

## **ЯК ВИМІРЯТИ ШІ-ГРАМОТНІСТЬ: ОГЛЯД НАЯВНИХ ІНСТРУМЕНТІВ ДЛЯ ОЦІНЮВАННЯ ШІ-ГРАМОТНОСТІ**

Штучний інтелект (ШІ) став частиною нашого повсякденного життя, саме тому грамотність у цій сфері або ШІ-грамотність є одним із ключових пріоритетів сучасної освіти, спрямованої на розвиток знань, навичок, ставлень і цінностей фахівця у XXI столітті. Визначення поняття ШІ-грамотності, запропоноване американськими дослідниками (Long, Magerko, 2020), є одним із найчастіше цитованих: «набір компетенцій, який дозволяє людям критично оцінювати технології штучного інтелекту, спілкуватися та ефективно співпрацювати з штучним інтелектом, а також використовувати штучний інтелект як онлайн-інструмент вдома та на робочому місці» [7].

Учені досягли певних успіхів у дослідженні поняття ШІ-грамотності: вивчають його ключові аспекти (розуміння основних концепцій ШІ, їхня практична реалізація, етичні зобов'язання), визначають структуру та можливі рівні ШІ-грамотності та створюють інструменти для її оцінювання. Лише за останні два роки дослідниками було розроблено низку інструментів для вимірювання ШІ-грамотності, які враховують різні контексти, аудиторії та аспекти.

В Україні поняття ШІ-грамотності вперше було запропоновано у 2024 році в рамках Проекту «Інструктивно-методичні рекомендації щодо ШІ в ЗЗСО». Робоча група українських дослідників розглядає ШІ-грамотність як «розуміння учасниками освітнього процесу основних принципів відповідального застосування систем штучного інтелекту, володіння навичками розпізнавання, коли використовується штучний інтелект, усвідомлення його обмежень та ризиків, пов'язаних з його невідповідальним використанням» [1]. Нажаль, на даний момент поки немає валідних україномовних інструментів для вимірювання ШІ-грамотності, зупинимось більш детально на наявних розробках іноземних дослідників.

Існуючі інструменти для оцінювання ШІ-грамотності можна умовно поділити на два типи: анкети самооцінки та тести, що передбачають виконання завдань. Аналіз літератури свідчить про те, що хоча суб'єктивні методи, такі як анкети самооцінки, можуть

вимірювати сприйняття здатності використовувати ІІІ, наукові та об'єктивні інструменти, такі як тести, повинні оцінювати рівень обізнаності студентів щодо ІІІ [2]. Учені підкреслюють, що оцінювання, засноване на виконанні завдань, безпосередньо аналізує компетентності, забезпечуючи надійніші результати. Це особливо важливо для ІІІ-грамотності, де існує розрив між уявленнями учасників про свої знання та їхньою реальною здатністю ефективно використовувати інструменти ІІІ [9].

Метою нашої доповіді є огляд найбільш вдалих наявних інструментів для вимірювання ІІІ-грамотності, розроблених сучасними дослідниками у 2023–2024 роках.

Так, наприклад, опитувальник SNAIL [6] оцінює знання респондентів, їхні уміння критично аналізувати і практично застосувати ІІІ. Він складається з 31 питання і може бути використаний як для індивідуальної, так і групової оцінки ІІІ-грамотності фахівців, які не є спеціалістами в сфері систем штучного інтелекту. Автори цього опитувальника підкреслюють, що він може застосовуватися для оцінки ефективності навчання на курсах із ІІІ.

Розроблена китайськими ученими анкета AILS-CCS [10] уявляє собою 12-завданнєвий інструмент для кількісної оцінки грамотності у сфері ІІІ. Вимірює обізнаність студентів, їхні уміння використовувати та оцінювати ІІІ, а також знання етичних аспектів. Ця анкета була розроблена для китайських студентів коледжу і апробована на цій аудиторії. Автори запропонували чотирьохкомпонентну модель, яка адекватно відображає ІІІ-грамотність, і довели, що вона значно корелює з цифровою грамотністю, ставленням до роботів і повсякденним використанням ІІІ користувачами.

Відмітимо тест MAIRS-MS, який було створено для студентів-медиків [5] з метою оцінювання їхніх когнітивних здібностей, навичок та етичних знань в сфері ІІІ. Результати апробації цього тесту серед студентів медичних факультетів двох університетів Туреччини доводять, що він є валідним і надійним інструментом для оцінки та моніторингу рівня сприйнятої готовності студентів до технологій ІІІ та їх застосувань. Як підкреслюють автори цього тесту, MAIRS-MS може стати цінним інструментом для розвитку навчальних програм у медичній та медико-науковій освіті, допомагаючи визначати потреби студентів та оцінювати їхній рівень готовності після закінчення курсів.

Варто ще згадати 32-завданнєву анкету для самооцінки AILQ [8], яка призначена для учнів середніх шкіл і охоплює афективні,

поведінкові, когнітивні та етичні аспекти грамотності в сфері ІІІ. Результати пілотного дослідження, у якому взяли участь 363 учні середніх шкіл у Гонконзі для аналізу психометричних характеристик інструменту, підтвердили доцільність чотирьохфакторної структури AILQ та довели високу надійність і валідність інструменту.

Хочеться відмітити опитувальник американських дослідників [3], який був створений для учителів нетехнічних спеціальностей і ґрунтується на системі з 17 компетенцій ІІІ, визначених Long та Magerko [7]. Інструмент оцінювання пройшов ретельну перевірку чотирма експертами з ІІІ та практичне тестування серед вчителів, що проходять підготовку та працюючих вчителів з мінімальними або взагалі без технічних знань. Остаточна версія інструменту складається з 25 пунктів, кожен з яких спрямований на вимірювання однієї або двох компетенцій. Цей інструмент може бути корисним для освіти в галузі ІІІ, зокрема для фахівців нетехнічних спеціальностей.

Психометрична шкала SAIL4ALL [9] була розроблена для комплексної оцінки ІІІ-грамотності серед дорослого населення, незалежно від середовища, в якому вона застосовується. Як наголошують автори, ця анкета може слугувати цінним інструментом для визначення як фактичних, так і суб'єктивних знань про ІІІ. Анкета містить 56 питань, розподілених за чотирма темами (Що таке ІІІ? Що може робити ІІІ? Як працює ІІІ? Як слід використовувати ІІІ?) які можуть використовуватися в комбінації або незалежно. Тест включає завдання у двох форматах відповіді («так/ні» та з 5-бальною шкалою оцінки), кожен з яких застосовується залежно від контексту

Дуже важливою є розробка тесту оцінки грамотності в сфері генеративного штучного інтелекту GLAT [4] групи австралійських учених. Він ураховує особливості генеративного штучного інтелекту і складається з 20 питань із множинним вибором і розроблений відповідно до встановлених процедур у психологічному та освітньому вимірюванні. Структурна валідність і надійність тесту були підтверджені на основі відповідей 355 студентів вищих навчальних закладів із використанням класичної теорії тестів та теорії відповідей на завдання. Результати експерименту свідчать про те, що GLAT пропонує надійний і більш об'єктивний метод оцінювання грамотності в сфері генеративного ІІІ.

Нами було зроблено огляд сучасних наявних розробок іноземних дослідників валідних інструментів для визначення рівня ІІІ-грамотності. До подальших напрямів наших досліджень належить переклад та адаптація одного з описаних валідних

тестів III-грамотності та апробація цього інструментарію у Мелітопольському державному педагогічному університеті серед майбутніх учителів математики.

### Література:

1. Проект «Інструктивно-методичні рекомендації щодо III в ЗЗСО» Міністерство освіти і науки України.(2024). URL: <https://mon.gov.ua/static-objects/mon/sites/1/news/2024/05/21/Instruktyvno.metodychni.rekomendatsiyi.shchodo.SHI.v.ZZSO-22.05.2024.pdf>
2. Chiu T., Chen Y., Yau K., Chai C., Meng H., King I., Wong S., Yamg Y. (2024). Developing and validating measures for AI literacy tests: from self-reported to objective measures. URL: <https://doi.org/10.1016/j.caeai.2024.100282>
3. Ding, L., Kim, S., & Allday, R. A. (2024). Development of an AI literacy assessment for non-technical individuals: What do teachers know?. *Contemporary Educational Technology*, 16(3), ep512. URL: <https://doi.org/10.30935/cedtech/14619>
4. Jina Y., Martinez-Maldonado R., Gašević and D, Yana L. (2024). GLAT: The Generative AI Literacy Assessment Test. URL: <https://arxiv.org/abs/2411.00283>
5. Karaca, O., Çalışkan, S. A., Demir, K., (2021). Medical artificial intelligence readiness scale for medical students (MAIRS-MS) – development, validity and reliability study. *BMC Medical Education* 21, 1–9.
6. Laupichler, M. C., Aster, A., Haverkamp, N., Raupach, T. (2023). Development of the “scale for the assessment of non-experts’ ai literacy”–an exploratory factor analysis. *Computers in Human Behavior Reports* 12, 100338.
7. Long, D., & Magerko, B. (2020). What Is AI Literacy? Competencies and Design Considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. URL: <https://doi.org/10.1145/3313831.3376727>
8. Ng, D. T.K., Wu, W., Leung, J. K.L., Chiu, T. K.F., Chu, S. K.W.(2024). Design and validation of the ai literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology* 55, 1082–1104
9. Soto-Sanfiel M., Angulo-Brunet A., & Lutz C. (2024). The Scale of Artificial Intelligence Literacy for all (SAIL4ALL): A Tool for Assessing Knowledge on Artificial Intelligence in All Adult Populations and Settings. URL: <https://doi.org/10.31235/osf.io/bvyku>
10. Wang, Y. Y., & Chuang, Y. W. (2023). Artificial intelligence self-efficacy: Scale development and validation. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12015-w>