

Н.А. Кузнецова, кандидат филологических наук
Московский городской педагогический университет,
Москва, Российская Федерация

N.A. Kuznetsova, Candidate of Philology
Moscow City University, Moscow, Russian Federation

ИНСТРУМЕНТЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ЛИНГВОКУЛЬТУРОЛОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

ARTIFICIAL INTELLIGENCE TOOLS IN LINGUOCULTURAL STUDIES

Рассмотрены современные методы и подходы с использованием технологий искусственного интеллекта, применяемые для лингвокультурологических исследований. Особое внимание уделено обработке текстов, алгоритмам машинного обучения и их применению для анализа художественных текстов.

Ключевые слова: концепт; искусственный интеллект; семантические поля; связи слов; лингвокультурологический анализ.

The modern methods and approaches employed in linguocultural studies that use artificial intelligence technologies have been examined. Special attention is given to text processing and machine learning algorithms, as well as their application for the analysis of literary texts.

Key words: concept; artificial intelligence; semantic fields; word relations; linguocultural analysis.

Ученые отмечают, что «цифровизация создает необходимость изменения отношений между словами, предложениями, текстами, заставляя язык подстраиваться под технологические условия» [Городищев, Ускова, Ходенкова 2023]. Эти изменения особенно актуальны для лингвокультурологического анализа сложных литературных текстов, где критически важна интерпретация глубоких философских и этических смыслов.

Цель работы заключается в исследовании возможностей применения технологий ИИ для анализа лингвокультурологических концептов.

Под лингвокультурологическим концептом, вслед за Г.В. Токаревым, мы понимаем многомерную единицу ментального уровня, структурированную более мелкими обыденными понятиями, представлениями, культурными установками, идеологемами, стереотипами, которые могут иметь разнообразные формы выражения и типы знаковых репрезентаций [Токарев 2009]. Метод концептуального анализа включает в себя анализ этимологии, парадигматических и синтагматических связей, а также деривационных аспектов слов, представляющих концепт. Подобное исследование предполагает обработку значительных объемов текстовой информации, и в оптимизации этого процесса важную роль играют технологии искусственного интеллекта (ИИ). Мы рассмотрели основные инструменты и платформы ИИ, которые могут быть полезны для описания лингвокультурологических концептов.

Обработка текстов на естественном языке (NLP) – одна из ключевых областей ИИ – широко интегрируется в филологических исследованиях. Среди основных методов обработки текстов можно выделить анализ частотности слов, определение контекста их употребления и построение семантических моделей. Эти методы открывают новые возможности для изучения лингвокультурологических концептов и их интерпретации.

Алгоритмы TF-IDF (Term Frequency-Inverse Document Frequency, частотность термина и обратная частотность термина) помогают выделить наиболее значимые слова в тексте, что особенно актуально при анализе художественных произведений, где отдельные слова могут нести глубокую смысловую нагрузку. При этом использование моделей на основе Word2Vec позволяет исследовать контекст слов и выявлять их семантические связи. ИИ способен помогать в построении семантических сетей, позволяющих изучить связи между словами и культурными концептами. С помощью программ визуализации данных, таких как Gephi или Cytoscape, можно наглядно представить семантические сети, показывая, как различные концепты пересекаются и взаимодействуют друг с другом.

Применение платформ ИИ для лингвокультурологического анализа требует адаптации алгоритмов обработки текстов к особенностям литературного языка. Инструменты SpaCy и Gensim, предлагают возможности для работы с большими массивами текстовых данных, но нуждаются в настройке параметров для анализа художественных текстов, где часто встречаются метафоры, эпитеты, архаизмы, сложные синтаксические конструкции и другие средства выразительности языка. Учет специфики литературного языка конкретного автора является ключевым для корректной интерпретации результатов анализа.

Будущее развитие инструментов для лингвокультурологического анализа связано с интеграцией более сложных алгоритмов искусственного интеллекта. Внедрение технологий глубокого обучения и нейронных сетей открывает новые горизонты для анализа текстов, позволяя учитывать не только лексические и синтаксические особенности, но и семантический контекст произведений.

Список литературы

Городищев А. В., Ускова С. В., Ходенкова Э. В. Обработка русского языка моделями искусственного интеллекта при формировании цифрового аватара // Мир науки. Социология, филология, культурология. 2023. Т. 14. № 3. URL: <https://sfk-mn.ru/PDF/54FLSK323.pdf> (дата обращения: 30.12.2024)

Токарев Г. В. Лингвокультурология: учеб. пособие. Тула : Изд-во Тул. гос. пед. ун-та им. Л. Н. Толстого, 2009.