

# ПРОЕКТИРОВАНИЕ И ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ПРОТОТИПА МУЛЬТИАГЕНТНОЙ СИСТЕМЫ ДЛЯ ПОИСКА И ОБОСНОВАНИЯ ЗАКОНОМЕРНОСТЕЙ В НЕСТРУКТУРИРОВАННЫХ ДАННЫХ

**СЛУЗОВА Наталья Александровна**

кандидат физико-математических наук, доцент

**СУХАНОВ Григорий Иванович**

студент

**ПАНКРАТОВ Максим Владимирович**

Студент

Северный (Арктический) федеральный университет

г. Архангельск, Россия

*Данное исследование посвящено задаче интерпретации и анализа итогов интеллектуального анализа данных (Data Mining) на фоне увеличения объемов неструктурированной информации. В этой статье предлагается концепция распределенной мультиагентной системы (МАС), которая синтезирует традиционные методы машинного обучения и генеративные способности больших языковых моделей (LLM) для реализации принципов объяснимого искусственного интеллекта (Explainable AI). Представлена практическая реализация прототипа: архитектура включает децентрализованную (P2P) кластеризацию, использование фреймворка SPADE для коммуникации агентов и пользовательский интерфейс.*

**Ключевые слова:** мультиагентные технологии, неструктурированные данные, Data Mining, LLM, объяснимый ИИ.

В современных реалиях критическим фактором конкурентоспособности становится способность эффективно извлекать знания из массивов Больших данных, где получение и обработки данных становится главным конкурентным преимуществом для бизнеса, госсектора и других сфер человеческой деятельности. Согласно актуальным метрикам, порядка 90% мирового объема информации приходится на неструктурированные данные [3]. Классические централизованные системы хранения и обработки, использующие жесткие реляционные модели, обнаруживают свою несостоятельность при работе с такими потоками информации.

Кроме того, существует проблема «черного ящика»: современные алгоритмы глубокого обучения, несмотря на высокую точность прогнозов, часто не позволяют интерпретировать логику принятия решений. Эксперты получают сухой статистический факт, но лишены контекстуального объяснения причин данного вывода [1].

Цель настоящего исследования – разработка

и апробация прототипа мультиагентной системы для поиска и обоснования закономерностей в неструктурированных данных с использованием больших языковых моделей (LLM).

Задачи исследования: провести анализ подходов к применению LLM и объяснимого искусственного интеллекта (ИИ) в задачах аналитики; спроектировать архитектуру агентов и протокол взаимодействия; реализовать прототип на базе SPADE и LangChain с пользовательским интерфейсом; продемонстрировать работу системы на примерном сценарии и сформулировать направления дальнейшего развития.

Предлагаемый подход направлен на устранение этого барьера путем создания системы, объединяющей поиск скрытых паттернов с их автоматизированной интерпретацией. В этой статье рассматривается архитектура МАС, в которой автономные агенты выполняют распределенные вычисления и, используя генеративный ИИ, пытаются обосновать полученные результаты.

Для обработки больших распределенных массивов данных была выбрана децентрализованная P2P-архитектура (Mesh). В отличие

от иерархических структур, где корневой узел становится «бутылочным горлышком», P2P-сеть обеспечивает высокую масштабируе-

мость и устойчивость к отказам. Глобальное знание в такой системе формируется эмерджентно через протоколы консенсуса.

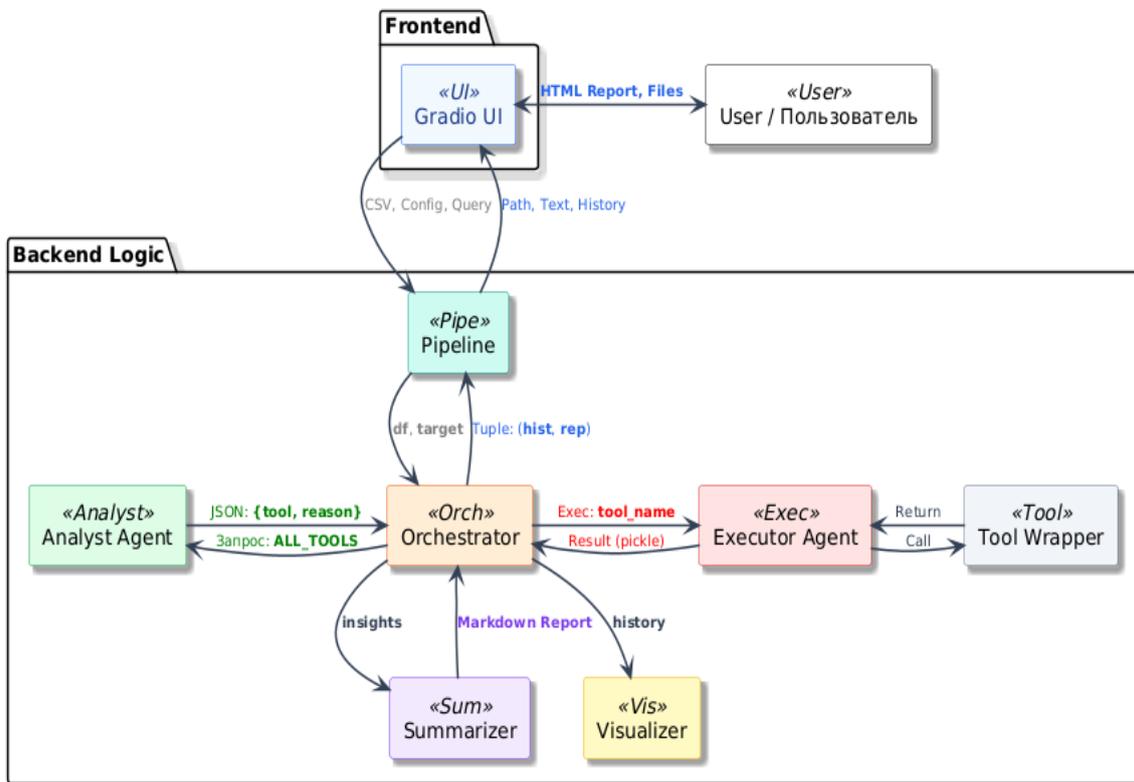


Рисунок 1. Концептуальная схема P2P сети агентов и пайплайн обработки данных

Архитектура системы (см. рисунок 1) включает четыре специализированных типа агентов:

1. Агент-Сборщик (Collector Agent): отвечает за мониторинг источников (файловые системы, API) и первичную очистку данных.

2. Агент-Векторизатор (Embedding Agent): преобразует неструктурированные данные в векторное представление. Для текстов используются трансформеры (BERT, RoBERTa), векторы сохраняются в локальных индексах FAISS [6].

3. Агент-Аналитик (Miner Agent): Ядро системы. Выполняет поиск закономерностей. Агенты реализуют протоколы Gossip для обмена параметрами моделей, достигая консенсуса без передачи сырых данных [2].

4. Агент-Аргументатор (Justification Agent): реализует функции ХАИ. Получая математи-

ческую модель паттерна, он использует LLM для генерации текстового обоснования по модели Тулмина [5].

Прототип системы реализован на языке Python. В качестве платформы для разработки агентов выбран фреймворк SPADE (Smart Python Agent Development Environment), основанный на протоколе XMPP [4]. Интеграция с LLM выполнена через библиотеку LangChain.

Для согласования параметров кластеризации используется итеративный алгоритм взвешенного среднего.

Для взаимодействия с системой разработан простой веб-интерфейс, который позволяет пользователям загружать данные, настраивать параметры подключения к LLM и запускать процесс анализа.

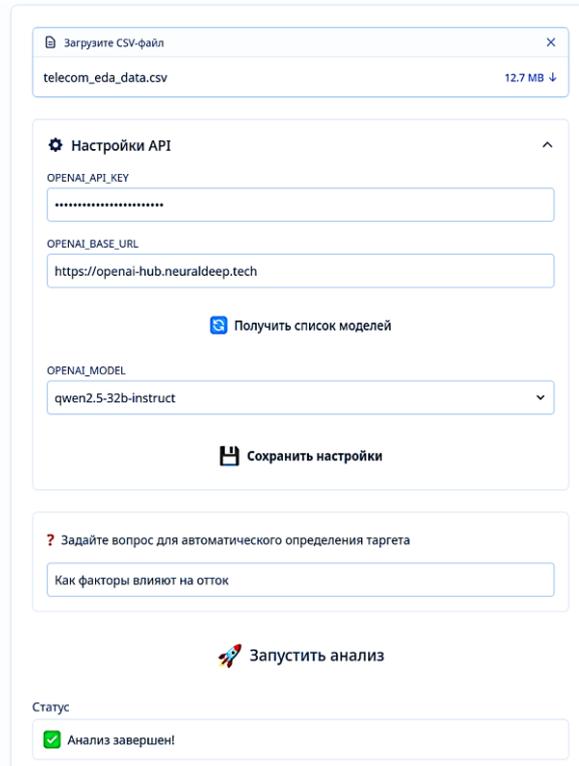


Рисунок 2. Пользовательский интерфейс прототипа: настройка API и загрузка датасета

Как показано на рисунке 2, пользователь имеет возможность:

- Загружать исходные данные (например, CSV-файлы телекоммуникационной статистики `telecom_eda_data.csv`).
- Настраивать ключи доступа к API LLM (OpenAI/Custom Endpoint) и выбирать модель (например, `qwen2.5-32b-instruct`).
- Формулировать исследовательские вопросы на естественном языке (например, «Как факторы влияют на отток?»).

Агенты работают асинхронно, если облачная или локальная система позволяет это.

Ключевой особенностью является то, что результат работы агента – это не просто число, а структурированный отчет. Агент-Аргументатор использует модель Тулмина: Тезис (Claim) – Данные (Data) – Основание (Warrant).

Пример работы системы:

1. Input: Пользователь загружает данные о сбоях оборудования.
2. Analysis: MinerAgent находит кластер документов с ключевыми словами «вибрация», «насос».
3. Reasoning: JustificationAgent через LLM

формирует вывод: «Выявлена закономерность: повышение вибрации на 15% коррелирует с отказами (Вероятность 85%). Рекомендуется превентивная проверка подшипников».

В рамках работы спроектирован и реализован прототип мультиагентной системы InsightFinder, ориентированный на поиск закономерностей в данных и формирование объяснений результатов анализа. Архитектура на базе взаимодействующих агентов позволяет разделить этапы планирования, выполнения вычислений и аргументации, что повышает интерпретируемость получаемых выводов и удобство их представления пользователю. Практический эффект заключается в снижении трудозатрат на первичную интерпретацию результатов анализа и повышении прозрачности выводов для пользователя за счёт объяснимого ИИ.

Дальнейшее развитие работы предполагает: проведение серии вычислительных экспериментов на нескольких датасетах и формализацию метрик качества объяснений; расширение набора инструментальных агентов и автоматизацию построения визуализаций; проработку вопросов безопасности и приватности при анализе чувствительных данных.

**СПИСОК ЛИТЕРАТУРЫ**

1. *Faisal B., Tunkel D.* Explainable AI in Multi-Agent Systems: Advancing Transparency with Layered Prompting // ResearchGate. 2025. – URL:[https://www.researchgate.net/publication/388835453\\_Explainable\\_AI](https://www.researchgate.net/publication/388835453_Explainable_AI) (дата обращения: 28.11.2025).
2. Gossip Protocol Explained // High Scalability. 2023. – URL:<https://highscalability.com/gossip-protocol-explained/> (дата обращения: 28.11.2025).
3. *Gubitosa B.* Unlocking the Power of Unstructured Data with AI // Rivery: Data Learning Center. 2024. – URL:<https://rivery.io/data-learning-center/unstructured-data-with-ai/> (дата обращения: 28.11.2025).
4. *Palanca J.* SPADE: Smart Python Agent Development Environment // GitHub. – URL:<https://github.com/javipalanca/spade> (дата обращения: 28.11.2025).
5. Toulmin Argument Model // Blinn College. 2025. – URL:<https://www.blinn.edu/writing-centers/pdfs/Toulmin-Argument.pdf> (дата обращения: 28.11.2025).
6. *Wolf T. [et al.]*. Transformers: State-of-the-Art Natural Language Processing // Hugging Face. 2024. – URL:<https://huggingface.co/docs/transformers> (дата обращения: 28.11.2025).

**DESIGN AND SOFTWARE IMPLEMENTATION OF A PROTOTYPE  
OF A MULTIAGENCY SYSTEM FOR SEARCHING FOR  
AND SUBSTANTIATING PATTERNS IN UNSTRUCTURED DATA****SLUZOVA Natalya Aleksandrovna**

Candidate of Sciences Physics and Mathematics, Associate Professor

**SUKHANOV Grigory Ivanovich**

Student

**PANKRATOV Maksim Vladimirovich**

Student

Northern (Arctic) Federal University  
Arkhangelsk, Russia

---

*This study addresses the challenge of interpreting and analyzing data mining results in the face of growing volumes of unstructured information. This article proposes the concept of a distributed multi-agent system (MAS) that synthesizes traditional machine learning methods and the generative capabilities of large-scale language models (LLM) to implement the principles of explainable artificial intelligence (Explainable AI). A practical implementation of the prototype is presented: the architecture includes decentralized (P2P) clustering, the use of the SPADE framework for agent communication, and a user interface.*

**Keywords:** multi-agent technologies, unstructured data, data mining, LLM, explainable AI.

---