

Наука и инновации в информационном обществе

**КРЕАТИВНОСТЬ НЕЙРОСЕТЕЙ В СОВРЕМЕННЫХ НАУЧНЫХ
ИССЛЕДОВАНИЯХ****Ивахненко Евгений Николаевич***Доктор философских наук, профессор**Московский государственный университет имени М.В.Ломоносова**Философский факультет**Кафедра философии гуманитарных факультетов**Москва, Российская Федерация**ivahnen@rambler.ru***Аннотация**

Тема, связывающая достижения науки с прорывами в области ИИ, за последнее десятилетие переместилась в фокус читательского интереса. Ее обсуждению посвящаются тематические выпуски высокорейтинговых отечественных и зарубежных журналов. Авторы Информационного общества, следуя избранной редакционной политике, вносят свой вклад в осмысление вопросов по данной проблематике. Основная идея статьи заключается в стремлении автора ответить на несколько вопросов, связанных с распространением влияния гигантского потенциала генеративного ИИ на весь ход современных научных исследований. Авторская аргументация креативности нейросетей в научных исследованиях подкрепляется конкретными примерами достижений в этой области. Наряду с этим в статье показано то, что достижения философии науки XX в. вовсе не выносятся за границы тематики ИИ, а наоборот, способствуют ее глубокому осмыслению в эволюционном ключе. В заключительной части статьи ставится вопрос о наполнении новыми смыслами понятия «научное понимание» в условиях роста креативного потенциала нейросетей.

Ключевые слова

наука; нейросеть; генеративный искусственный интеллект; органоподобная система; научное объяснение; научное понимание

Введение

В 2025 г. в двух крупнейших компаниях по разработке ИИ, OpenAI и Anthropic, созданы научные отделы, деятельность которых сосредоточена на создании нейросетей (НС) как базового условия развития современной науки [21]. Ежемесячно обновляются обзоры, в которых раздел, освещающий достижения и открытия с использованием ИИ в области математических и физических наук, регулярно пополняется [19, 22]. Так, например, физик-теоретик Эндрю Строминджер (Andrew Strominger), профессор Гарвардского университета, в течение года работал над одной из проблем энтропии чёрных дыр. GPT-5.2, решила эту задачу за 12 часов, предоставив формальные доказательства [13]. 19.02.2026 г. в «The Harvard Gazette» вышло пояснение под примечательным заголовком «Физики протестировали усовершенствованную версию ChatGPT и вернулись с важным открытием» [11]. В итоге, по признанию ученых, созданное НС открытие оказалось не только верным, но и было признано значительным достижением в теоретической физике.

Другой пример — решение GPT-5.5 Pro задачи Эрдёша, которое не поддавалось математикам всего мира в течение 60 лет. НС в апреле 2026 г. справилась за 80 минут, поскольку включила в поисковые решения новые методы, на которые математики все это время не обращали внимания [12]. Это достижение подвинуло теоретика чисел Тимоти Гаверса (Timothy Gowers) признать, что современные модели уже способны не только помогать исследователям, но и предлагать оригинальные идеи [5]. Здесь следует обратить внимание на то, что 23-летний Лиам Прайс (Liam Price), работавший с НС над задачей Эрдёша, не имел математического образования. Он выполнял

© Ивахненко Е. Н., 2026

Производство и хостинг журнала «Информационное общество» осуществляется Институтом развития информационного общества.

Данная статья распространяется на условиях международной лицензии Creative Commons «С указанием авторства - С сохранением условий версии 4.0 Международная» (Creative Commons Attribution – ShareAlike 4.0 International; CC BY-SA 4.0). См. <https://creativecommons.org/licenses/by-sa/4.0/legalcode.ru>

https://doi.org/10.52605/16059921_2026_04_98

роль инженера, то есть обладал операциональными навыками обслуживания НС. Причем воспользовался Прайс не какими-то специально созданными для данного случая нейросетевыми ресурсами, а ресурсами, доступными, по сути, каждому продвинутому пользователю. OpenAI признало, что инженеры не обучали систему отдельно и не создавали для неё особый поисковый инструмент. В итоге лауреат Филдсовской премии Теренс Тао (Táo Zhéxūān) назвал полученный НС результат важной вехой для математики ИИ [16]. Эта проблема нашла свое отражение в публикациях отечественных авторов [1, 2, 3].

1 Нейросеть — коллега по лаборатории?

Острая конкуренция государств в области научных, технологических и военных достижений, превращает ИИ в ключевой фактор завоевания мирового лидерства. И не только в науке. Данное обстоятельство не в последнюю очередь повлияло на то, что мы стали свидетелями экспоненциального роста научных открытий, достигнутых в соработничестве ученых с НС. В области же научных исследований этот фактор уже состоялся в полной мере, а НС превратилась в эффективного научного сотрудника — коллегу по лаборатории.

Оценка сложившейся ситуации приближается к тому, что НС для исследователя становится столь же незаменимым инструментом, как микроскоп для нейробиолога или телескоп для астронома. Если так, то что остается ученому, которому в прежние времена требовались годы и годы упорного труда, чтобы только выйти на передовые рубежи науки? Сохраняется ли место упорному теоретику, продолжающему игнорировать генеративный ИИ, на фронтире научных поисков и открытий? Похоже, что для точных формальных наук (логика высших порядков, математика, программирование) такое положение вещей маловероятно. На этом фоне как будто вырисовывается перспектива того, что задачи, которые будет решать отелесненный научный работник, будут походить на задачи инженера-оператора: определять проблему, задавать контекст для НС, критически оценивать полученные результаты и т. п. Если за основу взять производительность труда, которую уже сейчас демонстрируют НС, то, казалось бы, так тому и быть. Обрисованная ситуация на самом деле вовсе не выглядит как текст в жанре фантастики.

Однако, если вникнуть в сложившуюся ситуацию — ученый, решая научную проблему, привлекает ресурсы НС, — то однозначность такого вывода будет поставлена под сомнение. Научный сотрудник, который в своей работе задействует НС, вовсе не предлагает ей написать сочинение на вольную тему. В целом он переносит ИИ в им же очерченную и сконструированную проблемную область. И вот здесь вовсе не лишне будет обратиться к тому, что Имре Лакатос назвал отрицательной и положительной эвристикой.

Вступая в коммуникацию с НС, ученый по меньшей мере задает ей *направления и цели* поиска. Отрицательная эвристика, по Лакатосу, привлекает знания и интеллектуальные интуиции ученого к тому, каких тупиковых путей надо избегать. В противном случае, интеллектуальные и финансовые ресурсы (последнее чрезвычайно важно в современной науке) будут потрачены попусту. Положительная эвристика нацелена на решение конкретных проблем исследовательской программы. В естественных науках наиболее значимые теоретические проблемы решаются посредством создания математических моделей. Таким образом отрицательная эвристика остается главным образом в компетенции отелесненного исследователя, положительная — его нейросетевого коллеги. Здесь уместно вспомнить весьма примечательное высказывание Ричарда Фейнмана, одного из творцов квантовой электродинамики. «Единственное действительно физическое описание явления, — утверждал он, — это истолкования смысла величин в уравнении с точки зрения эксперимента... А раз так, то, наверное, наилучший способ создания новой теории — *угадывать* уравнения... (курсив мой, ЕИ)» [7, с. 229]. К этому «угадывать» мы еще вернемся.

Теперь же бросим взгляд на проблему с иной точки зрения. Прежде всего, подвижки в области ИИ заставляют критически пересмотреть устоявшиеся в прошлые столетия представления о научной деятельности ученого-исследователя. Чтобы это осмыслить, следует лучше разобраться в том, какие смыслы по инерции некритически мы вкладываем в такие слова-понятия, как «ИИ» и «НС». Чаще всего эти смыслы сводятся к двум крайностям: НС — инструмент, набор алгоритмов, или же НС — сверхчеловек. Такая крайность вполне объяснима. Ведь мы имеем дело с системой, которая демонстрирует поразительную продуктивность. Поэтому в академическом и публичном дискурсах ответы на вопрос «что есть НС?» колеблются между антропоморфизацией НС (сверхчеловек) и сведением ее к «стохастическому попугаю» (набор алгоритмов). Но все дело в том,

что НС не инструмент и не сверхчеловек. Если коротко, то она демонстрирует способность к рекурсивному процессированию текстов, из которых мы извлекаем новую информацию и новые смыслы. Однако, при этом она принципиально лишена феноменального опыта, то есть не обладает *человеческим* сознанием (умом) [4]. И хотя сказанное и написанное по этому поводу кое-что проясняет, все же мы еще не располагаем обстоятельно выверенной новой парадигмой, определяющей суть (механизм получения, природу, онтологию) *нового* научного знания, полученного с использованием НС, а по сути — встроенного в НС.

2 Нейросеть как органоподобная система, способная к предвосхищению

В связи со сказанным могут представлять интерес соображения, высказанные сооснователем компании Anthropic Крисом Ола (Chris Olah). Канадский специалист в области машинного обучения ассоциирует НС с *органоподобной системой*, порожденной инопланетной эволюцией. «Представьте, — поясняет он, — если бы какой-нибудь инопланетный организм приземлился на Земле и смог делать всё это. Все бы бросились выяснять, как этому инопланетному организму удаётся всё это осуществить» [20]. Такая и подобная ей апелляция к органологии нацеливает на то, чтобы избегать рассуждений о *равнозначности* машины и организма. Помимо того, органология подводит к идее об *эволюционном скачке* от механизма к организму. Это, если согласиться с позицией Юка Хуэя, оказалось возможным, когда удалось «освободить машину от предначертанных правил и функций» [9, с. 215]. В таких условиях органоподобная система, по выражению Бернара Стиглера (Bernard Stiegler), обретает способность не только восприятия и удержания информации, но и способность *предвосхищения*. Используя терминологию из гуссерлевской теории внутреннего сознания времени, Стиглер назвал это качество «третичной ретенцией» — способностью к искусственным воспоминаниям [18, р. 46–47]. О таком переходе от механо-подобия к органоподобию следует вспомнить, когда мы обнаруживаем то, что НС могут галлюцинировать. По сути — это состоявшийся факт: современные НС являются носителями одновременно и памяти, и предвосхищения. Сама природа организации памяти тесно увязана с искусственными воспоминаниями, а следовательно — со способностью к предвосхищению.

3 «Наука» и «научное исследование». Что изменилось?

Перед лицом таких, прежде не мыслимых, достижений по сближению науки с ИИ возникает вопрос: изменяется ли значение самого понятия «наука» на нечто отличное от того, что подразумевалось под этим словом в XVII, XVIII и XIX вв.? «Мы сталкиваемся, — писал Ян Хакинг (Ian Hacking) в 70-е гг. XX в., — с проблемой того, меняет ли теоретический прогресс само значение [наука], выдвигая новые критерии, или просто позволяет нам больше узнать...» [8, с. 146]. К этому мы можем добавить весьма существенное: изменилось ли что-то в этом отношении за прошедшие полвека в связи с победоносным шествием НС по всему миру? Рой Харрис (Roy Harris) в своей книге «Семантика науки» (2005) [14] показал то, *как* значение самого отношения «наука–ученый» («science–scientist») исторически переформатируется и обретает иные интерпретации в языке. Уместно предположить, что набирающее силу и масштаб задействования НС в исследовательских практиках в перспективе внесет коррективы и в наше устоявшееся представление о том, какие значения мы вкладываем в нашу речь, когда включаем в нее такие слова, как «наука» и «ученый». Похоже, ждать недолго. Уже в наши дни ставится вопрос о «новой научно-машинной» коллаборации (New scientist-machine collaborations). Причина та же: существенно изменились представления о роли ученого и научного метода в современных достижениях и открытиях [10].

Итак, какое содержание мы в изменившихся условиях должны вкладывать в понятия «наука» и «научное исследование»? По сути, это уже не только техническая, но и глубокая философская и концептуальная задача. Задача такого рода и такого масштаба породила целое направление современных гуманитарных исследований под общим названием «Философия автономной науки» (Philosophy of Autonomous Science — PAS). Основная деятельность набирающего силу движения заключается в обеспечении безопасного и успешного внедрения ИИ в научные исследования. В рамках PAS разработана программа с «Десятью вопросами для грядущей эры ученых, занимающихся искусственным интеллектом». К числу «неантропоцентрических задач» программы PAS отнесены: «понимание, любопытство, удивление, интерес, креативность и новизна так называемых „искусственных ученых“». Инициаторы движения PAS разделяют уверенность в том,

что идеи из философии науки могут напрямую влиять на разработку систем ИИ, призванных делать научные открытия [15].

О каких идеях философии науки может идти речь? К примеру, об «объективном знании» Карла Поппера. Поппер, как известно, выдвинул идею *существования* трех миров. Это: физический мир, мир нашего ментального опыта и мир «объективного знания». Третий мир, по Попперу, включает в себя содержание научных книг, библиотек, компьютерной памяти, а сегодня мы к нему можем добавить научную составляющую из всего массива Big Data. Поппер, как известно, искренне полагал, что третий мир, хоть и является продуктом человека, однако, в своем существовании он *автономен*. При этом с ним можно полностью согласиться в том, что большинство корпоративных продуктов (достижений) науки не могли быть созданы без участия автономного третьего мира. Эту мысль он проводит в своей получившей широкое признание книге «Объективное знание. Эволюционный подход» (1972). Если бросить свежий взгляд на то, что ее автор утверждает в третьей главе («Эпистемология без субъекта знания»), то мы можем уловить очевидное сходство с исследовательской программой PAS. Прежде всего — в признании *независимого существования третьего мира* [6, с. 110]. Здесь также следует обратить внимание на вторую часть названия книги — «Эволюционный подход». По сути НС, представленная Крисом Олой *органоподобной системой*, можно считать продолжением все того же эволюционного подхода Поппера, но уже с участием ИИ расширенного до космических масштабов.

Однако, все же произошли существенные сдвиги в наших представлениях о *научном исследовании и научном понимании*, когда позицию «коллеги по лаборатории» заняла автономная НС. Пока мы только подошли к постановке проблемных вопросов о существовании этих парадигмальных по существу изменений. Таковые, на мой взгляд, следует представить не в виде строго определенного перечня насущных задач, но — в виде проблемного поля, в которое помещен современный исследователь. Это поле все же должно быть четко очерчено.

4 Что происходит с «научным пониманием» на фоне роста креативности нейросетей?

Научные исследования и генеративный ИИ — это тот тип союза человеческих и нечеловеческих возможностей, с которым мы связываем не только наши надежды, но и наши тревоги. Достижения науки во все времена применялись с разными целями. Со времен создания атомной бомбы (1945) стало понятно, что ее массовое применение «по назначению» неизбежно приведет к катастрофическим последствиям. Принимая в расчет НС с их потенциальными возможностями, законно поставить вопрос: возросла ли угроза глобального масштаба? Консолидированный ответ со стороны интеллектуалов, погруженных в данную проблему, совсем не утешительный — однозначно ДА. Перечень рисков существенно пополнился. К примеру, обнаружение легко воспроизводимого биологического оружия или алгоритма, способного взламывать современные схемы шифрования и т. д.

По мере того, как НС становятся все более совершенными, можем ли мы по-прежнему оперировать понятиями «научное понимание» или «научное объяснение» в их устоявшемся человеческом измерении? Эти понятия прежде несли в себе скорее психологический и прагматический, чем логический или математический характер. Не все ученые разделяли эту точку зрения. Физики, в том числе лорд Кельвин, Эрвин Шрёдингер и Ричард Фейнман, разработали собственные модели научного понимания. Примечательно, что в 2005 году философы науки Хенк В. де Регт (Henk W. de Regt) и Деннис Дикс (Dennis Dieks) разработали новую теорию о том, как изменяется смысл понятия «научное» в словосочетании «научное понимание». Их точка зрения заключается в том, что явление можно понять из характерных следствий теории без опоры на точные вычисления. Позже де Регт опубликовал книгу «Понимание научного понимания» (2017) [17], выводы из которой научное сообщество до сих пор оценивает неоднозначно. Теория де Регта-Дикса является контекстуальной и прагматичной. Она относится к наблюдаемым *свойствам*, вскрываемым ИИ, а не к психологическим состояниям ума. Тем самым, отмечает в своей книге де Регт, «научное понимание» можно распространить на объекты, выходящим за рамки человека, включая машины. На этом фоне в последние годы разгораются дискуссии вокруг интерпретаций теории научного понимания де Регта-Дикса. До сих пор спорным остается вопрос: можем ли мы найти одно (или даже несколько) неантропоцентрических и вычислимых приближений научного понимания? До сих пор убедительных аргументов так и не было предъявлено, не сложилось на этот счет и конвенциональное согласие ученых.

Но ведь это только начало. Страшно подумать — во всех смыслах, — что еще нас ждет впереди.

Благодарности

Исследование выполнено за счет гранта Российского научного фонда № 26-18-00851, <https://rscf.ru/project/26-18-00851/>

Литература

1. 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.
2. Ефимов А.Р., Агеева А.В., Крайнов А.Г., Фёдоров А.К., Кардымон О.Л., Стариков П.П. Искусственный интеллект в науке: на пороге новой области знания? // Вопросы философии. 2024. № 4. С. 30–41.
3. Ивахненко Е.Н., Ковальзон М.М. Будет ли присуждена искусственному интеллекту Нобелевская премия в 2040 году? // Информационное общество. 2024. № 2. С. 2–10.
4. Ивахненко Е.Н., Янукович М.Ф. Аутопоэтический зомби: нейросеть как коммуникативная система без субъекта // Вопросы философии. 2026. № 7. С. 63–75.
5. OpenAI решил геометрическую тайну Эрмита через теорию чисел. URL: <https://www.securitylab.ru/news/572862.php>
6. Поппер К. Объективное знание. Эволюционный подход. Пер. с англ. Д.Г. Лахути. М.: Эдиториал УРСС, 2002.
7. Фейнман Р. Характер физических законов. М., 1968.
8. Хакинг Я. Почему язык значим для философии? / пер. с англ. В.В. Целищева. М.: Канон+ РООИ «Реабилитация», 2026.
9. Хуэй Ю. Рекурсивность и контингентность. М.: V-A-C Press, 2020.
10. AI & Science: What Is the Future of Discovery? URL: <https://www.amacad.org/daedalus/ai-science-what-is-the-future-of-discovery>
11. Can a Chatbot be a co-author? URL: <https://news.harvard.edu/gazette/story/2026/02/can-a-chatbot-be-a-co-author/>
12. Fields Medalist says ChatGPT 5.5 Pro delivered "PhD-level" math research in under two hours with zero human help. URL: <https://the-decoder.com/fields-medalist-says-chatgpt-5-5-pro-delivered-phd-level-math-research-in-under-two-hours-with-zero-human-help/>
13. GPT-5.2 derives a new result in theoretical physics. February 13, 2026. URL: <https://openai.com/index/new-result-theoretical-physics/>
14. Harris R. The semantics of science. London: Continuum International, 2005.
15. Krenn M., Champion H. Philosophy of Autonomous Science: Ten Questions for the Coming Age of Artificial Scientists. URL: <https://www.amacad.org/publication/daedalus/philosophy-autonomous-science-ten-questions-coming-age-artificial-scientists>
16. Primitive sets and von Mangoldt chains: Erdős Problem #1196 and beyond URL: <https://terrytao.wordpress.com/2026/05/03/primitive-sets-and-von-mangoldt-chains-erdos-problem-1196-and-beyond/>
17. Regt H.W. Understanding scientific understanding. Oxford University press 2017.
18. Stiegler B. The Theater of Individuation: Phase-shift and Resolution in Simondon and Heidegger // Pharrhesia, N. 7, 2009.
19. "The future of artificial intelligence and the mathematical and physical sciences" (workshop report, IOP, 2025). URL: <https://iopscience.iop.org/article/10.1088/2632-2153/ae3e4e>
20. The 80,000 Hours Podcast. Chris Olah on what the hell is going on inside neural networks (Episode 107). URL: <https://80000hours.org/podcast/episodes/chris-olah-interpretability-research/>
21. What is OpenAI for Science? // Early Experiments in Accelerating science with GPT-5. URL: <https://openai.com/index/accelerating-science-gpt-5/>
22. Wang H., Fu T., Du Y. et al. Scientific discovery in the age of artificial intelligence // Nature. 2023. Vol. 620. P. 47–60. URL: <https://www.nature.com/articles/s41586-023-06221-2>

CREATIVITY OF NEURAL NETWORKS IN MODERN SCIENTIFIC RESEARCH

Ivakhnenko, Eugene N.

*DSc in Philosophy, professor
Lomonosov Moscow State University
Faculty of Philosophy
Department of philosophy for the humanities
Moscow, Russian Federation
ivahnen@rambler.ru*

Abstract

The topic linking scientific achievements with breakthroughs in AI has moved into the focus of readers' interest over the past decade. Thematic issues of high-ranking domestic and foreign journals are devoted to its discussion. The authors of Information Society, following the chosen editorial policy, contribute to the comprehension of issues related to this problematic. The main idea of the article lies in the author's aspiration to answer several questions related to the spread of the influence of generative AI's enormous potential on the entire course of contemporary scientific research. The author's argumentation regarding the creativity of neural networks in scientific research is supported by concrete examples of achievements in this field. Along with this, the article demonstrates that the achievements of twentieth-century philosophy of science are by no means taken outside the boundaries of AI-related themes; on the contrary, they contribute to its deeper comprehension in an evolutionary key. In the concluding part of the article, the question is raised of filling the concept of "scientific understanding" with new meanings under conditions of the growing creative potential of neural networks.

Keywords

science; neural network; generative artificial intelligence; organ-like system; scientific explanation; scientific understanding

References

1. Anoxin K.V., Novoselov K.S., Smirnov S.K., Efimov A.R., Matveev F.M. Iskusstvenny`j intellekt dlya nauki i nauka dlya iskusstvennogo intellekta. Voprosy` filosofii. 2022. № 3. P. 93–105.
2. Efimov A.R., Ageeva A.V., Krajnov A.G., Fyodorov A.K., Kardy`mon O.L., Starikov P.P. Iskusstvenny`j intellekt v nauke: na poroge novoj oblasti znaniya? Voprosy` filosofii. 2024. № 4. P. 30–41.
3. Ivakhnenko E.N., Koval`zon M.M. Budet li prisuzhdena iskusstvennomu intellektu Nobelevskaya premiya v 2040 godu? Informacionnoe obshchestvo. 2024. N 2. P. 2–10.
4. Ivaxnenko E.N., Yanukovich M.F. Autopoe`ticheskiy zombi: nejroset` kak kommunikativnaya sistema bez sub`ekta . Voprosy` filosofii. 2026. № 7. P. 63–75.
5. [OpenAI reshil geometricheskuyu tajnu Erdős cherez teoriyu chisel. URL: https://www.securitylab.ru/news/572862.php](https://www.securitylab.ru/news/572862.php)
6. Popper K. Objective Knowledge. Oxford, Clarendon Press, 1979.
7. Feynman R. Xarakter fizicheskix zakonov. Moscow, 1968.
8. Hacking Ian. Pochemu yazy`k znachim dlya filosofii? Moscow, 2026.
9. Hui Yuk. Rekursivnost` i kontingentnost`. Moscow. V–A–C Press, 2020.
10. AI & Science: What Is the Future of Discovery? URL: <https://www.amacad.org/daedalus/ai-science-what-is-the-future-of-discovery>
11. Can a Chatbot be a co-author? URL: <https://news.harvard.edu/gazette/story/2026/02/can-a-chatbot-be-a-co-author/>
12. Fields Medalist says ChatGPT 5.5 Pro delivered "PhD-level" math research in under two hours with zero human help. URL: <https://the-decoder.com/fields-medalist-says-chatgpt-5-5-pro-delivered-phd-level-math-research-in-under-two-hours-with-zero-human-help/>
13. GPT-5.2 derives a new result in theoretical physics. February 13, 2026. URL: <https://openai.com/index/new-result-theoretical-physics/>
14. Harris R. The semantics of science. London: Continuum International, 2005.

15. Krenn M., Champion H. Philosophy of Autonomous Science: Ten Questions for the Coming Age of Artificial Scientists. URL: <https://www.amacad.org/publication/daedalus/philosophy-autonomous-science-ten-questions-coming-age-artificial-scientists>
16. Primitive sets and von Mangoldt chains: Erdős Problem #1196 and beyond URL: <https://terrytao.wordpress.com/2026/05/03/primitive-sets-and-von-mangoldt-chains-erdos-problem-1196-and-beyond/>
17. Regt H.W. Understanding scientific understanding. Oxford University press 2017.
18. Stiegler B. The Theater of Individuation: Phase-shift and Resolution in Simondon and Heidegger. Pharrhesia, N. 7. 2009.
19. "The future of artificial intelligence and the mathematical and physical sciences" (workshop report, IOP, 2025). URL: <https://iopscience.iop.org/article/10.1088/2632-2153/ae3e4e>
20. The 80,000 Hours Podcast. Chris Olah on what the hell is going on inside neural networks (Episode 107). URL: <https://80000hours.org/podcast/episodes/chris-olah-interpretability-research/>
21. What is OpenAI for Science? Early Experiments in Accelerating science with GPT-5. URL: <https://openai.com/index/accelerating-science-gpt-5/>
22. Wang H., Fu T., Du Y. et al. Scientific discovery in the age of artificial intelligence. Nature. 2023. Vol. 620. P. 47-60. URL: <https://www.nature.com/articles/s41586-023-06221-2>