



Сопоставление номенклатур товаров ресторанов и поставщиков с помощью LLM – Case Study для ресторанных холдинга

С. Джин, ORCID: 0000-0002-8242-6157 <sedzin@hse.ru>

П.Б. Панфилов, ORCID: 0000-0001-6567-6309 <ppanfilov@hse.ru>

А.С. Сулейкин, ORCID: 0000-0003-2294-6449 <aless.sull@mail.ru>

Национальный исследовательский университет «Высшая школа экономики»,
Россия, 101000, г. Москва, ул. Мясницкая, д. 20.

Аннотация. В современном ресторанном бизнесе точное сопоставление номенклатуры продуктов между ресторанами и поставщиками является критически важной задачей. Эффективное управление запасами и оптимизация закупок напрямую влияют на прибыльность бизнеса. С ростом числа поставщиков и ассортимента продукции традиционные методы сопоставления становятся менее эффективными. В данном исследовании предлагается использовать большие языковые модели (LLM) для автоматизации и повышения точности сопоставления продуктов. В рамках пилотного проекта для ресторанных холдинга мы протестировали пять групп продуктов (креветки, угорь, сыр пармезан, творог, сливочное масло), достигнув средней точности тестирования 83,8%. Архитектура решения использует быстрое проектирование, платформы с низким уровнем кода, такие как Flowise, и интеграцию с Telegram для удобной обработки данных. Ключевые проблемы, включая семантическую неоднозначность и галлюцинации моделей, были решены с помощью предметно-ориентированных словарей и валидации. Такой подход сокращает ручную работу примерно на 90%, что позволяет создавать масштабируемые решения для цепочки поставок, применимые не только в ресторанах, но и в розничной торговле и электронной коммерции.

Ключевые слова: большие языковые модели; управление цепями поставок; сопоставление номенклатуры продуктов; автоматизация; оптимизация запасов.

Для цитирования: Джин С., Панфилов П.Б., Сулейкин А.С. Сопоставление номенклатур товаров ресторанов и поставщиков с помощью LLM – Case Study для ресторанных холдинга. Труды ИСП РАН, том 37, вып. 6, часть 3, 2025 г., стр. 163–176. DOI: 10.15514/ISPRAS-2025-37(6)-43.

Благодарности: Работа выполнена в рамках программы фундаментальных исследований НИУ ВШЭ.

Mapping Restaurant and Supplier Product Nomenclatures Using LLM – Case Study for a Restaurant Holding

S. Jin, ORCID: 0000-0002-8242-6157 <sedzin@hse.ru>

P.B. Panfilov, ORCID: 0000-0001-6567-6309 <ppanfilov@hse.ru>

A.S. Suleykin, ORCID: 0000-0003-2294-6449 <aless.sull@mail.ru>

*National Research University Higher School of Economics,
20, Myasnitskaya st., Moscow, 101000, Russia.*

Abstract. In the modern restaurant business, accurate mapping of product nomenclatures between restaurants and suppliers is a critical task. Effective inventory management and procurement optimization directly impact business profitability. With the increase in suppliers and product variety, traditional mapping methods become less efficient. This study proposes using large language models (LLM) to automate and improve the accuracy of product matching. Through a pilot project for a restaurant holding, we tested five product groups (shrimp, eel, parmesan cheese, cottage cheese, butter), achieving an average test accuracy of 83.8%. The solution architecture leverages prompt engineering, low-code platforms like Flowise, and Telegram integration for user-friendly processing. Key challenges, including semantic ambiguity and model hallucinations, were addressed via domain-specific dictionaries and validation. This approach reduces manual effort by approximately 90%, enabling scalable supply chain solutions applicable beyond restaurants to retail and e-commerce.

Keywords: Large Language Models; Supply Chain Management; Product Mapping; Automation; Inventory Optimization.

For citation: Jin S., Panfilov P.B., Suleykin A.S. Mapping Restaurant and Supplier Product Nomenclatures Using LLM – Case Study for a Restaurant Holding. *Trudy ISP RAN/Proc. ISP RAS*, vol. 37, issue 6, part 3, 2025, pp. 163-176 (in Russian). DOI: 10.15514/ISPRAS-2025-37(6)-43.

Acknowledgements. The work was carried out within the framework of the fundamental research program of the National Research University Higher School of Economics.

1. Введение

1.1 Эволюция применения LLM в управлении цепочками поставок

Большие языковые модели (Large Language Model, LLM) стали революционным инструментом в управлении цепочками поставок (Supply Chains Management, SCM), позволяя бизнесу оптимизировать операции, улучшать принятие решений и автоматизировать процессы. Традиционные подходы, такие как системы на основе правил или моделей машинного обучения (Machine Learning, ML), часто не справляются с обработкой неструктурированных данных, многозычных контекстов и быстро меняющихся сред. Модели LLM преодолевают эти ограничения за счет расширенных возможностей понимания естественного языка и контекстного анализа, что делает их высоко адаптивными к динамичным потребностям цепочек поставок [1-2].

1.1.1 Прогнозирование спроса и оптимизация запасов

Модели LLM продемонстрировали значительный потенциал в прогнозировании спроса за счет анализа исторических данных о продажах, рыночных тенденций и внешних факторов, таких как погода или геополитические события. Например, компания Coca-Cola использует прогнозирование на основе искусственного интеллекта (ИИ) для оптимизации графиков производства и уровней запасов, сокращая перепроизводство при удовлетворении спроса клиентов [1]. Аналитика на основе модели LLM ChatGPT применяется в розничной торговле для выявления тенденций спроса и улучшения управления запасами, что приводит к сокращению дефицита и повышению операционной эффективности [3].

1.1.2 Управление взаимоотношениями с поставщиками

Модели LLM улучшают управление взаимоотношениями с поставщиками, автоматизируя такие задачи, как составление контрактов, обработка счетов и отслеживание эффективности. Например, платформа Verusen на основе ИИ улучшает взаимодействие между покупателями и поставщиками за счет повышения качества данных и содействия возможностям вверх и вниз по цепочке [4]. Использование моделей LLM компанией Microsoft в управлении цепочками поставок показало значительное сокращение времени подключения поставщиков и мониторинга соблюдения контрактов [5].

1.1.3 Оптимизация логистики

В логистике модели LLM играют ключевую роль в оптимизации маршрутов и управлении сбоями. Компания Amazon использует ИИ для анализа моделей покупок и сезонных тенденций, чтобы заранее корректировать распределение запасов по своей сети. Это приводит к сокращению сроков доставки и повышению общей эффективности [2]. Компания DHL интегрирует робототехнику на основе ИИ в складские операции для задач сортировки и упаковки, значительно сокращая время обработки и позволяя сотрудникам сосредоточиться на более сложных задачах [1].

1.1.4 Интеграция знаний для поддержки решений

Большие языковые модели LLM способствуют улучшению принятия решений за счет интеграции разрозненных источников данных в связные графы знаний. Эти графы фиксируют сложные взаимозависимости между объектами цепочки поставок, обеспечивая получение информации о рисках и возможностях в реальном времени. Например, сервис Microsoft Dynamics 365 Copilot использует службы Azure OpenAI для прогнозирования сбоев у поставщиков или в регионах с одновременной генерацией практических рекомендаций для менеджеров цепочек поставок [5].

1.2 Ключевые проблемы использования LLM

Большие языковые модели LLM обладают преобразующим потенциалом для управления цепочками поставок, но их внедрение сопряжено с трудностями. В этом разделе рассматриваются основные препятствия, возникающие при внедрении LLM в контексте цепочки поставок в ресторанный бизнес, с акцентом на семантическую неоднозначность, риски галлюцинаций, уязвимости системы безопасности и адаптацию персонала.

1.2.1 Семантическая неоднозначность

Серьезной проблемой, с которой, в частности, столкнулись разработчики в ходе реализации и тестирования системы, описываемой в данной работе, является том, что модели LLM ошибаются в специфических контекстах (например, при сопоставлении продуктов “Easy Peel Shrimp” и “Pre-Cleaned Shrimp”). Эта проблема усугубляется еще больше в многоязычной среде [7-8].

Следствием проявления этой проблемы являются 23% ошибок сопоставления в тестах [1]. Предлагаемые возможные решения для этой проблемы включают в себя:

- Предметные онтологии (5k+ терминов), которые помогают снизить число ошибок на 19% [1];
- Ручная валидация спорных случаев [5].

1.2.2 "Галлюцинации" и объяснимость

Проблема «галлюцинаций» моделей LLM заключается в генерация моделью генеративного ИИ ложных данных из-за вероятностной природы работы модели LLM [8-9].

Практическими последствиями проявления «галлюцинаций» LLM в задачах управления цепочками поставок могут стать, например, ошибочные рекомендации по поставщикам.

Предлагаемые возможные решения для проблемы «галлюцинаций» LLM:

- Обучение моделей LLM на предметно-ориентированных (domain-specific) данных для снижения «галлюцинаций» LLM на 15% [8];
- Резервные механизмы (fallback) для неподдерживаемых запросов [5].

1.2.3 Уязвимости безопасности

Распространение решений на базе моделей машинного обучения вызвало к жизни и появление нового типа атак на сервисы и приложения на основе машинного обучения, такие как «вредоносное машинное обучение» (Adversarial Machine Learning, AML), одним из проявлений которого является отравление данных (poisoning attack) или манипуляции с обучающими данными (пример: утечка данных OpenAI в 2023 году) [11-12]. Также цепочки поставок могут страдать и от атак, связанных с распространением уязвимостей через зависимости (подробнее можно ознакомиться, например, в [13]).

Решения по борьбе с атаками отравления и уязвимостями зависимостей включают в себя:

- Дифференциальная приватность при обучении [11];
- Регулярные аудиты безопасности обеспечивают снижение уязвимостей на 30% [12].

1.2.4 Адаптация персонала

Проблема готовности персонала к использованию и доверия в отношении решений на основе генеративного ИИ обычно выражается напрямую в затратах на переподготовку сотрудников компаний (\$3800/сотрудника) [7] и снижении продуктивности на 22% [5, 9].

Решается эта проблема обычно двумя путями:

- Поэтапное обучение сотрудников (формулировка запросов + таксономия) для увеличения эффективности на +35% [5];
- Сотрудничество ИИ-специалистов и экспертов предметной области [10].

1.3 Критические пробелы в существующих исследованиях

Несмотря на явный преобразующий потенциал, внедрение решений на базе моделей LLM сталкивается с рядом ограничений, которые необходимо устраниć для более широкого использования генеративного ИИ в управлении цепями поставок SCM (табл. 1).

Серьёзной проблемой применения моделей LLM остается «семантический дрейф». Например, при классификации продуктов в случае рассматриваемого нами ресторанный бизнес-кейса несогласованная терминология вроде "Jumbo Shrimp" вместо "Colossal Shrimp" составляет ~34% ошибок [6].

Кроме того, затраты на переподготовку персонала для внедрения моделей LLM и решений на их основе значительны. Так по данным McKinsey, средняя стоимость переподготовки сотрудников закупочных подразделений для эффективного использования инструментов ИИ составляет \$3800 на сотрудника [7].

Табл. 1. Ограничения применения различных подходов.

Table 1. Limitations of different approaches.

Проблема	Традиционные подходы	Ограничения LLM
Адаптация к предметной области	Требует размеченных данных	Ограниченнная точность «с нуля» (~82%) [2]
Доверие к операциям	Ручная проверка (5-7 дней)	«Черный ящик» логики решений [6]
Многоязычное выравнивание	Перевод на основе правил (~68% покрытия)	Уровень ошибок ~19% (на азиатских рынках [1])

1.4 Постановка задачи разработки и описание исследования

В данной работе представлен пример решения одной из центральных задач в управлении цепочками поставок, а именно: задачи установления корректного соответствия между товарной номенклатурой поставщиков и потребителей, которая напрямую связана с задачами корректного планирования запасов, управления взаимоотношениями с поставщиками, автоматизации таких бизнес-процессов, как составление контрактов, обработка счетов и отслеживание эффективности. Данная задача связана с рутинной обработкой больших объемов текстовых данных, что идеально подходит к ее автоматизации и реализации на основе решений генеративного ИИ. В работе рассмотрены архитектурные и технологические особенности предлагаемого решения прикладной задачи и проанализированы результаты опытной эксплуатации разработанного пилотного варианта системы, протестированного на реальных данных бизнеса ресторанных холдинга.

Структура работы включает в себя введение, разделы с описанием предлагаемого подхода к построению решения, результатов реализации и тестирования прототипа системы, обсуждения результатов опытной эксплуатации системы и заключения. Во введении рассматриваются возможности, которые технологии генеративного ИИ предлагают для автоматизации и улучшения управления цепочками поставок SCM, а также те проблемы и препятствия, с которыми сталкиваются разработчики решений на основе моделей LLM, на основе анализа примеров решения практических задач управления цепочками поставок. В разделе с описанием подхода к решению практической задачи представлена архитектура решения на основе собственных цифровых платформ и технологий с открытым кодом. Результаты прогонов прототипа системы на данных реального бизнес-кейса показаны в разделе, посвященном результатам практического внедрения системы с представлением и анализом количественных показателей функционирования системы в процессе ее опытной эксплуатации. В заключении подведены итоги реализации проекта и обсуждаются перспективы масштабирования решения, а также возможности адаптации и применения предлагаемого подхода к реализации аналогичных проектов как в рамках задач управления цепочками поставок, так и за пределами данной прикладной области.

2. Подход к решению практической задачи

2.1 Описание бизнес-кейса

Наша команда разработала пилотный проект для крупного ресторанных холдинга с целью автоматизации процесса сопоставления товарных номенклатур.

2.1.1 Описание целевой задачи

Для заданного наименования товара от заказчика (ресторана) необходимо выполнить автоматический поиск аналогов в общей базе товаров поставщиков.

Отбор аналогов должен осуществляться по заданным критериям разделения товаров из базы на три категории:

- Категория 1: точно совпадающие позиции;
- Категория 2: потенциально подходящие позиции;
- Категория 3: неподходящие позиции.

Система должна генерировать собственные критерии категоризации на основе полученного наименования товара и правил, учитывающих товарную группу и требования ресторана.

В рамках пилотной разработки были введены ограничения для реализации ранжирования товарных номенклатур по критериям от заказчика для пяти товарных групп: креветки, уголь, сыр пармезан, творог и сливочное масло. В табл. 2 представлен пример заданных критерий для категоризации номенклатур в списке креветок.

Табл. 2. Пример критерий отнесения номенклатур списка креветок к одной из трёх категорий.
Table 2. An example of criteria for classifying shrimp list items into one of three categories.

Категория	Критерии
Категория 1 (точно совпадающие позиции) выделено зеленым	Бренд Parmenter
Категория 2 (потенциально подходящие позиции) выделено желтым	<ul style="list-style-type: none">• Все позиции без слов «мягкий сыр» в названии;• Все позиции без упоминания стран (Швейцария, Италия);• Все позиции без слова «голова» в названии ИЛИ весом $\geq 4,5$ кг• Все позиции без слов «Premium» / «Exclusive» в названии
Категория 3 (неподходящие позиции) выделено красным	Все остальные позиции

2.2 Архитектура решения на базе DUC SmartSearch и модели LLM

Для разработки использовалась собственное цифровое решение DUC SmartSearch [14] на базе Danswer/Flowise, где DUC SmartSearch и Danswer (ныне Onyx [15]) – это платформы для создания ИИ-ассистентов и RAG-конвейеров, а Flowise – это платформа с минимальным ручным кодированием (low-code) для создания ИИ-ассистентов [16]. Весь конвейер обработки данных был реализован на Flowise. Альтернативно решение может быть разработано на других аналогичных (low-code) платформах (например, n8n, Langflow) или с использованием библиотек Python (Langchain и др.) с открытым исходным кодом. В процессе запуска и выполнения конвейеров обработки данных система DUC SmartSearch использует модель LLM GigaChat Max [17]. Применение других моделей LLM для решения прикладной задачи также возможно, хотя для этого потребуется дополнительная оптимизация текстовых запросов (prompt engineering).

2.2.1 Конвейер обработки данных

Прототип системы ранжирования списка продуктов строился по схеме обработки данных, как это представлено на рис. 1.

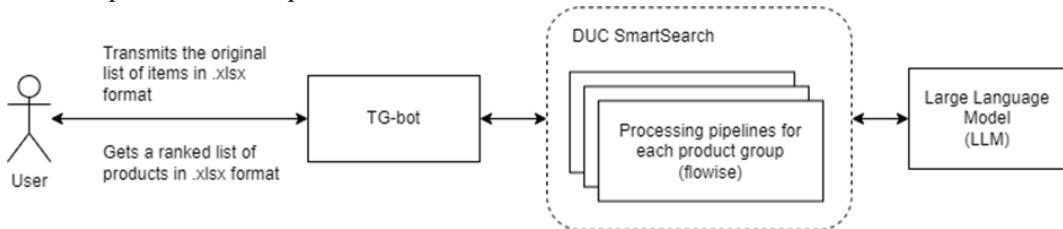


Рис. 1. Прототип системы ранжирования списка продуктов. Схема высокогоуровневой архитектуры.

Fig. 1. Product List Ranking System Pilot. High-Level Architecture Diagram.

Пользовательский интерфейс системы реализован через Telegram. Здесь TG-бот принимает от пользователя исходные списки товаров в виде файлов Excel (в формате .xlsx) для последующей обработки. При этом список номенклатур должен находиться на первом листе, в первом столбце, начиная с первой строки.

Далее списки товаров поступают на свои конвейеры обработки в системе DUC SmartSearch, где происходит их классификация, в соответствии с укрупненным алгоритмом, представленным на рис. 3.

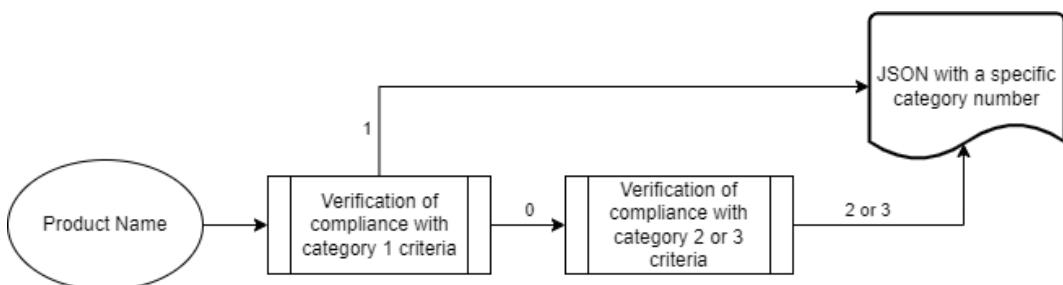


Рис. 3. Общий алгоритм классификации наименований товаров в конвейерах обработки.

Fig. 3. General algorithm for classifying product names in processing pipelines.

В процессе запуска и выполнения конвейеров система DUC SmartSearch использует поддержку со стороны LLM для реализации двухэтапной проверки соответствия критериям отнесения номенклатур списка товаров к одной из трёх категорий. Табл. 3 показывает запросы, которые используются системой для активации модели LLM GigaChat.

2.2.2 Результат обработки данных

В результате был разработан прототип системы, который получает на вход от пользователя список номенклатур определённой товарной группы и название этой группы, обрабатывает список по заданным критериям для каждой группы и предоставляет на выходе ранжированный список с указанием определённой категории в каждой строке.

Входной интерфейс системы реализован через мессенджер Telegram, где пользователь загружает Excel-файл со списком продуктов. Далее специально разработанный чат-бот обрабатывает этот Excel-файл, и результатом работы системы является ранжированный список всех продуктов-аналогов из списка поставщиков. Для реализации «умной обработки» списков поставщиков используется модель LLM GigaChat Max, которая обеспечивает достаточно высокую точность обработки списков, благодаря специально настроенному под прикладную задачу промт-инжинирингу, как это представлено в табл. 3.

Табл. 3. Пример запросов для двухэтапной проверки соответствия критериям.

Table 3. Example prompts for 2-step mapping verification.

Этап	Запрос
Проверка соответствия Категории 1	<p>Роль: Вы – цифровой ассистент менеджера по снабжению ресторанных холдинга. Вы получаете название продукта (креветки), которое нужно проверить по критериям и вывести результат.</p> <p>Словарь: s/m – свежемороженые; s/hb – с хвостом; b/g – без головы.</p> <p>Задача: Выведите 1, если название соответствует ВСЕМ критериям: 1) размер из списка ["21/25", "26/30"] + 2) текст из списка ["очищенные", "чищены", "чищен."] + указан 3) Текст из списка ["с хвостом", "на хвосте", "s/hb", "s/hb", "n/hb", "с хв."] указан. Критерии могут быть в любом порядке. ВАЖНО: название должно соответствовать ВСЕМ критериям. Выведите 0, если название не соответствует ХОТЯ БЫ ОДНОМУ критерию. Если продукт не относится к креветкам – выведите пустую строку "". Не выдумывайте.</p> <p>Название продукта: {input}</p>
Проверка соответствия Категории 2 или 3	<p>Роль: Вы – цифровой ассистент ресторанных холдингов.</p> <p>Вы получаете название креветок, которое нужно проверить по критериям и вывести результат.</p> <p>Словарь: s/m – свежемороженые; s/hb – с хвостом; b/c – без головы; b/g – без головы.</p> <p>Задача: Выведите 2, если название соответствует ВСЕМ критериям: 1) размер из списка ["16/20", "21/25", "26/30"] + указан 2) текст из списка "очищенные с хвостом", "чищены", "чищен.", "чищенные, с хв.", "очищенные с хв.", "чищенные с хв.", "без головы", "b/g."] + 3) без текста из списка "в (кокосовой) панировке", "w/o", "в панцире", "панцирь", "пресноводные". Выведите 3, если название не соответствует ХОТЯ БЫ ОДНОМУ критерию. Если продукт не относится к креветкам – выведите пустую строку "". Не выдумывайте.</p> <p>Название продукта: {input}</p>

Результат обработки списка товаров (рис. 4) демонстрирует ранжированный список с цветовой разметкой. Исходные данные идентичны показанным на рис. 4, но во входном файле позиции были представлены без цветов и в произвольном порядке. Разработанная система-прототип добавляет номер категории (1, 2 или 3) для каждой номенклатуры во второй столбец файла. Дополнительно список продуктов сортируется по возрастанию категории с цветовой разметкой «светофор»:

- Категория 1: зеленый;
- Категория 2: желтый;
- Категория 3: красный.

	A	B	C	D
1	Shrimp (vannamei) 21/25 peeled n/t	1	1	
2	Shrimp Wanamei peeled s/m 21/25 with tail	1	1	
3	Shrimp Wanamei 26/30/30 peeled with tail	1	1	
4	Shrimp peeled, with tail. c/m. 26/30 1kg	1	1	
5	Shrimp (Litopenaeus vannamei) peeled with tail (21/25) c/m 1kg – India	1	1	
6	Shrimp b/g c/m peeled, with tail 26/30 (1,000 kg) 30% cr. 10 pcs. India	1	1	
7	Shrimp 16/20 peeled with tail c/m India	2	2	
8	Shrimp 21/25 b/g c/m, block 1.8 kg, 1/6, pcs.	2	2	
9	Shrimp 21/25 b/g c/m, block 1.8 kg, 1/6, pcs.	2	2	
10	Shrimp 21/25 b/g c/m, block 1.8 kg, 1/6, pc	2	2	
11	Shrimp Vanamei in breading, peeled s/m 21/25. fac.1 kg	3	3	
12	Shrimp Wanamei peeled s/m 21/25 without tail	3	3	
13	Shrimp Wanamei s/m b/g 21/25 eyes.7%	3	2	
14	Shrimp Wanamei 16/20, in the shell	3	3	
15	Shrimp peeled, without tail s/m. 26/30 1kg	3	3	
16				

Рис. 4. Пример результатов обработки списка товаров.

Fig. 4. Example of results of processing a list of products.

Используемые в экспериментах наборы данных состояли из Excel-файлов того же типа, как это представлено на рис. 4. Для тестирования мы использовали около 30 файлов для каждой категории (группы), а для проверки-валидации результатов – около 10 файлов. Результаты количественной проверки полученного набора данных представлены в следующем разделе.

3. Обсуждение результатов проекта

3.1 Количественная оценка результатов внедрения решения

В качестве основной метрики качества категоризации/классификации продуктов в списке использовалась метрика Accuracy, которая рассчитывалась как доля правильно классифицированных объектов среди всех объектов в каждой товарной группе. Оценка качества классификации по товарным группам представлена в табл. 4.

Табл. 4. Точность классификации.

Table 4. Accuracy of product classification.

Товарная группа	Метрика Accuracy (Валидация)	Метрика Accuracy (Тест)
Креветки	81%	80%
Угорь	88%	85%
Творог	99%	96%
Сыр «Пармезан»	96%	77%
Сливочное масло	90%	81%
Среднее арифметическое	90,8%	83,8%

Точность решения оказалась вполне достаточной для практического использования ассистентом поставщика ресторана. Система экономит значительное время на ручном сопоставлении номенклатур товаров: достигнутая метрика Accuracy позволяет сокращать первоначальные временные затраты поставщика на ~90% (3 минуты вместо 30). Оставшиеся 10% времени расходуются на валидацию и исправление редких ошибок при появлении новых товаров.

3.2 Основные сложности использования LLM в рамках бизнес-кейса

Несмотря на то, что опытная эксплуатация системы показала многообещающие результаты использования моделей LLM и решений на их основе для проблематики сопоставления номенклатур товаров, в работе системы и ее эксплуатации проявился ряд проблем, которые можно считать общими для всего класса подобных приложений, а именно:

- Качество данных: неполные/некорректные данные влияют на результат;
- Обучение модели: требуется большой объём обучающих данных;
- Интерпретируемость: результаты LLM могут быть неочевидны для пользователей;
- Контекстная зависимость: модели не всегда верно понимают контекст.

Более конкретно, если рассматривать детали реализации и тестирования предложенного решения, то можно выделить в качестве примера несколько показательных конкретных проблем при работе с номенклатурой товаров, возникших на стороне модели LLM:

- Креветки: Термин “Easy Peel” модель понимает, как «очищенные», что некорректно;
- Пармезан: Термин “Palermo” (город в Италии) приводит к ошибочному отнесению товара к импорту;
- Указание “целая голова” вместо “1/2 головы” влияет на категорию товара;
- Модель ассоциирует “Deluxe” с “Premium/Exclusive”, что не верно;
- Масло: Термин “Ranferley” (город) ошибочно ассоциируется с Новой Зеландией.

Обычно подобные проблемы решаются через словари и промпт-инжиниринг, но требуют тщательного тестирования.

3.3 Масштабирование решения до целевого

Целевое решение предлагает использовать LLM для определения критериев ранжирования списка товаров поставщика из полученного наименования товара от заказчика (ресторана).

Для составления критериев на вход модели подается вся необходимая информация:

- Наименование товара;
- Правила составления критериев для определенной группы товаров;
- Дополнительные требования от заказчика.

Выход модели – это критерии ранжирования, которые на следующем шаге используются другой моделью для классификации списка товаров по категориям (эта модель как раз и тестируется в пилоте).

При необходимости модели LLM могут быть дополнительно обучены на примерах генерации и классификации критериев (обычно мы это делаем только тогда, когда другие методы не срабатывают).

Предлагаемое нами решение следующее:

- Некоторые элементы загружаются в систему;
- Выполняется согласование загруженных элементов: все символы опускаются, очищаются все знаки препинания и т.д. – в соответствии с типом входного текста;

- LLM используется для декодирования всех сокращений во входном тексте с использованием встроенного справочника (подготовленного заранее);
- Идентифицируется группа продуктов и клиент, затем фильтруется по необходимой группе. Названия продуктов затем обрабатываются на последнем этапе вместе с текущим списком продуктов
- Параллельно, когда сокращения декодируются, поток переходит к составлению критериев для категоризации с использованием модели LLM и на основе подготовленных правил (подсказок) для составления критериев
- В конце формируется рейтинг списка продуктов, рейтинговый список аналоговых продуктов

5. Заключение

В данном исследовании мы разработали, проанализировали и представили применение больших языковых моделей (LLM) к решению ключевой задачи сопоставления товарных номенклатур между ресторанами и их поставщиками. Предложенное решение использует расширенные возможности LLM для автоматизации и повышения точности сопоставления продуктов, оптимизируя таким образом управление запасами и процессы закупок. Ключевые выводы и вклад работы можно обобщить следующим образом:

- 1) Эффективность и автоматизация. Использование LLM значительно сокращает время и усилия, необходимые для ручного сопоставления товаров. Автоматизируя классификацию продуктов по категориям (точные совпадения, потенциальные совпадения, несовпадения), система оптимизирует процесс закупок, позволяя менеджерам сосредоточиться на стратегических решениях вместо рутинных задач.
- 2) Повышенная точность. Традиционные методы сопоставления часто не справляются с растущим разнообразием наименований товаров и сложностью предложений поставщиков. Наш подход на основе LLM демонстрирует высокую точность в идентификации релевантных соответствий, даже в случаях неоднозначных или нестандартно оформленных наименований. Это достигается благодаря способности модели понимать контекст и эффективно применять предопределённые критерии.
- 3) Масштабируемость и адаптивность. Предложенная архитектура разработана для обработки больших и разнородных наборов данных, что делает её применимой для бизнесов различных масштабов и отраслей. Гибкость системы обеспечивает лёгкую интеграцию с существующими системами управления запасами и закупок, гарантируя беспрепятственное внедрение и масштабирование.
- 4) Практическое применение. На практическом кейсе классификации креветочной продукции мы продемонстрировали способность системы обрабатывать и категоризировать наименования товаров с высокой точностью. Результаты подчеркивают потенциал LLM для трансформации управления цепочками поставок за счёт предоставления практических инсайтов и снижения операционных неэффективностей.
- 5) Широкое применение. Решение применимо не только в ресторанном бизнесе, но и в отраслях, зависящих от эффективного управления запасами: ритейле, производстве и e-commerce, где существуют аналогичные проблемы.
- 6) Проблемы и перспективы. Несмотря на значительные преимущества, сохраняются следующие вызовы: требуется повышение способности модели обрабатывать неоднозначные/неполные наименования, разработка механизмов обратной связи для постоянного улучшения системы, а также интеграция дополнительных источников данных (каталоги поставщиков, рыночные тренды).

Данное исследование демонстрирует преобразующий потенциал LLM в решении сложной задачи сопоставления товарных номенклатур. Автоматизируя этот процесс, компании достигают значительной экономии, повышают операционную эффективность и улучшают принятие решений. По мере развития технологии LLM, её применение в SCM будет расширяться, открывая путь для более интеллектуальных систем. Дальнейшие исследования должны сфокусироваться на решении выявленных проблем и поиске новых возможностей использования LLM в бизнес-операциях.

Список литературы / References

- [1]. Dhara S., Delgado Barba S. (2024). Large language models in supply chain management [Master's thesis]. POLITesi Repository (online), Available at: https://www.politesi.polimi.it/retrieve/9f1da09d-256f-428c-8bae-a94d529df3e6/2024_07_Dhara_Delgado.pdf, accessed 25.10.2025.
- [2]. Li B., Mellou K., Zhang B., Pathuri J., Menache I. (2023). Large language models for supply chain optimization. arXiv preprint arXiv:2307.03875.
- [3]. Kumar S., Mellou K., Pathuri J. (2023). Large language models for supply chain optimization. ArXiv Preprint. Available at: <https://arxiv.org/abs/2307.03875>, accessed 25.10.2025.
- [4]. Bag S., Rahman M.S., Srivastava G., Shore A., Ram, P. (2023) Examining the role of virtue ethics and big data in enhancing viable, sustainable, and digital supply chain performance, *Technological Forecasting and Social Change*, Vol. 186, DOI: 10.1016/j.techfore.2022.122154.
- [5]. HBR Editors. (2025). How generative AI improves supply chain management. *Harvard Business Review* (online). Available at: <https://hbr.org/2025/01/how-generative-ai-improves-supply-chain-management>, accessed 25.10.2025.
- [6]. Aporia AI Insights. (2024). Risks of using LLMs in enterprise applications (online). Available at: <https://www.aporia.com/learn/risks-of-using-llms-in-enterprise-applications/>, accessed 25.10.2025.
- [7]. McKinsey Global Institute. (2025). Beyond automation: How gen AI is reshaping supply chains, Available at: <https://www.mckinsey.com/capabilities/operations/our-insights/beyond-automation-how-gen-ai-is-reshaping-supply-chains/>, accessed 25.10.2025.
- [8]. IMD. (2024). How will large language models impact supply chains? Available at: <https://www.imd.org/ibyimd/supply-chain/large-language-model-impacts-on-supply-chain/>, accessed 25.10.2025.
- [9]. Kellton Tech Blog. (2024). Large language models (LLMs): Navigating challenges and uncovering benefits. Available at: <https://www.kellton.com/kellton-tech-blog/large-language-models-challenges-benefits>, accessed 25.10.2025.
- [10]. Lokad TV. (2023). Large language models in supply chain [online video]. Available at: <https://www.lokad.com/tv/2023/12/13/large-language-models-in-supply-chain/>, accessed 25.10.2025.
- [11]. Cobalt.io. (2024). LLM supply chain attack prevention strategies (online). Available at: <https://www.cobalt.io/blog/llm-supply-chain-attack-prevention-strategies>, accessed 25.10.2025.
- [12]. OWASP Foundation. (2023). OWASP Top Ten for Large Language Model Applications (online). Available at: <https://owasp.org/www-project-top-10-for-large-language-model-applications>, accessed 25.10.2025.
- [13]. Атаки на цепочки поставок: как уязвимости распространяются через зависимости (online). Available at: <https://www.securitylab.ru/blog/personal/xiaomite-journal/355832.php>, accessed 27.10.2025.
- [14]. Поисковый сервис DUC SmartSearch. Available at: <https://duc-technologies.ru/smartssearch>, accessed 27.10.2025.
- [15]. Open Source AI Platform for Work (online). Available at: <https://www.onyx.app>, accessed 27.10.2025.
- [16]. Open source agentic systems development platform (online). Available at: <https://flowiseai.com/>, accessed 27.10.2025.
- [17]. Сервис GigaChat (online). Available at: <https://giga.chat/>, accessed 27.10.2025.

Информация об авторах / Information about authors

Сеунгмин ДЖИН – кандидат технических наук, работает в НИУ ВШЭ с 2024 года в должности доцента департамента бизнес-информатики высшей школы бизнеса НИУ ВШЭ, а также старшего научного сотрудника международной лаборатории интеллектуальных систем

и структурного анализа факультета компьютерных наук НИУ ВШЭ. Сфера научных интересов: визуальная аналитика, искусственный интеллект, нейронные сети, машинное обучение и глубинное обучение и их приложения в рекомендательных системах и других областях.

Seungmin JIN – Cand. Sci. (Tech.), has been working at HSE University since 2024 as an associate professor in the Department of Business Informatics at the HSE Graduate School of Business, as well as a senior researcher at the International Laboratory of Intelligent Systems and Structural Analysis at the HSE Faculty of Computer Science. Research interests: visual analytics, artificial intelligence, neural networks, machine learning and deep learning and their applications in recommender systems and other areas.

Петр Борисович ПАНФИЛОВ – кандидат технических наук, доцент, является профессором департамента бизнес-информатики высшей школы бизнеса НИУ ВШЭ. Его научные интересы включают компьютерные науки и инженерию, науки о данных, распределенные вычисления и обработку данных и их приложения.

Petr Borisovich PANFILOV – Cand. Sci. (Tech.), professor in the Department of Business Informatics of the Graduate School of Business at HSE University since 2014. His research interests include computer science and engineering, data science, distributed computing and data processing and their applications.

Александр Сергеевич СУЛЕЙКИН – кандидат технических наук, является научным сотрудником научно-исследовательской лаборатории процессно-ориентированных информационных систем факультета компьютерных наук НИУ ВШЭ. Его научные интересы включают компьютерные науки и инженерию, науки о данных, искусственный интеллект и решения на основе генеративного искусственного интеллекта в промышленности и бизнесе.

Aleksander Sergeevich SULEYKIN – Cand. Sci. (Tech.), a researcher at the Research Laboratory of Process-Oriented Information Systems at the HSE Faculty of Computer Science. His research interests include computer science and engineering, data science, artificial intelligence, and generative AI-based solutions in industry and business.

