



КОНЦЕПЦІЯ HUMAN-IN-THE-LOOP В ІНТЕРФЕЙСАХ МОДЕЛЮВАННЯ ІГРОВИХ СЦЕНАРІЇВ ВАРГЕЙМІВ

Олійник В.М., аспірант, кафедра МСТ, ХНУРЕ
Бізюк В.М., професор, кафедра МСТ, ХНУРЕ

Abstract. *This paper proposes a Human-in-the-Loop architecture for wargame scenario design interfaces to mitigate Large Language Model hallucinations. Replacing error-prone iterative text prompting, the approach enables direct UI manipulation of generated JSON structures. Through a dual-layer data model and state locking, designers can efficiently validate parameters and ensure semantic consistency.*

Keywords: *Human-in-the-Loop, Large Language Models, Wargames, Interface Design, Hallucination Mitigation.*

Великі мовні моделі (ВММ) застосовуються для генерації контенту, зокрема трансформації природної мови у структуровані JSON-дані для гексагональних варгеймів [1]. Проте семантичні галюцинації моделей призводять до порушення ігрового балансу та виходу атрибутів за допустимі межі [2, 3].

Стандартний метод виправлення помилок – ітеративний промптинг (Prompt Chaining) – є неефективним. Кожна ітерація ризикує спотворити коректно згенеровані вузли даних, а словесний опис числових змін вимагає високих когнітивних витрат. Навіть формалізація промптів не усуває потреби у ручній валідації [3].

Метою роботи є обґрунтування архітектурної інтеграції підходу Human-in-the-Loop (HITL) у графічний інтерфейс проектування варгеймів для прямої маніпуляції об'єктами сценарію замість текстового діалогу з моделлю.

Архітектура та патерни HITL-взаємодії.

Базовий вектор взаємодії $\text{Text} \rightarrow \text{LLM} \rightarrow \text{Result}$ трансформується у цикл $\text{Text} \rightarrow \text{LLM} \rightarrow \text{UI_state} \rightarrow \text{Human_Override} \rightarrow \text{Result}$. Згенерований моделлю JSON-документ розкладається на дискретні UI-компоненти. Інтерфейс діє як семантичний дешифратор: текстовий вихід ВММ стає множиною атомарних редагованих елементів. Запропонована архітектура передбачає сувору типізацію згенерованих даних. Атрибутам бойових одиниць відповідають числові поля з вбудованими валідаторами діапазонів; гексагональним тайлам ландшафту – інтерактивна сітка з параметрами місцевості; модифікаторам кубиків – випадючі списки, що обмежені виключно дозволеними рядками граматики варгейму. Це перетворює загальний текстовий вивід на контрольоване середовище проектування.

Користувач напряму змінює некоректні параметри через UI. Ця зміна фіксується як Human Override у двошаровій моделі даних, де людське втручання має пріоритет під час серіалізації фінального сценарію. На відміну від повторного текстового запиту, який реініціалізує контекст моделі та створює ризик непередбачуваної зміни суміжних коректних вузлів, ручна правка ізолює модифікований елемент JSON-структури. HITL у цьому застосунку фіксує рівень автоматизації, на якому система лише пропонує базовий варіант сценарію,



допускає його довільну точкову модифікацію і зберігає фінальний стан виключно після явного затвердження людиною. Це реалізує принципи мішано-ініціативного співтворення [4] та Human-Centered AI [5].

Практична реалізація НІТЛ-взаємодії спирається на три патерни.

1. Трасування зв'язків (Traceability). Кожен параметр інтерфейсу зберігає посилання на фрагмент промпту, який його згенерував. Виділення тексту локалізує похідні UI-вузли і навпаки. Це безпосередньо усуває проблему непрозорості моделей [2].

2. Локальна маніпуляція станом. UI-компоненти підтримують операцію Lock (блокування). При повторному запиті до ВММ заблоковані вузли передаються як незмінний (read-only) контекст, що усуває ризик небажаної регресії раніше валідованих елементів під час часткової регенерації.

3. Візуальна диференціація. UI-вузли маркуються за походженням («згенеровано», «верифіковано», «модифіковано»). Стан зберігається у метаданих сценарію для можливості зовнішнього аудиту.

Семантична цілісність забезпечується дворівневою валідацією. Перший рівень – атомарні граничні умови, вбудовані в UI-компоненти (обмеження дб-таблиць), що гарантує дотримання формальної граматики. Другий рівень – автоматична перевірка крос-залежностей усього JSON-документа перед серіалізацією. Дельта змін між генерацією та фінальним станом зберігається для зворотного зв'язку та подальшого калібрування промптів [3].

Висновки.

Інтеграція підходу НІТЛ через двошарову модель даних і патерни прямої маніпуляції усуває ризики спотворення валідних даних, які притаманні текстовому промптингу, та мінімізує зусилля на специфікацію змін. Інструментальне середовище функціонує як засіб підтримки прийняття рішень (Decision Support Tool) – ВММ відповідає за чорнову генерацію структури, а людина – за точкову валідацію через графічний інтерфейс. Збереження дельти між початковою генерацією та ручним коригуванням створює базу для зворотного зв'язку та подальшого калібрування промптів з метою підвищення точності автоматизованої трансформації сценаріїв.

Підхід обмежений доменом гексагональних варгеймів на основі дб-таблиць. Подальші дослідження будуть спрямовані на кількісну оцінку скорочення ітерацій та часових витрат у циклі «генерація – верифікація – коригування».

Список літератури

1. Biziuk, A., & Oliinyk, V. (2025). El papel de la inteligencia artificial en la generación de contenido. *Jóvenes en la Ciencia*, 35. <https://www.jovenesenlaciencia.ugto.mx/index.php/jovenesenlaciencia/article/view/4730>.
2. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1-38. <https://doi.org/10.1145/3571730>.
3. Oliinyk, V., Biziuk, A., Deineko, Zh., Chelombitko, V. (2025). Formalization of text prompts to artificial intelligence systems. *Eastern-European Journal of Enterprise Technologies*, 5(2), 84-97. <https://doi.org/10.15587/1729-4061.2025.335473>.
4. Yannakakis, G.N., Liapis, A., & Alexopoulos, C. (2014). Mixed-initiative co-creativity. *Proceedings of the 9th International Conference on the Foundations of Digital Games (FDG 2014)*.
5. Shneiderman, B. (2022). *Human-centered AI*. Oxford University Press.