

доступности и предиктивной аналитике, не только повышают эффективность проектного управления, но и формируют условия для непрерывного профессионального образования лиц с особыми потребностями. Они позволяют таким пользователям развивать цифровые компетенции (digital-skills), участвовать в коллективной работе, осваивать современные инструменты и интегрироваться в профессиональные процессы на равных условиях [1].

Таким образом, предиктивная аналитика при корректной реализации позволяет превращать хаотичный проектный процесс в управляемую, предсказуемую и адаптивную систему, одновременно расширяя возможности профессионального роста и участия в социально-экономической жизни для лиц с особыми потребностями в условиях цифровой экономики.

Список использованных источников

1. Project Management Institute, Руководство к Своду знаний по управлению проектами (Руководство PMBOK®) – Седьмое издание // Project Management Institute. – 2021. – С. 370.
2. Shmueli G., Koppius O. R. Predictive analytics in information systems research / Shmueli G., Koppius O. R. // Mis Quarterly. – 2011. – С. 553-572.
3. UNESCO. A global framework of reference on digital literacy skills for indicator 4.4.2 [Электронный ресурс]. – Режим доступа: <https://unesdoc.unesco.org/ark:/48223/pf0000265403>. – Дата доступа: 10.11.2025.

УДК 004.8

Ю.Д. Лыкова

Белорусский государственный
технологический университет
Минск, Беларусь

ПРИЧИНЫ ГАЛЛЮЦИНАЦИЙ В СОВРЕМЕННЫХ ЯЗЫКОВЫХ МОДЕЛЯХ И МЕТОДЫ ИХ УСТРАНЕНИЯ С ПОМОЩЬЮ ФОРМАЛИЗОВАННЫХ МОДЕЛЕЙ ЗНАНИЙ

Аннотация. В работе рассматриваются причины галлюцинаций в современных больших языковых моделях и методы их снижения с использованием

семантических сетей и графов знаний. Показано, что интеграция ИИ с формализованными моделями знаний повышает точность и достоверность ответов на естественном языке.

Y.D. Lykova

Belarusian State Technological University
Minsk, Belarus

CAUSES OF HALLUCINATIONS IN MODERN LANGUAGE MODELS AND METHODS FOR THEIR ELIMINATION USING FORMALIZED KNOWLEDGE MODELS

Abstract. The paper explores the causes of hallucinations in modern large language models and methods to reduce them using semantic networks and knowledge graphs. Integration of AI with formal knowledge models improves the accuracy and reliability of natural language responses.

Современные большие языковые модели (LLM – *Large Language Model*) представляют собой статистические модели, обучающиеся на огромных корпусах текстов. Принцип их работы основан на вероятностном прогнозировании следующего токена на основе предыдущего контекста. Такой подход обеспечивает высокую языковую плавность, но не гарантирует соответствие генерируемого текста объективным фактам. Несмотря на использование в обучении колоссальных массивов данных, модели не обладают формализованным механизмом проверки истинности или логической непротиворечивости результата. Именно поэтому возникает феномен галлюцинаций – ситуаций, в которых модель уверенно создаёт ложные ответы, не имеющие связи с реальностью.

Причины таких ошибок лежат в природе архитектуры трансформеров. Память модели распределена по весам нейронной сети и представлена в виде статистических связей между токенами, а не в виде структурированных правил или онтологических зависимостей. Для LLM знание – это не факт, а распределение вероятностей. Если в данных встречались противоречивые формулировки или редкие события, модель может вывести усреднённый или вообще несуществующий факт. В отличие от формальных систем знаний, трансформер не различает логическую истинность и правдоподобность. Дополнительным фактором является отсутствие устойчивого механизма обращения к внешним базам знаний: всё, что

модель хранит, – это компрессия обучающего датасета внутри параметров.

Выходом из этой ситуации становится интеграция языковых моделей с формализованными моделями знаний. Семантические сети, графы знаний и онтологические модели предоставляют явное представление сущностей, их свойств и отношений. В отличие от LLM, такие структуры основаны на строгих правилах формальной логики, что позволяет обеспечивать проверку фактов и концептуальную непротиворечивость. Наиболее перспективным направлением является создание гибридных систем, в которых генеративная модель отвечает за понимание естественного языка, а граф знаний – за верификацию и уточнение содержания ответа.

*В современных интеллектуальных системах, относящихся к сфере искусственного интеллекта (ИИ), подобные гибриды уже начинают внедряться. Одним из наиболее известных примеров является система **Google Assistant**, использующая **Google Knowledge Graph** – крупнейший в мире граф знаний, содержащий более 500 млрд фактов. При формировании ответа на запрос пользователей, система ИИ сначала анализирует семантику запроса с помощью языковой модели, а затем проверяет факты и уточняет данные через Knowledge Graph. Такой подход позволяет значительно снизить вероятность ошибок и обеспечивает фактически гарантированную достоверность при работе со структурированной информацией. Аналогичные решения применяет и Microsoft Bing Chat, который объединяет LLM с Microsoft Satori Graph.*

Одним из ключевых элементов такой архитектуры является алгоритм автоматической генерации формализованных запросов к графу знаний. Языковая модель способна преобразовать пользовательский запрос в структурированную форму, например SPARQL или внутренний формат доступа к базе знаний. Для этого разрабатываются специальные математические модели, описывающие преобразование естественно-языковых конструкций в семантические структуры. Модель должна учитывать морфологические особенности языка, синтаксические зависимости и семантические роли слов в предложении. Наиболее эффективным подходом является построение промежуточного представления в виде семантического графа, который затем сопоставляется с онтологией предметной области.

Использование таких моделей знаний позволяет значительно снизить частоту галлюцинаций. Пока LLM формирует гипотезу ответа, она может сверяться с графиком знаний, уточнять значения сущностей и проверять фактическую корректность данных. В ряде современных

систем применяется механизм Retrieval-Augmented Generation (RAG), в котором языковая модель получает доступ к внешним документам. Однако тексты не являются формальными структурами, поэтому RAG не обеспечивает полной защиты от ошибок. В отличие от него, интеграция с онтологиями позволяет применять логический вывод, выявлять противоречия и давать строго проверенные ответы.

Показателен пример системы IBM Watson Discovery, которая используется в фармацевтической и юридической практике. Watson сочетает семантические модели с онтологиями предметной области и проверяет ответы через специализированные графы знаний. Это позволяет существенно сократить количество неверных интерпретаций медицинских терминов и юридических формулировок – областей, где ошибка ИИ недопустима. Такие решения стали одними из первых коммерческих примеров интеграции LLM с формальным знанием, подтверждая эффективность гибридного подхода.

Экспериментальные исследования показывают, что гибридная архитектура особенно эффективна для областей с высокой ценой ошибки – медицины, юриспруденции, технической документации и академической справочной информации. В системах, использующих графы знаний для проверки ответов, уровень галлюцинаций снижается в 2-4 раза. Дополнительным преимуществом является интерпретируемость: каждая часть ответа может быть сопоставлена с конкретной вершиной или отношением графа, что повышает доверие пользователей к системе.

Существенное внимание уделяется также разработке алгоритмов, позволяющих автоматически уточнять запросы пользователя. Такие алгоритмы анализируют структуру исходной фразы, устраниют неоднозначность, дополняют недостающие элементы и формируют корректные структурированные запросы к базе знаний. Пример запроса пользователя: «Кто изобрел телефон?» должен быть преобразован в конкретную форму, учитывая исторические данные, а также различия между разными этапами развития технологии. Генерация таких запросов представляет собой математическую задачу сопоставления лексико-синтаксических конструкций с формальными концептами онтологии.

Таким образом, переход от чисто статистических моделей к гибридным интеллектуальным системам является важнейшим направлением развития искусственного интеллекта. Эти системы объединяют плавность генерации LLM и точность логических моделей. Формальные модели знаний могут компенсировать фундаментальные ограничения нейросетей, обеспечивая достоверность и устойчивость к

ошибкам. В перспективе такие решения открывают возможность построения полноценных естественно-языковых интерфейсов нового поколения, где пользователи смогут получать точные и проверенные ответы, не сталкиваясь с искажением информации.

Список использованных источников

1. Баланов А. Н. – Искусственный интеллект. Понимание, применение и перспективы. – Санкт-Петербург: Лань, 2024.
2. Машнин Т. – Машинное обучение и искусственный интеллект – Москва: Литрес, 2022.
3. Агарвал О., Ге Х., Шакери С., Аль-Рфу Р. – Генерация синтетического корпуса на основе графа знаний для предварительного обучения языковой модели, обогащённой знаниями. – NAACL, 2021.
4. Мачадо М., Феррейра Лима Г.А., Соарес Э.Ф. – Расширяемый подход для пополнения мультимодального графа знаний на основе запросов – IBM Research, ISWC 2022.

УДК 37.015.31:502

А.А. Мазайхина, Т.А. Юрения, Ю.В. Чеботкова
Белорусский государственный технологический университет
Минск, Беларусь

ИГРА «СКАРБНІЧКА ВЕДАЎ» КАК ИНСТРУМЕНТ ЭКОЛОГИЧЕСКОГО ПРОСВЕЩЕНИЯ

Аннотация. В работе представлена настольная игра «Скарбнічка ведаў», направленная на экологическое просвещение школьников и студентов. Описаны цели, задачи, новизна и актуальность игры. Приведены правила и механизм игрового процесса, способствующие запоминанию информации о редких видах растений и животных. Игра сочетает в себе элементы викторины, стратегии и запоминания, что делает процесс обучения увлекательным и эффективным.

A.A. Mazaikhina, T.A. Urenya, Y.V. Chebotkova
Belarusian State Technological University
Minsk, Belarus