданов, можно выделить расширенную схему: $S \rightarrow T \rightarrow O \rightarrow T \rightarrow R \rightarrow T^* \rightarrow S$, где T – средства детектирования / сбора первичных данных, а T^* – средства интерпретации и представления результатов [Лепский, 2010]. Именно на этапах $T \rightarrow R \rightarrow T^*$ и происходит функциональная спецификация элементов КМА.

Микросервис выполняет роль преобразователя (transformer) потока данных / знаний. Он инкапсулирует логику применения конкретного средства (T, T^* или T^{**}) к конкретному объекту (данным, модели, запросу) для получения промежуточного или конечного результата (R). Например, сервис спектрального анализа (T) преобразует сырые данные (O) в спектр (R).

Кластер выступает как агрегатор результатов. Он объединяет выходы (R) множества кооперирующихся микросервисов, синтезируя их в комплексный, междисциплинарный результат более высокого порядка. Кластер реализует логику синергетического сложения, где целое больше суммы частей.

Научная платформа является мета-агрегатором или наблюдателем второго порядка. Она обеспечивает инфраструктуру для взаимодействия множества кластеров и микросервисов, аккумулирует мета-результаты (метаданные, показатели активности, карты знаний), обеспечивая наблюдаемость всей научной экосистемы в режиме, близком к реальному времени [Стёпин, 2003]. Платформа реализует функцию T в масштабе всей экосистемы, предоставляя интерфейсы для интерпретации и принятия решений субъектом (S^*).

Таким образом, деятельностный тетраэдр Буданова представляет не только философскую онтологию, но и структурную схему для проектирования элементов КМА, где четко разводятся функции преобразования, агрегации и мета-наблюдения.

Концепция кластерно-микросервисной архитектуры (КМА) научного знания

В рамках предлагаемой концепции постнеклассическая наука предстает как сложноорганизованная экосистема, структурно включающая в себя три основных уровня: микросервисы, кластеры и платформы. Каждый из этих уровней выполняет специфические функции, которые были детально рассмотрены выше.

Ключевым операциональным элементом данной модели выступает микросервис. В предлагаемой интерпретации микросервис представляет собой автономный узкоспециализированный модуль, предназначенный

для трансформации входных данных или запросов в определенный результат научной деятельности (данные, информация, знания).

Его работа основывается на строго заданной внутренней логике, которая может быть выражена в виде метода, алгоритма или протокола. Функционально микросервис соответствует реализации одного или нескольких элементов и связей в рамках деятельностного тетраэдра (например, пары «Теория–Объект» или функции T^*). Многообразие микросервисов можно классифицировать по их роли в исследовательском процессе.

Можно выделить следующие *основные типы*.

Алгоритмические преобразователи. К этой категории относятся инструменты, реализующие вычислительные алгоритмы: модели машинного обучения для анализа данных, симуляторы физических процессов или программное обеспечение для статистической обработки информации.

Трансляторы данных. В эту группу входят сервисы, отвечающие за преобразование форматов, очистку и верификацию данных, а также модули для извлечения информации из неструктурированных текстов.

Методические исполнители. Это сервисы, которые обеспечивают проведение стандартизированных протоколов – как экспериментальных, так и вычислительных, – гарантируя воспроизводимость процедур.

Интерпретаторы. К ним относятся средства визуализации, системы построения отчетов и сервисы семантического анализа, предназначенные для осмысления и наглядного представления полученных научных результатов.

Подобно элементам программной архитектуры, научные микросервисы-преобразователи обладают рядом *ключевых свойств*, которые обеспечивают их эффективное функционирование в составе единой экосистемы.

Инкапсуляция и стандартизация интерфейса. Каждый микросервис скрывает детали своей внутренней реализации, предоставляя для взаимодействия четко определенный программный интерфейс (API).

Абстракция входных и выходных данных. Сервис оперирует данными на уровне абстракций, что позволяет использовать его независимо от конкретного содержания запроса.

Полиморфизм. Способность микросервиса обрабатывать различные типы данных в рамках своей абстрактной модели, адаптируясь к меняющимся условиям задачи.

Наследование и версияность. Как и в случае с программным обеспечением, научные микросервисы могут развиваться: на основе существующих наработок создаются новые версии, что обеспечивает преемственность и прослеживаемость эволюции методов.

Кластеры и платформы как агрегаторы результатов вокруг проблемных аттракторов

Кластеры – это динамические объединения микросервисов, сформированные вокруг комплексной научной проблемы.

Функция кластера – агрегация и синтез результатов (R), полученных от множества микросервисов-преобразователей. Кластер обеспечивает их кооперацию, управляя потоками данных и обеспечивая согласованность онтологий. Он представляет собой пространство синергетической сборки междисциплинарного знания. Примеры: кластер по изучению изменения климата агрегирует результаты моделирования атмосферы, анализа спутниковых снимков, социологических опросов.

Центральным элементом инфраструктуры, обеспечивающим целостность и жизнеспособность всей экосистемы КМА, выступает исследовательская платформа. Ее фундаментальная роль заключается в реализации функций «наблюдателя второго порядка» (second-order observer). Иными словами, платформа осуществляет мета-наблюдение за деятельностью множества кластеров и микросервисов, аккумулируя метаданные об их текущем состоянии, производительности и о формирующихся между ними связях.

Выступая в роли мета-агрегатора, платформа создает единую операционную среду, предоставляя субъектам научной деятельности различные возможности, в частности, поддержание единых реестров сервисов и унифицированных стандартов интерфейсов (API), что обеспечивает совместимость различных компонентов научной экосистемы; внедрение систем мониторинга и наблюдаемости, позволяющих отслеживать работу распределенных элементов в реальном времени; создание межкластерных рабочих пространств для реализации кросс-дисциплинарных проектов; предоставление средств для интеграции так называемых «эпистемических микросервисов», отвечающих за философскую и этическую экспертизу проводимых исследований; развертывание дашбордов и аналитических инструментов (реализующих теоретическую функцию T), адресованных субъектам управления наукой (S*). Данные инструменты обеспечивают возможность оперативного мониторинга состояния науки в режиме «здесь и сейчас», что создает фундамент для принятия обоснованных управленческих решений.

Архитектурные критерии демаркации: как отличить науку от фейка в эпоху КМА

Традиционные критерии демаркации, ориентированные на статичную оценку готового знания, уступают место новым подходам в условиях динамично развивающихся научных экосистем. При опоре на модель

КМА и предложенную деятельностную схему представляется возможным сформулировать систему критериев для оценки научной достоверности, понимаемой не как свойство изолированного утверждения, а как качество архитектуры познавательной деятельности.

Критерий интероперабельности и четкости интерфейсов (API)

Научный статус нового познавательного элемента – будь то теория, метод или набор данных – определяется его способностью к интеграции в существующую экосистему знания. Истинно научный элемент может быть реализован в форме микросервиса с документированным и открытым интерфейсом (API), что позволяет другим исследователям подключаться к нему, использовать его и проверять. Псевдонаучные конструкции, напротив, тяготеют к созданию закрытых «монолитов» – самодостаточных систем, принципиально несовместимых со смежными областями знания.

Пример. Публикация набора генетических данных с открытым API, позволяющим биоинформатикам напрямую подключать эти данные к своим конвейерам анализа (микросервисам), соответствует научному критерию. Публикация же сенсационных результатов «торсионных исследований» с отказом предоставить сырые данные или описание методики, достаточное для независимого воспроизведения, представляет собой закрытый монолит.

Критерий наблюдаемости и рекурсивной проверяемости

В экосистеме КМА критически важна прозрачность процессов преобразования данных. Критерий требует, чтобы была доступна для анализа логика как первичного преобразования «теория–объект» ($T \rightarrow O$), так и процедуры интерпретации (T^{**}). Это обеспечивает возможность независимого аудита на всех уровнях – от эмпирической проверки до логической валидации вывода.

Пример. Воспроизводимый вычислительный эксперимент, где код, данные и спецификации окружения открыты и могут быть проверены другим микросервисом для валидации результатов, удовлетворяет данному критерию. Напротив, клиническое исследование нового препарата, результаты которого опубликованы без предоставления протокола и индивидуальных данных пациентов для стороннего мета-анализа, нарушает принцип наблюдаемости.

Критерий агрегируемости в кластеры (когерентность)

Научное знание обладает свойством когерентности. Новые результаты должны не просто коррелировать с существующими, но и способствовать их агрегации в более сложные кластеры, создавая синергетиче-

ский эффект. Если же результат остается изолированным артефактом, не встраиваясь в более широкие сети знания и не усиливая смежные направления, это ставит под сомнение его научную состоятельность.

Пример. Открытие нового фермента приобретает научный статус именно тогда, когда оно позволяет объяснить ранее разрозненные метаболические пути (агрегирует их в связный кластер) и предлагает инструмент для смежных областей – например, для разработки микросервисов в фармакологии. В отличие от этого, сообщение об открытии «биоэнергетического поля человека», которое никак не связано с существующей биохимией, физиологией или физикой и не помогает агрегировать знания этих дисциплин, остается изолированным артефактом.

Критерий платформенной наблюдаемости (включенность в мета-агрегатор)

На уровне исследовательской платформы, выполняющей функции «наблюдателя второго порядка», деятельность научных элементов должна поддаваться объективной оценке. Критерий предполагает возможность отслеживать по объективным метрикам качество данных, востребованность сервиса и устойчивость его связей с другими элементами экосистемы. Псевдонаучные конструкции либо не поддаются такому наблюдению, либо демонстрируют аномальные паттерны (например, замкнутость ссылок внутри узкой группы).

Пример. Наукометрическая платформа, анализирующая сети цитирования, может выявить кластер статей по новой методике секвенирования, который активно интегрируется с другими методами, востребован и расширяется. Это демонстрирует соответствие критерию. Если же алгоритмы платформы обнаруживают изолированный «пузырь» публикаций, которые ссылаются только друг на друга и игнорируются основным корпусом науки, это сигнализирует о возможном нарушении данного критерия.

Критерий устойчивости к архитектурным антипаттернам

Данный критерий оценивает «гигиену» исследовательской программы с точки зрения архитектурных рисков. Научная деятельность должна демонстрировать устойчивость к таким антипаттернам, как:

- гиперспециализация: создание микросервиса, настолько узкого, что он не имеет точек сопряжения с другими;
- монополизация интерфейсов: ситуация, когда критически важный сервис принадлежит одной группе, не документирован и блокирует развитие смежных направлений;
- хрупкость: чрезмерная зависимость целого кластера от одного элемента, отказ которого парализует работу.

Пример. В области машинного обучения научным стандартом становится публикация не только статьи, но и открытого кода модели с четким API (борьба с монополизацией и хрупкостью). Напротив, компания, которая утверждает, что создала «универсальный ИИ для диагностики всех болезней», но предоставляет к нему доступ только через «черный ящик» (закрытый API, скрытые данные обучения), демонстрирует сразу несколько антипаттернов: монополизацию интерфейса, отсутствие наблюдаемости и потенциальную хрупкость результата.

Философия как эпистемический микросервис: от метанаблюдения к проектированию

Важнейшим следствием предлагаемой концепции выступает трансформация роли самой философии науки. В рамках модели КМА философия перестает быть исключительно внешним мета-дискурсом, лишь ретроспективно осмысляющим результаты научной деятельности. Она обретает статус набора активных эпистемических микросервисов, встраиваемых непосредственно в ткань исследовательских платформ и функционирующих на различных этапах цикла $S \rightarrow T \rightarrow O \rightarrow R \rightarrow T^{**} \rightarrow S^*$.

На этапе перехода от теории к объекту ($T \rightarrow O$) философские сервисы могут осуществлять критический анализ концептуальной чистоты исходных посылок, проверяя корректность используемых категорий и операционализацию понятий – подобно тому, как инструменты линтинга проверяют качество программного кода.

В процессах агрегации микросервисов в кластеры философские компоненты способны выполнять координирующую функцию, обеспечивая согласованность разнородных методологических установок и способствуя формированию непротиворечивых междисциплинарных конфигураций.

На стадии интерпретации результатов (реализация функции T^{**}) философская экспертиза может выступать в роли фильтра, оценивающего этические импликации полученного знания, а также помогать встраивать результаты в более широкий смысловой контекст. Тем самым философия науки обретает возможность практически реализовать идеи В.И. Аршинова о понимании науки как коммуникативной деятельности в условиях гибридных человеко-машинных систем.

Встраиваясь в цифровую инфраструктуру исследований, философские микросервисы призваны выполнять критически важную функцию – создавать «островки рефлексии и смысла», удерживая гуманитарное измерение в непрерывно нарастающем потоке данных и алгоритмических операций.

Заключение

Проблема демаркации научного и ненаучного знания, традиционно решавшаяся через логические, методологические или социологические критерии, в условиях цифровой трансформации науки приобретает новое измерение.

Классические подходы, ориентированные на оценку статичных теоретических конструктов, обнаруживают свою ограниченность перед лицом постнеклассической науки, функционирующей как распределенная, технологически опосредованная и междисциплинарная экосистема.

В ответ на этот вызов в работе предложена концепция кластерно-микросервисной архитектуры (КМА) – не как техническая спецификация, а как эпистемическая модель организации научного знания и практический инструмент его демаркации.

Опираясь на философию сложности В.И. Аршинова, принципы постнеклассической рациональности В.С. Стёпина и деятельностьную модель обобщенного тетраэдра В.Г. Буданова, мы показали, что знание в современной науке производится не изолированными исследователями, а сетями интероперабельных микросервисов (человеческих, алгоритмических, инструментальных), агрегируемых в кластеры вокруг комплексных проблем и наблюдаемых на уровне платформ как мета-агрегаторов.

Ключевым результатом работы стала система архитектурных критериев научной достоверности, позволяющих отличать научное знание от псевдонаучного не по содержанию изолированных утверждений, а по качеству процесса его производства. К таким критериям отнесены: интероперабельность и открытость интерфейсов, рекурсивная проверяемость и наблюдаемость преобразований, агрегируемость результатов в кластеры, платформенная наблюдаемость (включенность в мета-агрегатор) и устойчивость к архитектурным антипаттернам (гиперспециализация, монополизация интерфейсов, хрупкость).

Важнейшим следствием предложенной концепции является трансформация роли самой философии науки. В рамках КМА философия перестает быть внешним мета-дискурсом и обретает статус активного эпистемического микросервиса, встраиваемого непосредственно в исследовательские платформы. На этапах концептуального анализа, координации методологий и этической интерпретации результатов философские сервисы призваны выполнять критическую функцию удержания рефлексии и смыслового измерения в непрерывно нарастающем потоке данных и алгоритмических операций.

Предложенная модель КМА не подменяет, но существенно дополняет классические критерии демаркации (верифицируемость, фальсифицируемость, парадигмальность), перенося фокус с изолированного утверждения на архитектуру познавательной деятельности. Тем самым

открывается возможность для эмпирической операционализации демаркационных процедур в цифровой среде – например, через разработку метрик интероперабельности, инструментов аудита воспроизводимости и алгоритмов детекции антипаттернов на научных платформах.

Перспективы дальнейших исследований связаны с практической реализацией КМА-подхода на существующих научных платформах (Open Science Framework, ResearchGate, институциональные репозитории), разработкой прототипов философских микросервисов (концептуальный линтинг, автоматизированная этическая экспертиза) и эмпирической верификацией предложенных критериев на кейсах демаркации спорных областей знания. Кроме того, требует отдельного осмысления вопрос о том, как предложенная архитектура соотносится с существующими системами научной коммуникации, экспертизы и оценки результативности.

Таким образом, концепция КМА предлагает не только теоретический инструмент для философии науки, но и практическую рамку для проектирования цифровых инфраструктур исследования, способных поддерживать высокие эпистемические стандарты в эпоху платформенной науки.

Список литературы

- Аршинов В.И. Синергетика как феномен постнеклассической науки. – М.: ИФ РАН, 1999. – 203 с.
- Аршинов В.И., Буданов В.Г. Синергетика: эволюционный аспект // Самоорганизация и наука: опыт философского осмысления. – М.: Фирма "АРГО", 1994. – С. 229–242.
- Буданов В.Г. Обобщенная научная рациональность: истоки, структура, перспективы в цифровую эпоху // Антропомерность как вызов и ответ современности. – Курск: Университетская книга, 2022. – С. 44–63.
- Буданов В.Г. Посткритическая рациональность: нейросетевой путь от мира истин к миру мнений // Сложность. Разум. Постнеклассика. – 2023. – №1. – С.58–63.
- Буданов В.Г. Техносубъект и антропосоциальные проблемы взаимодействия с искусственным интеллектом: синергия, демаркация, новая рациональность, риски // Философские науки. – 2024. – Т. 67. – № 3. – С. 27–52.
- Буданов В.Г., Румянцев В.Ю., Шохов А.С. Философия управления сверхсложными организациями: кибернетика второго порядка // Сложность. Разум. Постнеклассика. – 2024. – №2. – С.29–38.
- Лепский В.Е. Рефлексивно-активные среды инновационного развития. – М.: Когито-Центр, 2010. – 253 с.
- Стёпин В.С. Теоретическое знание. – М.: Прогресс-Традиция, 2003. – 744 с.

References

- Arshinov, V.I. (1999). *Synergetics as a phenomenon of post-nonclassical science*. Moscow: IPh RAS, 1999. 203 p. (In Russian)

- Arshinov, V.I., & Budanov, V.G. (1994). Synergetics: an evolutionary aspect. In *Self-organization and science: the experience of philosophical understanding*. Moscow: Firm "AR-GO". P. 229–242. (In Russ.)
- Budanov, V.G. (2022). Generalized scientific rationality: origins, structure, prospects in the digital age. In *Anthropomericity as a challenge and response of modernity*. Kursk: University Book. P. 44–63. (In Russ.)
- Budanov, V.G. (2023). Postcritical rationality: a neural network path from the world of truth to the world of opinions. *Complexity. Mind. Postnonclassics*, 1, 58–63. (In Russ.)
- Budanov, V.G. (2024). Technosubject and anthroposocial problems of interaction with artificial intelligence: synergy, demarcation, new rationality, risks. *Russian Journal of Philosophical Sciences*, 67(3), 27–52. (In Russ.)
- Budanov, V.G., Rumyantsev, V.Yu., & Shokhov, A.S. (2024). Philosophy of management of super-complex organizations: second-order cybernetics. *Complexity. Mind. Postnonclassics*, 2, 29–38. (In Russ.)
- Lepsky, V.E. (2010). *Reflexive-active environments of innovative development*. Moscow: Kogito-Center, 2010. 253 p. (In Russ.)
- Stepin, V.S. (2003). *Theoretical Knowledge*. Moscow: Progress-Tradition. 744 p. (In Russ.)

МЕЖДИСЦИПЛИНАРНОСТЬ КАК МЕТОДОЛОГИЧЕСКИЙ РЕЖИМ СОВРЕМЕННОЙ НАУКИ

Гребцов Максим Вячеславович – аспирант, Юго-Западный государственный университет, Курск, Россия; unforgiva@mail.ru.

Аннотация. В статье предпринимается философско-методологический анализ междисциплинарности как одного из значимых способов организации современного научного познания. Исследование не претендует на исчерпывающее описание всех предметных и организационных форм междисциплинарности. Правовое познание рассматривается как показательное поле, в котором особенно отчетливо проявляется необходимость согласования эмпирических, технических, нормативных, социальных и ценностных аспектов анализа. Показано, что обращение к междисциплинарному подходу обусловлено не формальным расширением исследовательского инструментария, а природой сложных объектов, которые не могут быть исчерпывающе описаны в пределах одной дисциплинарной оптики. Особое внимание уделяется тому, что междисциплинарность не сводится к простому соединению различных методов, понятий и научных языков. Ее эвристическая ценность раскрывается при условии строгого согласования исследовательских процедур, точного понятийного перевода и сохранения критериев доказательности. В работе подчеркивается, что современная наука нуждается не в механическом преодолении дисциплинарных границ, а в их осознанном и методологически контролируемом пересечении. На материале философии науки и правового познания обосновывается, что междисциплинарность позволяет раскрывать взаимосвязь различных измерений исследуемой проблемы. Делается вывод о том, что междисциплинарный подход продуктивен лишь тогда, когда он сохраняет дисциплинарное ядро исследования и обеспечивает проверяемость получаемых выводов.

Ключевые слова: методология науки; философия науки; междисциплинарность; научная рациональность; дисциплинарность; правовая наука; эпистемология.

Для цитирования: Гребцов М.В. Междисциплинарность как методологический режим современной науки // Научно-исследовательские исследования. – 2026. – № 3. – С. 83–96. – DOI: <http://www.doi.org/10.31249/scis/2026.03.07>

Статья поступила в редакцию: 19.05.2026

Принята к публикации: 18.06.2026

INTERDISCIPLINARITY AS A METHODOLOGICAL REGIME OF MODERN SCIENCE

Maxim V. Grebtsov – postgraduate student, Southwest State University, Kursk, Russia; unforgiva@mail.ru.

Abstract. The article undertakes a philosophical and methodological analysis of interdisciplinarity as one of the significant modes of organizing contemporary scientific knowledge. The study does not claim to provide an exhaustive description of all subject-specific and organizational forms of interdisciplinarity; legal inquiry is considered as an indicative field in which the need to coordinate empirical, technical, normative, social, and value-related aspects of analysis becomes especially visible. It demonstrates that the turn toward an interdisciplinary approach is determined not by the formal expansion of the research toolkit, but by the nature of complex objects that cannot be exhaustively described within a single disciplinary perspective. Particular attention is given to the fact that interdisciplinarity cannot be reduced to a simple combination of different methods, concepts, and scientific languages. Its heuristic value becomes evident under the condition of rigorous coordination of research procedures, precise conceptual translation, and preservation of standards of justification. The paper emphasizes that contemporary science requires not a mechanical overcoming of disciplinary boundaries, but their conscious and methodologically controlled crossing. Drawing on philosophy of science and legal inquiry, the article argues that interdisciplinarity makes it possible to reveal the interrelation of different dimensions of the problem under study. The conclusion is made that the interdisciplinary approach is productive only when it preserves the disciplinary core of the research and ensures the verifiability of the conclusions obtained.

Keywords: methodology of science; philosophy of science; interdisciplinarity; scientific rationality; disciplinarity; legal science; epistemology.

For citation: Grebtsov, M. V. (2026). Interdisciplinarity as a methodological regime of modern science. // *Naukovedcheskie issledovaniya = Science Studies*. – 2026. – N 2. – P. 83–96. – DOI: <http://www.doi.org/10.31249/scis/2026.03.07>

The paper was submitted: 19.05.2026

Accepted for publication: 18.06.2026

Введение

Современная наука существует в ситуации, когда предмет исследования все реже совпадает с границами одной дисциплины. Это не означает ослабления дисциплинарной организации знания. Напротив, высокая специализация наук сделала очевидным, что ряд проблем не может быть корректно решен путем простого накопления частных результатов. Научное знание сталкивается с объектами, в которых фактическое, техническое, социальное, нормативное и ценностное начала не расположены последовательно, а взаимно определяют друг друга. Поэтому междисциплинарность становится не формальным атрибутом исследования, а способом сохранения целостности сложного объекта.

Для российской науки данная проблема имеет не только теоретическое, но и нормативно-институциональное значение. Законодательное регулирование научной и научно-технической деятельности исходит из связи науки, государства, потребителей научной продукции и инновационной сферы¹. Стратегические документы научно-технологического развития также ориентируют исследование не на изолированное производство знания, а на решение сложных задач, требующих согласования научных, технологических и общественных последствий². Однако из самого факта нормативного признания комплексности еще не следует методологическая ясность. Главный вопрос состоит в том, каким образом объединять разные способы познания без утраты научной строгости.

При этом междисциплинарность не имеет единой формы для всех областей знания. В крупных физических, биомедицинских или инженерных проектах она может быть связана прежде всего с распределением экспериментальных, технических и вычислительных процедур. В социально-гуманитарном познании, и особенно в юриспруденции, центр тяжести смещается к согласованию факта, нормы, ценности, института и процедуры. Поэтому в настоящей статье правовое познание используется не как единственная модель междисциплинарности, а как методологически показательный случай: в этой области особенно заметно, что соединение разных научных языков требует не только расширения исследовательского инструментария, но и строгого контроля перехода от описания к нормативному выводу.

Междисциплинарность целесообразно рассматривать как особый режим исследования, при котором проблема формулируется на границе нескольких языков описания и требует согласования их объяснительных возможностей. Если дисциплина задает предмет, метод, критерии доказательства и допустимый тип результата, то междисциплинарное исследование вынуждено дополнительно обосновывать каждый из этих элементов. Оно не отменяет дисциплину, а выявляет пределы ее объяснительной достаточности. Поэтому междисциплинарность начинается с признания того, что объект не поддается однолинейному объяснению в пределах одной исследовательской оптики.

Междисциплинарность позволяет рассматривать объект в его целостности, соотнося данные и объяснительные модели, полученные в разных областях знания. Однако ее продуктивность не вытекает из самого факта привлечения нескольких дисциплин. Если междисциплинарность

¹ О науке и государственной научно-технической политике: Федеральный закон от 23.08.1996 № 127-ФЗ // Собрание законодательства Российской Федерации. 1996. № 35. Ст. 4137.

² О Стратегии научно-технологического развития Российской Федерации: Указ Президента Российской Федерации от 28.02.2024 № 145 // Официальный интернет-портал правовой информации. 2024. 28 февр. № 0001202402280003.

используется как риторическая декларация, возникают методологическая произвольность, неконтролируемый перенос понятий и утрата критериев проверки. Поэтому существенен не перечень названных дисциплин, а наличие между ними обоснованного перехода.

Философия науки должна анализировать практически сложившиеся формы междисциплинарных исследований с учетом результатов науковедения. В этой связи возникает вопрос о том, каким образом эти результаты включаются в философскую рефлексию о целях и границах научного познания, какие аспекты научной рациональности они позволяют уточнить и какие задачи философского анализа делают более предметными. Сближение с науковедением иногда дает основание рассматривать саму философию науки как междисциплинарное исследование.

Как утверждает Л.П. Клеева: «современные тенденции повышения доли и значимости междисциплинарных исследований, вызванные, с одной стороны, логикой развития науки, а с другой – настоятельной необходимостью отвечать требованиям, предъявляемым ей со стороны общества и народного хозяйства, актуализируют проблемы использования подходов одних отраслей науки к другим» [Клеева, 2025, с. 38].

Междисциплинарность в контексте современной научной рациональности

Такое понимание соответствует общей логике развития научной рациональности. В.С. Степин показывал, что тип научного познания изменяется вместе с характером исследуемых систем и формами включенности субъекта в познавательный процесс [Степин, 2003, с. 39–44]. Для классической науки была характерна ориентация на выделение объекта в максимально очищенном виде. Неклассическая и постнеклассическая рациональность усложняют эту схему: исследовательские средства, ценности, цели и последствия знания уже не могут быть вынесены за скобки. Поэтому современная методология вынуждена обсуждать не только вопрос о том, что познается, но и вопрос о том, в какой системе познавательных и социальных условий это знание производится.

Как отмечает В.С. Стёпин: «появление каждого нового типа рациональности не отбрасывало предшествующего, а только ограничивало сферу его действия» [Степин, 2003, с. 632]. Это не уничтожает дисциплинарную организацию науки, а выявляет границы ее достаточности.

В данном контексте важно различать несколько близких, но не тождественных форм научной кооперации. Мультидисциплинарность обычно означает параллельное рассмотрение одного объекта разными специалистами, когда результаты сопоставляются, но не обязательно образуют общий методологический механизм. Междисциплинарность предполагает более

глубокий уровень: она требует согласования понятий, процедур доказательства и исследовательских целей. Наконец, трансдисциплинарность связана с выходом за пределы академических дисциплин к таким проблемным полям, где в производство знания включаются практические, институциональные и общественные участники. В литературе эта градация неоднократно связывалась с изменением самой организации науки: от дисциплинарно замкнутого исследования к производству знания в контексте применения и общественной ответственности.

М. Гиббонс и соавторы в монографии, посвященной трансформации знания, называют междисциплинарный подход Mode 2 в отличие от традиционного дисциплинарного Mode 1. Они прямо указывают: «Mode 2 – это нечто большее, чем просто объединение широкого круга специалистов для совместной работы над проблемами в сложной среде...» [The New Production, 1994, p. 4]. Междисциплинарность не должна пониматься как простое сотрудничество нескольких специалистов или дисциплин.

Э. Янч характеризует движение к междисциплинарности как ситуацию, в которой «в конечном счете вся система образования / инноваций может быть скоординирована как многоуровневая многоцелевая иерархическая система через трансдисциплинарный подход, предполагающий обобщенную аксиоматику и взаимное усиление дисциплинарной эпистемологии» [Jantsch, 1970, с. 401]. Эта мысль показывает, что междисциплинарность связана не с внешним расширением тематики, а с координацией уровней знания и деятельности.

Однако указанное различие не образует простой шкалы прогресса, в которой каждая следующая форма автоматически выше предыдущей. Для философии науки принципиально важно определить, какой тип связи между дисциплинами соответствует природе исследуемого объекта. Если проблема может быть решена внутри одной науки, междисциплинарность остается внешней риторикой. Если же описание объекта неизбежно включает разные уровни реальности, исследование должно строиться как внутренне связанная конструкция, а не как совокупность комментариев. Поэтому междисциплинарность усложняет саму логику доказательства, а не только расширяет текст за счет новых терминов.

Междисциплинарность в юридической науке: возможности и методологические риски

Проблема междисциплинарности особенно отчетливо проявляется в юридической науке. Право нельзя свести ни к тексту нормативного акта, ни к поведению участников общественных отношений, ни к экономической эффективности регулирования. Юридическое познание работает одновременно с нормой, фактом, ценностью, процедурой и институтом. Поэтому правовая наука всегда содержит внутреннее напряжение между

догматическим анализом, философией права, социологией права, сравнительным правоведением и теорией правоприменения. Междисциплинарность здесь необходима, но она не должна разрушать собственно юридический способ постановки вопроса.

Первая методологическая опасность состоит в подмене сложности объекта произвольным смешением терминов. Исследователь может ссылаться на междисциплинарность, но фактически использовать слова разных наук без формирования единого доказательного пространства. В таком случае возникает не синтез знания, а риторическое расширение исследования. Научная ценность междисциплинарного подхода определяется не широтой привлеченных областей, а тем, насколько ясно показано, какую часть проблемы каждая дисциплина позволяет увидеть и какие ограничения при этом сохраняются.

Вторая опасность связана с методологическим империализмом одной науки. Экономический анализ может сводить право к критерию эффективности, социология – к поведению групп, психология – к индивидуальным установкам, информационные науки – к обработке данных. Каждый из этих подходов способен дать важный результат, но ни один из них не исчерпывает правовую реальность. Для юридической науки принципиально важно сохранять нормативность права как самостоятельный уровень анализа. В противном случае право превращается либо в социальный факт, либо в управленческую технологию, либо в массив данных, но перестает рассматриваться как форма должного.

Третья опасность состоит в утрате критериев проверки. Дисциплинарное исследование имеет устойчивые правила оценки: источники, методы, процедуры аргументации, допустимые выводы. Междисциплинарное исследование должно специально объяснять, почему вывод, полученный на пересечении разных подходов, является обоснованным. Критический рационализм К. Поппера важен здесь не как готовая методика для всех наук, а как напоминание о необходимости проверяемости и открытости знания для возможного опровержения [4, с. 60–64]. При отсутствии такого требования междисциплинарность легко превращается в форму интеллектуальной непроверяемости. К. Поппер подчеркивал, что «не верифицируемость, а фальсифицируемость системы следует рассматривать в качестве критерия демаркации» [Поппер, 2004, с. 63]. Так, соединение разных методов не освобождает исследование от требования проверяемости, а делает это требование более строгим.

Не менее значима и мысль Т. Куна о том, что научное сообщество действует внутри определенной парадигмальной рамки, задающей образцы постановки и решения проблем [Кун, 2009, с. 21–23]. Междисциплинарность усложняет данную ситуацию, так как в одном исследовании встречаются несколько рамок, каждая из которых считает свои методы естественными. Поэтому задача исследователя состоит не в формальном

перечислении подходов, а в выявлении их скрытых предпосылок. Иначе разные дисциплины будут говорить об одном объекте, но фактически иметь в виду разные предметы.

Дисциплинарные границы, понятийный перевод и доказательность

В связи с этим междисциплинарное исследование предполагает особую работу с границей дисциплин. Эта граница не является только препятствием. Она выполняет продуктивную функцию, поскольку заставляет исследователя уточнять, в каком смысле употребляется каждое понятие. Один и тот же термин может иметь различную смысловую функцию в правоведении, социологии, экономике или информатике. Например, понятие эффективности в экономике описывает соотношение затрат и результата, тогда как в праве оно не может быть отделено от законности, справедливой процедуры, равенства и предсказуемости решения. Поэтому простой перенос термина не формирует междисциплинарного знания, без перевода и ограничения смысла он создает методологический риск.

Для описания такого перевода полезна идея «пограничных объектов», разработанная в исследованиях науки и технологий. Подобные объекты позволяют разным профессиональным сообществам работать над общей проблемой, сохраняя собственные стандарты и языки описания [Star, Griesemer, 1989]. В юридическом исследовании такими пограничными объектами могут выступать цифровая платформа, экспертное заключение, статистический показатель, модель риска, медицинский протокол или экологический норматив. Они являются общими для нескольких дисциплин, но не имеют одного и того же значения в каждой из них. Поэтому задача методологии состоит в том, чтобы показать, каким образом такой объект получает юридически значимую форму и какие элементы его научного, технического или социального происхождения сохраняют значение для правового вывода. С.Л. Стар и Дж.Р. Грисемер пишут: «Пограничные объекты одновременно адаптируются к различным точкам зрения и достаточно устойчивы, чтобы сохранять свою идентичность при их рассмотрении» [Star, Griesemer, 1989, p. 387]. Для юридической науки это особенно важно: один и тот же объект может быть техническим, социальным, экономическим и правовым одновременно.

Методологически корректная междисциплинарность должна включать четыре обязательных элемента. Первый элемент – проблемное ядро. Исследование должно начинаться с вопроса, который нельзя удовлетворительно решить в пределах одной дисциплины. Второй элемент – понятийный перевод. Термины, заимствуемые из других наук, должны быть не перенесены буквально, а заново определены в отношении предмета исследования. Третий элемент – согласование доказательности. Необходимо по-

казать, какие данные, аргументы и процедуры признаются достаточными. Четвертый элемент – контроль вывода. Итоговое положение должно быть проверяемо с точки зрения тех дисциплин, которые участвовали в построении исследования, но при этом не распадаться на их отдельные результаты.

Практически это означает, что междисциплинарная работа должна строиться как последовательная процедура, а не как риторическое расширение исследования. На первом этапе необходимо установить, почему исходная дисциплина не обладает достаточной объяснительной силой. На втором этапе следует определить, какие дисциплины привлекаются и какую функцию выполняет каждая из них: описательную, объяснительную, нормативную, прогностическую или критическую. На третьем этапе требуется сформировать общий словарь исследования, где ключевые понятия не просто перечисляются, а получают рабочие определения. На четвертом этапе необходимо показать, каким образом соединяются результаты: через сопоставление, взаимную проверку, построение общей модели или выведение нормативного критерия. В современных учебниках по междисциплинарным исследованиям стадия интеграции результатов рассматривается как решающая, поскольку без нее работа остается совокупностью дисциплинарных фрагментов [Repko, Szostak, Buchberger, 2020].

Особую роль играет операция проблематизации. Она состоит в том, что исследователь не принимает готовую формулировку проблемы как самоочевидную, а выявляет, какие дисциплинарные предпосылки уже скрыты в этой формулировке. Так, вопрос о «регулировании искусственного интеллекта» может быть поставлен как техническая задача безопасности, как экономическая задача стимулирования инноваций, как правовая задача распределения ответственности или как философская задача сохранения человеческого контроля над решением. Все эти постановки правомерны, но они ведут к различным выводам. Междисциплинарность начинается тогда, когда эти постановки не конкурируют механически, а приводятся в такую связь, при которой каждая из них уточняет пределы другой.

В этом отношении полезна методология исследовательских программ И. Лакатоса, поскольку она позволяет рассматривать развитие знания не как ряд изолированных гипотез, а как устойчивую структуру, имеющую ядро, защитный пояс и правила движения проблемы [Огурцов, 1988, с. 135–140]. Междисциплинарная работа также нуждается в подобной структуре. У нее должно быть собственное ядро, иначе она становится набором внешних заимствований. Для области права таким ядром остается вопрос о нормативной организации общественных отношений, даже когда исследование использует данные социологии, экономики, психологии или цифровой аналитики.

Дисциплинарность при этом не является устаревшей формой организации знания. А.П. Огурцов убедительно показал, что дисциплинарная структура знания выполняет не только административную, но и познава-

тельную функцию: она удерживает предметность, нормирует язык и воспроизводит исследовательскую культуру [Огурцов, 1988, с. 112–118]. Поэтому противопоставление дисциплинарности и междисциплинарности методологически неверно. Междисциплинарность возможна только там, где существуют достаточно развитые дисциплины. Без дисциплинарной базы она превращается в произвольную универсализацию исследовательского языка.

Применительно к юридической науке это означает, что междисциплинарный подход не должен подменять юридическую аргументацию внешними объяснениями. Например, цифровизация правоприменения не может быть понята только как техническое внедрение алгоритмов. Она затрагивает вопросы ответственности, доказательств, процессуальной формы, правовой определенности и пределов усмотрения.

Если исследование ограничивается технической эффективностью, оно теряет собственно правовое измерение. Если оно ограничивается только догматикой, то не видит изменившихся условий функционирования нормы. Поэтому задача состоит в согласовании, а не в замещении методов.

Прикладные измерения междисциплинарного анализа права

Показательным примером является исследование цифрового правоприменения. На первый взгляд оно может быть описано как внедрение информационных систем в деятельность суда, администрации или правоохранительных органов. Но уже при ближайшем рассмотрении становится понятно, что техническая модель не совпадает с правовой моделью решения. Алгоритм работает с данными, вероятностями, классификациями и корреляциями; право работает с нормой, компетенцией, процедурой, доказательством и ответственностью. Поэтому вопрос состоит не только в том, насколько точно система прогнозирует результат, но и в том, может ли такой результат быть объяснен, оспорен, проверен и включен в процессуальную форму. Здесь информатика дает описание архитектуры системы, статистика – критерии надежности модели, социология – анализ институциональных последствий, а правовая наука устанавливает пределы допустимости автоматизированного решения. Междисциплинарность оказывается необходимой именно потому, что ни одна из этих оптик сама по себе не отвечает на вопрос о правомерности применения технологии.

Эта логика прослеживается на примере дела SyRI, рассмотренного Окружным судом Гааги 5 февраля 2020 г.¹ SyRI представляла собой систему риск-индикации, которую государство использовало для выявления возможного мошенничества в сфере социальных выплат, налогов и иных

¹ Rechtbank Den Haag. C/09/550982 / HA ZA 18-388, ECLI:NL:RBDHA:2020:865. 05.02.2020. URL: <https://uitspraken.rechtspraak.nl/details?id=ECLI%3ANL%3ARBDHA%3A2020%3A865> (дата обращения: 18.06.2026).

публичных обязательств. С философско-методологической точки зрения данный пример значим не только как случай применения цифровой технологии, но и как ситуация оценки объекта, который не принадлежит полностью ни праву, ни информатике, ни социальной политике. Техническая модель работала с массивами данных, корреляциями и риск-профилями; правовая оценка требовала ответа на другой вопрос: может ли такой способ обработки данных быть признан прозрачным, проверяемым, соразмерным и совместимым с правом на уважение частной жизни [Van Bekkum, Zuiderveen Borgesius, 2021].

Если применить к делу SyRI предложенную выше четырехэлементную процедуру, проблемным ядром становится не общий вопрос о пользе цифровизации, а допустимость превращения алгоритмического риск-сигнала в юридически значимый инструмент публичного управления. Понятийный перевод необходим уже потому, что «риск», «точность», «профиль» и «ошибка» имеют различный смысл в статистике, информатике и праве. Для технической системы ошибка может быть допустимым элементом вероятностной модели; для права она становится вопросом о пределах вмешательства государства, возможности оспаривания и защите лица от необоснованного подозрения. Согласование доказательности здесь означает различение технической надежности модели и юридической допустимости ее результата: даже высокая прогностическая эффективность не освобождает государство от обязанности раскрыть достаточные основания вмешательства. Контроль вывода требует, чтобы результат цифровой обработки мог быть объяснен, проверен и оспорен в правовой процедуре. Именно поэтому дело SyRI показывает: междисциплинарность возникает не при простом добавлении технологического материала к юридическому анализу, а там, где технический результат проходит через нормативный фильтр публичной ответственности.

Другой пример связан с судебной экспертизой, особенно с генетическими, медицинскими и психологическими исследованиями. Естественно-научный результат сам по себе еще не является юридическим фактом. Он должен пройти через процессуальные формы допустимости, относимости, достоверности и оценки доказательств. В этом случае биология или медицина дают материал, без которого невозможно установить значимые обстоятельства, но право определяет порядок превращения этого материала в доказательство. Ошибка возникает тогда, когда научная точность экспертного метода автоматически отождествляется с юридической достаточностью вывода. Междисциплинарный подход позволяет удержать различие между знанием о факте и правовым значением этого знания.

Третий пример дает экологическое и климатическое регулирование. Здесь юридическая норма не может быть адекватно понята без данных естественных наук, математического моделирования, экономики природопользования и социальной оценки рисков. Но и эти данные не превраща-

ются в правовое решение непосредственно. Между научным прогнозом и юридической обязанностью находятся процедуры нормотворчества, распределение ответственности, баланс публичных и частных интересов, а также вопрос о допустимом уровне риска. В подобных областях особенно заметно то, что современная наука производит знание в условиях неопределенности и общественного запроса, а потому нуждается в формах ответственности за последствия собственных рекомендаций [Nowotny, Scott, Gibbons, 2001].

Эти кейсы показывают, что методологически состоятельная междисциплинарность возникает там, где результаты дисциплин не просто подтверждают друг друга, а вступают в отношения взаимного ограничения. Техническая возможность ограничивается правовой допустимостью, экономическая целесообразность – нормативной справедливостью, эмпирическое наблюдение – процедурой доказательства, а ценностная оценка – требованием аргументированности. Поэтому междисциплинарное исследование не должно стремиться к устранению различий между науками. Напротив, оно должно выявлять эти различия и делать их методологически контролируемыми.

Критерии научной состоятельности междисциплинарного подхода

Содержательный критерий методологически состоятельной междисциплинарности состоит не в самом факте обращения к источникам из разных областей знания, а в том, что главный вывод исследования не может быть получен без их согласованного применения. Если искомый результат достигается средствами одной дисциплины, привлечение иных подходов остается вспомогательным и, по существу, иллюстративным. Иная ситуация возникает в тех случаях, когда вывод формируется за счет методологического соединения различных исследовательских оптик. В этом случае междисциплинарность приобретает не внешнее, а собственно научное значение.

Из этого вытекает специальное требование к научной рефлексии. Автор междисциплинарного исследования должен ясно понимать и обозначать пределы собственной компетенции, основания переноса понятий из одной области знания в другую, а также меру применимости полученных выводов. Для юридической науки это имеет принципиальное значение, поскольку некорректное заимствование понятий способно привести не только к теоретической неточности, но и к искажению понимания нормы, ответственности, процедуры либо правового статуса субъекта. В этом смысле философия науки выполняет функцию методологического контроля, она позволяет увидеть, где расширение предмета остается продуктивным, а где начинается утрата его собственных оснований.

Вместе с тем методологическая корректность междисциплинарного исследования сама по себе еще не устраняет его институциональных затруднений. Междисциплинарный проект нередко оказывается в промежуточном положении между сложившимися формами научного признания, его трудно соотносить с профилем кафедры, паспортом специальности, тематикой узкопрофильного журнала или критериями грантовой экспертизы. В результате исследование может быть достаточно строгим по существу, но восприниматься как не полностью относящееся к предметному полю каждой из участвующих дисциплин. Поэтому междисциплинарность требует не только понятийного, но и институционального перевода, автор должен показать, в каком научном сообществе, по каким критериям и в каком виде его результат может быть проверен и признан [Klein, 1990; Nowotny, Scott, Gibbons, 2001].

Следовательно, междисциплинарность не следует понимать как декларативный признак современной науки. Ее следует рассматривать как строгий способ работы со сложными объектами. Ее смысл заключается не в отказе от дисциплинарных границ, а в их осознанном преодолении тогда, когда сам предмет исследования требует более сложной исследовательской конструкции. Для современной науки это означает необходимость иной культуры аргументации, в которой фактическое описание, теоретическая модель, нормативная оценка и практические последствия рассматриваются не изолированно, а во взаимной связи.

Таким образом, философско-методологический анализ междисциплинарности важен еще и потому, что он позволяет отделить предметную сложность от неопределенности, прикрытой научной терминологией. Сложность требует ясного описания уровней объекта, способов связи между ними и процедуры перехода от частного результата к общему выводу. Неопределенность проявляется тогда, когда автор использует разные языки без объяснения их соотношения. В первом случае междисциплинарность усиливает исследование, во втором – ослабляет его, поскольку делает вывод менее проверяемым.

Для современной философии науки данная проблема имеет принципиальное значение. Она показывает, что научная рациональность не сводится к единообразию метода. Напротив, рациональность современного исследования заключается в способности сохранять различие методов и вместе с тем выстраивать между ними контролируемую связь. Поэтому вопрос о междисциплинарности является вопросом не только об организации науки, но и о культуре научного мышления. Чем сложнее объект, тем выше требование к методологической рефлексивности исследователя: необходимо ясно показать, в каком случае используется юридическая аргументация, где привлекаются данные другой науки, где производится философское обобщение и каким образом эти уровни связаны между собой.

Заключение

Рассмотрение правового познания в настоящей статье не исчерпывает многообразия междисциплинарных практик современной науки, однако позволяет выявить одну из наиболее методологически напряженных проблем: необходимость согласования эмпирических, технических, нормативных, социальных и ценностных аспектов анализа. Из изложенного следуют три вывода. Во-первых, междисциплинарность не является универсальным методом, она представляет собой особый режим исследования, возникающий при недостаточности одной дисциплинарной оптики. Во-вторых, ее научная состоятельность определяется сохранением дисциплинарного ядра, точностью понятийного аппарата и проверяемостью вывода. В-третьих, применительно к юридической науке речь идет не о растворении права в социальных науках, а о раскрытии права как нормативной формы, действующей в сложной социальной и технологической реальности.

Список литературы

- Клеева Л.П. Современные тенденции развития науки: проблемы междисциплинарности // Экономика науки. – 2025. – №2. – С. 29-40.
- Кун Т. Структура научных революций / пер. с англ. И.З. Налетова. – М.: АСТ, 2009. – 317 с.
- Лакатос И. Методология исследовательских программ / пер. с англ. – М.: АСТ; Ермак, 2003. – 380 с.
- Огуцзов А.П. Дисциплинарная структура науки. – М.: Наука, 1988. – 256 с.
- Поппер К. Логика научного исследования / пер. с англ. под общ. ред. В.Н. Садовского. – М.: Республика, 2004. – 447 с.
- Степин В.С. Теоретическое знание. – М.: Прогресс-Традиция, 2003. – 744 с.
- Jantsch E. Inter- and transdisciplinary university: A systems approach to education and innovation // *Policy Sciences*. – 1970. – Vol. 1. – P. 403–428. <https://doi.org/10.1007/BF00145222>
- Klein J.T. Interdisciplinarity: History, Theory, and Practice. – Detroit: Wayne State University Press, 1990. – 331 p.
- Nowotny H., Scott P., & Gibbons M. Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty. – Cambridge: Polity Press, 2001. – 278 p.
- Repko A.F., Szostak R., & Buchberger M.P. Introduction to Interdisciplinary Studies. 3rd ed. – Thousand Oaks: SAGE Publications, 2020. – 464 p.
- Star S.L., & Griesemer J.R. Institutional Ecology, “Translations” and Boundary Objects: Amateurs and Professionals in Berkeley’s Museum of Vertebrate Zoology, 1907-39 // *Social Studies of Science*. – 1989. – Vol. 19. – № 3. – P. 387-420.
- The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies / Gibbons M. et al. – London: SAGE Publications, 1994. – 179 p.
- Van Bekkum M., Zuiderveen Borgesius F. Digital welfare fraud detection and the Dutch SyRI judgment // *European Journal of Social Security*. – 2021. – Vol. 23. – № 4. – P. 323-340. – DOI: 10.1177/13882627211031257.

References

- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. SAGE Publications. 179 p.
- Jantsch, E. (1970). Inter- and transdisciplinary university: A systems approach to education and innovation. *Policy Sciences*, 1, 403–428. <https://doi.org/10.1007/BF00145222>
- Kleeva, L. P. (2025). Modern trends in the development of science: problems of interdisciplinarity. *Economics of Science*, (2), 29–40. (In Russ.)
- Klein, J. T. (1990). *Interdisciplinarity: History, theory, and practice*. Wayne State University Press. 331 p.
- Kuhn, T. (2009). *The structure of scientific revolutions* (I. Z. Naletov, Trans.). Moscow : AST. 317 p. (In Russ.)
- Lakatos, I. (2003). *The methodology of scientific research programmes*. Moscow : AST; Ermak. 380 p. (In Russ.)
- Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-thinking science: Knowledge and the public in an age of uncertainty*. Polity Press. 278 p.
- Ogurtsov, A. P. (1988). *The disciplinary structure of science*. Moscow : Nauka. 256 p. (In Russ.)
- Popper, K. (2004). *The logic of scientific research* (V. N. Sadovsky, Ed.). Moscow : Respublika. 447 p. (In Russ.)
- Repko, A. F., Szostak, R., & Buchberger, M. P. (2020). *Introduction to interdisciplinary studies* (3rd ed.). SAGE Publications. 464 p.
- Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, translations and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19(3), 387-420.
- Stepin, V. S. (2003). *Theoretical knowledge*. Moscow : Progress-Tradition. 744 p. (In Russ.)
- Van Bekkum, M., & Zuiderveen Borgesius, F. (2021). Digital welfare fraud detection and the Dutch SyRI judgment. *European Journal of Social Security*, 23(4), 323-340. <https://doi.org/10.1177/13882627211031257>

ПРЕДИКТИВНАЯ МЕТОДОЛОГИЯ МАШИННОГО ОБУЧЕНИЯ И ПОИСК МЕХАНИЗМОВ ПОЗНАНИЯ И ПОВЕДЕНИЯ

Сущин Михаил Александрович – канд. филос. наук, старший научный сотрудник, Институт научной информации по общественным наукам РАН (ИНИОН РАН), Москва, Россия; sushchin@bk.ru; ORCID 0000-0002-8805-6716

Аннотация. В статье рассматриваются возможности методологического влияния современной области машинного обучения на психологию и нейронауку для улучшения предсказывающего потенциала нейропсихологических моделей. Обсуждается аргумент сторонников предиктивных теорий, что преодоление проблем исследований, вызвавших кризис воспроизводимости в психологии, может быть достигнуто посредством использования больших датасетов, аккуратной оценки предсказательных ошибок модели, а также упрощения модели за счет устранения несущественных коэффициентов. В этой связи указывается, что критикуемые предиктивистами исследования с небольшой выборкой не являются образцом объяснительных психологических моделей. Теории в когнитивной психологии, когнитивной нейронауке, нейропсихологии и психофизиологии в большей степени имеют отношение к идентификации причинных механизмов познания и поведения. Вероятно, в большей степени точности и в больших возможностях манипуляции каузальными факторами познания и поведения кроется важнейшее различие между теориями механистических объяснений и вдохновленными машинным обучением предиктивными моделями. Последние в отрыве от поиска реальных причинных механизмов могут ухватывать только поверхностные корреляции между событиями. При принятии важных решений в контексте образовательной, биомедицинской и правовой деятельности лучше опираться на знание каузальных механизмов познания и поведения. Соответственно, предиктивные модели не могут вытеснить теории, ищущие механистические объяснения познания и поведения. Вместо выбора между предсказательной и объяснительной методологиями более продуктивным было бы нахождение оптимального баланса между ними.

Ключевые слова: психология; машинное обучение; предсказание; объяснение; методология науки.

Для цитирования: Сущин М.А. Предиктивная методология машинного обучения и поиск механизмов познания и поведения // Научоведческие исследования. – 2026. – № 3. – С. 97–112. – DOI: <http://www.doi.org/10.31249/scis/2026.03.08>

Статья поступила в редакцию: 09.06.2026

Принята к публикации: 26.06.2026

PREDICTIVE METHODOLOGY OF MACHINE LEARNING AND THE SEARCH FOR MECHANISMS OF COGNITION AND BEHAVIOR

Mikhail A. Sushchin – Candidate of Sciences (Philosophy), Senior Researcher, Institute of Scientific Information on Social Sciences of the Russian Academy of Sciences (INION RAN), Moscow, Russia; sushchin@bk.ru; ORCID 0000-0002-8805-6716

Abstract. This article considers the potential methodological impact of the modern field of machine learning on psychology and neuroscience to improve the predictive potential of neuropsychological models. It discusses the argument advanced by proponents of predictive theories that overcoming the research problems that have led to the reproducibility crisis in psychology can be achieved through the use of large datasets, careful evaluation of model prediction errors, and model simplification by dropping irrelevant coefficients. In this respect, it is noted that small-sample studies criticized by predictive theorists cannot be considered a paragon of explanatory psychological models. Theories in cognitive psychology, cognitive neuroscience, neuropsychology, and psychophysiology are more deeply involved in the business of identifying causal mechanisms of cognition and behavior. Theories of mechanistic explanations and machine-learning-inspired predictive models may be distinguished most importantly by the former's greater precision and greater potential for manipulating the causal factors of cognition and behavior. The latter, when disconnected from the search for real causal mechanisms, might only capture correlations between events. When making important decisions in the context of educational, biomedical, and legal practices, it is better to rely on knowledge of the causal mechanisms of cognition and behavior. Accordingly, predictive models cannot replace theories that strive for mechanistic explanations of cognition and behavior. Rather than choosing between predictive and explanatory methodologies, it would be more fruitful to search for an optimal balance between them.

Keywords: psychology; machine learning; prediction; explanation; methodology of science.

For citation: Sushchin, M.A. Predictive methodology of machine learning and the search for mechanisms of cognition and behavior // *Naukovedcheskie issledovaniya = Science Studies.* – 2026. – N 3. – P. 97–112. – DOI: <http://www.doi.org/10.31249/scis/2026.03.08>

The paper was submitted: 09.06.2026

Accepted for publication: 26.06.2026

О недавнем влиянии области искусственного интеллекта на науку

Недавние громкие успехи области искусственного интеллекта находят у всех на слуху. Менее чем за десять лет большие языковые модели (БЯМ) превратились из рудиментарных экспериментальных моделей в мощный инструмент, по своему влиянию на общественную жизнь сопоставимый со ставшими привычными поисковыми системами, электронной почтой, видеохостингами и т.д. Данное влияние, разумеется, не могло не распространиться и на науку и систему образования. За короткий срок обсуждение возможностей и рисков применения БЯМ в научном и образовательном процессах породило обширную литературу.

Так, почти сразу после ошеломляющего успеха ChatGPT и Github Copilot в 2022 г. многие авторы и авторские коллективы стали обсуждать те значительные возможности, которые открылись перед учеными, преподавателями и учащимися с появлением БЯМ. В этой связи чаще всего отмечается способность БЯМ моментально делать краткие обзоры каких-либо отдельных работ или даже целых областей знания, отталкиваясь от уровня пользователя, разъяснять простым языком сложные вещи (скажем, основы программирования, высшей математики или квантовой физики), корректировать тексты и переводить их на другие языки (а также давать консультации по тончайшим аспектам перевода), генерировать (и разъяснять) компьютерный код на всех основных языках программирования и др. Таким образом, у представителей научного и образовательного сообществ появился легкодоступный мощный персональный ассистент, способный облегчить выполнение многих видов работ.

Равным образом значительное внимание было приковано и к рискам, связанным с использованием БЯМ в науке и образовании. Корень проблем может быть во многом связан с природой БЯМ. Как известно, в своей основе БЯМ представляют собой автоматы, осуществляющие генерацию следующего наиболее *вероятного* в данном контексте символа. Ввиду этого с их ответами связана принципиальная ненадежность, невозможность гарантировать предоставление исключительно корректной информации, подчас проявляющаяся в том, что даже на элементарный вопрос машина может дать неверный ответ. (Часто генерация всецело вымышленной информации проявляется при запросе списка литературы по какой-либо теме.) Данная проблема усугубляется возрастающей степенью неотличимости ответов БЯМ от ответов простых людей и даже экспертов. Наконец, само обстоятельство активного распространения БЯМ ставит проблему излишней зависимости пользователей от такого рода интеллектуальных инструментов. Так, способность БЯМ в считанные мгновения делать обзор отдельных работ или целых областей знания создает искушение пренебрежения традиционными аналитическими видами работ с текстами, что может иметь губительные последствия для формирования навыков критического анализа информации и творческих способностей.

Впрочем, эти и иные схожие вопросы применения БЯМ в науке и образовании затрагивают лишь поверхностные аспекты влияния разработок искусственного интеллекта (ИИ) на эти области. В этом отношении наука испытывает такое же влияние, как другие области жизни современного человека. Между тем ажиотаж последних лет оттеснил на второй план гораздо более радикальные и глубокие попытки концептуального влияния современной области машинного обучения на научные исследования. Так, еще за некоторое время до возникновения БЯМ в психологии появился новый тип предсказывающих моделей, вооружившихся ключевыми методологическими принципами области машинного обучения. В этой связи один из главных сторонников моделей нового типа психолог

Т. Яркони выступил с чрезвычайно смелым тезисом, что психология должна отдать приоритет способности предсказания поведения человека, а не главенствующим уже долгое время попыткам его объяснения [Yarkoni, Westfall, 2018]. По мнению Яркони и других адептов этого вида предиктивизма, приоритет предсказаниям может быть отдан даже *в ущерб* объяснительной активности. Вдохновленное машинным обучением предиктивное моделирование уже было охарактеризовано как «поворот» в психологии и нейронауках [Weiskopf, 2022]. Вместе с тем обоснованность некоторых утверждений сторонников этого направления вызывает сомнения. В этой статье будет предпринята попытка оценить статус предиктивных моделей нового типа в связи с более традиционными глубокими причинными объяснительными теориями в психологии и нейронауке.

Современные предиктивные модели: теоретическая мотивация и конкретные примеры

Традиционно значительная часть психологических исследований (особенно в области психологии личности и социальной психологии) фокусировалась на нахождении статических показателей, факторов, которые позволяли бы предсказывать поведение человека (predictors). Например, как добросовестность может влиять на производительность труда или академическую успеваемость, как психологические травмы в детстве могут обуславливать депрессию во взрослом возрасте, как размер подкупа может влиять на самовнушение [Festinger, Carlsmith, 1959] и т.п.

Подобного рода исследования обычно проводились на основе небольшой выборки данных (небольшого количества участников) с использованием экспериментального протокола, именуемого проверкой значимости нулевой гипотезы. В основе этого протокола лежит принятие в качестве базовой так называемой нулевой гипотезы H_0 – предположения об отсутствии причинной связи между изучаемыми феноменами. Одновременно с этим выдвигается противоположная гипотеза H_1 о наличии связи между данными феноменами. После сбора данных проводится статистический тест для расчета вероятности нулевой гипотезы p . Если $p < 0.05$ (менее одного шанса из двадцати), заключают о верности гипотезы H_1 , а нулевую гипотезу отбрасывают. Как отмечают в этой связи Д. Иглесиас, М. Соррель и Р. Олмос: «Если обнаруживаются статистически значимые результаты (т.е. $p < 0.05$), делается вывод, что рассматриваемая теория прошла проверку на соответствие действительности (passed the truth test) и, таким образом, предлагает объяснение на основе эмпирических свидетельств» [Iglesias, Sorrel, Olmos, 2025].

В теории исследования, проведенные в соответствии с указанными стандартами, должны быть достаточны и для *объяснения*, и для *предсказания* поведения человека. «С этой точки зрения, объяснение необходимым

образом способствует предсказанию: модель, которая в наилучшей степени аппроксимирует ментальные процессы, порождающие наблюдаемое поведение, также лучше всего предсказывает будущее поведение» [Yarkoni, Westfall, 2017, p. 1]. Однако, как повсеместно отмечают критики, на практике исследования с небольшой выборкой столкнулись с серьезными проблемами в аспекте предсказания поведения человека.

Прежде всего, 2010-е гг. были отмечены масштабным кризисом воспроизводимости (*replication crisis*) в психологической науке, поставившим под сомнения основания исследований с небольшой выборкой. Так, по результатам работы проекта «Сотрудничество в сфере открытой науки» (*Open Science Collaboration*) были сделаны весьма неутешительные выводы. В ходе попыток воспроизвести результаты 100 известных исследований из областей социальной и когнитивной психологии, две трети попыток «оказались неуспешными, а когда они были успешными, масштаб эффектов часто оказывался меньше, чем в оригинальных исследованиях.... Как утверждает в комментарии в журнале *Nature* (Baker, 2015): “Не верьте всему, что Вы читаете в психологической литературе. В действительности, две трети [этой литературы], вероятно, следует поставить под сомнение”» [Smith, Little, 2018].

Это обстоятельство позволило критикам заключить, что, хотя объяснение и предсказание в психологии могут быть связаны между собой как две стороны одной монеты, на практике исследователям придется делать ставку либо на объяснительную, либо на предсказательную способность своих теорий. Так, в уже цитировавшемся манифесте Яркони и Вестфолла, громко названном «Выбирая предсказание вместо объяснения в психологии: уроки машинного обучения», указывается, что «со статистической точки зрения, просто не верно, что модель, лучше всего аппроксимирующая порождающий данные процесс (*data generating process*), в общем случае будет успешнее всего предсказывать результаты в реальном мире» [Yarkoni, Westfall, 2017, p. 1]. По их мнению, за образец в этом отношении стоит взять добившуюся резонансных успехов область машинного обучения.

Так, в то самое время, как в психологии разворачивался кризис воспроизводимости исследований с небольшой выборкой, вооруженная массивными датасетами и новыми вычислительными мощностями область машинного обучения поставила, по сути, на конвейер развитие успехов коннекционистских сетей 1980–1990-х гг. С учетом успехов близкой по духу области машинного обучения и ввиду появления больших данных в психологии, Яркони и Вестфолл предлагают реорганизовать психологическую науку по методологическим лекалам области машинного обучения: «с учетом относительно недавней революции в теории и методологии *машинного обучения* – в которой предсказание ненаблюдаемых данных рассматривается как золотой стандарт успеха, а объяснение обычно представляется малоинтересным или неинтересным – а также последовательно увеличивающейся доступностью масштабных датасетов, описывающих

поведение человека <...> возможно продвинуть психологию ближе к предсказывающей науке...» [Yarkoni, Westfall, 2017, p. 2].

В чем же именно с этой точки зрения может заключаться проблема исследований с небольшой выборкой, которая привела к ситуации, когда результаты двух третей известных исследований просто не были воспроизведены повторно? Как уже говорилось выше, в теории исследования, предлагающие убедительные объяснения психологических феноменов, также должны хорошо предсказывать их. Однако на практике этого часто не происходит. Как правило, объяснительные исследования оцениваются по тому, насколько хорошо они описывают исходные данные. Этот показатель известен как «качество соответствия» (goodness of fit). Естественно, что модель, хорошо описывающая свои данные, будет иметь низкую степень ошибок (т.н. ошибок обучения) при проверке этими же исходными данными. Как утверждают Яркони и Вестфолл, проблема заключается в том, что хорошая степень соответствия исходным данным не только не гарантирует того, что модель будет хорошо предсказывать при проверке новыми данными (out-of-sample data), но и зачастую этому откровенно препятствует.

Почему же модель, хорошо описывающая исходные данные, часто имеет слабые предсказательные характеристики? Так, данная проблема особенно остро стоит для моделей с *большим* количеством предикторов (тех самых факторов, позволяющих предсказывать поведение), подгоняемых к *небольшим* датасетам. Существо проблемы многие авторы связывают с тем, что в процессе обучения на небольших данных модель ухватывает не только паттерны для предсказания, но и характерный для этого конкретного небольшого датасета шум (скажем, имеющиеся у данной небольшой группы работников, а не населения в целом, особенности, позволяющие им показывать высокую производительность труда). Поэтому при тестировании такой модели на новом датасете ее предсказательный потенциал оставляет желать лучшего. Этот феномен плохой обобщаемости модели на новые данные получил название переподгонки (или переобучения, overfitting).

Кроме статистической переподгонки, положение дел усложняется еще и склонностью некоторых исследователей слишком вольно обращаться с данными, чтобы получить желаемое значение нулевой гипотезы (p -значение, p -value, ниже конвенционального порога 0.05). Например, распространена практика произвольной остановки исследования (optional stopping) во время сбора данных – для этого периодически подсчитывают p -значение и, если оно укладывается в 0.05, дальнейшие подсчеты просто останавливают. Такого рода практики вольного обращения с данными для получения нужного p -значения обрели широкую известность как p -хакинг (подгонка данных, p -hacking). По мнению Яркони и Вестфолла, p -хакинг можно рассматривать как особого рода форму переподгонки: «ее можно понимать как форму *процедурной переподгонки* (курсив авторов – М.С.), которая происходит до или параллельно с оценкой модели – например, во время очистки данных, отбора модели или выбора анализа для отчета. По-

добно тому, как статистическая модель может учитывать специфические закономерности в учебных данных, которых на самом деле нет в генеральной совокупности, исследователи-люди часто замечают и действуют на основе специфических различий в результатах, полученных при использовании различных аналитических методов» [Yarkoni, Westfall, 2017, p. 5].

Итак, для обуздания этой излишней гибкости психологических теорий и превращения психологии в успешную предсказывающую науку Яркони и Вестфолл предлагают обратиться к ключевым методологическим правилам современной области машинного обучения, позволившим ей добиться тех впечатляющих успехов, которые, собственно, находятся сейчас у всех на слуху. В общем смысле важнейшим фактором успешной минимизации ошибок предсказаний должно стать нахождение оптимального баланса между тем, что в статистике и машинном обучении именуется смещением (bias) и дисперсией (variance). Например, при игре в дартс может случиться так, что все дротики попадут только в какую-то одну половину мишени, будучи там относительно свободно распределены (ситуация высокого смещения и высокой дисперсии) или, скажем, тесно сгруппированы (высокое смещение и низкая дисперсия). Либо же они могут свободно распределиться относительно центра мишени (низкое смещение и высокая дисперсия) или вообще все попасть близко к «яблочку» (низкое смещение и низкая дисперсия). Чаще всего на практике исследователи сталкиваются с двумя средними случаями: высоким смещением / низкой дисперсией и низкой дисперсией / высоким смещением. В последнем случае модель может быть гибкой и хорошо описывать имеющиеся данные, но плохо предсказывать при проверке новым датасетом (характерная особенность многих психологических моделей, полученных на небольшой выборке). Либо же модель может пожертвовать свойством аккуратного соответствия тренировочным данным в пользу большей простоты (высокого смещения), но и одновременно лучшей обобщаемости и предсказательной способности. Именно данному подходу, отмечают Яркони и Вестфолл, чаще всего и отдается предпочтение в области машинного обучения.

В общем и целом, авторы манифеста предлагают три непосредственные меры для улучшения предсказательной способности психологических моделей, широко применяющиеся в области машинного обучения. Таково, во-первых, использование больших данных (больших датасетов) для улучшения обобщаемости модели. Во-вторых, требуется применение техник аккуратной оценки ошибок предсказаний моделей для их диагностики и улучшения. В-третьих, это способность желаемым образом настроить соотношение между смещением и дисперсией (преимущественно в пользу смещения), чтобы избежать переподгонки и улучшить предсказывающие характеристики модели.

Итак, первой непосредственной мерой в борьбе с переподгонкой и излишне гибкими исследовательскими практиками должно стать использование больших данных. Главное достоинство массивных датасетов – их

большая репрезентативность: чем больше выборка, тем меньше сюрпризов будет при проверке предсказаний модели новыми данными. При наличии больших датасетов даже гибкие полиномиальные модели могут выдавать меньшую ошибку, чем более простые линейные модели, что крайне маловероятно, когда исследователи оперируют небольшими данными. Эта особенность больших датасетов, по замечанию Яркони и Вестфолла, стала основанием для известного афоризма, а именно что «больше данных лучше, чем лучшие алгоритмы»: «Очень часто, – пишут они, – самая лучшая вещь, которую исследователь может сделать для улучшения обобщающей способности модели – и едва ли вообще не самая простая и дешевая – это собрать больше данных» [Yarkoni, Westfall, 2017, p. 9].

Одним из интересных и неожиданных следствий постепенного внедрения больших данных в психологическую науку стал пересмотр результатов, полученных в исследованиях с небольшой выборкой. Так, с увеличением датасетов оценки эффектов, полученных в прежних исследованиях, снизились. Предполагая, что сами психологические феномены в целом не изменились, Яркони и Вестфолл утверждают, что наиболее правдоподобным объяснением снижения оценок эффектов в эпоху больших данных является, то обстоятельство, что прежние исследования, вероятно, страдали от переподгонки. В силу своей природы исследования с небольшой выборкой могут давать более вариативные оценки тех или иных феноменов, которые имеют тенденцию сглаживаться при появлении более широкой перспективы. По Яркони и Вестфоллу, для преодоления кризиса воспроизводимости психология должна отдавать предпочтение небольшим эффектам из исследований с большой выборкой, а не наоборот [Yarkoni, Westfall, 2017, p. 11].

Второй важнейшей мерой для улучшения предсказывающего потенциала психологических теорий должно стать применение техник количественной оценки ошибок предсказаний, получаемых на новых датасетах. В области машинного обучения для этих целей чаще всего используется техника перекрестной проверки (кросс-валидации, cross-validation). Существо метода (в упрощенном виде) заключается в разделении датасета на две части (блока, folds), одна из которых используется для подгонки модели, а другая – для тестирования (имитируя таким образом проверку модели новыми данными). Затем, чтобы продуктивно использовать весь датасет и избежать недоподгонки (underfitting, поскольку изначально тренировочные данные были уменьшены в два раза), проводят еще один цикл подгонки и тестирования, меняя ролями те самые части датасета (блоки). В итоге состоятельность модели оценивается путем усреднения значений, полученных в каждом цикле (или итерации) подгонки и проверки. (В общем виде данный подход именуется k -кратной перекрестной проверкой, k -fold cross-validation, так как датасет разбивается на k равных частей – чаще всего от 3 до 10, – где одна часть используется для обучения, а все остальные – для тестирования.)

Перекрестная проверка представляет собой важнейший инструмент борьбы с переподгонкой моделей. При аккуратном использовании с ее помощью можно получить значения, близкие к тем, которые обычно наблюдаются при проверке модели новыми данными. Также она может помочь и в борьбе с *p*-хакингом. Например, исследователь может попробовать несколько методов обработки данных и выбрать в итоге тот, который лучше всего себя показывает с точки зрения перекрестной проверки [Yarkoni, Westfall, 2017, p. 12].

Наконец, в дополнение к указанным мерам исследователям также нужны непосредственные средства манипуляции моделями, если известно, что их предсказывающий потенциал может быть улучшен. На эту роль авторы манифеста предлагают технику, известную как регуляризация «лассо» (lasso regularization, least absolute shrinkage and selection operator). Суть регуляризации «лассо» заключается в уменьшении до нуля (и, по сути, элиминации) наименее существенных коэффициентов модели, что, собственно, и ведет к ее упрощению. Это достигается за счет добавления к функции ошибок (разнице между предсказаниями модели и наблюдаемыми результатами) так называемого штрафа (penalty term) – числа, пропорционального сумме абсолютных значений коэффициентов модели. Скажем, с увеличением суммы коэффициентов модели возрастает и ее «объяснительная» способность (т.е. ее соответствие данным), однако вместе с этим возрастает и штраф. В духе обсуждавшегося выше подхода к дилемме смещения и дисперсии метод лассо намеренно отдает предпочтение простоте моделей (высокому смещению) за счет обнуления несущественных коэффициентов. Такой подход может быть наиболее полезен, по замечанию Яркони и Вестфолла, как раз для моделей с большим числом предикторов по отношению к размеру тренировочных данных (которые, как мы видели выше, распространены в психологии образования, психологии личности и др.)

Таким образом, Яркони, Вестфолл, Вейскопф и другие адепты этого направления призывают к массовой переориентации психологических исследований с поиска каузальных механизмов поведения на создание предиктивно аккуратных моделей. Непосредственным следствием данной методологической установки оказывается рассмотрение причинных механизмов поведения в качестве своего рода черного ящика. Психолога, принявшего на вооружение методологические принципы области машинного обучения, в первую очередь интересует предсказательная точность его моделей, а не то, насколько хорошо они ухватывают причинные механизмы поведения. В этом отношении Яркони и Вестфолл ссылаются на идею двух статистических культур Бреймана: «Подавляющее большинство статистиков относится к “культуре моделирования данных”, в рамках которой предполагается, что данные являются следствием некоторого процесса порождения данных (data generating process), и главная цель заключается в оценке истинных параметров этого процесса. В противоположность, не-

значительное меньшинство статистиков (и большинство исследователей в области машинного обучения) принадлежит к “культуре алгоритмического моделирования”, в которой предполагается, что *данные представляют собой результат некоторого неизвестного и, возможно, непознаваемого процесса* [курсив мой – М.С.], а основная цель заключается в нахождении алгоритма, который при тех же входных данных дает те же результаты, что и этот процесс» [Yarkoni, Westfall, 2017, p. 19].

Помимо, собственно, теоретических манифестов, с этим направлением ассоциируется ряд конкретных экспериментальных исследований, выдержанных в духе обсуждавшейся выше методологии. Одним таким исследованием является экспериментальная работа Риссмана, Грили и Вагнера [Rissman, Greely, Wagner, 2010]. В этом исследовании была предпринята попытка оценить, является ли зафиксированная с помощью фМРТ активность мозга надежным маркером (предиктором) воспоминаний. Так, двадцати участникам предъявлялся ряд изображений лиц, а час спустя их способность распознать предъявлявшиеся изображения была оценена во время сканирования мозга с помощью фМРТ. Зафиксированные паттерны активности мозга субъектов с помощью алгоритма мультивоксельного анализа паттернов (МВАП) использовались для предсказания (1) субъективной уверенности участников (их убежденности, в том, видели ли они то или иное изображение ранее), а также (2) того, насколько они реально помнят то или иное изображение. Оказалось, что предиктивный алгоритм намного увереннее справляется с первой задачей, корректно предсказывая субъективную уверенность участников с вероятностью примерно в 90 процентов, тогда как действительные воспоминания успешно предугадывались с вероятностью чуть более шанса (около 56 процентов). Эти результаты дали основание авторам исследования заявить, что метод фМРТ имеет существенные ограничения для определения действительных воспоминаний субъектов (что крайне важно, например, в контексте рассмотрения использования фМРТ в судебных разбирательствах, так как с его помощью, вероятно, нельзя достоверно определить, видел ли свидетель подозреваемого или нет).

В другом известном исследовании паттерны нейронной активности отдельных индивидов использовались для предсказания поведения целых групп населения [Falk, Berkman, Lieberman, 2012]. Так, группе курильщиков, желавших покончить с этой вредной привычкой, были показаны видеоролики из трех рекламных кампаний против курения (по несколько роликов из кампаний, условно названных А, В и С). В это время активность их мозга (медиальной префронтальной коры) фиксировалась с помощью фМРТ. После этого им давали возможность субъективно оценить убедительность этих кампаний. В результате участники наиболее убедительной назвали кампанию В, затем кампанию А, а затем кампанию С. Между тем анализ данных нейронной активности показывал иной рейтинг убедительности: $C > B > A$. В итоге именно нейронная активность субъектов оказа-

лась более состоятельным предиктором эффективности рекламных кампаний в масштабе всего населения: по результатам звонков на горячую линию был получен аналогичный рейтинг эффективности – $C > B > A$.

Несмотря на использование отдельных предиктивных техник и алгоритмов, указанные исследования все-таки основывались на стандартных небольших выборках. Так, недавно созданная генеративная модель Delphi-2M является примером полноценной предсказывающей модели, созданной на основе больших данных [Shmatko et al., 2025]. Эта модель позволяет предсказывать индивидуальные риски развития более тысячи заболеваний на срок до двадцати лет вперед. Delphi-2M представляет собой нейросеть-трансформер на основе модели GPT-2, натренированную на анонимизированных данных 402799 пациентов из британской биомедицинской базы данных UK Biobank. Подобно тому, как большие языковые модели предсказывают последовательности символов (токенов) с учетом поступившего запроса, так и Delphi-2M предсказывает развитие того или иного заболевания на основе имеющейся клинической картины, где отдельными токенами выступают события из жизни пациента (пол, состояние здоровья в определенном возрасте, вредные привычки в определенном возрасте и т.п. – всего 1258 токенов; например, если на входе имеется «пациент 41 год, курит и употребляет алкоголь», то на выходе может быть получен ответ «44 года, варикозное расширение вен»). Модель получила независимую проверку на данных почти двух миллионов пациентов из датского национального реестра пациентов. В итоге Delphi-2M достигла предсказательной точности, сопоставимой с показателями предиктивных моделей для одного заболевания, воплощая в себе обсуждавшиеся методологические принципы и в то же время делая шаг на пути к прозрачности «черного ящика» внутренних вычислений [Learning the natural history, 2025].

Предиктивизм черного ящика и поиск причинных механизмов в науке

Итак, мы рассмотрели теоретическую мотивацию, методологические основания и конкретные примеры недавних предсказывающих моделей в психологии и нейронауке. Как можно оценить их статус в общем контексте развития и структурной организации психологии и нейронауки? Действительно ли появление предиктивной методологии является «поворотным моментом» в истории этих дисциплин? Действительно ли предиктивная методология обладает автономией по отношению к стандартным объяснительным практикам и должна выйти из их тени [Weiskopf, 2022]? Насколько оправданы призывы выбрать предсказание вместо объяснения в качестве ведущего методологического ориентира в психологии и нейронауке?

Прежде всего попробуем разобраться с уместностью альтернативы «предсказание или объяснение» для психологии и нейронауки. Здесь следует начать с того, что критикуемые адептами методологического предиктивизма распространенные исследования с небольшой выборкой в социальной психологии или психологии личности не исчерпывают собой все стили психологических исследований. Более того, было бы преувеличением говорить, что именно теории в этих областях служат парадигмой научных работ, направленных на поиск причинных механизмов поведения, психических функций и познания. Исследования в нейропсихологии, психофизиологии, когнитивной психологии и когнитивной нейронауке куда в большей степени связаны с поиском реальных причинных механизмов познания и поведения.

Говоря о причинных механизмах познания и поведения, я буду во многом опираться на концепцию механистических объяснений в нейронауке философа науки К. Крейвера [Craver, 2007]. Согласно Крейверу, объяснение представляет собой прежде всего описание причинных механизмов. Механизмы состоят из сущностей (entities), обладающих теми или иными свойствами, которые позволяют им принимать участие в определенных процессах, или активностях (activities). Например, механизм везикулярного высвобождения нейромедиаторов включает в себя множество компонентов (т.е. сущностей: ионы кальция, везикулы с нейромедиаторами, поры слияния, нейронные мембраны и др.) и их разнообразных активностей (открытие, фиксация, диффузия, слияние, прайминг и т.д.) [Craver, 2007, p. 5]. Коротко говоря, по Крейверу, механизм есть «ряд сущностей и активностей, организованных таким образом, что они демонстрируют (exhibit) подлежащий объяснению феномен» [Там же]. Используя терминологию философии науки, в концепции Крейвера описание механизма представляет собой эксплананс (объяснение), а порождаемый механизмом феномен – экспланандум (нечто, подлежащее объяснению).

При этом объяснение не сводится только к описанию активностей наиболее элементарных компонентов на клеточно-молекулярном уровне: психологи и нейрочлены в своей работе (а Крейвер в первую очередь настаивает на дескриптивной аккуратности своей концепции) принимают во внимание реальную каузальную значимость более высоких уровней, включая активность непосредственно на уровне психологических событий и процессов. В качестве примера многоуровневых объяснений Крейвер ссылается на исследования пространственной навигации у грызунов в водном лабиринте Морриса [Craver, 2007, pp. 165–170]. В этих исследованиях чаще всего различают четыре уровня. (1) На верхнем уровне пространственной памяти осуществляются задачи навигации и выучивания пути к скрытой под водой платформе. (2) На более низком уровне формирования пространственных карт ищут и исследуют непосредственный нейронный субстрат навигационной активности (связывая его с деятельностью гиппокампа и других височных и лобных областей коры мозга).

(3) «Этажом» ниже, на клеточно-электрофизиологическом уровне лежат механизмы долговременной потенциации (long-term potentiation), выступающие основой для навигационной активности гиппокампа. (4) Наконец, на нижнем молекулярном уровне находятся «механизмы, делающие возможными химические и электрические активности нервных клеток» [Craver, 2007, p. 169]. Именно все в совокупности эти уровни обеспечивают объяснение феномена пространственной навигации у грызунов.

В конечном счете нейронаука, по Крейверу, имеет в качестве фундаментальных целей как механистическое объяснение тех или иных психофизиологических феноменов (восприятие цвета, пространственная навигация, развитие нервной системы и т.д.), так и контроль над мозгом и нервной системой для диагностирования и лечения болезней, улучшения мозговых функций и др. [Craver, 2007, p. 1].

По-видимому, именно в большей степени точности и в больших возможностях контроля реальных каузальных факторов познания и поведения кроется важнейшее различие между теориями и моделями механистических объяснений, с одной стороны, и вдохновленных машинным обучением современными предиктивными моделями, с другой стороны. Последние при всех их достоинствах в борьбе с перепогонкой и *p*-хакингом могут ухватывать поверхностные корреляции между событиями, в действительности не имеющими причинных взаимосвязей. (Подобно тому, как, скажем, активное использование кондиционеров коррелирует, но не является причиной всплеска продаж мороженого.)

Например, авторы рассмотренной выше модели Delphi-2M открыто заявляют: «Несмотря на то, что Delphi-2M представляется способной моделировать направленные во времени зависимости, мы предостерегаем от интерпретации этих зависимостей в качестве причинно-следственных связей, которые можно было бы использовать для изменения будущего состояния здоровья» [Learning the natural history, 2025]. По их мнению, результаты работы с Delphi-2M говорят в пользу осторожного подхода «при использовании моделей ИИ для осуществления вывода и прогнозирования в гетерогенных наборах данных здравоохранения, потенциально указывая на то, что они могут быть полезным дополнением к существующим диагностическим конвейерам, а не заменой им» [Там же].

Представляется, что только что приведенные суждения содержат в себе, в целом, взвешенную оценку характера и особенностей практического использования результатов, полученных с помощью предиктивно-статистической методологии. Очевидно, что предельная осторожность требуется при отсылке к результатам таких исследований в областях образования и медицины, при рассмотрении правовых вопросов. Во всех этих контекстах гораздо более надежной опорой для принятия решений является знание конкретных каузальных механизмов поведения. Поэтому, вероятно, вместо альтернативы «предсказание или объяснение» более продуктивным выбором было бы нахождение оптимального баланса между

этими исследовательскими стилями с пониманием того, что исследования каузальных механизмов не могут быть вытеснены методологией, опирающейся на идею черного ящика.

Впрочем, это признают и сами авторы манифеста предсказательных моделей нового типа. Так, они пишут, что они вовсе не призывают к «полному отказу от попыток механистически объяснить поведение человека. <...> Мы, однако, охотно признаем, что предиктивно ориентированные (prediction-focused) подходы не подходят для всех видов исследовательских вопросов в психологии. В частности, хорошо организованные, масштабные, рандомизированные, контролируемые эксперименты есть и должны оставаться золотым стандартом для получения каузальных заключений о работе человеческого разума. <...> Таким образом, мы отстаиваем точку зрения, что даже в тех случаях, когда каузальное объяснение есть (надлежащим образом) главная цель, понятия и методология машинного обучения все же могут оказать неоценимую помощь при их инструментальном использовании – посредством минимизации *p*-хакинга, увеличения исследовательской эффективности, способствования оценке работы моделей и увеличения интерпретируемости» [Yarkoni, Westfall, 2017, p. 19].

Заключение

Итак, в данной статье был рассмотрен конкретный проект методологического влияния современных исследований искусственного интеллекта (машинного обучения) на такие научные дисциплины, как психология и нейронаука. Мы рассмотрели основные аргументы сторонников использования предиктивной методологии и предиктивных моделей, сводящиеся к тому, что преодоление основных проблем исследований с небольшой выборкой (переподгонка, *p*-хакинг и, как следствие, плохая воспроизводимость) может быть достигнуто посредством (1) использования больших датасетов, (2) аккуратной оценки предсказательных ошибок модели (перекрестная проверка, или кросс-валидация), а также (3) упрощения модели за счет обнуления несущественных коэффициентов (регуляризация лассо). Были также рассмотрены конкретные примеры таких моделей из областей психологии, нейронауки и биомедицины.

Оценивая этот проект, я обосновывал следующие положения. (1) Исследования с небольшой выборкой, критикуемые в связи с кризисом невоспроизводимости, не исчерпывают собой все методологические стили в психологии и нейронауке. Неверно рассматривать их в качестве образца объяснительных теорий и моделей в этих дисциплинах. Такие дисциплины, как когнитивная психология, когнитивная нейронаука, нейропсихология и психофизиология, в большей степени имеют отношение к установлению и использованию каузальных механизмов познания и поведения. (2) Вероятно, в большей степени точности и в больших возможностях ма-

нипуляции каузальными факторами познания и поведения кроется важнейшее различие между теориями механистических объяснений и вдохновленными машинным обучением предиктивными моделями. Последние в отрыве от поиска реальных каузальных механизмов могут ухватывать только поверхностные корреляции между событиями. (3) При принятии важных решений в контексте образовательной, биомедицинской и правовой практик лучше опираться на знание каузальных механизмов познания и поведения. Предельная осторожность требуется при отсылке к результатам предиктивных исследований во всех этих областях. (4) Соответственно, предиктивные модели не могут вытеснить теории, ищущие механистические объяснения познания и поведения. Вместо выбора между предсказательной и объяснительной методологиями более продуктивным было бы нахождение оптимального баланса между ними.

Список литературы

- Craver C. F. *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. – Oxford : Clarendon Press ; New York : Oxford University Press. – 2007. – 308 p.
- Iglesias D., Sorrel M. A., Olmos R. Cross-validation and predictive metrics in psychological research: Do not leave out the leave-one-out // *Behavior Research Methods*. – 2025. – Vol. 57. – No. 3. – P. 85.
- Falk E. B., Berkman E. T., Lieberman M. D. From neural responses to population behavior: neural focus group predicts population-level media effects // *Psychological science*. – 2012. – Vol. 23. – No. 5. – P. 439–445.
- Festinger L., Carlsmith J. M. Cognitive consequences of forced compliance // *The journal of abnormal and social psychology*. – 1959. – Vol. 58. – No. 2. – P. 203–210.
- Learning the natural history of human disease with generative transformers / Shmatko A. et al. // *Nature*. – 2025. – Vol. 647. – No. 8088. – P. 248–256.
- Rissman J., Greely H. T., Wagner A. D. Detecting individual memories through the neural decoding of memory states and past experience // *Proceedings of the National Academy of Sciences*. – 2010. – Vol. 107. – No. 21. – P. 9849–9854.
- Smith P. L., Little D. R. Small is beautiful: In defense of the small-N design // *Psychonomic bulletin & review*. – 2018. – Vol. 25. – No. 6. – P. 2083–2101.
- Weiskopf D. A. The Predictive Turn in Neuroscience // *Philosophy of Science*. – 2022. – Vol. 89. – No. 5. – P. 1213–1222.
- Yarkoni T., Westfall J. Choosing prediction over explanation in psychology: Lessons from machine learning // *Perspectives on Psychological Science*. – 2017. – Vol. 12. – No. 6. – P. 1100–1122.

References

- Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford : Clarendon Press ; New York : Oxford University Press. 308 p.
- Iglesias, D., Sorrel, M. A., & Olmos, R. (2025). Cross-validation and predictive metrics in psychological research: Do not leave out the leave-one-out. *Behavior Research Methods*, 57(3), 85.

- Falk, E. B., Berkman, E. T., & Lieberman, M. D. (2012). From neural responses to population behavior: neural focus group predicts population-level media effects. *Psychological science*, 23(5), 439–445.
- Festinger, L., Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The journal of abnormal and social psychology*, 58(2), 203–210.
- Rissman, J., Greely, H. T., & Wagner, A. D. (2010). Detecting individual memories through the neural decoding of memory states and past experience. *Proceedings of the National Academy of Sciences*, 107(21), 9849–9854.
- Shmatko, A., Jung, A. W., Gaurav, K., Brunak, S., Mortensen, L. H., Birney, E., Fitzgerald, T. & Gerstung, M. (2025). Learning the natural history of human disease with generative transformers. *Nature*, 647(8088), 248–256.
- Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design. *Psychonomic bulletin & review*, 25(6), 2083–2101.
- Weiskopf, D. A. (2022). The Predictive Turn in Neuroscience. *Philosophy of Science*, 89(5), 1213–1222.
- Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, 12(6), 1100–1122.

ЧТО ГОВОРИТ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Самохин Александр Александрович – д-р физ.-мат. наук, главный научный сотрудник, Институт общей физики им. А.М. Прохорова Российской академии наук, Москва, Россия; asam40@mail.ru; ORCID 0000-0002-5416-791X

Аннотация. В статье исследуются возможности имеющихся в открытом доступе вариантов искусственного интеллекта (ИИ) посредством анализа ответов, которые дает ИИ на ряд вопросов из области точных и гуманитарных наук. Задаваемые вопросы относятся к тем разделам науки, в которых по тем или иным причинам присутствуют определенные терминологические, методологические и другие неясности, в том числе и не всегда полностью осознаваемые на фундаментальной основе. Вопросы касались, в частности, представлений о спонтанном и вынужденном излучении, о соотношении смыслов слов «великоросс» и «русский» и другим близким понятиям. Анализ ответов ИИ «Нейро» показывает, что его работа слабо отличается от процесса подбора из доступной ему элементной базы некоторой совокупности утверждений, определений и т.п., выделяемых из этой базы на основе ключевых слов, содержащихся в исходном вопросе. ИИ оказывается не в состоянии обратить внимание на нарушение логики в рассматриваемых им конкретных примерах использования терминологии в физике или на обстоятельства и причины некоторых терминологических пертурбаций в области гуманитарных наук. Вопрос о полноте и логической непротиворечивости используемой им базы данных для формулировки окончательного ответа также остается нерешенной проблемой. По этим и другим подобным причинам рассмотренный вариант ИИ не может самостоятельно выполнять функции нормального научного исследователя. Вопрос о границах совершенствования ИИ в плане возможности его самостоятельного участия в научно-исследовательской работе как полноценного соавтора является частью большой фундаментальной проблемы, которая пока еще далека от своего решения.

Ключевые слова: искусственный интеллект; база данных; термины; логика; информационная война.

Для цитирования: Самохин А.А. Что говорит искусственный интеллект // Научно-ведческие исследования. – 2026. – № 3. – С. 113–129. – DOI: <http://www.doi.org/10.31249/scis/2026.03.09>

Статья поступила в редакцию: 24.11.2025

Принята к публикации: 20.02.2026

WHAT ARTIFICIAL INTELLIGENCE SAYS

Alexandr A. Samokhin – Doctor of Sciences (Physics and Mathematics), Chief Research Fellow, Prokhorov General Physics Institute of the Russian Academy of Sciences, Moscow, Russia; asam40@mail.ru; ORCID 0000-0002-5416-791X

Abstract. This paper explores the capabilities of publicly available artificial intelligence (AI) systems by analyzing their responses to a series of questions from the exact sciences and humanities. The questions relate to areas of science that, for various reasons, contain certain terminological, methodological, and other ambiguities, including those not always fully understood at a fundamental level. The questions touched, in particular, on concepts such as spontaneous and stimulated emission, the relationship between the meanings of the words “Great Russian” and “Russian”, and other related concepts. An analysis of the responses from the AI system named “Neuro” shows that its operation differs little from the process of selecting from the available elemental database a set of statements, definitions, etc., extracted from this database using the keywords contained in the original question. The AI is unable to detect logical violations in specific examples of terminology used in physics, or to identify the circumstances and causes of certain terminological perturbations in the humanities. The question of the completeness and consistency of the database it uses to formulate a final answer also remains unresolved. For these and other similar reasons, the AI under consideration cannot independently perform the functions of a normal scientific researcher. The question of the limits of AI improvement in terms of its ability to independently participate in scientific research as a full-fledged co-author is part of a larger fundamental problem that is still far from being resolved.

Keywords: artificial intelligence; artificial intelligence; database; terms; logic; information warfare.

For citation: Samokhin, A.A. (2026) What artificial intelligence says. // *Naukovedcheskie issledovaniya = Science Studies*. – 2026. – N 3. – P. 113–129. – DOI: <http://www.doi.org/10.31249/scis/2026.03.09>

The paper was submitted: 24.11.2025

Accepted for publication: 20.02.2026

Введение

В настоящее время об искусственном интеллекте (ИИ) пишут и говорят, наверное, почти столь же много слов, как и прежде когда-то о звездных войнах или о судьбоносной озоновой дыре. Эти слова сейчас звучат из уст первых лиц и лиц рангом ниже, когда они сообщают о своих успехах, достигнутых при исполнении руководящих указаний не только вследствие своих надлежащих качеств, но и благодаря использованию ИИ. В этом смысле роль ИИ для народного хозяйства не отличается от электрификации, химизации или цифровизации. Продолжением этого списка вполне могло бы быть слово «роботизация», но название ИИ или хотя бы нейросеть кажется более возвышенным и прогрессивным, чем робот, как это и должно быть у так называемого «нового слова» в науке и в жизни.

Одной из особенностей нынешней ситуации является тот факт, что обсуждаемый объект, т.е. ИИ, в отличие от вышеупомянутой озоновой дыры, может уже высказываться как бы самостоятельно, демонстрируя тем самым собственный интеллект неестественного (неорганического) происхождения. Даже не прибегая к строго логическим построениям, можно согласиться с тем, что интеллекты бывают разными, как это имеет место для различных людей, собак, кошек, дельфинов и прочих населяющих нашу планету существ. В связи с этим стоит обратить внимание, как рукотворный ИИ проявляет себя в ответах на вопросы, с которыми приходится сталкиваться в науке и в жизни нормальному, естественному интеллекту. В настоящей статье анализируются некоторые подобные ответы, получаемые с помощью ИИ сервисом «Нейро», на вопросы из области точных и других наук, характеризующие нынешнее состояние этого творения.

При таком подходе требуется достаточно подробное цитирование из Интернета, необходимое для формирования более полного представления об обсуждаемом материале. С учетом изменчивости этой информации указывается время ее извлечения.

В разделе два представлена информация, которую сервис «Нейро» дает о себе и об ИИ. В разделах три и четыре анализируются ответы ИИ на некоторые вопросы из области физики и гуманитарных наук. Основные выводы работы суммированы в заключительном разделе пять.

«Нейро» о себе (запрос от 16.05.2025 г.)

«Нейро» – сервис для поиска информации в интернете с помощью искусственного интеллекта, разработанный «Яндексом». Сервис запустили в апреле 2024 года¹.

При первом запросе «Нейро о себе» программа выделяет следующие свои возможности:

- *Ведение диалога с пользователем.* Сервис помнит предыдущий контекст беседы, не нужно каждый раз писать запрос, можно просто задавать уточняющие вопросы.
- *Приведение ссылок на источники.* «Нейро» оставляет ссылки на сайты, с которых брал информацию, и обозначает в тексте, что конкретно и откуда именно он взял.
- *Работа с изображениями.* Можно загрузить изображение и задать вопрос, что на нем изображено.

Отметим, как работает «Нейро»:

1. Пользователь пишет вопрос.
2. Нейросеть перефразирует вопрос в поисковый запрос.

¹ «Яндекс» запустил сервис поиска в интернете с помощью нейросети «Нейро» <https://trends.rbc.ru/trends/industry/6620b8fb9a794728a5557c07>

3. Поисковая система находит топ подходящих источников и выбирает из них пять.
4. На их основе нейросеть из семейства YandexGPT-3 генерирует ответ.
С октября 2024 года «Нейро» может находить ответы на любые вопросы во всех браузерах¹.

Однако ответы «Нейро», как и ответы других нейросетей, нужно перепроверять. Сервис может предоставлять некорректную и неактуальную информацию².

Повторение того же запроса («Нейро о себе») в тот же день – 16.05.2025 г. – привело к следующему ответу:

«Не удалось найти информацию по указанному запросу. Возможно, имелся в виду сервис «Нейро» от Яндекса, запущенный в апреле 2024 года».

«Нейро» – сервис для поиска информации в интернете с помощью искусственного интеллекта. В ответ на запрос пользователь получает не список ссылок на сайты, а ответ, сгенерированный моделью YandexGPT-3.

Некоторые возможности сервиса:

- *Ведение диалога с пользователем.* Сервис помнит предыдущий контекст беседы, не нужно каждый раз писать запрос, можно просто задавать уточняющие вопросы.
- *Приведение ссылок на источники.* «Нейро» оставляет ссылки на сайты, с которых брал информацию, и обозначает в тексте, что конкретно и откуда именно он взял.
- *Работа с изображениями.* Можно загрузить изображение и задать вопрос, что на нем изображено.

С октября 2024 года сервис доступен во всех браузерах».

Ввиду подобного «самопредставления» следует обратить внимание на один из его существенных недостатков. Здесь ничего не говорится о том, как поступает «Нейро» в тех случаях, когда сталкивается в обрабатываемом материале с явными проявлениями различных факторов информационной войны – нарушением логики аргументации, гипертрофированным вниманием к отдельным аспектам бытия или их полным замалчиванием. Важность адекватного учета таких факторов будет видна из дальнейшего изложения.

«Нейро» об ИИ (запрос от 16.05.2025 г.)

«Искусственный интеллект (ИИ) – это обширная область информатики, включающая множество различных подходов к созданию умных си-

¹ Что такое Яндекс Нейро и как с ним работать
<https://callibri.ru/blog/chto-takoe-yandex-neiro>

² «Яндекс» запустил сервис поиска в интернете с помощью нейросети «Нейро»
<https://trends.rbc.ru/trends/industry/6620b8fb9a794728a5557c07>

стем. Он использует не только нейронные сети, но и другие методы: экспертные системы, генетические алгоритмы, нечеткую логику и многие другие технологии.

Нейросети – это вычислительные модели, созданные по образу и подобию биологической нейронной сети человеческого мозга. Они представляют собой набор узлов (нейронов), объединенных слоями, и обрабатывают информацию, моделируя работу мозга.

Нейросети в ИИ отвечают за решение определенных задач: генерацию текста или изображений, поиск информации или анализ данных. Искусственный интеллект и нейросети применяются не только для создания текстов и решения математических задач. Они широко используются в медицине, научной деятельности, финансовом секторе, в маркетинге и рекламе».

С учетом всех этих предупреждений и прочих нюансов перейдем далее к непосредственному зондированию предлагаемых возможностей посредством ответов данного варианта ИИ на конкретные научные вопросы.

Вопросы из истории физической терминологии

Начнем испытание описанной выше системы «Нейро» с вопросов по физике. В этой точной теоретической и экспериментальной науке, как может показаться на первый взгляд, нет места для распространения недоразумений, связанных с терминологической нечеткостью, искажающей сущность обсуждаемых вопросов. Однако, к сожалению, такое предположение в действительности не подтверждается, на что уже в той или иной мере обращалось внимание, например, в работах [Гинзбург, 1983; Букин, 2006]. Академик В.Л. Гинзбург в статье «О природе спонтанного излучения», в частности, пишет: «Спонтанное излучение, наряду с поглощением и индуцированным (вынужденным) излучением, принадлежит к числу основных, наиболее элементарных процессов, и с ним приходится иметь дело, особенно в оптике, буквально на каждом шагу <...> Какова, однако, природа спонтанного излучения? Чем оно вызвано, является ли оно квантовым или классическим эффектом? <...> Для того, чтобы убедиться в существовании весьма распространенных недоразумений в вопросе о природе спонтанного излучения, достаточно, вероятно, будет «спросить соседа». <...> Именно, значительная часть опрошенных считает, что спонтанное излучение представляет собой результат существования нулевых колебаний электромагнитного поля, т.е. является квантовым эффектом. Между тем подобное мнение представляется неверным, поскольку спонтанное излучение света заведомо имеет место в классической теории и, вообще говоря, не может считаться квантовым явлением в большей мере, чем поглощение или индуцированное испускание. Недоразумения, с которыми приходится сталкиваться в вопросе о природе спонтанного излучения, по-

рождены ходом исторического развития квантовой теории в целом и квантовой теории излучения в частности...» [Гинзбург, 1983, с. 687].

А вот как отвечает ИИ на вопросы по этой теме.

Вопрос: Спонтанное излучение – классическое или квантовое явление? Ответ «Нейро» (22.04.2025 г.):

«Спонтанное излучение – *квантовое* явление. Оно не может быть объяснено классической электромагнитной теорией и связано с переходом квантовомеханических систем (молекул, атомов или субатомных частиц) из возбужденного энергетического состояния в более низкое и испусканием фотона».

На такой же вопрос 06.05.2025 г. был получен с такими же оговорками фактически аналогичный ответ:

«Спонтанное излучение *является квантовым явлением*, оно не может быть объяснено классической электромагнитной теорией».

Ответы дружные, но неверные, поскольку спонтанное излучение существует и в классической теории. В рамках классической электромагнитной теории такое излучение можно рассматривать как процесс «отрыва» от движущейся заряженной частицы волновой составляющей ее электромагнитного поля, который был исследован на рубеже XIX и XX веков с помощью «запаздывающих потенциалов».

Следующий вопрос: Вынужденное излучение – классическое или квантовое явление?

Ответы от 06.05.2025 г. (с отмеченными уже выше и повторяющимися далее оговорками) выглядят так:

«Вынужденное излучение (индуцированное излучение) – *квантовое явление*. Это генерация нового фотона при переходе квантовой системы (атома, молекулы, ядра и т.д.) между двумя состояниями (с более высокого на более низкий энергетический уровень) под воздействием индуцирующего фотона».

«Вынужденное излучение связано с *квантовыми свойствами излучения* и не может быть описано с точки зрения классической электродинамики. При вынужденном излучении атом переходит из возбужденного состояния в основное, но не самопроизвольно, а под воздействием внешнего электромагнитного поля. При этом испускаемый фотон неотличим от тех, которые вызвали его испускание».

«Вынужденное излучение может описываться как с точки зрения *классической электродинамики*, так и с точки зрения *квантовой теории*. С одной стороны, для описания испускания света классическим осциллятором используют классическую электродинамику. С другой стороны, вынужденное излучение связано с квантовыми свойствами излучения, так как при нем новый фотон неотличим по своим свойствам от фотона, вызвавшего переход. Таким образом, вынужденное излучение может рассматриваться в рамках как классических моделей, так и квантовой теории».

Ответы разные, но только последний из приведенных, в котором вопрос формулировался как «вынужденное излучение – классический или

квантовый эффект», является правильным, хотя и не совсем точным. Действительно, вынужденное излучение – это просто эффект усиления электромагнитного поля, взаимодействующего при определенных условиях с возбужденной системой. Такое усиление действующего поля, как и спонтанное излучение, существует в классической и квантовой области параметров процесса. Фактически в этих областях для одного и того же физического явления (возможность усиления электромагнитного поля при его взаимодействии с материальной системой) используются разные названия, и это элементарное обстоятельство остается не вполне осознанным до настоящего времени. В том числе, столь модным сейчас ИИ.

Именно этот аспект неявно отражается в прилагаемом ниже ответе «Нейро» на вопрос о вынужденном излучении (без акцента на квантовую или классическую природу этого явления). В этом ответе нет даже намека на общность понятий усиления и вынужденного излучения:

«Вынужденное излучение (индуцированное излучение) – генерация нового фотона при переходе квантовой системы (атома, молекулы, ядра и т.д.) между двумя состояниями (с более высокого на более низкий энергетический уровень) под воздействием индуцирующего фотона, энергия которого равна разности энергий этих состояний.

Созданный фотон имеет ту же энергию, импульс, поляризацию, а также направление распространения, что и индуцирующий фотон (который при этом не поглощается).

На вынужденном излучении основан принцип работы квантовых усилителей, лазеров и мазеров».

Заметим, что замена слов о «вынужденном излучении» на слова об «эффекте усиления» в заключительном предложении из данного ответа приводит к выявлению очевидной элементарности исходного утверждения и дополнительной специфичности рассматриваемой ситуации, например, в отношении слова «лазер».

На вопрос об истории создания лазера «Нейро» отвечает так (запрос от 16.05.2025 г.):

«История создания лазера началась в 1916 году, когда Альберт Эйнштейн предсказал существование явления вынужденного излучения – физической основы работы любого лазера¹.

В 1951 году профессор Колумбийского университета Чарльз Таунс решил воплотить теорию вынужденного излучения на практике и создать такой прибор. В 1954 году он представил первый в мире реально работающий лазер, правда, тогда он назывался «мазер» – от английского Microwave Amplification by Stimulated Emission of Radiation, что в переводе означает «усиление микроволн с помощью вынужденного излучения»².

¹ Лазер <https://znanierussia.ru/articles/лазер>

² История открытия: от мазера к лазеру <https://rostec.ru/media/news/istoriya-otkrytiya-ot-mazera-k-lazeru/>

В 1960 году американский физик Теодор Майман создал первый лазер. Это был импульсный рубиновый лазер, который состоял из кристалла рубина (сантиметром в диаметре и около двух в длину) с посеребренными торцами, а также лампы-вспышки.

Примерно через год первый лазер был запущен в СССР. Это произошло 2 июня 1961 года в ГОИ, старшим научным сотрудником Л.Д. Хазовым с участием И.М. Белоусовой».

Приведенная цитата содержит, как минимум, две существенных неточности. Эйнштейн не «предсказал существование явления вынужденного излучения», а применил в своем рассмотрении двухуровневой квантовой системы известный из классической физики эффект усиления поля при его взаимодействии с движущейся (возбужденной) частицей, обозначив это в своей работе с достаточной ясностью: «Для этих переходов установим простые гипотезы, причем будем руководствоваться предельным случаем классической теории, кратко охарактеризованном выше» [Эйнштейн, 1966, с. 386].

Однако приведенное там же сравнение процесса спонтанного излучения с радиоактивным распадом («Едва ли можно представить себе, что он аналогичен чему-либо другому, кроме радиоактивного распада») является не совсем точным из-за возможности его сравнения также и с обычным излучением при ускоренном движении заряженной частицы. Влияние такой исторически обусловленной неточности усугубляется добавлением к русскому переводу анонимного примечания: «В этой работе высказаны идеи, которые впоследствии привели к возникновению и развитию электроники. В ней впервые были введены коэффициенты Эйнштейна A и B » [Там же].

Эта вторая неточность, связанная со стилем изложения, способствует формированию у читателя, считающего вынужденное излучение квантовым эффектом, неверного представления о том, что любой лазер является квантовым прибором. На существование такого заблуждения обратил внимание, в частности, нобелевский лауреат по физике 1955 г. Уиллис Юджин Лэмб в статье 1972 года «Классический лазер» в солидном американском журнале *Physical Review*. В заключительной части этой статьи написано: «Рассмотрена полностью классическая модель лазера, в которой никакие упоминания о фотонах или индуцированном излучении не используются» [Borenstein, 1972, с. 1298]. На работу Лэмба 1972 года ссылается в 1977 году Уильям Колсон в статье о классической теории лазера на свободных электронах (ЛСЭ) [Colson, 1977, с. 192]. Факт существования ЛСЭ свидетельствует о явной некорректности утверждения о том, что «история создания лазера началась в 1916 году», когда в работе Эйнштейна эффект усиления поля рассматривался как «вынужденное излучение». Возникающее из-за неадекватного использования терминологии недоразумение выразительно подчеркивается в названии статьи «Парадокс ЛСЭ как

лазера без лазера» («The paradox of a free-electron laser without the laser»), напечатанной в июльском за 2021 год выпуске е-журнала Phys. Org.¹

Об этом обстоятельстве напоминает также в недавней статье «Когда можно обойтись без квантовой теории. Родословная лазера гораздо старше, чем зачастую принято считать»². Если процесс усиления и генерации когерентных электромагнитных колебаний с изобретением триода был освоен в самом начале XX века, то с генерацией когерентных механических колебаний это случилось примерно на два века раньше при изобретении хронометра.

На запрос «триод история» информация от «Нейро» (19.05.2025 г.) начинается следующим образом:

«Триод – электронная лампа с тремя электродами: термоэлектронным катодом, сеткой и анодом.

История изобретения: триод был изобретен и запатентован в 1906 году американским инженером Ли де Форестом. В качестве эксперимента он поместил между анодом и катодом диода сетку из проволоки.

Вначале триод использовали только как приемно-усилительную лампу. В 1913 году немецкий ученый А. Мейснер установил возможность применения триода в качестве генераторной лампы...».

Поскольку лазеры и мазеры являются генераторами когерентных электромагнитных волн, то их естественными предшественниками являются усилители и генераторы электромагнитных колебаний радио и СВЧ диапазона. Это очевидное обстоятельство осталось за пределами данного ответа ИИ, полученного с помощью «Нейро».

На два варианта прямого вопроса (Можно ли считать генераторы когерентных электромагнитных колебаний (радио и СВЧ диапазона) предшественниками лазеров?), соответственно, без внутренних скобок и со словами «радио и СВЧ диапазона» «Нейро» дает 19.05.2025 г. два таких ответа: *«Да, генераторы когерентных электромагнитных колебаний можно считать предшественниками лазеров, в частности мазеры (молекулярные генераторы)»*³.

Основные принципы работы лазера были сформулированы и реализованы первоначально в микроволновом диапазоне при создании мазеров. В основе работы мазеров и лазеров лежит один и тот же принцип, сформулированный в 1951 году В.А. Фабрикантом:

«Да, генераторы когерентных электромагнитных колебаний радио- и СВЧ-диапазона можно считать предшественниками лазеров, так как путь к созданию последних был найден радиофизиками, которые умели строить

¹ The paradox of a free-electron laser without the laser <https://phys.org/news/2021-07-paradox-free-electron-laser.html>

² Когда можно обойтись без квантовой теории https://www.ng.ru/science/2024-10-08/13_9110.html

³ Лазер <https://bigenc.ru/c/lazer-430c3c>

генераторы и усилители электромагнитных колебаний, использующие резонаторы и обратную связь»¹. Например, первая квантовая генерация электромагнитного излучения была получена не в оптической области спектра, а в радиодиапазоне, в области так называемого СВЧ-излучения. Эти молекулярные квантовые генераторы были созданы в 1954 году почти одновременно Н.Г. Басовым и А.М. Прохоровым в Физическом институте им. П.Н. Лебедева АН СССР и независимо от них Ч. Таунсом, Дж. Гордоном и Х. Цайгером в Колумбийском университете (США).

Эти два вроде бы совпадающих, в основном, ответа в то же время имеют существенные различия, свидетельствующие, в частности, о неустойчивости процесса их формирования в зависимости от незначительных вариаций исходного вопроса. Кроме того, их содержание не согласуется с приведенным выше утверждением «*История создания лазера началась в 1916 году*, когда Альберт Эйнштейн предсказал существование явления вынужденного излучения – физической основы работы любого лазера». Такое же утверждение «Нейро» выдает в ответ на заданный 19.05.2025 г. вопрос «откуда можно начинать историю лазеров»: «Историю лазеров можно начинать с 1917 года, когда Альберт Эйнштейн впервые ввел представление о вынужденном испускании»². Эта теория стала основой для создания первого лазера».

Возникновение противоречивых ответов является прямым следствием наличия противоречивых элементов в исходной информационной базе и нынешней неспособности ИИ, действующего на основе такой базы, явно подчеркнуть и прокомментировать это обстоятельство (кроме общего предупреждения о возможных неточностях). Существенным недостатком приведенных ответов является также отсутствие в них какого-либо намека на логические противоречия и ограничения использования понятий о спонтанном и вынужденном излучении, на которые уже обращалось внимание [Букин, 2006].

Выявление подобных недостатков работы ИИ в ближайшее время вряд ли возможно без прямого участия естественного интеллекта, непосредственно формирующего используемую искусственным интеллектом базу данных. При этом необходимо учитывать, что эта область деятельности одновременно является также полем сражений информационной войны со всеми ее специфическими особенностями, которые обусловлены самыми различными интересами и обстоятельствами. Подобный фактор способствует продолжительному существованию различных «недоразумений», которые своей массовостью затрудняют восприятие элементарных научных истин даже в области точных наук.

¹ История одного из самых важных изобретений XX века – лазера <https://www.techinsider.ru/technologies/5466-kvantovyy-svetoch-istoriya-odnogo-iz-samykh-vazhnykh-izobreteniy-xx-veka-lazera/>

² История изобретения и развития лазера <https://npf-specmash.ru/stati/laser-history>

Вопросы к ИИ из области гуманитарных наук

На примерах из этнографии покажем неточность ответов ИИ на чрезвычайно важные, с нашей точки зрения, вопросы и далеко идущие следствия данной проблемы. Прежде всего обратимся к основным понятиям раздела.

«Русские» – восточнославянский народ, самый многочисленный этнос в России и Европе¹. Согласно поправкам к Конституции России 2020 года (статья 68.1), «государственным языком Российской Федерации на всей ее территории является русский язык как язык государствообразующего народа, входящего в многонациональный союз равноправных народов Российской Федерации»². Только таким образом в Конституции определен государствообразующий народ России.

Некоторые характеристики:

- *Язык*: родной и национальный – русский.
- *Религия*: верующие русские – в основном православные.
- *Этнические группы*: среди русских выделяют, например, севернорусскую, южнорусскую, казаков, сибиряков и другие.
- *Проживают*: русские составляют большую часть населения Российской Федерации, а также значительную часть населения в странах бывшего СССР. Крупные общины есть в США, Канаде, Бразилии, Германии и других странах».

Очевидная неточность этого ответа связана с его неполнотой из-за отсутствия, в частности, какой-либо информации о слове «великоросс», которое имеет непосредственное отношение к слову «русские». На следующий вопрос (Великороссы кто это такие?) «Нейро» дает такой ответ (запрос от 12.05.2025 г.):

«“Великороссы” (также встречаются варианты написания: великорусы, великоруссы, великороссияне) – *этноним, который использовался в отношении русских* в качестве их самоопределения либо внешнего представления на протяжении XVIII – начала XX века. Согласно официальной государствообразующей концепции Российской империи, великороссы представляли собой одну из трех ветвей общерусского народа (наряду с малороссами и белорусами). В XX веке, с распадом Российской империи и образованием СССР, этноним «великоросс» утратил правомерность, и в широком употреблении остался только этноним “русские”».

Последнее предложение в этом ответе озадачивает утверждением «утратил правомерность», которое не подкрепляется ссылками на соответствующие правовые или иные регулирующие акты. Этим предложением как бы обосновывается естественность широкого употребления слова «русские» вместо термина «великороссы». Такая подмена носит

¹ Русские <https://ru.wikipedia.org/wiki/Русские>

² Конституция Российской Федерации <http://kremlin.ru/acts/constitution/item>

характер специфической диверсии, опасность которой осознана еще далеко не полностью.

На негативные последствия такой подмены уже обращалось внимание, в частности, в моей статье «Неизвестная информационная война» [Самохин, 2022], где подчеркивалась целесообразность и важность различения смыслов слов «великороссы» и «русские» как конкретно этнического и суперэтнического, наднационального понятий. Целесообразность и необходимость такого различения можно подтвердить множеством примеров, из которых здесь ограничимся тремя.

«Я не был никогда аскетом

И не мечтал сгореть в огне.

Я просто русским был поэтом

В года, доставшиеся мне», – это начало стихотворения Наума Коржавина (Манделя).

«В Дагестане – я аварец, в России – дагестанец, а за границей – я русский» (Расул Гамзатов).

«Наш долг – отстаивать честь бойцов и командиров Красной армии, великий подвиг представителей разных национальностей, которые навсегда останутся в мировой истории Русскими солдатами» (из выступления В. Путина на военном параде 9 мая 2025 г.).

Очевидно, что здесь слово «русский» должно восприниматься не как «великоросс», а в более широком, общем смысле, например, как в понятии «Русский мир». В этом случае, как будет показано ниже при анализе получаемых от ИИ ответов, употребление слова «русский» вместо «великоросс» может приводить к далеко идущим последствиям.

В ответ на запрос «русский мир это что?» «Нейро» сообщает (запрос от 20.05.2025):

«Вот что удалось найти среди результатов поиска:

- ru.wikipedia.org: «Русский мир» – это культурно-цивилизационная, геополитическая и религиозная концепция, подразумевающая объединение русскоязычного населения по всему миру. Концепция подразумевает объединение России, Белоруссии и Украины, а также усиление влияния в странах с большой долей русскоязычного населения, например, Прибалтике¹.
- vesti.ru: «Русский мир» – это Древняя Русь, Московское царство, Российская империя, Советский Союз и современная Россия. Он объединяет всех, кто чувствует духовную связь с Родиной, считает себя носителем русского языка, истории и культуры, независимо от национальной или религиозной принадлежности².

¹ Русский мир https://ru.wikipedia.org/wiki/Русский_мир

² От Древней Руси до СССР: что такое Русский мир <https://www.vesti.ru/article/3677636>

- ru.ruwiki.ru: «Русский мир» – это культурно-цивилизационная, геополитическая и религиозная концепция, подразумевающая единство русского / русскоязычного населения, проживающего по всему миру. В 2008 году термин «Русский мир» впервые был включен во внешнеполитическую доктрину Российской Федерации.
- [otvet.mail.ru](mailto:otvet@mail.ru): Понятие «русский мир» имеет разные значения в разных контекстах. В общем смысле, «русский мир» может относиться к культурному, историческому и географическому пространству, где проживают русские люди или те, кто считает себя частью русской культуры. В политическом контексте понятие «русский мир» может использоваться для обозначения стран, где русский язык и русская культура имеют значительное влияние, а также для обозначения стран, которые поддерживают близкие отношения с Россией.
- cyberleninka.ru: В основе концепции «Русского мира» лежит православная вера, русская культура и язык, а также общая историческая память и взгляды на общее развитие. Ядром «Русского мира» являются Россия, Украина и Белоруссия.

В такой форме ответа трудно усмотреть особые проявления какого-либо интеллекта, кроме способности подбора материала по заданным ключевым словам из имеющейся базы. Отдавая должное необходимости и полезности такой работы, ее все-таки следует отличать от иных форм интеллектуальной деятельности, например, от проведения сравнительного анализа материалов подобной подборки с той или иной точки зрения. В рассмотренном случае с «Русским миром» такой интеллектуальной деятельности со стороны «Нейро» не было продемонстрировано.

Однако примерно через час в тот же день 20.05.2025 информация от «Нейро» по этому же запросу выглядела несколько более обработанной:

- *«Русский мир» – термин, который используется для обозначения культурно-цивилизационной, геополитической и религиозной концепции, подразумевающей объединение русскоязычного населения по всему миру. Понятие «Русский мир» не имеет однозначного толкования и юридического определения.*

Некоторые трактовки термина:

- *Культурное, историческое и географическое пространство, где проживают русские люди или те, кто считает себя частью русской культуры.*
- *Страны, где русский язык и русская культура имеют значительное влияние, а также страны, которые поддерживают близкие отношения с Россией.*
- *Идеологический концепт, который направлен на объединение русскоязычных народов вокруг общих ценностей и культурных традиций.*

В 2023 году Владимир Путин заявлял, что «Русский мир» – это Древняя Русь, Московское царство, Российская империя, Советский Союз и современная Россия¹.

Основами концепции «Русского мира» могут считаться русский язык, русская культура, православие, а также вытекающие из этого традиционные русские ценности» [Назарова, 2015].

Этот вариант, за исключением некоторых пояснений и примерно полуторного сокращения, по сути, не отличается от исходного, то есть влияние какого-либо самостоятельного интеллекта по сравнению с имеющейся базой здесь остается минимальным. Не затрагиваются, в частности, вопросы о том, как и при каких условиях возможно формирование объединенной концепции, или какими причинами определяется отношение к этому термину.

Читатель, осознающий специфическую размытость и недосказанность таких ответов, может самостоятельно обратиться к многочисленным интернетным первоисточникам и убедиться в наличии алогичности, политической ангажированности и прочих особенностей при обсуждении терминологических аспектов в подобных вопросах.

В отличие от ИИ, естественный интеллект не может не обратить внимание на смысловые манипуляции с такими терминами как «великоросс» и «русский», «вынужденное излучение» и «усиление». Как уже было отмечено выше, научные обсуждения здесь фактически попадают в «туман» информационной войны.

К сожалению, наши исторические уроки, в том числе, в области ментального и информационного противодействия, усваиваются далеко не сразу и не полностью. Преодолению подобных барьеров может способствовать интенсивное развитие всех видов интеллектуальной деятельности, при которой ни один осмысленный и значимый вопрос не должен оставаться без адекватного ответа.

Заключение

Проведенный в настоящей работе анализ полученных от ИИ ответов на ряд конкретных вопросов из области физики и гуманитарных наук дает основание для следующих выводов.

1. Рассмотренный вариант ИИ проявляет себя в полученных ответах в основном как подборщик из определенной базы данных конкретного материала, характеризующегося некоторым набором ключевых слов. Особого интеллекта для такой работы не требуется, но она, несомненно, полезна

¹ От Древней Руси до СССР: что такое Русский мир
<https://www.vesti.ru/article/3677636>

при достаточной оперативности исполнения и фундаментальной полноты используемой базы данных.

2. Недостатком такой работы ИИ является отсутствие собственных комментариев в случае разночтений или противоречивости отобранного материала. Такие «детали» или вообще не замечаются, или без какого-либо дополнительного анализа отмечаются словами типа «не имеют однозначного толкования» с перечислением этих «толкований». В случае конкретного нормального интеллекта это могло бы служить признаком его ущербной ограниченности, отсутствием необходимой любознательности и т.п.

3. Аналогичным признаком ущербности обсуждаемого ИИ является его неспособность отметить тот факт, что два различных термина, фигурирующие в его ответах, обозначают, по существу, одно и то же физическое явление – усиление излучения за счет взаимодействия с веществом, находящимся в неравновесном (возбужденном, ускоренном) состоянии.

4. Очевидна зависимость ответов ИИ от используемой базы данных, которая определяется и самой базой во всей ее полноте, и способом отбора материала для конкретного ответа. При отсутствии достаточно четких и обоснованных критериев такого отбора потребитель ответа ИИ вполне может оказаться жертвой пиара, специфической пропаганды и прочих проявлений вездесущей информационной войны.

5. Связанная с ИИ пиар-кампания не должна отвлекать от решения актуальных проблем нашего бытия, обусловленных, в частности, неадекватным, с точки зрения логики, использованием обычного, естественного интеллекта для достижения специфических целей. Проблему нарушения логики в общественно значимых областях обмена и получения информации следует поднять на такой же информационный уровень, каким обеспечена проблема «нарушения прав человека» [Самохин, 2023].

Приведенные замечания не представляют собой, разумеется, исчерпывающий список. У разных ИИ подобные списки могут различаться, меняясь при этом со временем. Предложенный здесь метод анализа ответов ИИ на специальные проблемные вопросы дает возможность не только оценить возможности различных вариантов ИИ, но и выявить особенности используемых баз данных в отношении затрагиваемых вопросов.

Сейчас ИИ еще далек от способности самостоятельно выявлять подобные специфические особенности используемых баз данных, связанные с нарушениями логики, противоречиями в терминологии и другими недостатками в их материалах. Вопрос о достижимом уровне приближения ИИ к этому состоянию фактически эквивалентен вопросу о возможности ИИ самостоятельно вести исследовательскую научную работу, начиная с инициативной постановки исходной конкретной задачи. Такой вопрос в настоящее время остается открытым. Академик РАН Хохлов А.Р. в своем телеграмм-канале¹ 25 июля 2025 г. подчеркнул: «Для любого научного

¹ Алексей Хохлов – Telegram <https://t.me/s/khokhlovAR>

исследования важно то, что мы называем критическим мышлением. А ИИ-инструменты на это неспособны (по крайней мере, пока)». С учетом динамичного развития событий в этой сфере у нормального интеллекта практически не остается сомнений в необходимости постоянно следить за тем, что происходит непосредственно с ИИ и вокруг него.

В дополнение к отмеченным в тексте особенностям ответов от ИИ «Нейро» на предложенные вопросы еще раз подчеркнем важность внимательного к ним отношения. Рассмотренные здесь примеры представляют не только научно-исторический интерес, но и вполне определенную, актуальную общественно-политическую значимость, в частности, как элементы непрерывной, широко ведущейся информационной войны, которая продолжает оставаться не полностью осознаваемым, а порой и как бы совсем не замечаемым процессом. Например, на вопрос «Вынужденное излучение – классический или квантовый эффект?» П.А. Пивоваров получил от сервиса *liner.com* двухстраничный ответ, отмечающий классические и квантовые проявления этого эффекта. Однако при этом почему-то не замечается, что вынужденное излучение – это просто эффект усиления электромагнитного поля, взаимодействующего с определенным состоянием среды («отрицательное поглощение»). Обсуждение подобных проблем, определивших выбор задаваемых вопросов, заслуживает отдельного рассмотрения, продолжающего тематику процитированных работ 2022 и 2023 годов.

Благодарность. Автор выражает благодарность старшему научному сотруднику ИОФ РАН Пивоварову Павлу Александровичу за результативное внимание к данной работе.

Список литературы

- Букин В.В., Гарнов С.В., Самохин А.А. О спонтанном и вынужденном излучении // Краткие сообщения по физике Физического института им. П.Н. Лебедева Российской Академии Наук. – 2006. – № 3. – С. 3–12.
- Гинзбург В.Л. О природе спонтанного излучения // УФН. – 1983. – Т. 140, вып. 4. – С. 687–698. DOI: 10.3367/UFGN.0140.198308e.0687
- Назарова Г.Ф., Фокина А.В. Русский мир: обновление подходов к концепции // Ученые записки Орловского государственного университета. – 2015. – №6, (69). – С. 338–343.
- Самохин А.А. Неизвестная информационная война // Россия: тенденции и перспективы развития. Ежегодник / РАН. ИНИОН. – 2022. – Вып. 17, ч. 1. – С. 260–265.
- Самохин А.А. Информационное гражданское взаимодействие // Россия: тенденции и перспективы развития. Ежегодник / РАН. ИНИОН. – 2023. – Вып. 18, ч. 1. – С. 540–245.
- Эйнштейн А. Работы по кинетической теории, теории излучения и основам квантовой механики. 1901–1955 // Собрание научных трудов : в 4-х т. – Москва : Наука, 1966. – Т. 3. – С. 386–392.
- Borenstein M., Lamb W.E.Jr. Classical Laser // Physical Review A. – 1972. – Vol. 5, N 3. – 1298. DOI: 10.1103/PhysRevA.5.1298

Colson W.B. One-body electron dynamics in a free electron laser // *Physics Letters A*. – 1977. – Vol. 64, N 2. – P. 190–192. DOI: 10.1016/0375-9601(77)90712-5
<https://phys.org/news/2021-07-paradox-free-electron-laser.html>

References

- Borenstein, M., Lamb, W.E.Jr. (1972). Classical Laser. *Physical Review A*, 5(3), 1298.
DOI: 10.1103/PhysRevA.5.1298
- Bukin, V.V., Garnov, S.V., Samokhin, A.A. (2006). On spontaneous and stimulated emission. *Bulletin of the Lebedev Physics Institute*. 3, 3–12. (In Russ.)
- Colson, W.B. (1977). One-body electron dynamics in a free electron laser. *Physics Letters A*, 64(2), 190–192. DOI: 10.1016/0375-9601(77)90712-5. <https://phys.org/news/2021-07-paradox-free-electron-laser.html>
- Einstein, A. (1966). Works on kinetic theory, radiation theory and the foundations of quantum mechanics. 1901–1955. Collected Scientific Papers: in 4 volumes. Moscow: Nauka, 3, 386–392 (In Russ.)
- Ginzburg, V.L. (1983). The nature of spontaneous radiation. *Sov. Phys. Usp.* 26(8), 713–724.
DOI: 10.1070/PU1983v026n08ABEH004484
- Nazarova G.F., Fokina A.V. (2015) Russian world^ the renewal of approaches to the conception. *Scientific notes of Orel State University*. 6(69), 338-343.
- Samokhin, A.A. (2022). Unknown information war. *Russia: trends and development prospects. Yearbook*, 17(1), 260–265. (In Russ.)
- Samokhin, A.A. (2023). Information civil interaction. *Russia: trends and development prospects. Yearbook*, 18(1), 540–245. (In Russ.)

НАУКОВЕДЧЕСКИЕ ИССЛЕДОВАНИЯ

научный журнал
2026. № 3

Дизайнер (художник) О.Е. Максимов
Компьютерная верстка Д.А. Ушакова
Корректоры Д.Д. Корнейчук, Д.А. Ушакова

Подписано в свет 13 / VIII– 2026 г.
Формат 70×100/16 Бум. офсетная № 1
Усл. печ. л. 7,5 Уч.-изд. л. 7,25

**Институт научной информации
по общественным наукам
Российской академии наук**
Нахимовский проспект, д. 51/21, Москва, 117418
<http://inion.ru>

Отдел печати и распространения изданий
Тел.: +7 (499) 124-32-15
e-mail: izdat@inion.ru

Отпечатано в типографии
АО «Т8 Издательские Технологии»,
109316, г. Москва, Волгоградский проспект,
д. 42, корп. 5, к. 6