

Н. А. Бондарева¹✉

Гибридный методологический подход к автоматизированному проектированию в среде big data

¹Институт Прикладной Математики им. М. В. Келдыша РАН,
г. Москва, Российская Федерация
e-mail: nicibond9991@gmail.com

Аннотация. Представлена концепция гибридного методологического подхода к автоматизированному моделированию, сочетающего обработку естественного языка с верифицированными инженерными программными комплексами. В отличие от прямого использования генеративных нейросетей, предложенный подход предполагает применение больших языковых моделей для генерации управляющих скриптов, исполняемых проверенными САПР-системами и геоинженерными комплексами. Это обеспечивает смещение контроля качества с проверки готовых результатов на верификацию параметров программного кода, минимизируя риски скрытых ошибок. Разработан алгоритм функционирования гибридной системы, включающий этапы формулировки задачи, генерации скрипта, обработки в инженерной системе и верификации с возможностью итеративной коррекции. Показана универсальность подхода и перспективы применения в геотехническом моделировании, геоинформационном анализе, обработке данных изысканий и мониторинге геотехнических процессов. Методология обеспечивает повышение производительности, демократизацию доступа к инженерным инструментам при сохранении надежности расчетов и прозрачности процессов проектирования в условиях работы с большими геопространственными данными.

Ключевые слова: машинное обучение, компьютерное зрение и распознавание образов, нейронная сеть, компьютерная графика, GeoAI

N. A. Bondareva¹✉

Hybrid methodological approach to automated design in big data environment

¹Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences,
Moscow, Russian Federation
e-mail: nicibond9991@gmail.com

Abstract. This paper presents a concept of a hybrid methodological approach to automated modeling that combines Natural Language Processing (NLP) with verified engineering software systems. Unlike the direct application of generative neural networks, the proposed approach employs Large Language Models (LLMs) for generating control scripts executed by verified CAD systems and geoen-gineering software packages. This enables a shift in quality control from validation of final results to verification of source code parameters, thereby minimizing the risks of hidden errors. An algorithm for the hybrid system operation has been developed, comprising the stages of problem formulation, script generation, processing in the engineering system, and verification with iterative correction capability. The universality of the approach and prospects for its application in geotechnical modeling, geoinformation analysis, survey data processing, and geotechnical process monitoring are demonstrated. The methodology ensures enhanced productivity, democratization of access to engineering

tools while maintaining computational reliability and transparency of design processes in the context of working with large geospatial datasets.

Keywords: machine learning, computer vision and pattern recognition, neural network, computer graphics, GeoAI

Введение

Современная геоинженерия переживает период фундаментальной трансформации, обусловленный стремительным ростом объемов геопро пространственных данных и развитием технологий их интеллектуальной обработки. По различным оценкам, объем геопро пространственной информации, получаемой с помощью постоянно развивающихся и совершенствующихся технологий, вроде спутниковых систем дистанционного зондирования Земли, или систем лазерного сканирования, с каждым годом лишь нарастает. Эта экспоненциальная динамика формирует типичную среду «больших данных» (Big Data), требующую принципиально новых подходов к анализу и интерпретации, что создает объективные предпосылки для применения современных технологий искусственного интеллекта.

В ответ на этот вызов сформировалось междисциплинарное направление GeoAI (Geospatial Artificial Intelligence) – интеграция методов искусственного интеллекта с геопро пространственными науками для решения задач анализа, моделирования и прогнозирования пространственно распределенных процессов [1, 2]. Особый импульс развитию GeoAI придало появление foundation models – больших предобученных моделей искусственного интеллекта, демонстрирующих беспрецедентные способности к обобщению знаний и адаптации к различным прикладным задачам [3]. Современные большие языковые модели (LLM) открывают качественно новые возможности для геоинженерной практики: автоматизацию обработки разнородных геоданных, генерацию аналитических отчетов, создание интеллектуальных систем поддержки принятия решений и, что особенно важно, формирование интуитивно понятных интерфейсов взаимодействия на естественном языке.

Потенциал применения таких систем в геопро пространственном обеспечении очевиден: это и автоматизация обработки геоданных, и возможность генерации аналитических отчетов, и определенное эффективное ускорение процессов проектирования.

Однако критический анализ современного состояния технологий искусственного интеллекта выявляет существенные ограничения их применения в геоинженерных задачах, где цена ошибки может измеряться человеческими жизнями и катастрофическими экономическими потерями. Ключевая проблема заключается в явлении «галлюцинаций» нейросетевых моделей – генерации правдоподобных, но фактически некорректных результатов. В работе [4] демонстрируется, что даже наиболее продвинутые LLM систематически создают ложную информацию в 5-15% случаев при решении задач, требующих точных вычислений. В контексте геопро пространственного обеспечения, где проектные решения

напрямую влияют на устойчивость инженерных сооружений, корректность расчета осадок фундаментов или прогноза оползневых процессов, подобные неточности категорически недопустимы.

Дополнительные вызовы применения «чистых» LLM-подходов в геоинженерии включают [2-5]:

- отсутствие гарантированной воспроизводимости, когда стохастическая природа генерации может давать различные результаты при идентичных входных данных;
- ограниченная точность математических операций, так как современные языковые модели демонстрируют неудовлетворительную производительность в задачах, требующих точных численных расчетов и геометрических построений;
- проблема «черного ящика»: непрозрачность механизмов принятия решений препятствует верификации корректности результатов и их сертификации для использования в критически важных проектах;
- сложность соответствия нормативным требованиям;
- юридические риски: при чрезвычайных происшествиях возникает неопределенность в вопросах ответственности за проектные решения, полученные с использованием недетерминированных алгоритмов.

Таким образом, формируется определенное противоречие: с одной стороны, системы искусственного интеллекта обладают мощным потенциалом автоматизации обработки больших геопространственных данных и повышения эффективности инженерной деятельности; с другой стороны, они не могут гарантировать уровень надежности, точности и юридической обоснованности результатов, необходимый для практического применения в геоинженерии. Разрешение этого противоречия требует разработки принципиально новых методологических подходов, способных синтезировать преимущества современных технологий ИИ с проверенными инженерными практиками, что и определяет актуальность настоящего исследования.

Концепция гибридного методологического подхода

Для преодоления указанных ограничений предлагается концепция гибридного методологического подхода, основанного на синергетической интеграции систем обработки естественного языка [7] и верифицированных инженерных программных комплексов. Фундаментальная идея заключается в рациональном распределении функций между компонентами системы: NLP-модуль (Natural Language Processing, обработка естественного языка) отвечает за интерпретацию запросов специалиста и формирование параметров задачи, тогда как критически важные расчеты проводятся сертифицированными программными средствами.

Ключевое отличие предлагаемого подхода состоит в принципиальном изменении парадигмы применения искусственного интеллекта в инженерных процессах. Вместо непосредственного использования нейросетей для генерации конечного результата предлагается применять искусственный интеллект для создания управляющего скрипта, исполняемого специализированной системой (рис. 1).

Такое решение позволяет сместить фокус контроля качества с проверки готового результата на верификацию параметров программного кода, обеспечивая раннее выявление потенциальных ошибок, прозрачность процесса и воспроизводимость решений при сохранении гарантированной точности расчетов.

Для валидации концепции выполнена практическая реализация методологии в области автоматизированного трехмерного моделирования с использованием САПР-систем. Архитектура решения включает три взаимосвязанных компонента: интеллектуальный генератор скриптов на основе больших языковых моделей, исполнительную среду в виде верифицированных программных комплексов автоматизированного проектирования и контур верификации с контролем качества и механизмом итеративной коррекции на основе обратной связи от специалиста.

Предложенная архитектура обеспечивает ряд принципиальных преимуществ по сравнению с альтернативными подходами. Прозрачность процесса достигается за счет того, что каждый шаг построения модели явно задан в скрипте и доступен для инспекции. Воспроизводимость результатов гарантируется возможностью сохранения и повторного использования скриптов для решения аналогичных задач. Масштабируемость подхода проявляется в возможности пакетной генерации серий моделей с вариацией параметров, а естественная интеграция с существующими процессами обеспечивается совместимостью генерируемых скриптов с привычными технологическими процессами.

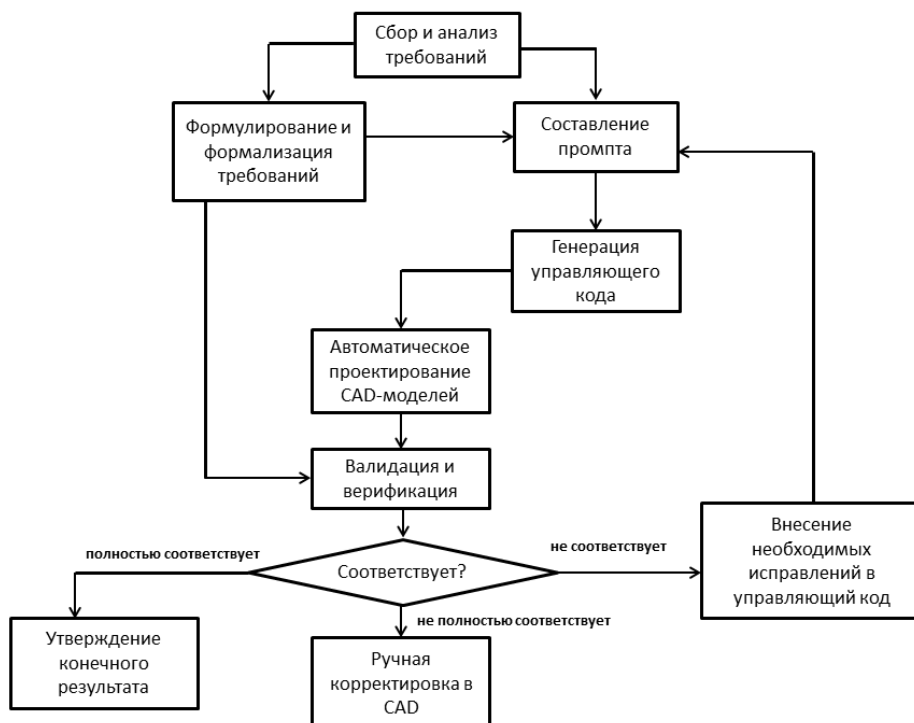


Рис.1. Схема гибридного методологического подхода применительно к CAD системам

Разработанный алгоритм реализации гибридного подхода представляет собой последовательность из шести взаимосвязанных этапов. Первый этап включает формулировку задачи с четким определением требований к конечному результату. Вторым этапом является составление структурированного промпта для языковой модели с описанием задачи в терминах предметной области и указанием технических параметров. На третьем этапе большая языковая модель осуществляет генерацию управляющего кода, совместимого с выбранной инженерной системой. Четвертый этап предполагает автоматическое выполнение скрипта в целевой системе, где критические вычисления производятся детерминированными алгоритмами, минимизируя риски нейросетевых генераций. Пятый этап образует ядро системы обеспечения качества и включает процедуры верификации и валидации. Заключительный этап включает финализацию результатов с сохранением в стандартных форматах и пополнением корпоративной базы знаний.

Области применения в геопространственном обеспечении и перспективы развития

Несмотря на то, что практическая апробация предложенной методологии была выполнена в области трехмерного моделирования в САПР-системах [6], концептуальная основа подхода носит универсальный характер и может быть распространена на широкий спектр задач геоинженерии и геопространственного обеспечения. Принципиальная переносимость методологии обусловлена тем, что любая инженерная система, обладающая программным интерфейсом или возможностью скриптового управления, может быть интегрирована в гибридную архитектуру при условии наличия формализованного языка команд и детерминированных алгоритмов обработки данных.

В области геотехнического моделирования методология применима для работы со специализированными расчетными комплексами, обеспечивая автоматизированное формирование расчетных схем на основе текстового описания инженерно-геологических условий. Ключевым преимуществом является возможность верификации параметров на уровне скрипта до запуска ресурсоемких вычислений, что минимизирует риски ошибок в исходных данных.

Применительно к геоинформационным системам подход открывает перспективы автоматизированной обработки больших массивов геопространственных данных по естественным языковым запросам, позволяя специалистам формулировать сложные аналитические задачи без детального знания синтаксиса геообработки.

Адаптация базового алгоритма для решения геоинженерных задач сохраняет общую логику шести этапов при варьировании их содержательного наполнения. Если в реализованном варианте для трехмерного моделирования этап формулировки задачи фокусируется на требованиях к геометрии детали, то в геоинженерном контексте акцент смещается на описание инженерно-геологических условий и требований к расчету или анализу. Промпт трансформируется из описания формы и размеров в описание грунтовых условий, нагрузок и конструктивных особенностей. Генерация скрипта адаптируется под специфику целевой системы, будь то Python для

QGIS и Plaxis или специализированные языки иных комплексов. Этап обработки включает не построение трехмерной геометрии, а формирование расчетной схемы или выполнение геопространственного анализа. Верификация при этом фокусируется не на геометрической корректности, а на соответствии нормативным требованиям и физической адекватности модели. Финализация расширяется за счет сохранения не только модели, но и результатов анализа и отчетной документации.

Вместе с тем необходимо признать наличие ряда ограничений и вызовов, связанных с применением предлагаемой методологии. Технологические ограничения проистекают из зависимости точности генерируемых скриптов от возможностей используемой большой языковой модели. Требования к формулировке промпта предполагают необходимость четкого, структурированного описания задачи, поскольку неоднозначность запроса приводит к ошибкам в коде. Ограниченность текстовой модальности создает сложности при передаче визуальной или пространственной информации. Сохраняется также риск генерации синтаксически корректного, но логически ошибочного кода, что обуславливает критическую важность этапа верификации.

Направления минимизации выявленных ограничений включают разработку специализированных NLP-моделей через дообучение на инженерной документации, создание библиотек шаблонных промптов для типовых задач, развитие мультимодальных интерфейсов и формирование экосистемы автоматизированной верификации сгенерированных скриптов на соответствие нормативам..

Заключение

Предложенная концепция гибридного методологического подхода представляет собой сбалансированное решение проблемы интеграции технологий искусственного интеллекта в геоинженерную практику. Синергия обработки естественного языка и верифицированных вычислительных комплексов позволяет совместить преимущества современных LLM – интуитивность взаимодействия, гибкость, способность к автоматизации рутинных операций – с гарантированной точностью и надежностью инженерных расчетов.

В контексте геопространственного обеспечения данный подход открывает качественно новые возможности эффективной обработки больших данных, автоматизации проектирования и повышения производительности труда специалистов без компромиссов в области безопасности и качества результатов. Дальнейшее развитие методологии будет способствовать формированию интеллектуальной экосистемы геоинженерных инструментов, соответствующей вызовам цифровой трансформации отрасли.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Abdelkhalek I. Alastal, Ashraf Hassan Shaqfa GeoAI Technologies and Their Application Areas in Urban Planning and Development: Concepts, Opportunities and Challenges in Smart City (Kuwait, Study Case) // Journal of Data Analysis and Information Processing – 2022 – Vol.10 – No.2.
2. Mai, G., Huang, W., Sun, J., et al. (2022). On the opportunities and challenges of foundation models for geospatial artificial intelligence. arXiv preprint arXiv:2304.06798.

3. Bommasani, R., Hudson, D. A., Adeli, E., et al. (2021). On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258.
4. Ji, Z., Lee, N., Frieske, R., et al. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, 55(12), 1-38.
5. Phoon, K. K., & Zhang, W. (2023). Future of machine learning in geotechnics. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 17 (1): 7–22.
6. N.A. Bondareva, A.E. Bondarev, S.V. Andreev, I.G. Ryzhova. Development of a Methodology for the Application of Generative Neural Networks in Creating 3d Models (2025). *Scientific Visualization* 17.3: 25 - 34, DOI: 10.26583/sv.17.3.034.
7. Radford A., Jong W.K., Hallacy C., Ramesh A., Goh G., Agarwal S., Sastry G., Aspell A., Mishkin P., Clark J., Krueger G., Sutskever I. Learning Transferable Visual Models From Natural Language Supervision. – 2021 – arXiv preprint – arXiv:2103.00020 – <https://doi.org/10.48550/arXiv.2103>.

© *Н. А. Бондарева*, 2026