

RU

## Англоязычная терминология в цифровой среде: неологизация и полисемия в эпоху искусственного интеллекта (на примере сферы информационных технологий)

Рогова М. Ю.

**Аннотация.** Цель настоящего исследования – выявить основные изменения и проблемные аспекты развития современной лексикографии и терминологии в условиях цифровизации, междисциплинарности научного знания и полисемии терминологических единиц. В статье рассмотрено современное состояние словарного дела, проанализировано влияние цифровых технологий и определены границы применимости искусственного интеллекта в терминографическом описании. Исследование показывает, что активное расширение словарного состава за счет неологизмов в цифровой среде и детерминализации лексики приводит к ослаблению прескриптивных установок и переходу к более гибкому отражению динамики живой речи. Научная новизна исследования заключается в систематизации механизмов семантической трансформации терминов высокотехнологичных отраслей и в обосновании необходимости сочетания корпусных методов, больших данных и лингвистического анализа в лексикографической обработке лексики. Полученные результаты показали технологические ограничения полной автоматизации лексикографических процессов. Нейросеть склонна к предоставлению ложной информации (галлюцинациям), плохо справляется с многозначностью, ошибается в частях речи и приводит однотипные примеры. Это обуславливает необходимость сохранения ведущей роли терминолога при проектировании словарей.

EN

## English terminology in the digital environment: neologization and polysemy in the era of artificial intelligence (by the example of the information technology sphere)

M. Y. Rogova

**Abstract.** The aim of the present study is to identify the main changes and problematic aspects of the development of modern lexicography and terminography in the context of digitalization, the interdisciplinarity of scientific knowledge, and the polysemy of terminological units. The article examines the current state of lexicography, analyzes the impact of digital technologies, and defines the limits of applicability of artificial intelligence in terminographic description. The study shows that the active expansion of vocabulary due to neologisms in the digital environment and the determinologization of vocabulary lead to a weakening of prescriptive attitudes and a transition toward a more flexible reflection of the dynamics of live speech. The scientific novelty of the study lies in systematizing the mechanisms of semantic transformation of terms in high-tech industries and in substantiating the need to combine corpus methods, big data, and linguistic analysis in the lexicographic processing of vocabulary. The obtained results showed the technological limitations of full automation of lexicographic processes. A neural network is prone to generating false information (hallucinations), handles polysemy poorly, makes errors in parts of speech, and provides repetitive examples. This necessitates maintaining the leading role of the terminologist when designing dictionaries.

### Введение

Актуальность исследования обусловлена тем, что развитие цифровых технологий, массовое появление неологизмов и размывание границ между узкоспециализированной и общеупотребительной лексикой требуют пересмотра традиционных прескриптивных подходов к кодификации терминов. Традиционные печатные и академические словари из-за долгого цикла подготовки не успевают отражать реальные языковые изменения, в результате чего актуальные определения профессиональных понятий чаще фиксируются на онлайн-

платформах и в сетевых ресурсах. В данной работе предпринята попытка анализа этой проблемы и обоснования необходимости перехода к более гибким моделям терминографирования, способным оперативно фиксировать динамику живой специальной (профессиональной) речи.

Для достижения цели были сформулированы следующие задачи:

- изучить теоретико-методологические основы и современные направления развития лексикографии и терминографии в цифровую эпоху;
- проанализировать влияние процессов цифровизации и междисциплинарности на динамику неологизации и детерминализации специальной лексики;
- исследовать специфику проявления полисемии при функционировании терминологических единиц в межотраслевом контексте;
- оценить возможности и ограничения применения инструментов искусственного интеллекта в автоматизации терминографических процессов.

Материалом для исследования послужили терминологические единицы и неологизмы сферы информационных технологий, цифровой экономики и искусственного интеллекта, тексты новостных сообщений, данные лексикографических источников (Cambridge Dictionary, Большого Оксфордского словаря, онлайн-словаря Gramota), зафиксированные в профессиональном дискурсе, примеры контекстуального употребления и семантической трансформации специальных лексем.

Теоретическую базу исследования составляют работы в области терминоведения, терминографии и специальной лексикографии (Гринев-Гриневиц, 2008; 2021; Гринев-Гриневиц, Сорокина, 2015; Гринев-Гриневиц, Сорокина, Молчанова, 2022; Дубичинский, 2020; Лейчик, 2009; Комарова, 2017; Алексеева, Мишланова, 2021), посвященные изучению и описанию фундаментальных закономерностей функционирования терминотерминологических систем и методологии составления специализированных словарей. В исследовании учитывались труды в области общей лексикологии, практической лексикографии и корпусных исследований (Арнольд, 2012; Головкин, 2020; Рябцева, 2022; Воронцова, 2022), в которых освещаются вопросы теоретических и практических аспектов лексикологии, рассматриваются характерологические особенности явления полисемии, анализируются приемы функционального машинного перевода, тенденции современных онлайн-лексикографии и терминографии.

Теоретической основой исследования послужили: учебное пособие О. М. Карповой и Л. А. Маник (2026), посвященное английской лексикографии в цифровую эпоху; публикация G.-M. de Schryver (2023), представляющая анализ влияния цифровых технологий и автоматизации на современные лексикографические процессы; исследование M. Jakubíček, M. Rundell (2023), в котором авторы оценивают потенциал ChatPGT в лексикографической практике.

Для проведения исследования применяются методы теоретического анализа лингвистической и научной литературы, которые способствуют систематизации понятийного аппарата и определению современных направлений развития лексикографии и терминографии; методы дефиниционного анализа использованы для фиксации проявления полисемии в современных ИТ-терминотерминологических системах. Сравнительно-сопоставительный метод был применен для оценки влияния экстралингвистических факторов на скорость формирования понятийно-терминологической базы высокотехнологичной сферы.

Практическая значимость исследования предполагается в возможности применения полученных результатов в практике преподавания теоретических и прикладных дисциплин, связанных с информационными технологиями в лингвистике и переводом в сфере профессиональной коммуникации. Результаты данного исследования могут быть использованы в профессиональной деятельности лексикографа/терминографа.

## Обсуждение и результаты

Лексикография, возникшая как практический инструмент толкования сложного лексического материала, в ходе эволюции трансформировалась в нормативно-кодификационную и исследовательскую дисциплину, где словари закрепляют языковую норму и систематизируют языковой материал, стремясь к его максимально полной фиксации (Арнольд, 2012, с. 347).

«Сегодня русская лексикография включает прескриптивное (предписывающее) и дескриптивное (описывающее) направления, которые реализуют разные подходы к словарному представлению лексики», – пишет М. В. Копотев, научный консультант корпуса русского языка и руководитель ИИ-направления онлайн-словаря «Грамота» (Словари). Первый подход (прескриптивный) обусловлен заинтересованностью государства в регулировании языка, что напрямую касается и словарей. Так, Закон о государственном языке, принятый в 2005 году, предписывает фиксировать языковые нормы, в том числе в нормативных словарях (О государственном языке..., 2005). Второе направление (дескриптивное) сочетает в себе фундаментальные труды с информационными технологиями, которые работают с обширными данными текстов. В этой парадигме разрабатывается большая часть авторитетных академических словарей (Копотев, 2025).

В условиях стремительной цифровой трансформации, глобализации информационных потоков и междисциплинарного характера современного научного знания лексикография сталкивается с необходимостью решения некоторых проблемных вопросов. Классические методы создания словарей и глоссариев зачастую перестают удовлетворять запросы пользователя, требующего оперативного доступа к актуальным, многоаспектным и регулярно обновляемым данным. Так, О. М. Карпова и С. А. Маник отмечают, что «уходят в прошлое

старые традиционные типы словарей, уступая место инновационным англоязычным лексикографическим продуктам» (2026, с. 37). Исследователи также говорят о необходимости создания Интерактивного указателя общедоступной библиографической базы данных по словарям, который будет «включать ссылки на оцифрованные словари в Интернете с гибким поиском для исследователей различных профилей».

Н. К. Рябцева отмечает развитие нового направления – электронной лексикографии, которая помимо составления традиционных и электронных словарей играет роль в лингвистическом обеспечении систем искусственного интеллекта и смежных ресурсов. В качестве основного вектора развития данного направления исследователь выделяет создание «активных» многофункциональных платформ (электронных словарей, лексикографических баз данных, онтологий и т. д.). При этом основной вызов связан с «необходимостью разработки нового формата представления лексикографических данных, обусловленного инновационными возможностями цифровых технологий» (Рябцева, 2022, с. 76).

Известно, что развитие лексикографии не ограничивается общей нормой и описанием узуса: оно закономерно привело к появлению специализированных направлений, отвечающих запросам отдельных профессиональных сфер. Так, терминография, которая, согласно С. В. Гринева-Гриневичу (2008, с. 11), является наукой о составлении словарей специальной лексики, оформилась в самостоятельную отрасль. Одной из предпосылок формирования и становления данной научной отрасли, по мнению терминоведа З. И. Комаровой (2017, с. 15), стала необходимость преодоления отрицательных последствий информационного, в том числе и терминографического, «взрывов», которые нужно было превратить в управляемые процессы за счет сознательного регулирования специальной лексики.

Л. М. Алексеева и С. Л. Мишланова связывают развитие современного терминоведения с внедрением цифровых технологий и автоматизированных ресурсов, в результате которых «происходит расширение пространства терминологии, благодаря пониманию того, что специальный язык – это не только социолект как субнормативный вариант языка, но и элитарный языковой идиом, “клубный”, преодолевающий ограничение нормы; за счет исчезновения границы между общеязыковым и терминологическим (демаркации между научным и обыденным), образования трансязычного пространства, увеличения ресурсной базы терминотворчества и неологизации» (2021, с. 24).

Взгляды исследователей на понимание того, как развивающаяся действительность преобразует словарное дело, сходятся. Так И. А. Воронцова, отмечает, что «словарь выходит на уровень динамически меняющегося, персонально конфигурируемого, инновационного синергетического ресурса, становясь полноценным партнером пользователя в решении большого числа задач, связанных с рецепцией и генерированием информации» (2022, с. 192).

Неоспоримым представляется и утверждение Н. К. Рябцевой о трансформации лексикографических потребностей общества: «Спрос на традиционные бумажные словари значительно снизился, вызвав нерентабельность их выпуска и обновления, а спрос на (бесплатные) цифровые лексикографические ресурсы резко увеличился, вызвав потребность в их коренной перестройке, особенно для стремительно развивающихся приложений класса искусственный интеллект» (2022, с. 88).

Возросшая потребность в новых цифровых форматах обусловлена не только сменяемостью технологий, но и быстрым расширением самого понятийного аппарата науки. Закономерно, что новые профессиональные языки на начальных этапах заимствуют термины из уже существующих, смежных дисциплин, что создает сложность в упорядочении и классификации лексических единиц (Гринева-Гриневич, 2008, с. 6).

На современном этапе развития науки данный процесс усложняется за счет нового фактора – междисциплинарной интеграции, которая, по данным опроса Института статистических исследований и экономики знаний НИУ ВШЭ, лидирует в рейтинге значимости трендов развития науки и перспектив до 2030 г., опережая цифровизацию исследований и разработок, стремительно развивавшуюся в последние годы (Китова, 2025).

Выявленная закономерность подчеркивает значимость семантической точности научной терминологии, поскольку термин является основой научного дискурса. Как справедливо отмечает С. В. Гринева-Гриневич, «удачные термины могут способствовать ее развитию, а неудачные – тормозить процесс развития научного знания» (Гринева-Гриневич, 2008, с. 6).

Следующая проблема проявляется в переводческой практике: эффективность работы переводчика напрямую зависит от качества словарей. Так, Э. А. Сорокина и С. В. Гринева-Гриневич (2015, с. 58) пишут, что во многих случаях переводимому термину в словарях соответствует несколько терминов, не являющихся эквивалентами исходному термину, что приводит к ошибкам перевода.

Именно поэтому требования к качеству специальной лексикографии (терминографии) предполагают строгую упорядоченность понятийного аппарата.

М. В. Копотев полагает, что возникает противоречивый эффект: современные технологии и их развитие обеспечивают равный доступ к языковым ресурсам всем пользователям, но одновременно порождают рост количества текстов, что ведет к стремительной трансформации языковой реальности. В результате, с точки зрения языковой политики наблюдается устойчивая тенденция к снижению роли традиционных институтов в процессах языковой кодификации и нормирования (Словари).

В настоящее время отмечается, что даже современные системы машинного перевода сталкиваются с существенными ограничениями при обработке полисемантических единиц, и это нередко приводит к недостаточной степени качества перевода. К подобным проблемам относятся грамматические и лексические ошибки, нарушение связности изложения. Н. К. Рябцева указывает, что «автоматическая обработка естественного языка, ввиду чрезвычайной сложности его устройства, пока еще значительно ограничена всего несколькими

возможностями, которые, правда, постепенно и постоянно расширяются, в первую очередь, благодаря совершенствованию технологий» (2022, с. 78).

Основная сложность заключается в том, что существующие алгоритмы не могут интерпретировать и воспроизводить скрытые смыслы исходного текста, такие как культурные особенности мышления, в некоторых случаях – сленг или эмоциональную окраску. Одним из потенциальных решений данной проблемы может стать активное использование контекстологических словарей, где для каждого полисемантического термина заранее определяется предпочтительный переводной эквивалент в рамках конкретной предметной области (Головкин, 2020, с. 28).

Цифровая среда усугубляет и проблему полисемии, возникающей по причине того, что интернет является глобальным пространством, где пересекаются люди, имеющие языковую культуру, соответствующую разным профессиональным сообществам, культурам и языковым группам. Специальные термины из сферы информационных технологий (ИТ), медицины и других научных областей активно используются сегодня не только в сугубо профессиональном, но и в бытовом общении. Для специалиста подобные термины имеют точное значение, в то время как для обычного пользователя – размытое и метафоричное. Так, например, термин *sandbox* обозначает не только песочницу, но и отдельную часть компьютерной системы, где программное обеспечение может быть протестировано без какого-либо риска нанесения вреда всей системе (Cambridge Dictionary); термином *pipeline* обозначается не только понятие «трубопровод», но и «план/продукт, которые находятся в стадии разработки, обсуждаются или подготавливаются и будут произведены или завершены в будущем». В ИТ-сфере термином «пайплайн» номинируется «последовательность шагов, которые проходит программный код перед тем, как стать приложением в смартфонах»; а в производственной сфере – это то, что необходимо сделать, чтобы получить определенный продукт; в продажах – это последовательность действий, которые используются для привлечения и удержания клиентов (Лебедев, 2024). Термин *scaffolding* означает не только возвышение, служащее опорой для рабочих и материалов во время проведения строительных работ (Cambridge Dictionary), но и метод, который расширяет возможности программистов по эффективной настройке и изменению программного обеспечения, в частности в отношении структурированных данных (Scaffold (programming), 2018). Лексема *orphan* может быть соотнесена не только с понятием «сирота», но и с данными, которые не связаны с соответствующей записью или объектом в базе данных, в хранилище данных или в другой информационной системе (Komprise).

1. Данные языковые единицы наглядно демонстрируют расхождение между обыденной и терминологической картинами мира. Для неспециалиста метафорическая связь остается прозрачной и образной, тогда как в профессиональной коммуникации происходит концептуализация конкретного технического механизма. При формировании термина из исходного понятия выделяется ключевой структурно-функциональный признак, определяющий его дальнейшую терминологическую специализацию.

Язык, особенно в цифровой среде и в некоторых профессиональных сферах (ИТ, маркетинг, финансы, наука о данных и др.), меняется быстрее, чем создаются традиционные словари. Возникают новые реалии, для которых требуются новые названия. Накапливаемая обществом информация непрерывно преломляется через призму человеческого знания, в результате чего отражается в языке через формирование новых понятийных смыслов (Комарова, 2017, с. 5).

Например, термин *bootstrap*, согласно словарю Cambridge Dictionary, имеет три разных значения:

1) a piece of leather or other strong material at the back of a boot that you use to help you pull the boot on (*кусок кожи или другого прочного материала на задней части ботинка, который вы используете для надевания ботинка*);

2) to improve your situation or become more successful, without help from others or without advantages that others have (*улучшить свое положение или стать более успешным, без помощи других или без преимуществ, которыми обладают другие*);

3) to start a business, especially an internet one, with little money (*начать бизнес, особенно в Интернете, с небольшим количеством денег*) (Cambridge Dictionary).

В сфере ИТ термин *bootstrap* обозначает команду ввода, исходную команду, команду самозагрузки (Merriam-Webster), а также – готовую модель, шаблон программной платформы для создания сайтов и веб-приложений с открытым исходным кодом (Getbootstrap). Термин демонстрирует пять разных толкований, которые могут быть соотнесены с разными областями знаний.

Семантическая деривация термина *bootstrap* обусловлена когнитивной метафорой, в основе которой находится концептуальный инвариант «автономность / запуск процесса за счет внутренних ресурсов без внешнего импульса» (от идиомы “pull oneself up by one’s bootstraps” – дословно «вытащить себя за ремешки обуви»). Если в общезыковой сфере эта идиома передает идею достижения успеха или преодоления трудностей исключительно своими силами, то в ИТ-дискурсе происходит метафорический перенос от физического действия к процессу автономного запуска системы, где исходная команда запускает последующую цепочку вычислительных операций.

Приведенный пример демонстрирует, как общеупотребительная лексема посредством метафорической семантической деривации приобретает специализированные значения в ИТ-дискурсе, сохраняя при этом связь между исходным и терминологическим значениями.

С точки зрения особенностей семантики, представляет интерес и термин *agile*, который в своем первом значении связан с пониманием способности быстро и легко двигаться (*able to move your body quickly and easily*) (Cambridge Dictionary). Применительно к умственной деятельности данный термин приобретает переносное

значение – *able to think quickly and clearly* (способный мыслить быстро и ясно), то есть обладать гибкостью и быстротой ума, способностью мгновенно адаптироваться к новой информации. Однако наиболее значимым в современном профессиональном мире стало третье значение. Согласно словарю Cambridge Dictionary, термин *agile* также используется для описания способов планирования и выполнения работы, предполагающих внесение изменений по мере необходимости как часть процесса (Cambridge Dictionary), где вместо строгого следования изначальному плану, приоритет отдается адаптивности и непрерывному получению обратной связи. Наиболее ярко эта философия воплотилась в сфере ИТ, где появилось понятие *agile manifesto* – «манифест гибкой разработки программного обеспечения», содержащий описание ценностей принципов работы (The Agile Manifesto).

В данном случае проблема заключается в необходимости разграничения общеупотребительного прилагательного *agile* и его терминологизированного употребления для обозначения методологии организации разработки программного обеспечения. Данный пример иллюстрирует процесс терминологизации общеупотребительной лексемы, в результате которого прилагательное приобретает специализированное значение и становится обозначением подхода к организации профессиональной деятельности в ИТ-сфере.

Термин *NFT* (non-fungible token) закрепился в специальном профессиональном лексиконе, отражая цифровую революцию в области права собственности. По определению словаря Cambridge Dictionary, это уникальная единица данных, которая привязана к конкретному цифровому объекту (произведению искусства, музыке, видео и т. д.) и может быть куплена или продана (*non-fungible token: a unique unit of data (the only one existing of its type) that links to a particular piece of digital art, music, video, etc. and that can be bought and sold*) (Cambridge Dictionary).

Ключевая характеристика NFT – его уникальность и невзаимозаменяемость: каждый NFT существует в единственном экземпляре и не может быть равноценно заменен другим (Cambridge Dictionary).

Интересно отметить, что один из составляющих термина *NFT* – «токен» – является полисемантическим компонентом и представляет три значения в различных технологических сферах:

- *токеном* является минимальная единица текста, которая обрабатывается искусственным интеллектом (Губайловский, 2024);
- в области кибербезопасности *токен* – это средство идентификации пользователя или отдельного сеанса работы в компьютерных сетях и приложениях (Токен), то есть цифровой ключ, обеспечивающий безопасный доступ к ресурсам;
- в области финансов термин обозначает единицу учета в распределенном реестре, выражающую ценность того или иного объекта и/или объем прав на них, то есть финансовый инструмент (Большая российская энциклопедия).

Объединяя эти понятия, *NFT* становится цифровым сертификатом собственности, основанным на технологии блокчейн, которая представляет собой распределенную базу данных, хранящую информацию обо всех транзакциях участников в виде последовательной цепочки блоков (Генкин, 2025). Он превращает произведение, изначально легко копируемое, в уникальный, верифицируемый и отслеживаемый актив. На цифровом рынке появляется новая модель собственности: пользователь владеет не самим файлом, а подписанным автором и защищенным криптографически доказательством своего уникального права на него. Одним из примеров адаптации традиционных институтов к новому рынку стала инициатива Государственного Эрмитажа в 2021 году о выпуске и продаже NFT-токенов (цифровых копий) знаменитых картин из своей коллекции (Эрмитаж продал NFT-токены..., 2021).

Так, пример NFT демонстрирует полисемию компонента *token*, значение которого конкретизируется в зависимости от профессиональной сферы его употребления. При этом различные значения сохраняют общий семантический компонент, связанный с представлением определенной цифровой единицы, объекта, права или доступа. Развитие цифровых технологий расширяет семантический потенциал уже существующих терминологических компонентов. Для терминографии это создает необходимость точного разграничения значений полисемантических единиц, функционирующих одновременно в нескольких технологических сферах, и их корректной репрезентации в цифровых терминологических ресурсах.

Расширение семантической структуры под влиянием развития технологий демонстрирует и термин *prompt*, первым значением которого является «побуждать, подталкивать» или в театральной сфере – «дать подсказку» актеру. С развитием информационных технологий термин получил значение инструкции для ИИ при описании задачи (Cambridge Dictionary).

Термин *agent*, согласно словарю Cambridge Dictionary, имеет четыре значения:

- человек, действующий от имени или в интересах другого лица;
- кто-то тайно работающий на правительство или другую организацию (*secret agent, foreign agent*);
- человек или предмет, который вызывает результат или изменение (*a powerful cleaning agent* – мощное чистящее средство; *a raising agent for cakes* – разрыхлитель для теста; (литературный вариант) *He was the agent of their destruction. / Он был виновником их гибели*);
- система ИИ, которая может выполнять задачи без участия человека и обычно используется для принятия решений (*AI agent*) (Cambridge Dictionary).

Помимо этого в программировании существует понятие *User Agent*, термин, «обозначающий строку HTTP-заголовка, которую браузер, мобильное приложение, поисковый робот или другой клиент отправляет серверу при каждом запросе страницы. Она сообщает базовую информацию о программе, с помощью которой пользователь открывает сайт» (Бухтеев, 2021).

Термин демонстрирует смещение от значений, связанных с обозначением человека или объекта, выступающих в качестве посредника либо причины определенного действия или результата, к специализированным

значениям, сформировавшимся в цифровой среде. Значение *AI agent* отражает современное переосмысление лексемы *agent* применительно к системам ИИ, способным самостоятельно выполнять задачи и принимать решения.

Подобно тому, как с развитием компьютерных сетей термин *spam* перешел от первоначального наименования мясных консервов к обозначению массовых нежелательных сообщений, в том числе по электронной почте, примером современного семантического развития является термин *slop*, получивший новое значение с развитием ИИ.

Изначально данная лексема обозначала пищевые отходы, особенно те, что скормливают животным (Cambridge Dictionary). В современном переосмыслении термин обозначает «материал в Интернете очень низкого качества, особенно сгенерированный с помощью ИИ» (Cambridge Dictionary). В основе данного семантического сдвига лежит образно-оценочная аналогия: признаки исходного значения (безликость, низкое качество, отсутствие ценности) переносятся на информационный продукт. Следовательно, лексема начинает обозначать обесцененный, низкокачественный материал – «информационный мусор». Появление данного значения отражает существующие ограничения в области ИТ и негативные последствия полной автоматизации.

Термин *jailbreak*, первоначально означавший «побег из тюрьмы», с развитием технологий приобрел значение «снятия встроенных ограничений с электронного устройства» (Merriam-Webster). В основе указанной семантической деривации лежит концептуальный инвариант «преодоление внешне наложенных барьеров / защитных механизмов». Более того, на современном этапе наблюдается метонимический перенос (от процесса к инструменту), в результате которого «джейлбрейком» может называться не только факт обхода ограничений, но метод его реализации (Encyclopedia by Kaspersky).

О многозначности лексических единиц как о «свойстве большинства слов самых разнообразных языков» пишет И. В. Арнольд, проводя аналогию между часто используемым инструментом, который приобретает множество функций, и словом, регулярно встречающимся в разнообразных контекстах и оказывающимся полисемичным.

Универсальное свойство языка – порождать все новые значения у часто употребляемых слов – становится особенно острой проблемой в современной терминографии. Стремительное появление неологизмов и расширение семантики существующих терминов в высокотехнологичных сферах представляет сложности их систематизации. Многие методологические проблемы терминографии, выделенные В. В. Дубичинским (2020, с. 550) и обозначенные им как разработка инвариантного проекта словаря для описания различных специальных пластов лексики, исследование макро- и микроструктуры словаря, анализ путей отбора терминологического словника, являются в настоящее время весьма актуальными и для общей лексикографии, и для терминографии. Традиционные методы кодификации в условиях цифровой среды не успевают за скоростью языковых изменений. Так, «применение компьютеризации в создании терминологических словарей» и «создание терминологических баз данных и автоматизированных рабочих мест терминоведа», о которых пишет В. В. Дубичинский, перешли в вопрос о разработке полностью автоматизированных систем сбора и анализа «больших данных» (Big Data).

Важно отметить, что сама лексема «термин», по определению Э. А. Сорокиной, С. В. Гринева-Гринева и М. А. Молчановой (2022, с. 723), как «специальная лексема, именуемая общее понятие», изначально тяготеет к сфере профессионального знания. Следовательно, длительный процесс перехода термина из узкопрофессионального обихода в общеупотребительный язык, где он начинает фиксироваться в общих словарях, свидетельствует о глубоком усвоении соответствующего концепта культурой и его прочном внедрении в повседневную практику.

Но при этом следует заметить, что многие новые термины могут долгое время отсутствовать в авторитетных словарях, создавая лауну для пользователей. Известно, что фиксация любого неологизма требует определенного количества времени: лексикографам необходимо проанализировать значение новой лексемы, ее грамматические особенности и ее стилистическую окраску, отметить устойчивые примеры ее речевого употребления. Этапы этой работы, как отмечал В. М. Лейчик, включают: 1) изучение литературы; 2) составление картотеки терминов; 3) экспертный отбор и редактирование; 4) построение понятийной схемы; 5) создание алфавитной картотеки; 6) подбор эквивалентов (Дубичинский, 2020, с. 551).

Цифровая эпоха требует пересмотра формата и функций словаря. Как указывает С. В. Гринева-Гринева, «создание словарей является чрезвычайно трудоемким делом – для разработки дескриптивного словаря объемом в 1 тыс. лексических единиц одному человеку потребуется не менее двух лет, а для составления толкового или переводного словаря такого же объема – еще больше времени» (2021, с. 124). Подобная временная затратность традиционных методов приводит к тому, что академические издания неизбежно отстают от динамики живой речи. В результате под влиянием новых технологий традиционное понимание языковой нормы сменяется более гибкими, вероятностными моделями, ориентированными на живое употребление языка, что ослабляет прескриптивные установки лексикографов.

Цифровая лексикография формирует новую модель справочных ресурсов, которая основана на открытом доступе и привлечении пользователей к их составлению. Результаты деятельности цифровой лексикографии способны ускорить обновление лексикографических данных и усилить влияние инновационных цифровых проектов на развитие языковой практики (Карпова, Маник, 2026, с. 39).

Новым этапом в автоматизации словарей становится концепция «искусственного лексикографа», разрабатываемая по аналогии с системами искусственного интеллекта. Данный интеллектуальный процессор автоматически выбирает наиболее подходящий вариант информации и выводит его в виде оперативной подсказки

пользователю (Рябцева, 2022, с. 88). Но при этом необходимо заметить, что цифровизация словарного дела связана с существенными технологическими ограничениями.

Уже имеющийся опыт использования нейросетей достаточно уверенно опровергает гипотезу о возможности полной автоматизации словарей.

Результаты исследования, полученные Вендалин Николс, главой управления по изданию Кембриджского словаря, подтверждают, что, несмотря на ряд преимуществ, ИИ склонен к приведению ложной информации, созданию избыточно сложных формулировок и требует навыков работы с запросами. Более того, система допускает синтаксические ошибки, плохо справляется с синонимией и ограничена актуальностью своих данных (De Schryver, 2023, p. 369).

Интересно отметить, что феномен создания искусственным интеллектом ложной информации, обозначаемый английским термином “hallucinate”, не только является ключевой проблемой, но и нашел отражение в лексикографии – Кембриджский словарь обновил определение лексической единицы “hallucinate”, добавив специализированное значение, описывающее способность ИИ создавать несуществующие данные: “when an artificial intelligence (a computer system that has some of the qualities that the human brain has, such as the ability to produce language in a way that seems human) hallucinates, it produces false information” (Cambridge Dictionary). / Когда искусственный интеллект (компьютерная система, обладающая некоторыми качествами человеческого мозга, такими как способность воспроизводить речь так, чтобы она казалась человеческой) «галлюцинирует», он выдает ложную информацию. Значимость этого процесса была подтверждена присвоением термину “hallucinate” статуса «Слово года» в Cambridge Dictionary в 2023 году.

Наглядной иллюстрацией подобных ограничений служит эксперимент создания английского мини-словаря из 99 заголовочных слов на английском языке с помощью ChatGPT, результаты которого сравнивались с эталонными, академическими словарями, составленными людьми – Oxford Dictionary of English и Macmillan English Dictionary. Исследователи выяснили, что нейросеть плохо справляется с многозначностью: она создает «ложную полисемию», разбивая одно значение на несколько дублирующих синонимичных формулировок (так, для слова *climate* система предложила шесть практически одинаковых вариантов и не менее пяти – для *potato*), полностью упуская контекстные и переносные значения.

Исследователи объясняют это тем, что система не обладает представлением о сущности «словарного значения» в человеческом понимании. Приводя пример слова *beach*, авторы отмечают, что система описывает его как «место отдыха, куда люди приходят, чтобы поплавать, позагорать и заняться другими видами активного отдыха на свежем воздухе», не отмечая его близости к морю или озеру (Jakubíček, Rundell, 2023).

Кроме того, ChatGPT допускает грубые ошибки в частях речи и переходности глаголов, но при этом довольно хорошо справляется со стилистическими и оценочными пометами (верно определяет оскорбительные слова и некоторые устаревшие, предлагая замену). Однако отмечается, что предлагаемые системой примеры употребления выглядят искусственно и однотипно, в них не отражается живое употребление из корпусов, преимущественно они построены по шаблону «подлежащее в третьем лице с глаголом в простом прошедшем времени».

Тем не менее отмечается то, что нейросеть показала и сильные стороны: она хорошо справилась с формулировкой определений (технической области) и успешно разграничила синонимы и антонимы (Jakubíček, Rundell, 2023).

Рассмотренные ситуации демонстрируют риски автономного использования ИИ. Установлено, что в задачах, которые требуют высокой точности и достоверности данных, требуется дополнительная оценка экспертов, поскольку полностью полагаться на искусственный интеллект сегодня нельзя. Однако эти ограничения не отменяют стратегического потенциала технологий. Думается, что в перспективе развитие технологий ИИ позволит объединить существующие лексикографические ресурсы английского языка, «открывая новые возможности сотрудничества научных сообществ лексикографов, терминоведов, экспертов с пользователями для совершенствования словарной культуры, позволяя по-новому взглянуть на основополагающие проблемы лексикографии как своего рода справочно-информационных услуг» (Карпова, 2026, с. 40).

## Заключение

Полученные результаты проведенного исследования позволяют сформулировать такие выводы.

Эволюция лексикографии и терминографии, продолжающаяся в цифровую эпоху, характеризуется смещением парадигмы от прескриптивного нормирования к дескриптивному моделированию живой речи. Традиционные бумажные словари постепенно уступают место интерактивным платформам, корпусным инструментам и специализированным порталам, которые развивают словарное дело как современную справочно-информационную службу.

Современное состояние действующей междисциплинарности научных отраслей и процессы цифровизации науки значительно меняют профессиональный язык, вызывая активное расширение словарного состава за счет неологизмов и ускоряя детерминологизацию: термины IT-сферы и цифровых экосистем переходят в смежные дисциплины и повседневный дискурс, посредством прохождения пути семантической трансформации и когнитивного постепенного усвоения культурой.

Функционирование терминологических единиц высокотехнологичных отраслей в межотраслевом контексте сопровождается усилением процессов усложнения семантической структуры термина расширением

явления полисемии, при котором одна лексема становится способной сочетать в себе узкоспециальные значения и общеупотребительные, метафорические смысловые оттенки. А это значит, что усиливается проблема точного семантического разграничения терминов-полисемантов и терминов-омонимов, стоящая в центре внимания терминографии.

Уже имеющийся опыт использования инструментов искусственного интеллекта в автоматизации терминографических процессов позволяет продемонстрировать их эффективность при формулировании определений, особенно в технической сфере, и обработке стилистических помет, при этом выявляет значительные технологические ограничения, например, склонность нейросетей к предоставлению некорректной информации (галлюцинация) и «ложной полисемии» исключает полную автоматизацию лексикографических процессов и обосновывает необходимость сохранения ведущей роли терминоведа при сочетании больших данных, корпусных методов и лингвистического анализа.

В качестве перспектив исследования можно выделить дальнейшую разработку рассматриваемой проблематики и анализ терминологических систем других профессиональных областей.

## Материалы исследования | Research materials

1. Большая российская энциклопедия // АНО «Интернет-энциклопедия «РУВИКИ». <https://old.bigenc.ru/>
2. Бухтеев В. Что такое User agent и как его изменить // Timeweb. 2021. 24 июня. <https://timeweb.com/ru/blog/authors/viktor-buhteev/articles/chto-takoe-user-agent/>
3. Генкин А. Технология блокчейн: что нужно знать в 11 карточках // РБК Тренды. 2025. 12 августа. <https://trends.rbc.ru/trends/industry/5f05c0a79a7947a9a5c7577a?from=copy>
4. Губайловский В. Что такое токены, как с ними работает ИИ, и почему вам стоит это узнать // Techinsider. 2024. 7 марта. <https://www.techinsider.ru/news/news-1635093-chto-takoe-tokeny-kak-s-nimi-rabotaet-ii-i-pochemu-vam-stoit-eto-uznat/>
5. Китова Г. Будущее науки: междисциплинарность // ИСИЭЗ НИУ ВШЭ. 2025. <https://issek.hse.ru/news/1013764486.html>
6. Копотев М. Русская страсть к составлению словарей: достижения прошлого и цифровое будущее // Грамота.ру: справочно-информационный портал. 2025. 26 февраля. <https://gramota.ru/journal/stati/nauka/russkaya-strast-k-sostavleniyu-slovaray-dostizheniya-proshlogo-i-tsifrovoe-budushchee>
7. Лебедев А. Что такое pipeline и как его использовать в продажах // РБК Отрасли. 2024. 5 августа. <https://www.rbc.ru/industries/news/66aba8639a79475540b778dd>
8. О государственном языке Российской Федерации: Федеральный закон от 20 мая 2005 г. // Электронный фонд правовых и нормативно-технических документов. <https://docs.cntd.ru/document/901935235?marker=>
9. Словари // Грамота.ру: справочно-информационный портал. <https://gramota.ru/biblioteka/slovari>
10. Токен // Энциклопедия «Касперского». <https://encyclopedia.kaspersky.ru/glossary/token/>
11. Эрмитаж продал NFT-токены картин из своей коллекции на 32 млн рублей // РБК. 2021. 7 сентября. <https://www.rbc.ru/crypto/news/61376bd99a7947185cd5d967>
12. Bootstrap // Merriam-Webster.com Dictionary. <https://www.merriam-webster.com/dictionary/bootstrap>
13. Cambridge Dictionary. <https://dictionary.cambridge.org/ru/>
14. Getbootstrap. <https://getbootstrap.com/>
15. Scaffold (programming) // Computer Science: Research Starters / EBSCOhost. 2018. <https://www.ebsco.com/research-starters/computer-science/scaffold-programming>
16. The Agile Manifesto // Agile Alliance. <https://agilealliance.org/agile101/the-agile-manifesto/>

## Источники | References

1. Алексеева Л. М., Мишланова С. Л. Вековой путь российского терминоведения // Научный диалог. 2021. № 9. <https://doi.org/10.24224/2227-1295-2021-9-9-34>
2. Арнольд И. В. Лексикология современного английского языка. Изд-е 2-е. М.: Флинта; Наука, 2012.
3. Воронцова И. А. Современные практики англоязычной онлайн-лексикографии // Верхневолжский филологический вестник. 2022. № 1 (28). <https://doi.org/10.20323/2499-9679-2022-1-28-186-194>
4. Головкин Д. Р. Особенности и виды машинного перевода // Вестник Московского информационно-технологического университета – Московского архитектурно-строительного института. 2020. № 4.
5. Гринев-Гриневич С. В. Введение в терминографию: Как просто и легко составить словарь. М.: Либроком, 2021.
6. Гринев-Гриневич С. В. Терминоведение. М.: Академия, 2008.
7. Гринев-Гриневич С. В., Сорокина Э. А. Полисемия в общеупотребительной и специальной лексике // Вестник Московского государственного областного университета. Серия: Лингвистика. 2015. № 4. <https://doi.org/10.18384/2310-712X-2015-4-51-64>
8. Гринев-Гриневич С. В., Сорокина Э. А., Молчанова М. А. Еще раз к вопросу об определении термина // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2022. Т. 13. № 3. <https://doi.org/10.22363/2313-2299-2022-13-3-710-729>
9. Дубичинский В. В. Некоторые современные вопросы терминографии // Rasprave. 2020. № 46 (2).

10. Карпова О. М. Англоязычная терминография в электронно-цифровую эпоху // Вестник Ивановского государственного университета. Серия: Гуманитарные науки. 2026. № 1. <https://doi.org/10.46726/H.2026.1.22>
11. Карпова О. М., Маник С. А. Современная английская лексикография и корпусные технологии. Иваново: Иван. гос. ун-т, 2026.
12. Комарова З. И. Информационные вызовы современного общества: лингвотерминоведческие и терминографические процессы // Политическая лингвистика. 2017. № 5 (65).
13. Лейчик В. М. Терминоведение. Предмет, методы, структура. Изд-е 4-е. М.: Издательство ЛКИ, 2009.
14. Рябцева Н. К. Современная лексикография, дополненная лингвистическая реальность и проблемы терминологии // Лингвистика и методика преподавания иностранных языков. 2022. № 1 (16). <https://doi.org/10.37892/2218-1393-2022-16-1-75-110>
15. De Schryver G.-M. Generative AI and Lexicography: The Current State of the Art Using ChatGPT // International Journal of Lexicography. 2023. Vol. 36. № 4. <https://doi.org/10.1093/ijl/ecad021>
16. Jakubíček M., Rundell M. The end of lexicography? Can ChatGPT outperform current tools for post-editing lexicography? // Electronic lexicography in the 21st century: proceedings of the eLex 2023 conference. Brno: Lexical Computing, 2023.

#### Информация об авторах | Author information



**Рогова Маргарита Юрьевна<sup>1</sup>**

<sup>1</sup> Московский государственный технический университет им. Н. Э. Баумана



**Margarita Yurevna Rogova<sup>1</sup>**

<sup>1</sup> Bauman Moscow State Technical University

<sup>1</sup> [rogova.margo\\_rita@mail.ru](mailto:rogova.margo_rita@mail.ru)

#### Информация о статье | About this article

Дата поступления рукописи (received): 25.06.2026; опубликовано online (published online): 08.09.2026.

**Ключевые слова (keywords):** современное состояние лексикографии и терминографии; полисемия терминов; искусственный интеллект в терминографии; неологизмы цифровой сферы; детерминологизация лексики; current state of lexicography and terminography; polysemy of terms; artificial intelligence in terminography; neologisms of the digital sphere; determinologization of vocabulary.