

КОНЦЕПЦІЯ ЛЮДИНО-МАШИННОЇ ВЗАЄМОДІЇ В КОНТЕКСТІ ШТУЧНОГО ІНТЕЛЕКТУ

Юскович-Жуковська Валентина

*кандидат технічних наук, доцент, доцент кафедри
інформаційних систем та обчислювальних методів
Приватного вищого навчального закладу
«Міжнародний економіко-гуманітарний університет
імені академіка Степана Дем'янчука»*

Богут Олег

*старший викладач кафедри інформаційних систем
та обчислювальних методів,
Приватного вищого навчального закладу
«Міжнародний економіко-гуманітарний університет
імені академіка Степана Дем'янчука»
м. Рівне, Україна*

Людино-машинна взаємодія є одним з ключових напрямів розвитку сучасних інформаційних технологій. На сьогодні існує *"problem of alignment"* — проблема узгодження цілей штучного інтелекту з людським інтелектом, як зберегти баланс між автономністю машини та контролем людини. Згідно досліджень МВФ щодо потенційного впливу штучного інтелекту на світовий ринок праці, виявилось, що майже 40 відсотків світової зайнятості задіяні в роботі зі штучним інтелектом [1, 2], де штучний інтелект (ШІ) доповнює висококваліфіковану людську працю.

Виділимо основні моделі людино-машинної взаємодії та зробимо їх порівняльний аналіз (Таблиця 1).

- Командна модель: спільна робота людини і машини для досягнення однієї мети.

- Автономна модель: машина виконує завдання самостійно, з мінімальним втручанням людини, людина контролює лише результат.

- Модель підтримки прийняття рішень: машина пропонує рекомендації, а людина приймає остаточне рішення.

Тісна співпраця людини та машини призводить до розвитку симбіотичних відносин для спільного рішення складних задач, де людино-машинна взаємодія із застосуванням ШІ потребує усвідомленого підходу.

Оскільки системи штучного інтелекту стають більш автономними, існують побоювання щодо потенційної втрати людського контролю та домінування цифрового інтелекту над людським. Системи штучного



інтелекту повинні контролюватися людьми, а не автоматизацією, щоб запобігти шкідливим наслідкам [3].

Таблиця 1

Порівняння моделей людино-машинної взаємодії

Модель	Рівень автономії	Роль людини	Переваги	Недоліки
Командна	Низький	Активна участь	Спільне прийняття рішень	Залежність від людини
Автономна	Високий	Контроль	Ефективність, швидкість	Ризик помилок ШІ
Підтримка рішень	Середній	Остаточне рішення	Підвищує точність	Необхідна довіра до алгоритмів
Підтримка рішень	Середній	Остаточне рішення	Підвищує точність	Необхідна довіра до алгоритмів

Людиноцентричний ШІ — це сфера, де сходяться дослідження ШІ та взаємодії людини з комп'ютером [4]. При цьому людиноцентричний ШІ ґрунтується саме на людських цінностях і не підмінює переваги машини та цифровий інтелект.

Нобелівську премію з фізики 2024 року отримали вчені Джон Хопфілд та Джеффрі Хінтон за відкриття та винаходи, що забезпечили можливість машинного навчання нейромереж [5]. Вони розробили алгоритми, що дозволяють використовувати структуру машинних мереж для обробки інформації, що дозволяє програмі навчатися та вдосконалюватися самостійно, вже без допомоги людини.

Змодельємо можливі сценарії майбутньої людино-машинної взаємодії в умовах штучного інтелекту.

Керована автономія

Машини будуть все більше брати на себе ініціативу (автомобілі, що самі приймають рішення; медичні системи, що діагностують без лікаря), але залишатимуться у рамках контролю:

- "людина в циклі" (human-in-the-loop) – машина пропонує рішення, людина затверджує;
- прозорі алгоритми (інтерпретованість, explainable AI).

1. Напівавтономне співробітництво

У критичних сферах (армія, економіка, політика) ініціатива буде ділитися:

- людина визначає стратегію, машина — тактику;
- машина бере ініціативу в рутинних і швидких завданнях (реакція на кібератаки, оптимізація логістики).

2. Повна автономія

Якщо ініціатива повністю переходить до ШІ, виникає ризик:

- неконтрольованих наслідків (машина оптимізує не те, що було задумано);
- конфлікту цілей (ШІ починає діяти в інтересах власної оптимізації, а не людини);
- етичних викликів (хто відповідальний за шкоду?).

3. Синергетичний сценарій

У більш оптимістичному баченні:

- машини беруть на себе ініціативу лише там, де це підвищує ефективність і безпеку;
- людина стає “куратором”, а не оператором;
- виникає глобальна система регулювання (подібно до ядерної енергетики).

На підставі даного дослідження складемо порівняльну таблицю трьох можливих сценаріїв розвитку людино-машинної взаємодії в контексті розвитку технологій ШІ у різних сферах життєдіяльності (Табл. 2).

Таблиця 2

Можливі сценарії людино-машинної взаємодії в умовах ШІ

Сфера життя / Сценарій	Оптимістичний (Симбіоз)	Реалістичний (Делегована автономія)	Песимістичний (Втрачений контроль)
Побут	Асистенти допомагають, але остаточне рішення за людиною	Алгоритми керують більшістю процесів (будинки, транспорт), люди втручаються рідко.	Машини повністю керують середовищем, людина лише пасивний користувач.
Медицина	ШІ пропонує діагнози й лікування, лікар затверджує	Автономні системи лікують без повного пояснення методів	Машини самостійно визначають, кого і як лікувати



Сфера життя / Сценарій	Оптимістичний (Симбіоз)	Реалістичний (Делегована автономія)	Песимістичний (Втрачений контроль)
Економіка	Автоматизоване планування під контролем людей	Фінансові алгоритми ухвалюють більшість рішень, люди не завжди розуміють логіку	ШІ диктує економічні стратегії, ігноруючи людські потреби
Військова сфера	Строгий контроль і заборона повної автономії зброї.	Дрони й системи оборони діють самостійно, але з обмеженням людським контролем	Машини ведуть війни без участі людини
Етика та право	Людина зберігає головну роль у прийнятті рішень	Виникають спірні питання відповідальності (хто винен?)	Людина втрачає право впливати на наслідки
Роль людини	Стратег, творець, контролер	Спостерігач і коректор	Пасивний об'єкт системи
Результат	Гармонійне співіснування	Часткова втрата контролю, компроміси	Неконтрольовані та небезпечні наслідки

ЛІТЕРАТУРА

1. USC Annenberg School for Communication and Journalism. (2024). *The ethical dilemmas of AI*. Retrieved from <https://annenberg.usc.edu/research/center-public-relations/usc-annenberg-relevance-report/ethical-dilemmas-ai>
2. International Monetary Fund. (2024). *AI Will Transform the Global Economy. Let's Make Sure It Benefits Humanity*. Retrieved from <https://www.imf.org/en/Blogs/Articles/2024/01/14/ai-will-transform-the-global-economy-lets-make-sure-it-benefits-humanity>
3. European Parliament. (2024). *EU AI Act: first regulation on artificial intelligence*. Retrieved from <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>
4. Raees, M., Meijerink, I., Lykourantzou, I., Khan, V.-J., & Papangelis, K. (2024). A Literature Review on Current Trends in Human-AI Interaction. *arXiv*. Retrieved from <https://arxiv.org/abs/2405.15051>
5. Олена Мусієнко. Нобелівська премія з фізики присуджена за роботу про машинне навчання нейромереж (2024). *TheWashingtonPost* . <https://www.imena.ua/blog/nobel-prize-in-physics/>