

сти различных языковых моделей при генерации структурированных запросов для промышленного планирования. Экспериментальное тестирование на наборе сценариев управления, включая проверку технологических аксиом и расчет показателей валового производства, продемонстрировало, что использование стратегии обогащения с помощью протокола Model Context и интерфейса GraphQL значительно снижает уровень семантических галлюцинаций по сравнению со стандартными методами генерации подсказок. Модели демонстрируют более высокую устойчивость к сложным синтаксическим конструкциям и точное соблюдение типов данных, что критически важно при работе с финансовыми и количественными показателями производственных цепочек. Наличие жесткого математического контекста в системных инструкциях позволяет системе не только извлекать данные, но и выполнять предварительную логическую проверку, блокируя вывод заведомо ошибочных результатов, противоречащих физическим законам функционирования кластера.

ЛИТЕРАТУРА

1. Универсальность естественно языкового интерфейса для взаимодействия с различными типами программных интерфейсов приложений / Евченко И.В. // <https://cyberleninka.ru/article/n/universalnost-estestvenno-yazykovogo-interfeysa-dlya-vzaimodeystviya-s-razlichnymi-tipami-programmnyh-interfeysov-prilozheniy>. Дата доступа: 09.01.2026.

2. Естественно-языковые интерфейсы интеллектуальных вопросно-ответных систем / Житко В.А., Вяльцев В.Н., Гецевич Ю.С., Кузьмин А.А. // <https://cyberleninka.ru/article/n/estestvenno-yazykovye-interfeysy-intellektualnyh-voprosno-otvetnyh-sistem>. Дата доступа: 11.01.2026.

УДК 004.8

Н.И. Белодед, доц., канд. техн. наук;
Ю.Д. Лыкова, преп.-стажер
(БГТУ, г. Минск)

МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ ДЛЯ ЕСТЕСТВЕННО ЯЗЫКОВОГО ИНТЕРФЕЙСА

Развитие современных интеллектуальных информационных систем неразрывно связано с совершенствованием способов взаимодействия человека и вычислительной техники. Естественно-языковые интерфейсы являются приоритетным направлением, так как позволяют минимизировать порог вхождения для конечного пользователя и

обеспечить получение структурированной информации в доступной форме. Основу функционирования таких систем составляет модель представления знаний, которая определяет не только структуру хранения фактов, но и логику их извлечения при обработке произвольных запросов. Традиционно в интеллектуальных системах применялись логические модели и фреймовые структуры, обеспечивающие единое описание свойств объектов и ситуаций [1]. Однако с появлением больших языковых моделей возникла необходимость в интеграции классических методов хранения данных с современными нейросетевыми технологиями обработки текста.

Современные нейросетевые модели обладают значительной лингвистической гибкостью, что позволяет им имитировать человеческий диалог и адаптировать стиль общения. Тем не менее, их практическое применение в экспертных областях – таких как юриспруденция, медицина или корпоративное управление – ограничено критическим риском генерации фактически недостоверной информации, известным как эффект галлюцинации. Для решения данной проблемы в работе рассматривается архитектура, основанная на методе дополненной генерации. Данный подход позволяет разграничить функции лингвистического синтеза и хранения верифицированных данных. Модель представления знаний в рамках такой архитектуры строится на четком разделении между вычислительным ядром и контролируемым вектором внешних знаний. Информация предметной области формализуется в виде структурированных объектов, что формирует жесткий базис достоверных фактов и позволяет актуализировать базу знаний без внесения дорогостоящих изменений в параметры самой нейронной сети.

Важным аспектом модели является использование семантических сетей и графов знаний, которые позволяют формально описывать структуру предметной области и навигацию по связям между понятиями [2]. В данной работе семантическая сеть интегрируется в векторное пространство, где каждый узел данных получает свое уникальное числовое представление. Это позволяет объединить классическую строгость структурированных данных с мощностью семантического анализа, обеспечивая высокую точность сопоставления запроса пользователя с имеющимися фактами. В отличие от простых поисковых систем, такой подход позволяет учитывать иерархию понятий и сложные отношения между объектами предметной области [3].

Связующим звеном между пользовательским запросом на естественном языке и формализованной базой фактов выступает слой векторизации. Процесс векторизации преобразует неструктурированный

текст в многомерные числовые массивы, которые фиксируют семантические зависимости и контекстуальные особенности слов. В результате каждый текстовый объект – будь то вопрос пользователя или описание факта в базе знаний – получает координаты в семантическом пространстве. Поиск релевантных сведений осуществляется на основе метода косинусного сходства, который признан одной из наиболее эффективных метрик для определения степени близости между текстовыми объектами в задачах обработки естественного языка.

Математический смысл метода косинусного сходства заключается в вычислении косинуса угла между вектором поискового запроса и векторами информационных объектов, хранящихся в базе. В отличие от измерения евклидова расстояния, данный метод фокусируется на направлении векторов, что позволяет нивелировать разницу в длине текстов и сосредоточиться на их смысловой идентичности. Минимизация угла между векторами свидетельствует о максимальном семантическом соответствии, что дает системе возможность эффективно распознавать синонимичные конструкции и сложные речевые обороты, не ограничиваясь простым поиском буквальных совпадений символьных строк. Это особенно важно при работе с технической или нормативной документацией, где одно и то же понятие может быть сформулировано различными способами.

Для обеспечения абсолютной достоверности формируемых ответов в алгоритм работы системы включена процедура обязательной логической фильтрации. На этапе сопоставления векторов вычисляется коэффициент релевантности: если максимальное значение косинусного сходства оказывается ниже установленного порогового значения, запрос классифицируется как выходящий за рамки имеющейся базы знаний. В такой ситуации процесс генерации свободного текста принудительно прерывается программным фильтром. Вместо формирования вероятностного ответа система генерирует стандартное сообщение об отсутствии подтвержденных данных и предлагает пользователю скорректировать параметры запроса. Данный механизм является ключевым инструментом верификации, предотвращающим подмену реальных фактов ложными утверждениями языковой модели.

Дополнительным преимуществом предлагаемой архитектуры является возможность полной локализации вычислительного процесса. Использование современных платформ для запуска языковых моделей и специализированных векторных баз данных позволяет обрабатывать конфиденциальную информацию внутри защищенного контура организации. Это исключает необходимость передачи данных на внешние облачные серверы, обеспечивая соблюдение требований ин-

формационной безопасности. Таким образом, интеграция структурированных моделей представления знаний, векторного поиска на основе косинусного сходства и механизмов дополненной генерации позволяет создавать масштабируемые и безопасные интеллектуальные интерфейсы, пригодные для эксплуатации в профессиональных экспертных системах.

ЛИТЕРАТУРА

1. Методы представления знаний и фреймовые структуры: лекция. URL: http://www.machinelearning.ru/wiki/images/1/1d/Lect_8_ai_md_v.pdf (дата обращения: 25.11.2025).
2. Семантические сети и обработка естественного языка; Knowledge Graph. URL: <https://www.osp.ru/os/2017/02/13052229> (дата обращения: 11.12.2025).
3. Методы представления информации в простых семантических сетях; классификация семантических сетей / О. В. Мосин [и др.] // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20, № 3. С. 382–393. URL: <https://cyberleninka.ru/article/n/metody-predstavleniya-informatsii-v-prostyh-semanticheskikh-setyah> (дата обращения: 16.12.2025). DOI: 10.17586/2226-1494-2020-20-3-382-393.

УДК 004.77; 004.72

Е.А. Гончар ассист.
(БГТУ, г. Минск)

ОБЗОР ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ СИМУЛЯЦИИ NDN СЕТЕЙ

NDN (Named Data Networking) [1–5] – это архитектура сети, которая ориентирована на данные, а не на их местоположение. В отличие от традиционных IP-сетей, где коммуникация основана на адресах устройств (IP-адресах), в NDN фокус смещается на сами данные и их имена. Это позволяет создавать более гибкие, безопасные и эффективные сети, особенно в условиях современного интернета, где запросы на контент играют ключевую роль.

Основные цели NDN:

1. Ориентация на контент: В NDN пользователи запрашивают данные по их именам, а не по адресам устройств. Это упрощает доступ к информации, особенно в условиях, когда данные могут быть доступны из множества источников.