

УДК 004.891.2:004.942

Азеева Ануза Анваровна, магистрант кафедры Механики и процессов управления, РУДН им. Патриса Лумумбы, г. Москва

**МЕТОДЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В УПРАВЛЕНИИ
СЕМАНТИЧЕСКИМИ
ЦИФРОВЫМИ ДВОЙНИКАМИ ЧЕРЕЗ ЕСТЕСТВЕННЫЙ ЯЗЫК:
КЛАССИФИКАЦИЯ ПОДХОДОВ И ПРОБЛЕМА
СЕМАНТИЧЕСКОГО РАЗРЫВА**

Аннотация

В статье проводится систематический анализ методов искусственного интеллекта, применяемых к задаче управления семантическими цифровыми двойниками посредством команд на естественном языке. Предложена классификация трёх групп подходов: генерация трёхмерных форм из текста, генерация управляющего кода языковыми моделями, онтологические методы на основе графов знаний. Для каждой группы выявлены ключевые ограничения применительно к промышленным инженерным моделям. Показано, что ни одна из групп не обеспечивает одновременно высокого понимания текста, соблюдения физических ограничений и возможности изменения существующих моделей в нейтральных форматах. Проанализирован семантический разрыв — несоответствие между намерением пользователя и формальной спецификацией CAD/BIM-систем. Сформулированы пять требований к методу надёжного NL-управления цифровым двойником.

Annotation

The paper presents a systematic analysis of artificial intelligence methods applied to natural-language-driven semantic digital twin management. A classification of three groups is proposed: text-to-3D shape generation, code generation via language models, and ontological knowledge graph approaches. Critical limitations of each group for industrial engineering models are identified. None of the groups simultaneously provides high text understanding, constraint satisfaction, and modification of existing models in neutral formats. The semantic gap between user intent and CAD/BIM formal specifications is analysed. Five requirements for a reliable NL-driven digital twin management method are formulated.

Ключевые слова: цифровой двойник, граф знаний, большие языковые модели, BIM, CAD, управление на естественном языке, семантический разрыв.

Keywords: digital twin, knowledge graph, large language models, BIM, CAD, natural language management, semantic gap.

Введение

Концепция семантического цифрового двойника предполагает представление физического объекта в виде онтологически структурированного графа знаний, хранящего не только геометрию, но и инженерный смысл элементов: типы конструкций, физические свойства, нормативные требования и логические связи между компонентами [1, 2]. В отличие от традиционных CAD/BIM-моделей, такие двойники поддерживают автоматическое машинное рассуждение и непрерывную валидацию ограничений. Промышленные открытые форматы IFC (ISO 16739-1) и STEP (ISO 10303-21) являются шагами в направлении семантического обогащения геометрических данных [3].

Развитие больших языковых моделей поставило на повестку вопрос об управлении цифровыми двойниками через произвольные текстовые запросы.

Перспектива очевидна: вместо точных API-вызовов система должна понимать запросы вида «сдвинуть несущую стену, сохранив опирающиеся перекрытия» и автоматически транслировать их в корректные изменения модели [4]. Однако прямое применение языковых моделей к CAD/BIM-системам наталкивается на принципиальные ограничения, которые не удаётся преодолеть без специальной архитектуры [4, 5].

Цель настоящей работы — систематический анализ существующих методов ИИ применительно к задаче изменения семантических цифровых двойников через естественный язык, выявление ключевой проблемы (семантического разрыва) и формулировка требований к архитектуре метода её преодоления.

1. Классификация методов ИИ

Анализ литературы позволяет выделить три группы методов, различающихся по природе представления объекта и способу работы с текстовыми запросами.

1.1. Генерация трёхмерных форм из текста.

Диффузионные вероятностные модели в сочетании с архитектурой CLIP (Contrastive Language-Image Pre-training) [6] обеспечили прямую генерацию трёхмерных объектов по текстовому описанию. Система DreamFusion [7] впервые продемонстрировала Text-to-3D-синтез без использования трёхмерных обучающих данных через оптимизацию нейронного поля излучения (NeRF) с дистилляцией оценки вероятности (Score Distillation Sampling). Последующие работы — Magic3D, 3D Gaussian Splatting — существенно повысили скорость и визуальное качество результатов.

Ограничение применительно к цифровым двойникам принципиальное: генерируемые объекты представляют собой монолитные полигональные оболочки без инженерной семантики. Такая «стена» лишена материала,

нормативных требований и связей с другими конструктивными элементами. Модификация без регенерации невозможна.

1.2. Генерация управляющего кода языковыми моделями.

Вторая группа использует способность языковых моделей к написанию кода: запрос транслируется в Python-скрипт для API параметрической CAD-системы (FreeCAD, CadQuery, Revit). Системы Text2CAD, CAD-Llama обучаются на массивах CAD-скриптов [4]. При корректной генерации результатом является полноценная параметрическая B-Rep модель с историей построения.

Фундаментальный изъян — «геометрические галлюцинации» языковых моделей [4, 5]: генерация синтаксически корректного, но математически невыполнимого кода (незамкнутые контуры, недопустимые радиусы скруглений, несуществующие функции API). По данным бенчмарков 2024–2025 годов, корректный код генерируется не более чем для 40–60% простых запросов [4]. Для уже скомпилированных файлов IFC и STEP метод неприменим в принципе.

1.3. Онтологические методы на основе графов знаний.

Третья группа вводит промежуточный уровень — типизированный граф знаний, где узлами служат инженерные концепты («несущая колонна», «тепловой контур»), а рёбрами — их отношения. Система StructureNet [8] предложила иерархическое графовое представление форм. В задачах BIM система BIMCoder [7] транслирует NL-запросы в структурированные запросы к платформе, достигая точности совпадения 87,16%.

Достоинства подхода: высокий уровень семантического понимания, возможность формального кодирования инженерных ограничений (стандарт SHACL W3C), независимость от конкретного программного вендора. Критическое ограничение для нашей задачи: существующие реализации

поддерживают только чтение данных, но не изменение модели. Задача произвольной NL-модификации при сохранении физических ограничений остаётся нерешённой.

Сравнительный анализ трёх групп представлен в таблице 1.

Таблица 1 — Сравнение методов ИИ применительно к управлению цифровыми двойниками

Критерий	Генерация форм (Text-to-3D)	Генерация кода (LLM→CAD)	Граф знаний / онтология
Понимание текста	Высокое (CLIP, языковые модели).	Высокое. Трудность — трансляция в математику.	Среднее. Требуется явный маппинг на онтологию.
Соблюдение инженерных ограничений	Отсутствует. Нет инженерной семантики.	Высокое при корректном коде. Ненадёжное из-за галлюцинаций.	Высокое. Ограничения закодированы в рёбрах графа.
Изменение существующей модели (IFC/STEP)	Не поддерживается.	Только при наличии исходного кода. Не работает со скомп. форматами.	Чтение — есть [7]. Модификация не реализована.
Независимость от платформы	Полная. Экспорт в .obj, .stl.	Низкая. Жёсткая привязка к API вендора.	Высокая. Граф не зависит от платформы.

Ни одна из трёх групп не обеспечивает одновременного выполнения ключевых требований: высокого понимания текста, соблюдения инженерных ограничений и изменения существующих моделей в нейтральных форматах (см. таблицу 1).

2. Семантический разрыв: природа проблемы

Несоответствие между возможностями методов ИИ и требованиями промышленных систем обусловлено фундаментальным различием двух пространств.

Интенциональное пространство пользователя — высокоуровневое, неточное, контекстно-зависимое: «расширить дверной проём, не нарушая несущую способность». Операциональное пространство CAD/BIM API —

строго детерминированное: конкретные координаты, идентификаторы топологических элементов, упорядоченные последовательности операций. Разрыв между ними определяется не только лингвистической двусмысленностью, но и принципиальной неспособностью языковых моделей к геометрическим вычислениям [5].

Исследования пространственного мышления языковых моделей [5] демонстрируют систематические ошибки при определении пространственных отношений и поддержании ментальных трёхмерных моделей: языковые модели воспринимают пространство через статистические закономерности текста, а не через геометрические инварианты. Это не технологический изъян текущих реализаций — это архитектурное несоответствие.

Важно разграничить два класса задач. Извлечение данных («какова толщина стены C5?»): языковые модели вполне применимы как переводчики запросов в SQL/SPARQL [7]. Модификация («изменить высоту всех колонн оси A с перерасчётом примыкающих конструкций»): требует понимания каскадных зависимостей и гарантий физической корректности. Этот класс остаётся нерешённым.

3. Требования к методу NL-управления цифровым двойником

На основании анализа сформулированы пять требований к архитектуре метода.

Разделение семантического и геометрического уровней.

Языковая модель должна быть изолирована от прямых геометрических вычислений: её задача — извлечение намерения и его отображение на формальное описание изменения. Все пространственные вычисления делегируются отдельному детерминированному компоненту. Это устраняет основной источник галлюцинаций.

Промежуточное абстрактное представление.

Необходим уровень между текстовым намерением и конкретной командой API, понятный модели (работает с концептами «колонна», «нагрузка») и формально строгий (допускает математическую реализацию). Граф знаний с онтологической типизацией является наиболее подходящей структурой для этой роли.

Формальные гарантии целостности.

Любая модификация должна сопровождаться автоматической проверкой физических и нормативных ограничений. Эмпирическое тестирование недостаточно: при разнообразии возможных запросов необходимы формальные гарантии — например, на основе алгебраических систем преобразования графов [9].

Замкнутый цикл самокоррекции.

При выявлении нарушений верификатор формирует описание ошибки на естественном языке и автоматически подаёт его обратно модели для пересмотра решения. Такой цикл принципиально повышает долю успешных транзакций без участия пользователя.

Независимость от программных платформ.

Метод должен работать с открытыми нейтральными форматами (IFC, STEP, CityGML), а не с закрытыми API конкретных систем. Это обеспечивает применимость к существующим промышленным моделям.

Заключение

Проведённый анализ показал, что ни одна из трёх существующих групп методов ИИ не обеспечивает надёжного управления семантическими цифровыми двойниками через произвольные NL-запросы. Методы генерации форм лишены инженерной семантики; методы кодогенерации склонны к галлюцинациям и не работают с уже скомпилированными моделями; онтологические методы реализуют лишь чтение данных.

Ключевой причиной ограничений является семантический разрыв между намерением пользователя и математическими требованиями CAD/BIM-систем, обусловленный принципиальной неприспособленностью языковых моделей к геометрическим вычислениям.

Преодоление разрыва требует гибридной архитектуры, сочетающей стохастическую языковую модель в роли семантического парсера с детерминированным механизмом алгебраических преобразований промежуточного графового представления. Сформулированные пять требований составляют основу для разработки такой архитектуры, детальное описание которой является предметом отдельного исследования.

Литература

1. Tao F., Zhang H., Liu A., Nee A. Y. C. Digital Twin in Industry: State-of-the-Art // *IEEE Transactions on Industrial Informatics*. 2019. Vol. 15, № 4. P. 2405–2415. DOI: 10.1109/TII.2018.2873186.
2. Semeraro C., Lezoche M., Panetto H., Dassisti M. Digital Twin Paradigm: A Systematic Literature Review // *Computers in Industry*. 2021. Vol. 130. Article 103469. DOI: 10.1016/j.compind.2021.103469.
3. Deng M., Menassa C. C., Kamat V. R. From BIM to Digital Twins: A Systematic Review // *ITcon*. 2021. Vol. 26. P. 58–83. DOI: 10.36680/j.itcon.2021.005.
4. Zhang Y., Le X. et al. Large Language Models for Computer-Aided Design: A Survey. arXiv:2505.08137. 2025. URL: <https://arxiv.org/abs/2505.08137> (дата обращения: 01.04.2026).
5. Puy G. et al. Evaluating and Enhancing Spatial Cognition Abilities of Large Language Models // *International Journal of Geographical Information Science*. 2025. DOI: 10.1080/13658816.2025.2490701.
6. Radford A., Kim J. W. et al. Learning Transferable Visual Models from Natural Language Supervision // *Proc. ICML 2021*. Vol. 139. P. 8748–8763.
7. Liu B., Chen H. BIMCoder: A Comprehensive Large Language Model Fusion Framework for Natural Language-Based BIM Information Retrieval // *Applied Sciences*. 2025. Vol. 15, № 14. Article 7647. DOI: 10.3390/app15147647.
8. Mo K., Guerrero P. et al. StructureNet: Hierarchical Graph Networks for 3D Shape Generation // *ACM Transactions on Graphics*. 2019. Vol. 38, № 6. Article 242. DOI: 10.1145/3355089.3356527.
9. Ehrig H., Ehrig K., Prange U., Taentzer G. *Fundamentals of Algebraic Graph Transformation*. Berlin: Springer, 2006. 390 с. ISBN 978-3-540-31187-4.

10. ГОСТ Р 57700.37-2021. Компьютерные модели и моделирование. Цифровые двойники изделий. М.: Стандартиформ, 2021. 18 с.

Literature

1. Tao F., Zhang H., Liu A., Nee A. Y. C. Digital Twin in Industry: State-of-the-Art // IEEE Transactions on Industrial Informatics. 2019. Vol. 15, № 4. P. 2405–2415.
2. Semeraro C., Lezoche M., Panetto H., Dassisti M. Digital Twin Paradigm: A Systematic Literature Review // Computers in Industry. 2021. Vol. 130. Article 103469.
3. Deng M., Menassa C. C., Kamat V. R. From BIM to Digital Twins: A Systematic Review // ITcon. 2021. Vol. 26. P. 58–83.
4. Zhang Y., Le X. et al. Large Language Models for Computer-Aided Design: A Survey. arXiv:2505.08137. 2025.
5. Puy G. et al. Evaluating and Enhancing Spatial Cognition Abilities of Large Language Models // International Journal of Geographical Information Science. 2025.
6. Radford A., Kim J. W. et al. Learning Transferable Visual Models from Natural Language Supervision // Proc. ICML 2021. Vol. 139. P. 8748–8763.
7. Liu B., Chen H. BIMCoder: A Comprehensive LLM Fusion Framework for Natural Language-Based BIM Information Retrieval // Applied Sciences. 2025. Vol. 15, № 14. Article 7647.
8. Mo K., Guerrero P. et al. StructureNet: Hierarchical Graph Networks for 3D Shape Generation // ACM Transactions on Graphics. 2019. Vol. 38, № 6. Article 242.
9. Ehrig H., Ehrig K., Prange U., Taentzer G. Fundamentals of Algebraic Graph Transformation. Berlin: Springer, 2006. 390 p.
10. GOST R 57700.37-2021. Computer Models and Simulation. Digital Twins of Products. Moscow: Standartinform, 2021. 18 p.