

УДК 004.85

## ОСОБЕННОСТИ ИМПОРТА И ТИПИЗАЦИИ ГЕОДАНЫХ В ПРОЦЕССЕ РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В НАУКАХ О ЗЕМЛЕ

**Д. Л. Творонович-Севрук**

*Белорусский государственный университет, пр. Независимости, 4,  
220030, г. Минск, Беларусь, [seuruk@ya.ru](mailto:seuruk@ya.ru)*

Рассматриваются актуальные вопросы интеграции искусственного интеллекта в процесс генерации научных знаний, в частности, проблемы типизации геоданных, их дифференциации для совершенствования цифровых систем. Освещаются вопросы различий между традиционными программными продуктами и системами искусственного интеллекта, проблемы выявления и устранения «белого шума» из геоданных для улучшения качества информации, экспортируемой цифровыми системами.

**Ключевые слова:** искусственный интеллект; геоданные; науки о Земле; субъектность искусственного интеллекта; импорт данных, типизация геоданных.

## FEATURES OF IMPORTING AND TYPING GEODATA IN THE PROCESS OF ARTIFICIAL INTELLIGENCE DEVELOPMENT IN EARTH SCIENCES

**D. L. Tvaranovich-Seuruk**

*Belarusian State University, Nezavisimosti Ave., 4,  
220030, Minsk, Belarus, [seuruk@ya.ru](mailto:seuruk@ya.ru)*

The article discusses topical issues of integrating AI into the process of generating scientific knowledge, in particular, the problems of typing geodata, their differentiation for improving digital systems. The issues of differences between traditional software products and artificial intelligence systems, the problems of identifying and eliminating «white noise» from geodata to improve the quality of data exported by AI systems are also considered.

**Keywords:** artificial intelligence; geodata; Earth sciences; AI subjectivity; data import; geodata typing.

В настоящее время активное применение технологий искусственного интеллекта (ИИ) получило широкое распространение в большинстве областей познавательной деятельности человека [1], рассмотрим некоторые фундаментальные проблемы внедрения и использования данного инструмента в науках о Земле. Процесс работы технологий ИИ непосредственно

связан с постоянным развитием алгоритмов обработки геоданных, их накоплением, систематизацией, циклической обработкой и др. В настоящей статье под технологиями ИИ автор понимает не традиционные программные продукты, обладающие хотя и высоким функционалом [2], но не являющимися отдельными суверенными субъектами научного поиска (в несколько ограниченном и управляемом формате) в отдельной предметной области или смежных областях под определенным контролем исследователей – людей, – а именно такие вычислительные системы, которые способны «генерировать» некоторую информацию, являющуюся научным знанием. Такие системы могут быть основаны как на технологиях нейросетей, так и развитых компонентах машинного сознания. Естественно, результаты научного поиска, полученные при помощи ИИ, должны верифицироваться и проверяться на соответствие актуальным научным парадигмам.

*Отличительной особенностью систем ИИ является приобретаемый ими некоторый элемент субъектности в процессе научного поиска.* Выдвинутое положение в настоящее время может быть пока дискуссионным, однако, диффузия таких цифровых технологий в науку неизбежна и неотвратима. Рассмотрим же тезисно такую базовую проблему [3, 4] интеграции систем ИИ в науки о Земле, решение которой при обработке малых объемов информации для людей – исследователей не представляет значительных сложностей, как типизация импортируемых геоданных в контексте развития систем ИИ.

Импортируемые данные в науках о Земле [2], являющиеся исходным материалом для обработки системами ИИ, можно, с одной стороны, условно подразделить на растровые и векторные. С другой стороны, геоданные дифференцируются как условно одномерные, представленные слоями геоданных, например, в виде топоосновы, мониторинга глубины грунтовых вод, гидрохимии и др., так и многомерные, например, спектрально-временная и псевдоакустическая параметризации волнового поля на временных разрезах и кубах в нефтегазовой геологии. Естественно, геоданные, импортируемые в систему ИИ, дифференцируются по степени «сортированности», наиболее высокой степенью обладают упорядоченные массивы данных в виде разнообразных баз данных, хорошим примером которых являются метеорологические наблюдения, показания дебита скважин, сейсмического мониторинга и др., и «несортированные» в виде кластеров несистемных единичных наблюдений – составных компонентов «озер данных», сложенных, например, отдельными нецифрованными фондовыми материалами такими как геологические отчеты, в частности, по региональной геологии, в «растровом» формате.

Помимо указанных типов информации геоданные, содержащиеся в традиционных источниках, дифференцируются по специфике содержания, в частности, текстовой и графической составляющим, особенностям визуализации контента, что требует отдельной адаптации процесса импорта исходных материалов в среду ИИ. В частности, отдельным направлением обучения нейросетей [5] в науках о Земле является создание разнонаправленных алгоритмов обеспечения процесса ввода-вывода и многоуровневой верификации неоднородной информации (геоданные).

Следующей важной составляющей функционирования ИИ в контексте рассматриваемой проблемы типизации импортируемых данных является обеспечение адекватного сопоставления новой информации с ранее сохранёнными единичными и несистемными материалами, незадействованными ранее. Упомянутая операция важна для составления прогнозов, моделей и целей картирования в широком спектре прикладных и фундаментальных проблем в науках о Земле.

Для ИИ в настоящее время характерны сложности «встраивания» единичных несистемных данных в научную картину мира. Указанная особенность основана на имманентной составляющей архитектур систем обработки данных, распространённых в первой четверти XXI в. Следует отметить важность для субъектов исследования условно единичных несортированных данных, которые всегда будут встречаться в науках о Земле, по причинам присущей для них особенности, проявляющейся в «непредсказуемости» появления отдельных событий и феноменов, доступных для исследователя. Однако, указанный тип данных не является «потерянным» для ИИ в науках о Земле. При определённых сценариях обработки указанной информации, она требует отдельной «маркировки» как незадействованная в конкретных операциях анализа геоданных и должна «храниться» для последующего использования. Такой тип материалов «встраивается» в состав новых последующих цифровых алгоритмов обработки геоданных.

Отдельной субпроблемой типизации геоданных является процесс выделения и элиминации истинного «белого шума» среди импортированных материалов в системы ИИ и акцентировании на актуальной информации. Первичные материалы часто, на первый взгляд, по своей структуре являются хаотическими и несистемными, например, кривые каротажных диаграмм, распределение минералов в шлифах, особенности литологии определённого геологического разреза и др. Для человека-исследователя импорт в цифровые системы, сортировка и интерпретация больших объемов такой информации является затратной по времени и ресурсам операцией, а системы ИИ при соответствующей настройке способны оптимально решать указанные задачи [3–6].

## Библиографические ссылки

1. Chatgpt [Электронный ресурс]. URL: <https://chatgpt.org/> (дата обращения: 10.02.2024).
2. ГИС справочник [Электронный ресурс]. URL: <https://gis4arch.blogspot.com/2019/01/blog-post.html> (дата обращения: 10.02.2024).
3. Нейросеть для полевых работ [Электронный ресурс]. URL: <https://stimul.online/articles/innovatsii/neyroset-dlya-polevykh-rabot/> (дата обращения: 10.02.2024).
4. Применение технологий искусственного интеллекта для задач геологоразведки [Электронный ресурс]. URL: [https://nedra.gazprom.ru/d/textpage/da/218/bondarev\\_evgeniy\\_ai.pdf](https://nedra.gazprom.ru/d/textpage/da/218/bondarev_evgeniy_ai.pdf) (дата обращения: 10.02.2024).
5. Эксперименты с нейронными сетями на данных сейсморазведки [Электронный ресурс]. URL: <https://habr.com/ru/companies/ods/articles/482780/> (дата обращения: 10.02.2024).
6. ГенИИ в образовании. Как новое поколение нейросетей изменило EdTech [Электронный ресурс]. URL: <https://www.skolkovo.ru/expert-opinions/genii-v-obrazovanii-kak-novoe-pokolenie-nejrosetej-izmenilo-edtech/> (дата обращения: 10.02.2024).