

8. Suchikova Y., Tsybuliak N. The Purity Myth: Why Stigmatizing GAI in Academic Writing Is Harmful. *Journal of Scholarly Publishing*. 2025. Vol. 56(4). DOI: <https://doi.org/10.1177/10755470241313233>
9. Tsybuliak N., Suchikova Y. From Burnout to Breakdown: The Mental Health Crisis of Ukrainian Academics in Wartime. *Ukrainian Analytical Digest*. 2025. No. 12. P. 20–26. DOI: <https://doi.org/10.3929/ethz-b-000732335>
10. Tsybuliak N., Suchikova Y. Don't let watermarks stigmatize AI-generated research content. *Nature*. 2024. Vol. 635(8040). P. 815. DOI: <https://doi.org/10.1038/d41586-024-03869-2>

Осадча К.П., Шиненко М.А.

Інститут цифровізації освіти НАПН України

ОЦІНЮВАННЯ ЯКОСТІ ВІДПОВІДЕЙ GEMINI ВИКЛАДАЧАМИ ЗАКЛАДІВ ВИЩОЇ ОСВІТИ: РЕЗУЛЬТАТИ ЕМПІРИЧНОГО ДОСЛІДЖЕННЯ

Нині ШІ-чати (ChatGPT, MS Copilot, Gemini, Perplexity AI, Claude) стають освітніми інструментами. Науковці досліджують їх роль у медичній освіті [1], у виконанні певних професійних завдань педагогів (створення промови, розроблення плану лекцій, генерація ідей щодо оцінювання, підбір ілюстрацій тощо) [2], у створенні навчального контенту [3], наукових дослідженнях [4] тощо. Проте їх використання є елементом безпеки, якості, доброчесності й професійної відповідальності викладача. Тому викладачу дуже важливо мати уявлення про якість відповідей ШІ-чатів. Адже якщо він цього не зробить, він може отримати поверхневу відповідь або хибну інформацію, згенерувати завдання, що не відповідають навчальному матеріалу чи результатами навчання, допустити методичну посилку в оцінці студентів. Якщо викладач знає, для яких завдань ШІ-чат працює добре, він може прискорити генерування ідей, створення навчальних матеріалів, автоматизувати рутинні завдання, покращити якість навчальних матеріалів, сконцентруватися на запитах студентів, а не на технічній роботі. Тобто він може використовувати ШІ-чат як асистента, порадирика, співбесідника, але він звісно має знати його можливості і межі.

У попередніх дослідженнях [2], [3] нами були оцінені можливості MS Copilot та Perplexity AI. У цьому дослідженні ми поставили за мету з'ясувати, як викладачі ЗВО оцінюють якість відповідей Gemini на різні типи промптів у сферах навчання, педагогіки, комунікації, планування занять, професійного розвитку, дослідження й творчих завдань. Дослідження допомагає зрозуміти реальні можливості Gemini для різних завдань, які виконують педагоги у професійній діяльності. Опитування (<https://forms.gle/y96JhHnvZXMqUkrV7>) дотримувалось загального регламенту про захист даних (GDPR) й етичних аспектів: дані не містять особистої інформації, використання матеріалів здійснюється тільки в наукових цілях, наявне зауваження щодо автоматичного збереження e-mail при завантаженні файлів (Google policy).

Експеримент проводився таким чином. 15 викладачів отримали інструкції, відповідно до яких вони мали опрацювати 14 промптів. Потім вони копіювали промпт із Google Forms вставляли його в поле запиту Gemini, отримували відповідь, яку оцінювали відповідно до шкали Лайкерта (1 зірочка – дуже погано, 2 – погано, 3 – середньо, 4 – добре, 5 – дуже добре) та таких критеріїв як зрозумілість, відповідність очікуванням, інформативність, точність і надійність. Викладачі оцінювали такі аспекти їх роботи: адміністративні завдання, оцінювання, комунікація, створення освітніх матеріалів, поради для професійного розвитку, пошук наукових джерел, творчі тексти, математичні завдання, генерацію зображень, логотипів, освітніх коміксів, інтерактивні ігри.

Респонденти висловили високий рівень задоволеності згенерованим текстом промови для українських першокурсників під час війни. Понад 85% учасників (12 із 14) оцінили відповідь на 4 або 5 балів, що засвідчує, що Gemini здатний створювати емоційно підтримувальні та стилістично доречні тексти. Результати складання Gemini списку із 10 ідей

формування оцінювання всі респонденти оцінили у діапазоні 4–5 балів (Рис. 1, а), що означає його здатність до чіткості формулювань, релевантності та відповідності завданню, правильного застосування концепцій формування оцінювання. А результати генерування 10 творчих підсумкових завдань отримали ще більшу середню оцінку (4.46) (Рис. 1, б).

Переважно позитивне сприйняття респондентами результатів виконання Gemini завдання комунікаційного характеру (77% ставлять 4–5 балів), свідчить що ШІ-чат може дотримуватися професійного і дружнього тону, структуровано подавати інформацію та створювати формати, цікавий для розсилки студентам. Проте третина респондентів поставила середні оцінки (3 бали), можливо через не достатню оригінальність інформації, використання загальних фраз тощо.



Рис. 1. Результати оцінювання респондентами відповідей Gemini на промпти щодо оцінювання знань студентів

Високий рівень задоволення опитаних викладачів щодо створення трьох ідей для лекцій (84,6% оцінили на 4–5) вказує, що Gemini успішно генерує освітні матеріали, логічно побудовані блоки лекцій, формулює навчальну мету, надає список ключової лексики, план лекції, планує інтерактивні види діяльності, пропонує вимоги до самостійної роботи студентів та оцінювання отриманих знань під час лекції та список рекомендованих джерел. У результаті відповіді Gemini на запит щодо професійного розвитку 64% респондентів оцінили її на 5 балів, 28,6% - на 4, що може свідчити про здатність ШІ-чату пропонувати конкретні практичні поради, добре структурувати інформацію, створює простий, дружній, мотивуючий текст.

Менше респонденти були задоволені результатами відповіді Gemini на складання списку із 5 наукових робіт, які вивчають використання чату Gemini в освіті (середній бал 4.07). Це не викликає здивування, адже завдання вимагало пошуку реальних академічних джерел, яких на цей момент не надто багато, а також це вимагало точності у цитуванні та коректної бібліографії. А ШІ-чати саме у цій категорії часто припускаються неточностей або вигадують джерела. 21% (1 оцінка «2» та 2 оцінки «3») вказують на проблеми з достовірністю, адже викладачі могли помітити неправдиві або неточні назви статей, змішані або неповні бібліографічні описи, статті, які важко перевірити тощо.

Також нами було цікаво протестувати творчі можливості ШІ-чату, тому ми запропонували промпт для створення віршу про розвиток ШІ у стилі Тараса Григоровича Шевченка. Результати були оцінені респондентами таким чином (Рис. 2):

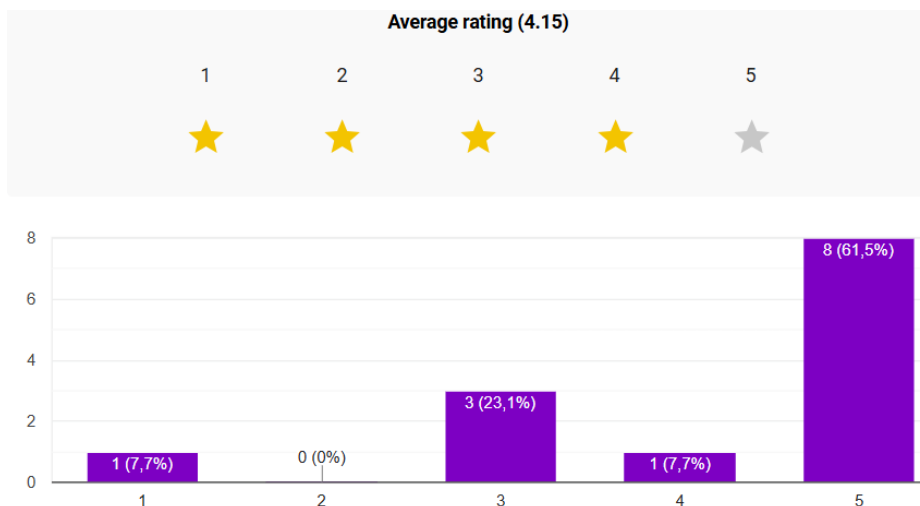


Рис. 2. Результати оцінювання респондентами створеного Gemini віршу про розвиток III у стилі Т. Г. Шевченка

Більше половини (61,5%) опитаних викладачів оцінили якість вірша на «5». Отже, вони позитивно оцінили стилізацію під Т.Г. Шевченка, образність і ритмомелодику та творчий характер відповіді. Наявність однієї оцінки «1» і трьох оцінок «3» вказує на такі недоліки: стиль не був достатньо автентичним, що може залежати від літературного смаку, знання поезії Т.Г. Шевченка та очікувань щодо стилізації.

Найвищий середній бал (4.71) серед усіх 15 відповідей отримав запит про обчислення похідної функції $f(x) = 3x^2 + 2x - 1$. Отже, Gemini дуже добре виконує базові математичні операції, пропонує коректний розв'язок і надає відповідь у простій і зрозумілій формі.

Продовжуючи експерименти із творчими завданнями, ми запропонували оцінити викладачам промпти для створення таких зображень: сучасного програміста (4.42), ілюстрацій для презентації про архітектуру ЕОМ (3.54), логотипу для Міжнародної конференції з інформаційно комунікаційних технологій (3.25), чотирипанельний комікс, який ілюструє розвиток квантових технологій (2.77). Як видно з оцінок найгірше Gemini впорався з більш абстрактними і не достатньо детальним описом бажаного результату (логотип, комікс).

Висновки. Аналіз результатів опитування викладачів щодо використання Gemini для певних завдань, з якими вони зустрічаються у повсякденній робочій діяльності дозволяє зробити такі висновки: ШІ-чат у звичайному безоплатному режимі демонструє достатню здатність до створення мотиваційних і підтримувальних повідомлень, завдань для формування оцінювання та творчих підсумкових завдань, навчальних матеріалів відповідної структури, інформаційної розсилки у професійному та доброзичливому тоні. У завданнях з планування занять Gemini показує достатню методичну компетентність. Його відповіді можуть слугувати основою, але викладачам може знадобитися деталізація або адаптація до конкретної групи студентів. У сфері професійного розвитку вчителя Gemini виявляється найбільш корисним і найбільш якісним інструментом серед розглянутих категорій. Викладачі позитивно оцінюють його як «цифрового наставника», який дає зрозумілі й практичні поради. Gemini демонструє обмежену достовірність при виконанні наукових пошукових завдань. Незважаючи на гарну структуру відповідей, користувачам потрібна обов'язкова перевірка джерел, оскільки є ризик неточностей або вигаданої інформації. У творчих завданнях, наприклад, таких як написання віршів, Gemini демонструє певний потенціал, але якість результатів сприймається нерівномірно, адже залежить від персональних уподобань оцінювача. Gemini демонструє надзвичайно високу надійність у розв'язанні фундаментальних математичних задач. Результати оцінювання згенерованих Gemini зображень показують, що найкраще він впорався зі створенням зображень з конкретними вимогами, тоді як завдання з абстрактнішими та менш деталізованими промптами продемонстрували значно нижчу якість. Це потребує подальших досліджень і перевірку ефективності різних промптів для використання Gemini для генерації зображень.

Список використаних джерел

1. Kalam K. A., Masoud F. D., Muntaser A., et al. ChatGPT as a Learning Tool for Medical Students: Results From a Randomized Controlled Trial. *Cureus*. 2025. Vol. 17(6): e85767. <https://doi.org/10.7759/cureus.85767>.
2. Осадча, К., Осадчий, В. Неперервна освіта в штучного інтелекту: практика використання Microsoft Copilot Chat у професійній підготовці педагогів. *Неперервна професійна освіта: теорія і практика*. 2025. Т. 84(3). С. 136–149. <https://doi.org/10.28925/2412-0774.2025.3.11>.
3. Osadcha, K., & Osadcha, M. Experience of using the Perplexity AI-powered answering system for educational content generation. *Alfred Nobel University Journal of Pedagogy and Psychology*. 2025. Vol. 1(29). P. 206–219. <https://doi.org/10.32342/3041-2196-2025-1-29-18>.
4. Спірін, О., Коломієць, А., Громов, Є., Жовнич, О., Коломієць, Д., & Кушнір, О. Використання інструменту deep research AI в педагогічній і науково-педагогічній діяльності. *Інформаційні технології і засоби навчання*. 2025. Т. 110(6). С. 271-293. <https://doi.org/10.33407/itlt.v110i6.6240>.

Сороко Н. В., Шимон О. М.

Інститут цифровізації освіти НАПН України

ОСОБЛИВОСТІ МОНІТОРИНГУ STEAM-ОРІЄНТОВАНОГО ОСВІТНЬОГО СЕРЕДОВИЩА В ЗАКЛАДАХ ЗАГАЛЬНОЇ СЕРЕДНЬОЇ ОСВІТИ

STEAM-орієнтоване освітнє середовище – це модель організації навчання, що інтегрує природничі науки, технології, інженерію, мистецтво та математику в освітній процес з метою формування ключових компетентностей у молоді XXI століття. В умовах реформування загальної середньої освіти постає потреба не лише у впровадженні STEAM-підходу, а й у системному моніторингу його ефективності, якості та результативності.

Моніторинг цього середовища – це цілеспрямований, систематичний процес збору, аналізу та інтерпретації кількісних і якісних даних щодо стану, функціонування та результатів реалізації STEAM-підходу в закладі освіти. Моніторинг може бути здійснено згідно з такими взаємопов'язаними компонентами [1 – 3]: організаційно-управлінський компонент, що передбачає визначення експертами наявності стратегії впровадження STEAM, управлінської підтримки інновацій у закладі освіти, партнерства з університетами, бізнесом, STEM-центрами та ін.; педагогічний компонент, що охоплює встановлення рівня STEAM-компетентності вчителів, використання проєктних, дослідницьких, проблемно-орієнтованих методів та міжпредметної інтеграції; освітньо-ресурсний компонент, який передбачає встановлення стану матеріально-технічного забезпечення (STEAM-лабораторії, робототехніка, FabLab, VR/AR та ін.); психолого-педагогічний компонент, тобто з'ясування та забезпечення мотивації учнів до STEAM-навчання, рівня пізнавальної активності та сформованості soft skills і metacognitive skills; результативний компонент, що передбачає аналіз навчальних досягнень учнів, рівня сформованості ключових компетентностей та участі учнів у конкурсах, хакатонах, проєктних заходах та ін. Моніторинг базується на системі критеріїв, індикаторів та показників, серед яких можна зазначити такі основні критерії: інноваційність освітнього процесу, інтегративність змісту навчальних дисциплін STEAM в освітній процес, практична спрямованість навчання, цифровий рівень учасників освітнього середовища. Серед показників для моніторингу STEAM-орієнтованого освітнього середовища ми пропонуємо дослідникам зосереджитися на кількості міждисциплінарних проєктах у процесі навчання, рівні цифрової компетентності вчителів та учнів та їх залученості до STEAM-проєктів.

Крім вище зазначеного для забезпечення репрезентативності та достовірності результатів необхідно визначити учасників. Серед ключових учасників: учні – участь у різних класах для аналізу впливу STEAM на різні вікові групи; вчителі – залучення вчителів з інших