

*На правах рукописи*



Шадрина Олеся Владимировна

**СТРУКТУРНО-СЕМАНТИЧЕСКИЕ ОСОБЕННОСТИ  
ТЕРМИНОЛОГИИ ПРЕДМЕТНОЙ ОБЛАСТИ  
«ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ»  
(КОРПУСНОЕ ИССЛЕДОВАНИЕ)**

Специальность 5.9.8. – Теоретическая, прикладная и сравнительно-сопоставительная лингвистика

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени  
кандидата филологических наук

Москва – 2026

Работа выполнена на кафедре английского языка и лингводидактики Государственного автономного образовательного учреждения высшего образования города Москвы «Московский городской педагогический университет».

**Научный  
руководитель:**

**Петрова Инна Михайловна,**  
доктор филологических наук, доцент, профессор кафедры  
английского языка и лингводидактики

**Официальные  
оппоненты:**

**Иконникова Валентина Александровна,**  
доктор филологических наук, доцент, декан лингвистики и  
межкультурной коммуникации ФГАОУ ВО «Московский  
государственный институт международных отношений  
(университет) Министерства иностранных дел Российской  
Федерации»

**Горохова Наталья Вячеславовна,**  
кандидат филологических наук, доцент, доцент кафедры  
иностраных языков ФГАОУ ВО «Российский  
государственный университет нефти и газа  
(национальный исследовательский университет) имени  
И. М. Губкина»

**Ведущая  
организация:**

Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Ивановский государственный университет»

Защита состоится «19» мая 2026 г. в 13:30 на заседании диссертационного совета 72.2.007.09 на базе ГАОУ ВО города Москвы «Московский городской педагогический университет» по адресу: 105064, г. Москва, Малый Казенный пер., д. 5Б.

С диссертацией можно ознакомиться в библиотеке ГАОУ ВО города Москвы «Московский городской педагогический университет» по адресу: 129626, г. Москва, 2-й Сельскохозяйственный проезд, д. 4 и на сайте ГАОУ ВО МГПУ [www.mgpu.ru](http://www.mgpu.ru).

Автореферат разослан « \_\_\_\_ » \_\_\_\_\_ 2026 г.

Ученый секретарь  
диссертационного совета



Л. А. Борботько

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Реферируемое диссертационное исследование выполнено в русле приоритетных направлений современного языкознания – терминоведения и корпусной лингвистики. Работа посвящена комплексному анализу структурно-семантических особенностей терминологии предметной области «искусственный интеллект» (далее – ИИ), включая морфологические, синтаксические и семантические механизмы терминообразования, их корпусную репрезентацию и функционирование в научном дискурсе.

**Актуальность** исследования обусловлена сложившейся научной парадигмой, в рамках которой структурно-семантические свойства терминологии предметной области «искусственный интеллект» исследуются преимущественно на материале английского языка. Являясь языком-донором и *lingua franca* международного научного сообщества, английский язык во многом определяет направления формирования терминосистемы ИИ [Smith 2024; Dabbara 2025]. Кроме того, активно развиваются исследования, посвященные анализу особенностей перевода терминологии ИИ на русский, китайский, испанский, немецкий и другие языки [Gómez-Pérez, Machado 2023; Кузнецова, Фуканчик 2023; Баланчуков 2024], а также изучению семантических процессов привлечения терминов и транстерминологизации [Мусаева 2022; Термины и понятия 2024]. В условиях нарастающих процессов глобализации и цифровизации науки изучение механизмов формирования, адаптации и стандартизации терминологии ключевой технологической области приобретает особое теоретическое и прикладное значение.

Вместе с тем термины ИИ ранее не рассматривались комплексно, поэтому целостная картина терминологии этой предметной сферы, учитывающая все многообразие явления, еще формируется. Кроме того, отсутствуют фундаментальные работы, посвященные структурно-семантическим особенностям терминосистемы ИИ, что создает определенные трудности в раскрытии ее функционального и деривационного потенциала [Фомина 2022; Гришин 2023].

В этом контексте в современном терминоведении особую значимость приобретает изучение потенциала цифровых методик, применяемых при анализе письменного академического дискурса. Настоящее исследование проведено с опорой на цифровой инструмент AntConc (версия 4.3.1), позволяющий оптимизировать процесс формирования и обработки эмпирической базы исследования [Сулейманова 2020; Anthony 2023]. Использование подобных технологий не только повышает точность и эффективность исследования, но и способствует более глубокому пониманию структуры и динамики терминологических систем. В связи с этим **актуальность** настоящей работы значительно возрастает благодаря демонстрации синергии классических лингвистических методов и передовых цифровых технологий, что отражает современные тенденции цифровизации в науке и обеспечивает более глубокий и точный анализ терминологии предметной области «искусственный интеллект».

**Объектом** исследования является терминология предметной области «искусственный интеллект» в научном дискурсе, репрезентированная в корпусе высокорейтинговых научных публикаций и соотнесенная с действующими международными стандартами ISO/IEC.

**Предметом** исследования выступают структурно-семантические особенности терминологических единиц предметной области «искусственный интеллект», проявляющиеся в их морфологической, синтаксической и морфолого-синтаксической организации, а также в процессах семантической деривации (эпонимизация, метонимизация, метафоризация, формирование синонимии и полисемии), выявляемые средствами корпусной лингвистики на материале английского языка.

**Цель** исследования – выявить и систематизировать структурно-семантические особенности терминологии предметной области «искусственный интеллект» в научном дискурсе на основе разработанной методики корпусного анализа.

Достижение поставленной цели потребовало решения следующих исследовательских **задач**:

– установить ключевые интралингвистические (морфологические, синтаксические, семантические) и экстралингвистические (технологические, институциональные, социокоммуникативные) факторы развития терминологии ИИ в научном дискурсе;

– изучить существующие подходы к анализу терминологии ИИ;

– разработать и апробировать корпусно-ориентированную методику исследования структурно-семантических особенностей терминологии ИИ;

– провести анализ структурно-семантических особенностей терминосистемы ИИ на материале английского языка;

– выявить и систематизировать деривационные механизмы и модели терминообразования в исследуемой области;

– обобщить результаты и представить системное описание структурно-семантической специфики терминосистемы ИИ на материале английского языка.

**Методологическую основу** исследования составляют положения работ отечественных и зарубежных лингвистов, связанные со следующими аспектами:

– определением особенностей термина (слова или словосочетания для наименования специальных понятий) [Винокур 1939; Wüster 1979; Головин, Кобрин 1987; Sager 1990; Cabré 1999; Авербух 2004; Лейчик 2022];

– применением методов систематизации термина как языка науки [Даниленко 1977; Felber 1984; Гринев-Гриневиц 2008; Буянова 2012; Лейчик 2022];

– присвоением наименования (термина) понятию на основе когнициии в ее лингвистических проявлениях [Lakoff 1987; Cabré 1999; Temmerman 2000];

– анализом научного дискурса и принципов отбора материала исследования на основе жанровых характеристик (исследовательская статья как ключевой жанр институционального научного дискурса; матрица IMRaD) [Карасик 2000; 2004; Чернявская 2001; 2010; 2016];

– анализом метонимии как системного механизма семантической деривации в научно-техническом дискурсе, включая типовые внутримоделные схемы и их операционализацию в корпусе через коллокационные профили и секционно-жанровые распределения IMRaD [Minsky 1975; Fillmore 1976; Peirsman, Geeraerts 2006; L’Homme 2020];

– ролью метафоры в формировании новых значений термина [Lakoff, Johnson 1980; Stefanowitsch, Gries 2006; Barsalou 2008; The Cambridge Handbook 2008];

– эпонимическими номинациями специальных понятий [Лейчик 2011; Новинская 2013; Какзанова 2019; Кузнецова, Петрулевич 2021];

– введением цифровых технологий в корпусные исследования языка [Suleimanova, Petrova 2020; Suleimanova 2020; Петрова 2022];

– особенностями терминологической лексикографии [Dubichynskyi, Dąbrowska 2021; Полубиченко, Алхастова 2024; Fuertes-Olivera 2025].

**Материал исследования** опирается на две взаимодополняющие эмпирические компоненты: 1) авторский корпус, включающий 897 рецензируемых англоязычных статей из журналов «Белого списка», отобранных по критериям тематической релевантности и жанровой однородности; 2) терминологическая выборка из 1648 единиц, извлеченная с помощью корпусного менеджера AntConc (версия 4.3.1) и представленная в Приложении № 1 в виде 1648 словарных статей.

Источником верификации полученных результатов выступают:

– толковые и этимологические словари английского языка: Longman Dictionary of Contemporary English URL, Merriam-Webster Dictionary URL, Macmillan Dictionary URL, Cambridge Dictionary URL, Online Etymology Dictionary URL, Oxford Learner’s Dictionaries URL, Oxford English Dictionary URL, The Concise Oxford Dictionary of English Etymology (2003);

– терминологические словари и справочники: The International Dictionary of Artificial Intelligence (1999), Толковый словарь по искусственному интеллекту (1992), Глоссариум по искусственному интеллекту: 2500 терминов (2023; 2024), Искусственный интеллект: электронный терминологический словарь (2023), Англо-русский толковый словарь по искусственному интеллекту и робототехнике (2022), Machine Learning Glossary URL;

– международные стандарты по искусственному интеллекту, разработанные совместной технической комиссией ISO/IEC JTC 1/SC 42 «Искусственный интеллект» Международной организации по стандартизации (ISO) и Международной электротехнической комиссии (IEC).

**Методика исследования** определялась его целью и поставленными задачами и включала несколько последовательно реализуемых этапов. На первом этапе сформирован авторский корпус из 897 рецензируемых англоязычных научных статей и материалов ведущих конференций по искусственному интеллекту (отбор источников осуществлялся с опорой на «Белый список» научных журналов), объемом 13 695 508 токенов, который был загружен в корпусный менеджер AntConc (версия 4.3.1). Автоматизированный частотный анализ выделил около 153 000 словоформ; в рабочий перечень включались единицы с

нормализованной частотностью  $\geq 10$  PMW, более редкие отсекались. Далее выполнена ручная фильтрация «шумовых» единиц – нерелевантных для терминографического анализа токенов (математико-технические маркеры и фрагменты кода; оценочная и общеупотребительная лексика, не имеющая отраслевой семантики; общеязыковые глаголы; имена собственные без терминологической функции и др.). В результате отбора получены 250 тематически релевантных токенов – «кандидатов» в термины.

На втором этапе осуществлялся углубленный статистический анализ с использованием показателей Mutual Information (MI) – меры ассоциативной силы слов относительно их независимого распределения, а также Log-Likelihood (LL) – критерия статистической значимости отклонения наблюдаемых совместных частот от модели независимого совместного распределения. Это позволило выделить 2300 релевантных слов и устойчивых словосочетаний (коллокаций). Полученная выборка прошла верификацию по авторитетным глоссариям и международным стандартам ISO/IEC, в результате чего в окончательный перечень вошли 1648 терминологических единиц.

На третьем этапе проводился лингвистический анализ отобранных терминов, включая выделение структурно-морфологических и семантических маркеров, специфичных для научно-технического дискурса ИИ. Дополнительно применялись метод лингвистического моделирования для идентификации ключевых признаков терминов; ономаσιологический подход для анализа формально-структурных моделей и оценки продуктивности словообразовательных механизмов (морфологические, синтаксические, морфолого-синтаксические); семасиологический анализ для исследования семантических процессов (метонимизация, метафоризация, эпонимизация, полисемия, синонимия); а также статистический анализ на всех этапах исследования для обеспечения точности, воспроизводимости и семантической достоверности результатов.

**Научная новизна** исследования состоит в том, что впервые на основе специально созданного репрезентативного корпуса научных статей с применением инструментария AntConc был проведен комплексный частотно-коллокационный анализ терминологии предметной области «искусственный интеллект», позволивший автоматизировать ее инвентаризацию и детально описать морфемно-словообразовательную организацию (аффиксальные, сложные, сложносокращенные и композитные модели) с оценкой деривационной продуктивности. На этой основе были системно исследованы семантические механизмы адаптации заимствованных и транстерминологизированных единиц, а также выявлены и типологизированы явления полисемии и синонимии с выделением критериев разграничения значений и управления эквивалентностью. Практическим результатом стала разработка и верификация двух взаимосвязанных классификаций терминов ИИ – деривационной и семантической, которые формируют научно-методическую основу для создания корпусно-ориентированного глоссария и последующей системной работы по унификации и стандартизации отраслевой терминологии.

**Теоретическая значимость** исследования обусловлена разработкой и обоснованием комплексных подходов к изучению структурно-семантических механизмов формирования и функционирования терминологии предметной области «искусственный интеллект», отражающей динамику обогащения новой терминосистемы под влиянием стремительного развития цифровых технологий. Особое внимание уделено взаимодействию внутренних лингвистических факторов, включая деривационные механизмы и семантическое терминообразование, а также внешних экстралингвистических условий (технологических инноваций, международного статуса английского как *lingua franca*, норм публикации и стандартизации). Полученные результаты расширяют теоретические границы терминологии, лексикологии и корпусной лингвистики, одновременно формируя методологическую базу для будущих исследований в области автоматизации процессов извлечения и унификации терминов ИИ.

**Практическая ценность** работы заключается в возможности использования ее результатов при проведении исследований в области терминологии ИИ, в курсах лекций и семинаров по корпусной и компьютерной лингвистике, терминоведению и терминографии, в спецкурсах по когнитивистике и прикладной лингвистике, в написании учебных пособий, подготовке курсовых, проектных и выпускных квалификационных работ, а также исследований по смежной тематике. Онлайн-гlossарий, составленный по результатам исследования, может найти применение в переводческой деятельности, при составлении научной и технической документации, написании научных статей. Кроме того, в работе предлагается описание алгоритма использования цифровых инструментов, что может быть полезно для оптимизации процесса сбора материала при создании эмпирической базы в сходных исследованиях на материале письменного академического дискурса.

**На защиту выносятся следующие положения:**

1. Современное терминоведение представляет собой комплексную междисциплинарную область, методологический аппарат которой для исследования динамичных терминосистем, подобно терминологии искусственного интеллекта, интегрирует классические лингвистические подходы с корпусно-вычислительными методами, социокогнитивной и функционально-коммуникативной парадигмами, а также нормами международной стандартизации (ISO 704, ISO 1087).

2. Терминосистема предметной области «искусственный интеллект», функционирующая в научном дискурсе, соотнесена с понятийной структурой соответствующей области знания и поддается многоуровневому описанию по функционально-дискурсивным, семантическим, формально-структурным, мотивационным и нормативным параметрам; данная классификация обеспечивает операциональную основу для корпусного анализа и терминографического моделирования (на материале английского языка).

3. Стандартизация выполняет роль ключевого механизма концептуализации и гармонизации терминологического аппарата ИИ. Основываясь на методологических стандартах терминологической работы (ISO 704, ISO 1087) и их национальных

версиях, такие международные нормы, как ISO/IEC 22989, 23053 и 23894, обеспечивают однозначность, воспроизводимость и междисциплинарную интероперабельность терминов, формируя нормативный каркас для профессиональной коммуникации и регуляторики.

4. Многоэтапная процедура выделения терминологических «кандидатов», включающая автоматизированное извлечение единиц из корпуса, последующую фильтрацию «шумовых» токенов и верификацию терминологической релевантности по конкордансам (KWIC), частотным показателям и статистическим мерам сочетаемости (Mutual Information, Log-Likelihood), формирует воспроизводимую эмпирическую базу для дальнейшего структурно-семантического анализа и терминографического описания терминологии предметной области «искусственный интеллект».

5. Формирование и пополнение терминологии ИИ осуществляется преимущественно через интегрированную модель терминообразования, сочетающую семантические, морфологические, синтаксические и морфолого-синтаксические механизмы. Данная модель отражает преобладание многословных номинаций (A+N, N+N, A+N+N) и высокую продуктивность морфологической деривации, системно кодирующей типы концептов (процесс/результат/инструмент/сущность), что вместе с привлечением общеупотребительной лексики и транстерминологизацией создает условия для оперативного и нормативно управляемого закрепления новых наименований.

6. Терминосистема ИИ демонстрирует управляемые проявления синонимии и полисемии, которые диагностируются корпусными методами (KWIC, распределительные профили, MI/LL) и регламентируются через стандартизованные дефиниции и установление предпочтительных форм. Методика корпусного исследования, основанная на построении репрезентативного специализированного корпуса и статистической триангуляции (частотность, Mutual Information, Log-Likelihood) в связке с качественным анализом, позволяет объективно идентифицировать и разграничивать ядро, периферию и инновационный слой терминов, а также подтвердить их релевантность и устойчивость.

**Структура работы.** Диссертационное исследование состоит из введения, трех глав, заключения, библиографического списка, включающего 198 источников (94 – на иностранном языке), списка использованных словарей и нормативных документов, 3 приложений. Общий объем работы составляет 201 страницу.

Во **Введении** представлено обоснование актуальности работы, определяются объект и предмет исследования, обозначаются его цель и задачи, теоретическая и практическая значимость, получают описание методы и эмпирический материал, излагаются основные положения, выносимые на защиту.

В **Главе 1** «Теоретико-методологические основы исследования терминологии предметной области “искусственный интеллект”» формируются концептуально-методологические основания работы; определяется место терминоведения в современной лингвистике, разграничиваются понятия «терминология», «терминосистема» и уточняется рабочая дефиниция термина; систематизируются механизмы терминообразования.

Представляется диахронический обзор эволюции терминологии ИИ – от кибернетики до трансформерных архитектур; устанавливаются нормативно-стандартные ориентиры гармонизации (ISO 704/1087, ISO/IEC 22989 и др.; ГОСТ; EU AI Act), задающие метаязык описания и принципы последующего корпусного анализа и лексикографической репрезентации.

В **Главе 2** «Методология проведения корпусного исследования терминологии предметной области “искусственный интеллект”» рассматриваются методологические основы корпусного анализа терминологии ИИ. Формулируются критерии отбора научных текстов и дается обзор ключевых инструментов платформы AntConc. Описываются принципы применения статистических метрик (Frequency, Mutual Information, Log-Likelihood) для выявления терминологических единиц. Детализируются этапы работы с корпусом – от подготовки и предобработки данных до извлечения и интерпретации терминов, и предлагается методика разработки онлайн-гlossария ИИ с учетом технических требований, предъявляемых к цифровым ресурсам.

В **Главе 3** «Структурно-семантические особенности терминосистемы предметной области “искусственный интеллект”» раскрываются основные параметры структурно-семантического профиля терминологии ИИ. Характеризуются морфологические и морфолого-синтаксические модели образования однословных терминов, такие как аффиксация (*binning* (биннинг), *pretraining* (предобучение)), конверсия (to mark up → *markup* (разметка)); словосложение (*keypoints* (ключевые точки)) и др.; описывается синтаксическая деривация двух- и многокомпонентных именных групп с учетом правоголовности, ограниченной вариативности и критериев устойчивости: A+N (*artificial agent* (искусственный агент), N+N (*data science* (наука о данных), V-ed+N (packed data (упакованные данные), N+prep+N (*Internet of Things* (интернет вещей) и др. Анализируются семантические способы номинации: заимствование (из греч. *autonomy* (автономность)); привлечение общеупотребительной лексики (*prompt* (подсказка) → инструкция/вход для LLM); транстерминологизация (из лингвистики в NLP: *parsing* (синтаксический разбор)), а также механизмы адаптации лексических единиц в терминосистеме ИИ: метонимизация (*DALL-E* → система и изображение/выход), метафоризация (*black box* (черный ящик)), эпонимизация (*Turing test* (тест Тьюринга)), сужение/расширение значения (*agent* (агент), *token* (токен)). Особое внимание уделяется выявлению полисемии (например, *model*: 1) математическая модель, 2) алгоритмическая модель, 3) предобученная нейросеть) и синонимии (например, *strong AI* (сильный ИИ) / *full AI* (полный ИИ)), что позволяет проследить многослойность значений и конкуренцию обозначений в терминологии ИИ.

В **Заключении** подводятся итоги исследования; полученные результаты систематизируются и интерпретируются в контексте структурно-семантических характеристик терминосистемы ИИ на материале английского языка; формулируются теоретические и прикладные выводы и намечаются перспективные направления дальнейшей работы.

**Приложения** содержат QR-код для доступа к онлайн-гlossарию терминов ИИ; QR-код для доступа к перечню лексикографических источников (включая онлайн-платформы и репозитории), а также к списку нормативно-правовых актов и стандартов; QR-код, обеспечивающий доступ к материалам, дополняющим основные положения исследования, включая примеры метонимических переносов; классификацию популяционных алгоритмов по метафорическим источникам и когнитивным эффектам; реестр терминов-эпонимов; таксономию синонимии в терминологии ИИ; описание полисемии ключевых терминов ИИ (семантические ядра, значения, источники и корпусные примеры).

**Апробация работы.** Основные положения и результаты диссертационного исследования были представлены в виде докладов в рамках следующих научных конференций: Международная научно-практическая конференция «Развитие современной науки и технологий в условиях трансформационных процессов» (18 октября 2024 г., Москва); Международная научная конференция «V Казанский международный лингвистический саммит – 2024» (29-31 октября 2024 г., Казань); Международная научно-практическая конференция «Искусственный интеллект и духовная культура» (30 октября 2024 г., Москва); VI Международная конференция «Human – AI Synergy for Effective Language Teaching of Science Communication in a Multidisciplinary, Multicultural Context» (21-23 ноября 2024 г., Долгопрудный); Международная молодежная конференция «В многомерном пространстве современной лингвистики: искусственный интеллект vs традиция» (10 декабря 2024 г., Москва); III Всероссийская конференция «Педагогический дискурс в современной научной парадигме и образовательной практике» (3-5 марта 2025 г., Москва); Научно-практическая конференция «Научный старт – 2025» (с элементами научной школы (3 апреля 2025 г., Москва); Всероссийская конференция «Взаимодействие мыслительных и языковых структур» (18-19 сентября 2025 г., Тамбов).

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Структура диссертации подчинена логике решения исследовательских задач и включает введение, три главы, заключение, список литературы и приложения.

В Главе 1 «**Теоретико-методологические основы исследования терминологии предметной области “искусственный интеллект”**» терминология ИИ трактуется как комплексный лингвокогнитивный феномен, изучение которого требует применения структурно-системного, когнитивного и дискурсивного подходов, а также методов корпусной лингвистики и стандартологического описания.

Становление терминоведения как самостоятельной области научного знания связывают прежде всего с работами О. Вюстера и Венской школы, утвердившими ономаσιологический принцип, связку «концепт – термин – дефиниция» и нормативную установку на однозначность [Wüster 1931; 1979; Felber 1984]. Дальнейшее развитие дисциплины обеспечили институционализация и стандартизация (INFOTERM, TermNet, DIN 2330, ISO 704, ISO 1087) и формирование франко-канадской и англо-американской

традиций [Rondeau 1984; Dubuc 1992; Sager 1990; Handbook of Terminology 1997]. В отечественной лингвистике терминоведение формировалось в русле системно-нормативного подхода [Лотте 1961; Реформатский 1967] и комплексной концепции термина и терминологии [Гринева-Гринева 2008; Буянова 2012; Лейчик 2022]. С 1990-х гг. к нормативной парадигме добавились когнитивный и социокогнитивный подходы [Cabré 1999; Temmerman 2000], а также корпусно-вычислительные методы [Kageura 2002; Bowker, Pearson 2002; Evert 2009].

Терминология рассматривается, с одной стороны, как совокупность специальных единиц определенной области знания, с другой – как научная дисциплина (терминоведение). В зарубежной традиции терминология понимается как организованный корпус наименований, отражающий систему понятий и поддерживаемый процедурами нормирования и стандартизации [Wüster 1979; Sager 1990; Arntz, Picht 1991; Depecker 2002; Diki-Kidiri 2022], тогда как коммуникативно-когнитивные подходы акцентируют ее динамичность и зависимость от дискурсивных задач и процессов категоризации в научных сообществах [Cabré 1999; Temmerman 2000; 2022].

В настоящем исследовании принимается структурно-системная позиция, согласно которой терминология понимается как автономная, но «проницаемая» подсистема лексики, отражающая систему понятий и взаимодействующая с общим языком [Реформатский 1967; Головин, Кобрин 1987; Лейчик 2022]; ее анализ опирается на комплекс лексикологических, семантических, когнитивных и корпусных процедур, дополняемых концептуальным моделированием и сочетанием ономастического и семасиологического подходов, что позволяет рассматривать термин одновременно как языковой знак и как элемент понятийной системы [Sager 1990; Cabré 1999; Temmerman 2000; Гринева-Гринева 2008; Лейчик 2022]. При этом в аналитическом плане разводятся понятия «терминология» как широкий терминный массив (кодифицированные термины, номены, пограничные единицы) и «терминосистема» как его концептуально упорядоченная проекция [Авербух 2004; Berger 2021; Лейчик 2022]. С учетом подвижности границ между терминами и номенами в работе не вводится жесткого противопоставления этих уровней, а вариативность (синонимия, полисемия) рассматривается как системный фактор, подлежащий описанию через отношения внутри массива и создающий основу для последующей стандартизации и корпусного анализа терминологии.

Определение термина опирается на широкий спектр дефиниций, отражающих междисциплинарный статус этого понятия и эволюцию терминоведческих подходов. В философско-логической традиции ключевым признаком термина выступает понятийность и его роль в структурировании суждения [Флоренский 1990; Лосев 2009]. Вюстеровская модель исходит из принципа «один термин – один концепт» и первичности понятийной системы [Wüster 1931; 1979]. Более широкие определения Л. Делекера и К. Кагеуры включают символы, формулы и номенклатурные обозначения как репрезентанты специальных понятий [Depecker 2002; Kageura 2002], а когнитивное и социокогнитивное терминоведение описывает термин как информационно-когнитивную структуру и

инструмент категоризации знания, для которого характерны прототипность, вариативность и дискурсивная обусловленность значения [Cabré 1999; Temmerman 2000; Thoiron, Béjoint 2010]. В отечественной традиции термин определяется как слово или устойчивое словосочетание профессионального языка, называющее научное понятие и встроенное в систему понятий соответствующей области [Реформатский 1967; Суперанская 2012; Лейчик 2022].

В рамках настоящего исследования термин – это единица языка для специальных целей (слово, устойчивое словосочетание, при необходимости – символослово), соотношенная с конкретным понятием и встроенная в систему понятий соответствующего домена; обладающая дефинированным содержанием (нормативно закрепленной дефиницией); стремящаяся к однозначности и высокой точности референции; стилистически нейтральная; а также нормативно кодифицируемая и воспроизводимая по форме (устойчивость графики и морфосинтаксической модели), что обеспечивает ее надежное функционирование в профессиональной коммуникации.

Способы терминообразования описываются в рамках двухэтапной модели терминотворчества, включающей первичную номинацию понятия в соответствии с понятийной структурой и языковыми законами и последующую деривационно-семантическую эволюцию терминологической единицы, отражающую углубление и развитие концепта [Татаринов 1996]. Динамика жизненного цикла термина («рождение – зрелость – угасание») [Guilbert 1975; Гак 1998] в быстро развивающихся междисциплинарных областях, подобных ИИ, характеризуется сокращением стадий [Головин, Кобрин 1987; Лейчик 2022]. Классическая традиция, исходящая из принципов системности и моносемии, трактует терминообразование преимущественно как использование внутренних ресурсов языка (аффиксация, словосложение, конверсия), синтаксических моделей терминологических сочетаний, семантического переноса, аббревиации и заимствования [Лотте 1961; Wüster 1979; Суперанская 2012; Лейчик 2022]. В качестве методологической основы принимается интегрированная классификация С. В. Гринева-Гриневица, сочетающая формальные и когнитивно-прагматические параметры и адаптированная к терминосистеме ИИ [Гринева-Гриневиц 2008].

Эволюция терминологии ИИ, соотношенная с этапами развития самой дисциплины, демонстрирует поступательное усложнение терминологического инвентаря – от ранних интернациональных однословных наименований к многокомпонентным структурам с развитой деривационной и семантической организацией. В 1940–1950-е гг. формируется концептуальный фундамент (нейросетевые и кибернетические модели, принцип обучения Хебба, поведенческий критерий Тьюринга) и закрепляются базовые единицы *neuron* (нейрон), *learning* (обучение), *feedback* (обратная связь), *Turing test* (тест Тьюринга) и название самой дисциплины *artificial intelligence* (искусственный интеллект) [McCulloch, Pitts 1943; Turing 1950; McCarthy 1955; Wiener 1961; Hebb 2002].

Далее терминология отражает переход от символического ИИ и экспертных систем к статистическому и глубокому обучению (*convolutional neural network* (сверточная

нейронная сеть), *deep reinforcement learning* (глубокое обучение с подкреплением)), а на современном этапе – к трансформерной архитектуре и парадигме фундаментальных моделей (*self-attention* (самовнимание), *pre-training* (предварительное обучение) и *foundation models* (фундаментальные модели)). Данная терминологическая динамика сопровождается формированием тематических кластеров, связанных с этикой, доверием и регулированием ИИ (*Explainable AI*, *XAI* (объяснимый ИИ), *responsible AI* (ответственный ИИ) и правовые понятия EU AI Act) [Vaswani 2017; Regulation (EU) 2024].

Гармонизация терминологии ИИ осуществляется в полицентричной нормативной рамке стратегий России, ЕС, США и Китая (нацстратегии [О стратегии НТР... 2016; О развитии ИИ... 2019], AI Act [Regulation (EU) 2024], NIST AI RMF, Executive Order 14110 [Artificial Intelligence RMF 2023; Safe, Secure... 2023], национальные программы КНР [State Council 2017]), где стандартизация выступает ключевым механизмом формализации требований и терминологической унификации. Определяющую роль играет подкомитет ISO/IEC JTC 1/SC 42, разработавший стандарты ISO/IEC 22989, 23053, 23894 на основе ISO 704 и ISO 1087 и их национальные эквиваленты (в том числе ГОСТ Р, GB/T 41867-2022, документы ТК 164), которые формируют ядро терминосистемы ИИ и служат точкой отсчета для сопоставления нормативной фиксации с фактическим употреблением терминов в корпусе [Гринева-Гринева 2008; Супранская 2012].

Глава 2 «Методология проведения корпусного исследования терминологии предметной области “искусственный интеллект”» посвящена разработке комплексной методики эмпирического анализа на основе корпусного подхода. Корпусная лингвистика рассматривается как одно из ведущих направлений современного языкознания, опирающееся на репрезентативные электронные собрания текстов и их формализованную обработку [Sinclair 1991; O’Keeffe, McCarthy 2010]. Корпус понимается как структурированная, документированная и размеченная коллекция текстов, сформированная на основе прозрачных принципов репрезентативности и предназначенная для решения конкретных лингвистических задач [Sinclair 1991; Захаров 2005]. Центральным элементом методики выступает корпусный менеджер – программная среда, обеспечивающая стандартизированный доступ к данным, выполнение поисковых запросов, расчет частот, коллокаций и показателей ключевости (keyness), а также экспорт результатов [Brezina 2021].

Развитие инструментов корпусного анализа демонстрирует переход от пакетной обработки данных без участия исследователя к интерактивным десктоп-системам (компьютерным программам) и веб-платформам (WordSmith Tools, Sketch Engine, CWB и др.), а корпусные менеджеры рассматриваются как ключевой элемент исследовательской инфраструктуры [Suleimanova 2020; Петрова 2022]. Выбор AntConc для проведения настоящего исследования обусловлен сочетанием функциональной достаточности и методологической прозрачности, поскольку данный ресурс представляет собой бесплатный кросс-платформенный корпусный менеджер с модулями Word List, KWIC, Collocate, Clusters/N-grams, Keyword и др., обеспечивающими частотный анализ, выявление коллокаций и ключевой лексики [Anthony 2023; 2024], и рекомендован как надежный

инструмент для составления авторских корпусов [Suleimanova 2020]. Статистические методы трактуются как ключевой механизм объективного описания языковых единиц, повышающий воспроизводимость и верифицируемость результатов [Теоретические и прикладные 2016; Gries 2024]. Современные корпусные менеджеры используют спектр показателей: абсолютную и нормализованную частотность, Log-Likelihood (LL), Mutual Information (MI), T-score, Z-score, Log Ratio и др. [Evert 2009; Anthony 2023]. В работе комбинируется частотный анализ (включая нормализацию и учет дисперсии) со статистическими метриками LL и MI. Такой подход позволяет выделять как статистически значимые высокочастотные сочетания, так и редкие, но устойчивые терминологические коллокации [Church, Hanks 1990; Evert 2009].

В качестве базового источника терминологического материала выбрана исследовательская статья как узловой жанр институционального научного дискурса, в котором формируются, определяются и аргументируются термины и фиксируются таксономические связи [Hyland 2004; Чернявская 2010]. Нормативно-стандартизационная линия (ISO 1087:2019; ISO 704:2022) задает связку «объект – понятие – определение – обозначение», но фактически опирается на реальное употребление терминов в научных публикациях. В функционально-коммуникативном и социокогнитивном подходах термин рассматривается как дискурсивная единица, форма и семантика которой зависят от коммуникативных задач и когнитивных моделей домена [Cabré 1999; Temmerman 2000]. Корпус специализированных статей при этом выступает оптимальной базой для анализа функционирования, вариативности и полисемии терминов. Корпусно-ориентированная терминография развивает эти подходы через автоматическое извлечение терминов (ATE) и концепт *knowledge-rich contexts* (KRC) – контекстов, в которых эксплицитно или имплицитно задаются определения и отношения между понятиями [Kageura, Umino 1996; Meyer 2001; Faber 2022]. Исследовательские статьи демонстрируют высокую концентрацию KRC, структурно распределенных по секциям IMRaD и маркированных устойчивыми шаблонами (*is defined as*, *refers to* и др.), что делает их идеальной средой для формализуемого поиска терминов и реконструкции фреймов.

Процедура проведения корпусного анализа с использованием инструмента AntConc включает подготовку корпуса англоязычных научных статей, отбор ключевой лексики, коллокационный и семантический анализ терминов. Эмпирический материал составляет корпус из 897 рецензируемых статей 2020–2024 гг. из ведущих журналов (*Artificial Intelligence Review*, *Journal of Artificial Intelligence Research* и др.), обеспечивающих тематическую репрезентативность и актуальность. Подготовленные тексты в формате .txt (UTF-8) загружаются в AntConc 4.3.1 как основной корпус (AI Target Corpus); справочный корпус (AI Reference Corpus) сформирован на основе «Глоссариума по искусственному интеллекту: 2500 терминов» [Баканач 2023; 2024]. Репрезентативность и «стабилизация частот» обеспечиваются соблюдением «закона текстового блока» и фильтрацией служебной и «шумовой» лексики с помощью расширенного списка стоп-слов [Палийчук 2022]. Корпус объемом 13 695 508 токенов подвергается двухуровневой обработке в

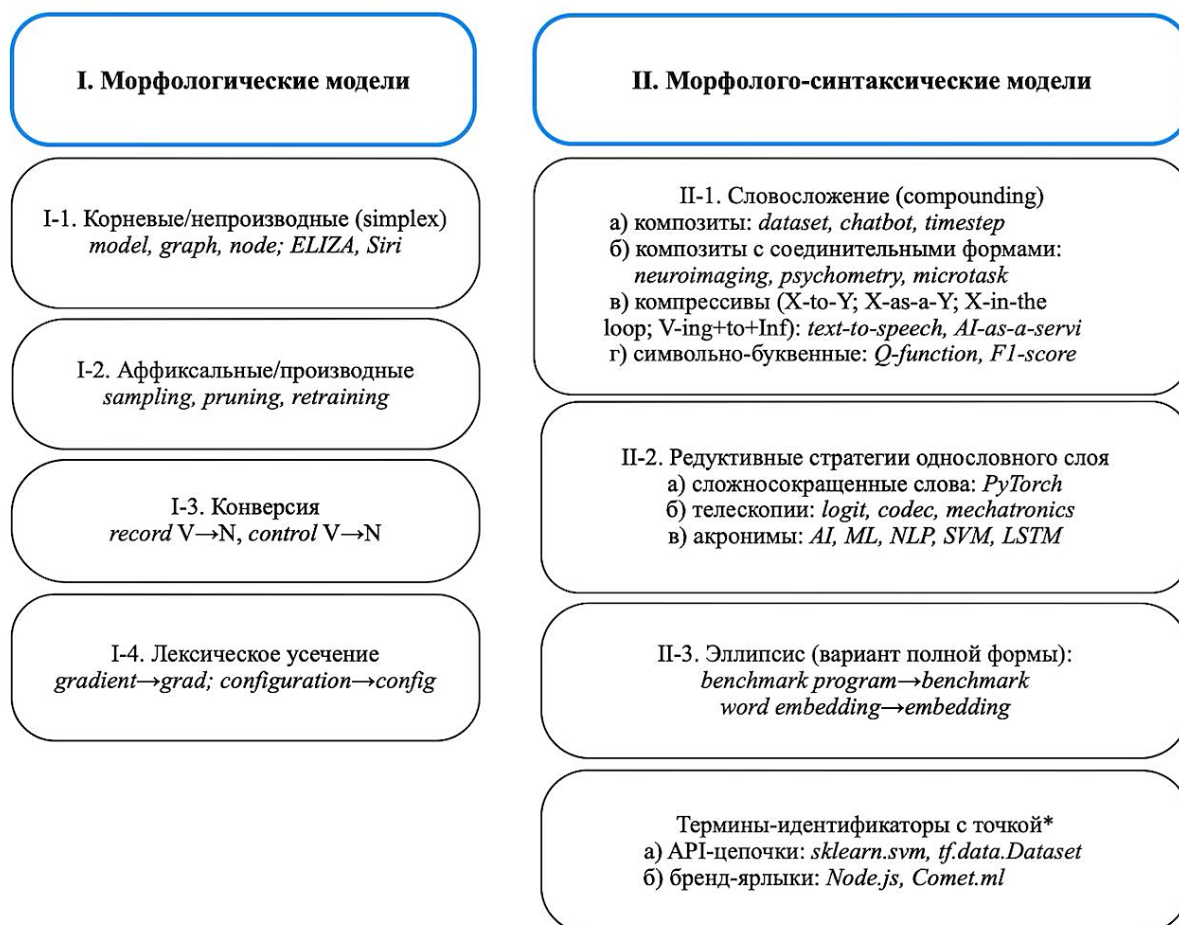
AntConc. В модуле Word List формируется список словоформ с абсолютной и нормализованной частотностью; в дальнейший анализ включаются только единицы с нормализованной частотностью  $\geq 10$  PMW, более редкие формы отсеиваются. Омонимия и варианты написания уточняются по конкордансам (KWIC), лемматизация осуществляется через шаблоны с подстановочным символом \*. В модуле Keyword List основной корпус сопоставляется со справочным, отбираются лексемы с наибольшими значениями, формирующие ядро терминологического поля научного дискурса ИИ (*learning, data, model* и др.) [Anthony 2023].

На следующем этапе высокочастотные однословные лексемы используются в модуле Collocates для выявления устойчивых терминологических сочетаний. Для каждой единицы рассчитываются две взаимодополняющие метрики коллокационной связанности – *Log-Likelihood* (LL) и *Mutual Information* (MI). Статистически значимыми считаются сочетания с  $LL > 10.83$  ( $p < 0.001$ ) и  $MI > 3$ , при этом дополнительно учитывается нормализованная частотность для исключения случайных и нерелевантных комбинаций. Совместное использование LL и MI позволяет, с одной стороны, отфильтровать частотные, но семантически слабо связанные сочетания, с другой – выделить редкие, но устойчивые терминологические коллокации. Итоговый перечень коллокаций подвергается внешней терминологической верификации, в результате которой формируется корпус из 1648 терминов (AI Verified Corpus). Словообразовательный и семантический анализ выполняются с использованием модулей Clusters, N-grams, Word List + Regex, KWIC и Collocates. Выявляются деривационные и композиционные модели (в том числе дефисные и CamelCase-композиции, акронимы) с оценкой их продуктивности по показателю Range, описываются эпонимические, метафорические и метонимические номинации, полисемия и синонимия, а также осуществляется тематическая кластеризация.

В завершение исследования результаты репрезентируются в форме электронного глоссария, в котором на основе KWIC и KRC-шаблонов (*is defined as, refers to* и др.) извлекаются дефиниционные и иллюстративные контексты и формируются статьи с определениями, переводами, пометами и примерами. Глоссарий реализован как онлайн-ресурс на платформе WordPress (WP Glossary) с алфавитно-тематической навигацией, перекрестными ссылками и возможностью обновления и выступает динамической моделью тематического поля терминологии ИИ.

Глава 3 «**Структурно-семантические особенности терминосистемы предметной области “искусственный интеллект”**» посвящена корпусному описанию структурно-семантических характеристик терминов ИИ. На первом этапе, в соответствии с дихотомией «термины-слова / термины-словосочетания» [Даниленко 1977; Гринев-Гриневиц 2008], выделено 450 однословных терминов (27% корпуса) и 1198 терминов-словосочетаний (73%).

В однословном слое выделяются морфологические и морфолого-синтаксические модели, операционализация которых обеспечивается единой схемой, представленной на Рисунке 1.



Примечание. Термины-идентификаторы с точкой – тип записи, не способ терминообразования.

Рисунок 1 – Структурные модели однословных терминов ИИ

Количественное распределение терминов-слов представлено на Рисунке 2. При подсчете не учитывались случаи эллипсиса и лексического усечения, рассматриваемые в работе как варианты полной терминологической формы.

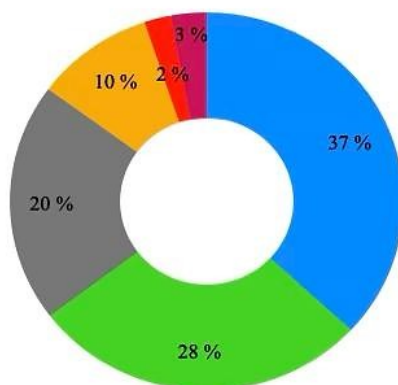
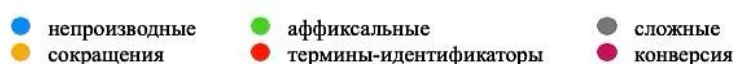


Рисунок 2 – Количественное распределение терминов-слов по типам морфологического и морфолого-синтаксического терминообразования

Непроизводные термины представлены интернациональными терминами-словами (*robot* (робот), *parameter* (параметр), *corpus* (корпус) и др.), а также номенами-идентификаторами (*ELIZA*, *Siri*, *Kaggle*, *Python*). В зоне аффиксального словообразования доминируют суффиксы -ing, -ity, -er/-or, -ation/-ization (*scoring* (оценивание), *sparsity* (разреженность), *quantifier* (квантификатор), *visualization* (визуализация)), обеспечивающие номинализацию процессов, их результатов, агентов/инструментов и свойств. Префиксы *re-*, *self-*, *hyper-*, *post-* (*retraining* (дообучение), *self-attention* (механизм самовнимания), *hyperplane* (гиперплоскость)) выполняют дифференцирующую функцию, модифицируя базовое терминологическое значение.

Морфолого-синтаксический блок однословного слоя представлен пограничными моделями: эллипсисом, расширенным компаундингом и редуктивными стратегиями, которые в совокупности дают 32% терминов-слов. Корпус показывает особую роль формульных компрессивов типа X-to-Y, X-as-a-Y, X-in-the-loop и V-ing+to+Infinitive (*text-to-speech* (синтез речи), *AI-as-a-service* (ИИ как сервис)), которые в сочетании с аббревиатурами и неоклассическими формами функционируют как лексикализованные композиты и демонстрируют черты терминологической грамматикализации. Символьно-буквенные композиты (*Q-function* (функция Q), *F1-score* (оценка F1)) и точечные термины-идентификаторы (*torch.nn*, *Node.js*, *Wit.ai*) оформляют специализированные зоны номинации, связывающие терминологию с программными экосистемами, тогда как разграничение диахронической и синхронической аббревиации позволяет избежать двойного счета и описать редуктивные модели как согласованную подсистему формальной экономики.

В результате однословный слой терминологии ИИ представляет собой результат взаимодействия интернационального ядра, продуктивной суффиксации, расширенного компаундинга и редуктивных механизмов.

Корпусный анализ подтверждает ведущую роль терминов-словосочетаний (около 73%, 1198 единиц), что согласуется с данными о доминировании составных терминов в европейских терминосистемах [Bowker, Pearson 2002; Гринев-Гриневиц 2008]. Аналитический строй английского языка и тенденция к «синтаксической компрессии» приводят к разворачиванию правоголовных именных групп, где левосторонние модификаторы компактно сворачивают дефиницию [Biber, Gray 2016]. Например, многоступенчатая конструкция *explainable ensemble predictive model* (объяснимая ансамблевая предиктивная модель) последовательно наращивается из базового термина *model* добавлением признаков предсказательности, ансамблевости и интерпретируемости.

Типология атрибутивной двухкомпонентной модели определяется морфологической природой и синтаксической ролью левого компонента и включает регулярные шаблоны N+N и A+N, на которые приходится около 70% двухкомпонентных терминологических словосочетаний (например, *error backpropagation* (обратное распространение ошибки), *neural network* (нейронная сеть), *probabilistic model* (вероятностная модель)). В конструкции N+N правая «голова» (*model* (модель), *network* (сеть), *algorithm* (алгоритм)) задает

таксономический класс, тогда как левый компонент специфицирует доменную, структурную или функциональную характеристику, часто образуя серийные гнезда типа *batch/online/stochastic backpropagation* (пакетное/онлайн/стохастическое обратное распространение ошибки). В модели A+N дифференцирующий признак кодируется относительным прилагательным (*neural* (нейронный), *convolutional* (сверточный), *generative* (генеративный)), конкурирующим с субстантивным препозитивным определением (*language model* (языковая модель) vs *linguistic model* (лингвистическая модель)).

Менее многочисленные, но систематические группы образуют конструкции с глагольной морфологией в головной позиции – N+V-ing (gerund) и A+V-ing (gerund) (*reinforcement learning* (обучение с подкреплением), *representation learning* (обучение представлений), *incremental learning* (инкрементное обучение)), а также в атрибутивной позиции – V-ed+N, V-ing (gerund)+N, V-ing (participle)+N и V-ed+V-ing (*trained model* (обученная модель), *learning algorithm* (алгоритм обучения), *streaming data* (поточные данные), *supervised learning* (обучение с учителем)), обеспечивающие компактную номинацию процессов, режимов и результатов. Смешанные схемы N+abb и abb+N (*game AI* (игровой ИИ), *CNN model* (модель сверточной нейронной сети)) отражают высокую частотность аббревиатур в ИИ-дискурсе и задают серию наименований моделей и метрик. Предлоговые группы N+prep+N (*area under the curve* (площадь под кривой)) и редкие притяжательные форматы N's+N (*experimenter's bias* (предвзятость экспериментатора)) фиксируются как периферийные и используются преимущественно там, где требуется явная экспликация ролевых отношений или устранение двусмысленности.

Трех- и многокомпонентные термины (28% терминов-словосочетаний) обычно формируются на базе стабилизированных двухкомпонентных конструкций, обладающих более тесными структурно-семантическими связями. Двухкомпонентные единицы функционируют как самостоятельные термины и служат опорным ядром для дальнейшего наращивания – либо посредством добавления «головы», либо за счет добавления еще одного модификатора при сохранении правоголовной структуры: *error propagation model* (модель распространения ошибки) ← *error propagation* (распространение ошибки), *propagation model* (модель распространения). В зависимости от положения исходного двучленного сочетания в трехчленной конструкции выделяются формализуемые структурные схемы: 1) A>(B>C) – централизованная модель, при которой к уже существующему двухкомпонентному термину добавляется препозитивный уточняющий элемент: *deep learning method* (метод глубокого обучения) ← *learning method* (метод обучения) + *deep* (глубокий); 2) (A,B)>C – параллельная модель, при которой два однородных атрибутивных компонента подчиняются одному головному элементу: *visual semantic embedding* (визуально-семантическое встраивание) ← *visual embedding* (визуальное встраивание), *semantic embedding* (семантическое встраивание); 3) (A>B)>C – децентрализованная модель, при которой двухкомпонентное сочетание выполняет функцию сложного определяющего элемента: *natural language processing* (обработка естественного языка) ←

*natural language* (естественный язык) + *processing* (обработка); 4) {A>(B)>C} – смешанный вариант из двух пересекающихся сочетаний, в котором промежуточный компонент – головной в одном и определяющий в другом: *deep feature extraction* (глубокое извлечение признаков) ← *deep feature* (глубокий признак) и *feature extraction* (извлечение признаков).

Многочленные именные группы с числом терминологических элементов более трех служат зонами максимальной спецификации понятий и источником последующей морфологической компрессии (*POMDP*, *partially observable Markov decision process* (частично наблюдаемый марковский процесс принятия решений)) [Cabré 1999; Temmerman 2000; Лейчик 2022].

Таким образом, синтаксическая деривация в терминологии ИИ описывается как системообразующий механизм, обеспечивающий точное и экономное представление сложных концептов за счет атрибутивного уточнения базовых терминов. Высокочастотные однословные единицы *model* (модель), *network* (сеть), *learning* (обучение), *agent* (агент) функционируют и как «головы», и как модификаторы в N+N-конструкциях, что повышает деривационную продуктивность лево-ветвящихся структур и поддерживает динамику терминосистемы.

Объединение морфологического и синтаксического анализа с коллокационными метриками (MI, LL) позволяет стратифицировать терминологию ИИ на устойчивое ядро и периферию. К ядру относятся частотные однословные непроеизводные и производные термины (*algorithm*, *data*, *learning*), а также термины-словосочетания, прежде всего классические конструкции A+N и N+N, включая модели с герундиальной головой (*artificial intelligence* (искусственный интеллект), *deep learning* (глубокое обучение), *machine learning* (машинное обучение), *neural network* (нейронная сеть), *reinforcement learning* (обучение с подкреплением)), которые характеризуются высокой частотностью, семантической цельностью и коллокационной устойчивостью и служат конструктивными опорами дискурса. Периферию составляют более узкоспециализированные, предлоговые и формульные модели; при этом многословность не тождественна периферийности, поскольку часть стабилизированных многокомпонентных единиц входит в устойчивый слой наравне с однословными терминами. Результаты структурного анализа позволили зафиксировать общий профиль терминологии ИИ в двух верхних классах – термины-слова и термины-словосочетания – и предложить авторскую классификацию, представленную на Рисунке 3.

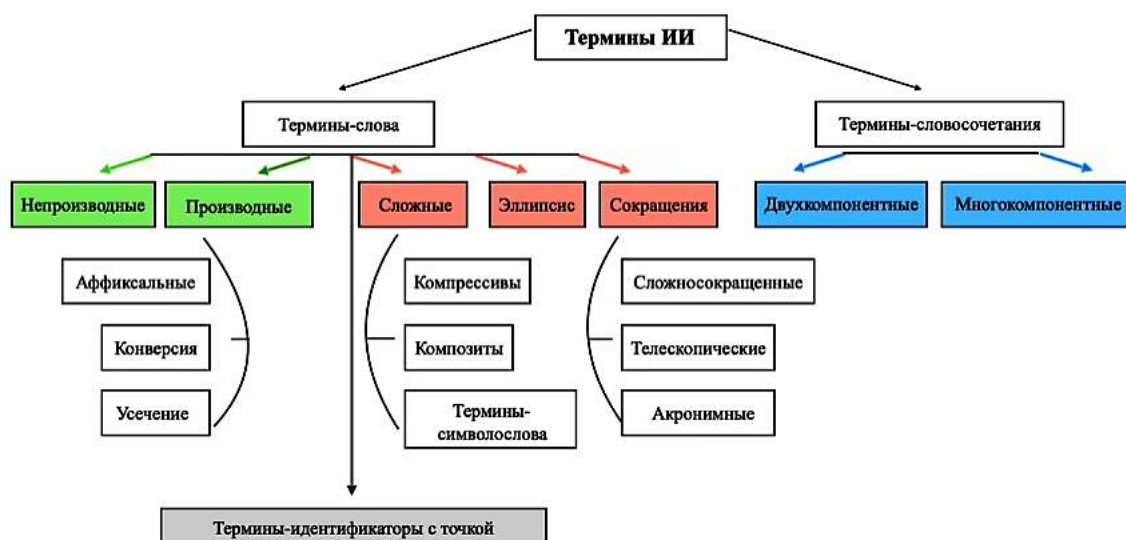


Рисунок 3 – Структурный профиль терминосистемы ИИ

Семантические особенности образования терминов ИИ описываются через совокупность процессов, при которых уже существующие лексемы – общеупотребительные, заимствованные или ранее терминологизированные – получают специализированные значения без изменения формы [Гринев-Гриневиц 2008; L’Homme 2020; Лейчик 2022]. Вслед за А. С. Мусаевой заимствование трактуется как межъязыковой переход терминов и терминологических элементов в терминосистему ИИ [Мусаева 2022], в русле контактной лингвистики – с разграничением прямых заимствований (*algorithm* (← лат. *algorismus* от имени аль-Хорезми)), калькирования (*principal components* (← нем. *Hauptkomponenten*)) и частичного заимствования формантов (*neuro-symbolic* (нейросимволический)) [Loanwords 2009]. Отдельную группу образуют слабо интегрированные, культурно маркированные единицы (*Ansatz* (анзац), *Gestalt* (гештальт), *Umwelt* (умвельт) и др.), функционирующие главным образом как дискурсивные маркеры научного стиля (не как именованья сущностей).

Привлечение в данном исследовании рассматривается как процесс терминологизации единиц общеупотребительного языка с сохранением формы слова, но с приобретением нового специализированного значения в профессиональном дискурсе (*prompt* (подсказка/стимул) → инструкция/вход для LLM (ср. *prompt engineering*, *prompt tuning*), что отражает динамику терминологической лексики и ее системное варьирование в научном языке [Мусаева 2022; Термины и понятия 2024].

Транстерминологизация понимается как «горизонтальный» перенос готовой терминологической единицы из одной терминосистемы в другую с последующей семантической адаптацией в новом профессиональном контексте; в отличие от межъязыкового заимствования, источник здесь – иная предметная область того же языка, а не другой язык [Гринев-Гриневиц 2008; Мусаева 2022]. Для терминологии ИИ типичны траектории из математики и статистики (*kernel* (ядро) → *kernel trick* (трюк ядра), *kernel method* (ядерный метод)), физики (*annealing* (отжиг) → *simulated annealing*

(симулированный отжиг)), когнитивных наук и лингвистики (*attention* (внимание) → *attention mechanism* (механизм внимания)), биологии/эволюционной теории (*population* (популяция)), экономики / теории игр и управления (*payoff* (выигрыш)).

В формировании терминосистемы ИИ привлечение общеупотребительных единиц и транстерминологизация из смежных дисциплин образуют единую нелинейную сеть, в которой такие лексемы, как *memory*, *neuron*, *attention*, *reward*, *policy*, проходят многоступенчатую семантическую переработку и превращаются в операциональные термины с разветвленными гнездами. Термины мигрируют между пластами (например, *reinforcement learning* (обучение с подкреплением) – от психологии к ИИ и обратно к когнитивной нейронауке), образуют гибриды (*convolutional neural network* (сверточная нейронная сеть) как сочетание математического, нейробиологического и общезыкового по происхождению компонента) и демонстрируют семантическую дивергенцию, при которой вторичные значения все более отрываются от исходной бытовой семантики (*attention* (внимание) → *attention mechanism* (механизм внимания) → *self-attention* (самовнимание), где семантика последнего понятия целиком определяется операциональным контекстом машинного обучения и не может быть выведена из бытового значения исходной единицы). Вследствие этого в работе нет жесткого противопоставления общеупотребительных и междисциплинарных источников, а акцент переносится на описание процессов их семантической адаптации в динамичной терминосистеме ИИ: 1) расширение значения – добавление специализированного, но по объему более широкого применения без вытеснения общеупотребительного смысла (например, *agent* (действующее лицо) → *agent* (автономная программная сущность / агент в ИИ среде) [L’Homme 2020; ISO 704 2022]; 2) специализацию (сужение) – закрепление узкого операционального содержания (например, *reward* (вознаграждение) → *reward* (скалярная обратная связь в RL) [Cabré 1999; Temmerman 2000; L’Homme 2020]; 3) метафоризацию – перенос по сходству образа/структуры (например, *hallucination* (галлюцинация) → *hallucination* (генеративная ошибка LLM)), [Lakoff, Johnson 1980; Temmerman 2000]; 4) метонимический перенос – замещение по смежности (например, «процедура → результат»: *loss* (функция потерь) → *loss* (ее численное значение); *minimize loss* = уменьшать величину потерь; *define the loss* = задавать функцию потерь) [Temmerman 2000; L’Homme 2020; ISO 704:2022]; 5) эпонимизацию – номинацию от имени ученого / автора / топонима / мифонима / брендониима: *Bellman equation* (уравнение Беллмана); *Markov chain* (марковская цепь) и др.

В настоящем исследовании фрейм выступает методологической основой интерпретации метафоры и метонимии и рассматривается как когнитивная структура, репрезентирующая стереотипное знание о типичной ситуации и ее участниках [Minsky 1975; Fillmore 1976]. Фреймовая семантика рассматривает значение лексической единицы как функцию активируемого ею фрейма, который связывает языковое и энциклопедическое знание и организует доступ к релевантным «слотам» ситуации [Lakoff, Johnson 1980; Barcelona 2000]. В этой перспективе метафора трактуется как междоменный перенос между различными фреймами, тогда как метонимия – как внутридоменное перераспределение

салиентности между элементами одного фрейма, обеспечивающее доступ к целому через наиболее значимый компонент [Lakoff 1987; Peirsman, Geeraerts 2006].

Метонимические употребления терминов в терминосистеме ИИ выявляются в когнитивно-фреймовой перспективе с опорой на корпусные методы (KWIC, LL, MI). Анализ лексико-синтаксических шаблонов (*improves X by N, achieves N on X, optimize X* с участием названий датасетов, метрик и моделей (*ImageNet, BLEU, BERT* и др.)) позволяет интерпретировать *X* как метонимический маркер фрейма «эксперимент / модель / оценка», обозначающий не объект данных, а достигнутый показатель или процедуру оптимизации. Конструкции *optimize BLEU* (оптимизировать метрику BLEU), *fine-tune BERT* (дообучить модель BERT) репрезентируют метрику или модель через их результаты или конфигурации, что дает возможность системно описать внутридоменные метонимические сдвиги и их роль в формировании терминологической структуры ИИ.

Метафора в настоящем исследовании анализируется на материале популяционных алгоритмов ИИ [Beheshti 2013] и понимается как междоменный перенос эволюционных, природных и социальных фреймов в их понятийное пространство (*particle swarm optimization* (оптимизация роя частиц), *simulated annealing* (имитация отжига), *genetic algorithm* (генетический алгоритм) и др.), благодаря которому сложные вычислительные процедуры оформляются как интуитивно интерпретируемые модели и формируют семантически связанные кластеры терминов. Корпусные процедуры (KWIC, LL, MI) фиксируют устойчивые метафорические коллокации и позволяют проследить, как эти проекции стабилизируются в научном дискурсе.

Эпонимизация в терминологии ИИ балансирует между критикой ее «непрозрачности» и признанием номинативного и мемориального потенциала [Лобач 1986; Суперанская 2023]. Вслед за Е. М. Какзановой эпоним трактуется как термин с проприальным компонентом (антропоним, топоним, мифоним) и терминологической «головой» (*theorem, test, model, algorithm* и др.) [Какзанова 2019]. В исследуемом корпусе при помощи регулярных выражений и порогов MI/LL выявлено 140 терминов-эпонимов ( $\approx 8,5\%$ ), преимущественно антропонимических (*Turing test* (тест Тьюринга), *Kalman filter* (фильтр Калмана)) и выделено 5 групп по основной семантической функции: алгоритм/модель (42%) – *Dijkstra's algorithm* (алгоритм Дейкстры); вычислительная архитектура (21%) – *von Neumann architecture* (архитектура фон Неймана); концептуальная/теоретическая рамка (14%) – *Nash equilibrium* (равновесие Нэша); математический инструмент/оператор (11%) – *Koopman operator* (оператор Купмана), название бренда (12%) – *Prometheus* (Прометей) и др., формирующие образ системы и задающие ожидания аудитории.

Расширение и сужение значения в терминообразовании ИИ выступают взаимодополняющими механизмами, опосредующими привлечение и транстерминологизацию и структурно объясняющими рост полисемии и конфигурации синонимии [Sager 1990; Cabré 1999]. Расширение увеличивает объем референции и наслаивает новые контекстные профили без устранения старых, порождая межрегистровую и межподдоменную полисемию и условную/аспектную синонимию. Так, *model* может

обозначать математическую модель, архитектуру нейросети и обученный экземпляр системы. Сужение, напротив, закрепляет вход в терминосистему, усиливает интенционал и снижает кросс-доменную полисемию, как в случае термина *token* (токен), сужающегося в NLP до «единицы субсловной сегментации» [Kageura 2002]. Метонимия, метафоризация и эпонимизация функционируют как дополнительные семантические механизмы, модулирующие эти базовые процессы, приводя, в частности, к появлению синонимических рядов типа *Bayesian network / belief network / probabilistic network* и многозначных терминов *model, training, artificial intelligence*. В социокогнитивном терминоведении такие эффекты рассматриваются как нормальное проявление динамичной терминосистемы и подлежат нормализации с опорой на корпусные процедуры (AntConc, KWIC, LL, MI) и онтологическое моделирование [Cabré 1999; Kageura 2002; L’Homme 2020]. Результаты анализа обобщены в семантической классификации способов терминообразования (см. Рисунок 4).



Примечание. ----- возможная внутридоменная синонимия и полисемия терминов ИИ.

Рисунок 4 – Семантическая классификация способов терминообразования в терминологии ИИ

В **Заключении** подводятся итоги исследования, обобщаются результаты и намечаются перспективы дальнейшего изучения данной проблематики.

В настоящей работе представлено комплексное описание структурно-семантических особенностей терминологии предметной области «искусственный интеллект», разработанное на основе корпусно-ориентированной методики. В рамках когнитивно-дискурсивного и корпусно-ориентированного подходов обоснована значимость применения воспроизводимых процедур анализа, основанных на сочетании статистических метрик, коллокационного профилирования, дефиниционно-компонентного описания и верификации по международным стандартам (ISO 704/1087, ISO/IEC 22989 и др.).

Предложенная методология позволила выявить ключевые закономерности терминообразования, включая преобладание многокомпонентных номинативных моделей, системную связь между структурной сложностью термина и концептуальной насыщенностью обозначаемого понятия, а также роль семантических и деривационных механизмов в формировании терминологического фонда. Установлено, что терминосистема ИИ обладает выраженной стратифицированной организацией, где ядро составляют устойчивые однословные единицы и классические конструкции A+N (*deep learning, artificial intelligence*) и N+N (*machine learning, neural network*), выступающие «строительными блоками» дискурса, а периферия представлена инновационными сложными терминами, отражающими динамику развития области. Особое внимание уделено вариативности и полисемии терминов, трактуемым как неизбежные, но управляемые явления, поддающиеся нормированию через корпусную диагностику, контекстную аннотацию и стандартизованный выбор предпочтительных наименований. Применение корпусного менеджера AntConc в сочетании со статистической триангуляцией и лингвистическим анализом позволило не только верифицировать репрезентативный массив из 1648 терминологических единиц, но и выявить продуктивные модели их функционирования в профессиональном дискурсе.

Перспективы дальнейшей работы связаны, на наш взгляд, с применением разработанного алгоритма для сравнительного анализа терминосистем смежных технологических областей, а также с интеграцией методов искусственного интеллекта для автоматического мониторинга терминологической динамики. Это позволит создать основу для прогнозирования развития профессиональной лексики и разработки эффективных инструментов терминологического менеджмента в условиях цифровой трансформации научного дискурса.

**Основные положения диссертации отражены в следующих публикациях автора:**

#### **Издания, рекомендованные ВАК РФ**

1. Шадрина, О. В. Структурно-семантические особенности эпонимов терминосистемы «Искусственный интеллект» в английском языке / О. В. Шадрина // Филология и культура. – 2024. – № 4(78). – С. 175-186. – DOI 10.26907/2782-4756-2024-78-4-175-186. (1,3 п.л.)
2. Шадрина, О. В. Лексические аспекты разработки онлайн-гlossария по искусственному интеллекту / О. В. Шадрина // Гуманитарные и социальные науки. – 2025. – Т. 108, № 1. – С. 105-114. – DOI 10.18522/2070-1403-2025-108-1-105-114. (0,85 п.л.)
3. Шадрина, О. В. Корпусный анализ репрезентации терминологии искусственного интеллекта в русском языке с использованием инструмента AntConc (на материале альманаха «Искусственный интеллект») / О. В. Шадрина, О. В. Маруневич // Научный диалог. – 2025. – № 14(7). – С. 133-160. – DOI 10.24224/2227-1295-2025-14-7-13-160. (1,7 п.л.)

4. Шадрина, О. В. Англоязычные термины-метафоры в описании популяционных алгоритмов искусственного интеллекта / О. В. Шадрина // Когнитивные исследования языка. – 2025. – № 2-1 (63). – С. 382-387. (0,3 п.л.)

#### **Сборники статей**

#### **и материалы научно-практических конференций**

5. Шадрина, О. В. История развития и основные этапы формирования терминологии ИИ / О. В. Шадрина // Современные тенденции развития науки, общества и образования : сборник материалов VIII Международной научно-практической конференции (Москва, 14 октября 2024 г.). – Москва : Академическая среда, 2024. – С. 225-229. (0,3 п.л.)
6. Шадрина, О. В. Термины-эпонимы англоязычной терминологии «искусственный интеллект» / О. В. Шадрина // Вызовы и тренды современной лингвистики : Казанский международный лингвистический саммит 2024 (Казань, 29-31 октября 2024 г.) : труды и материалы : в 3 т. / под общей редакцией И. Э. Ярмакеева, Ф. Х. Тарасовой. – Т. 1. – Казань : Отечество, 2024. – С. 261-264. (0,2 п.л.)
7. Shadrina, O. V. Bridge the communication gap with AI e-glossary / O. V. Shadrina // Искусственный интеллект и духовная культура : сборник материалов Международной научно-практической конференции (Москва, 30 октября 2024 г.). – Санкт-Петербург : Сциентиа, 2025. – С. 223-225. (0,25 п.л.)
8. Шадрина, О. В. Использование платформы AntConc для исследования модели корпусной репрезентации терминов искусственного интеллекта в русском языке (на основе анализа альманаха «Искусственный интеллект») / О. В. Шадрина // В многомерном пространстве современной лингвистики : сборник материалов Международной молодежной конференции (Москва, 10 декабря 2024 г.). – Москва : Р. Валент, 2025. – С. 87-90. (0,3 п.л.)
9. Шадрина, О. В. Обучение составлению онлайн-гlossария по искусственному интеллекту: этические нормы применения генеративных моделей / О. В. Шадрина // Педагогический дискурс: вызовы цифровизации и трансформация образовательных практик : материалы IV Всероссийской конференции (Москва, 3-5 марта 2025 г.). – Москва : Языки народов мира, 2025. – С. 403-409. (0,47 п.л.)