

**Исследование и разработка автоматизированной технологии проведения
экспертизы информационной продукции с интеллектуальным ядром на основе
искусственных нейронных сетей**

**Черников К. М. (ИТМО), Вяхирев И.О (ИТМО), Смирнов А.В. (ИТМО)
Научный руководитель – кандидат технических наук, доцент Федоров Д. А.
(ИТМО)**

Введение. На данный момент процесс получения и формирования экспертизы информационной продукции требует значительных временных и ресурсных затрат [1]. Подобные анализы всегда занимают много времени, так как требуют детального изучения произведения. Однако существуют методические рекомендации, описывающие ключевые характеристики текста, на которые следует обращать внимание [2]. Традиционные методы экспертизы, основанные на ручном анализе, становятся все менее эффективными в условиях роста объемов данных и необходимости оперативного принятия решений [3]. В условиях, когда объемы информации измеряются терабайтами, а требования к оперативности анализа возрастают, возникает необходимость в разработке автоматизированных технологий, способных проводить экспертизу с высокой точностью и минимальными временными затратами [4]. В связи с этим актуальной задачей становится создание автоматизированных систем, способных анализировать информационную продукцию с высокой скоростью и точностью [5].

Основная часть.

Целью данной работы является исследование и разработка автоматизированной технологии проведения экспертизы информационной продукции с использованием интеллектуального ядра на основе искусственных нейронных сетей и вспомогательных методов.

Для автоматизации экспертизы информационной продукции было принято рассмотреть две технологии, а именно Named Entity Recognition (NER): для извлечения именованных сущностей и классификации ключевых элементов текста (имена, даты, организации), а также эвристические методы, которые основаны на правилах и шаблонах, необходимых для обнаружения специфического контента, например, нецензурной лексики и описывающих определенные правила, которые текст должен соблюдать. На основе объединения NER и эвристик создано программное решение, включающее:

Модуль извлечения сущностей: автоматически выделяет и классифицирует именованные сущности с помощью NER, что в дальнейшем может помочь при анализе принадлежности реплик персонажам

Модуль эвристического анализа: применяет правила и шаблоны для обнаружения нежелательного контента, такого как нецензурная лексика или непотребного контента.

Выводы. Объединение технологии NER и эвристических методов позволяет частично автоматизировать экспертизу информационной продукции, обеспечивая структурирование данных и обнаружение нежелательного контента. Этот подход может быть применен в медиа аналитике, контроле качества контента и других областях, где требуется обработка и анализ текстовой информации.

Список использованных источников:

1. Методические рекомендации по экспертизе информационной продукции / Под ред. А.Н. Петрова. — М.: Изд-во «ИнфоПресс», 2020. — 120 с.
2. Смирнов И.В. Анализ текстовых данных: методы и инструменты. — СПб.: Питер, 2018. — 256 с.
3. Борисов Е.А. Проблемы ручного анализа больших данных // Вестник информационных технологий. — 2021. — № 4. — С. 45–52.
4. Гусев Д.А. Автоматизация экспертизы контента: подходы и перспективы // Искусственный интеллект и системы управления. — 2022. — № 3. — С. 12–19.
5. Loper E., Bird S. NLTK: The Natural Language Toolkit // arXiv. — 2002. — URL: <https://arxiv.org/abs/0205028> (дата обращения: 13.02.2025).