

ВІД СКРИПТІВ ДО НАВИЧОК: СТАНДАРТ AGENT SKILLS

Розвиток технологій штучного інтелекту досяг етапу, на якому традиційні підходи, орієнтовані виключно на генерацію текстового контенту, вже не відповідають сучасним вимогам ринку. Натомість спостерігається зміщення фокусу у бік інтелектуальних агентів, здатних до автономного виконання дій у реальному та цифровому середовищах [1]. Водночас це зумовлює виникнення суттєвої інженерної проблеми, яка полягає у відсутності уніфікованих механізмів забезпечення взаємодії між когнітивним ядром агента та зовнішніми інструментами і сервісами. Тому концепція стандарту Agent Skills виступає як перспективний підхід до формалізації такої взаємодії. Вона передбачає впровадження додаткового рівня абстракції, що забезпечує уніфікацію представлення функціональних можливостей агентів, підвищення безпеки їх використання, а також створення передумов для ефективного масштабування інтелектуальних систем.

Метою роботи є аналіз та обґрунтування застосування стандарту Agent Skills для підвищення ефективності, безпеки та масштабованості інтелектуальних агентних систем, а також визначення його ролі у формуванні уніфікованих підходів до інтеграції та взаємодії компонентів штучного інтелекту. Об'єктом дослідження є процеси побудови та функціонування інтелектуальних агентних систем на основі сучасних технологій штучного інтелекту. Предметом - методи, моделі та засоби стандартизації взаємодії між когнітивним ядром ШІ-агентів і зовнішніми інструментами, зокрема концепція та архітектурні принципи стандарту Agent Skills [2].

Основні завдання: проаналізувати сучасні підходи до побудови ШІ-агентів та виявити їх основні обмеження; дослідити концепцію та архітектурні принципи стандарту Agent Skills; оцінити переваги стандартизації навичок з точки зору масштабованості, безпеки та ефективності; визначити роль middleware та механізмів ізоляції у забезпеченні надійності агентних систем; проаналізувати можливості композиції навичок для формування складних сценаріїв роботи агентів; обґрунтувати перспективи впровадження стандарту Agent Skills у сучасних IT-системах та сформулювати напрями подальших досліджень у сфері стандартизації ШІ-агентів.

У сучасних архітектурах ШІ-агентів домінує підхід «жорсткого зв'язування», що передбачає пряму інтеграцію функціональності через ручне описання JSON-схем, системні промпти та окремі обробники для різних моделей [3]. Такий підхід характеризується низькою портативністю рішень,

перевантаженням контексту через зростання обсягу описів інструментів, а також крихкістю логіки, що потребує постійного оновлення у разі змін API.

Стандарт Agent Skills пропонує альтернативну архітектурну парадигму, засновану на уніфікованому представленні навичок у вигляді маніфестів із розширеним семантичним описом (рис.1). Це забезпечує можливість динамічного пошуку інструментів відповідно до намірів користувача без необхідності попереднього завантаження всіх функцій.

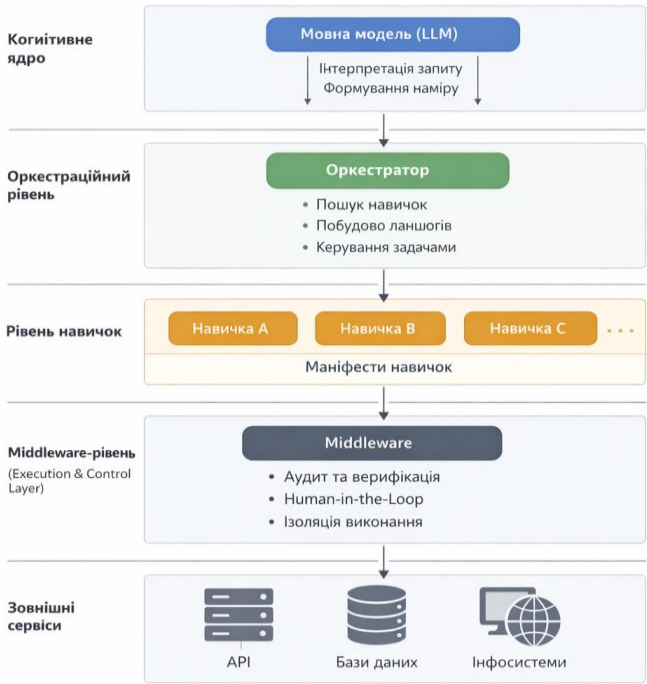


Рисунок 1 - Архітектура ШІ-агента на основі стандарту Agent Skills

Когнітивне ядро (LLM) інтерпретує запит і формує наміри; оркестратор здійснює пошук навичок та побудову ланцюгів виконання; рівень навичок реалізує функціональність через стандартизовані маніфести; middleware забезпечує безпечне та контрольоване виконання; зовнішні сервіси включають API, бази даних та інформаційні системи. Важливою складовою є впровадження проміжного шару (middleware), який забезпечує ізоляцію виконання, підтримку механізмів аудиту, реалізацію підходу human-in-the-loop та підвищення безпеки за рахунок контрольованого середовища виконання. Додатковою перевагою є композитність систем, яка дозволяє

формувати складні робочі процеси шляхом поєднання базових навичок у єдиному середовищі взаємодії. Це сприяє підвищенню автономності агентів і гнучкості їх поведінки.

Впровадження стандарту Agent Skills створює передумови для формування ринку уніфікованих навичок, зниження витрат на розробку за рахунок повторного використання компонентів та досягнення модельної агностичності. Водночас його поширення стримується зацікавленістю компаній у закритих екосистемах і відносною складністю впровадження для малих проєктів. Однак досвід розвитку інформаційних технологій засвідчує, що масштабування складних систем неможливе без стандартизації та уніфікації підходів.

Сучасний етап розвитку штучного інтелекту характеризується переходом від простих діалогових систем до автономних агентів, здатних виконувати прикладні завдання. Водночас існуючі підходи до їх реалізації часто призводять до формування складних, ресурсозатратних і нестабільних систем, що ускладнює їх масштабування та супровід. Запропонований стандарт Agent Skills формує концептуальну основу для подолання зазначених обмежень шляхом уніфікації представлення функціональних можливостей агентів. Його використання сприяє зниженню обчислювальних витрат завдяки оптимізації контексту, підвищенню безпеки через впровадження ізоляції та проміжних механізмів контролю, а також забезпеченню стабільності за рахунок модульності та повторного використання компонентів. Таким чином, стандартизація навичок є ключовим чинником переходу до надійних і масштабованих інтелектуальних систем. Також перспективним напрямом є розроблення формальних моделей опису навичок та механізмів їх семантичної інтероперабельності у гетерогенних середовищах. Доцільним також є дослідження методів автоматизованого виявлення, композиції та оркестрації навичок з урахуванням контексту задачі. Крім того, актуальним є аналіз ефективності модельно-агностичних архітектур та оцінювання їх впливу на зниження витрат і підвищення гнучкості систем штучного інтелекту.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Kostetskyi D.V., Tiahunova M., Kyrychek H.H., Computer system for distance learning with integrated artificial intelligence. *CEUR Workshop*, 2024. Vol-3781. P. 160–174.
2. Liu Y., Li Y., Xu X., Dou Y., Liu D. Heterogeneous skill learning for multi-agent tasks. *Advances in neural information processing systems*, 2022. 35: 37011-37023.
3. Gronauer S., Diepold K. Multi-agent deep reinforcement learning: a survey. *Artificial Intelligence Review*, 2022. 55.2: 895-943.