

УДК 004.8

DOI 10.36910/10.36910/6775-2313-5352-2025-26-11

Сахнюк І. О., Федосєєва І. К., Битков М. Х.

Технічний центр Національної академії наук України

СТАНДАРТИЗАЦІЯ ПЕВНИХ ПОЛОЖЕНЬ У СФЕРІ ТЕХНОЛОГІЇ ШТУЧНОГО ІНТЕЛЕКТУ

У статті розглянуто та проаналізовано міжнародну нормативну базу стандартів Міжнародної організації зі стандартизації (ISO) і Міжнародної електротехнічної комісії (IEC), пов'язаних із технологією AI, яка налічує майже 400 стандартів. Зосереджено увагу на декількох ключових стандартах штучного інтелекту, що забезпечують основу для відповідального та етичного використання технологій штучного інтелекту. Наведено огляд сфери їхнього застосування. Зазначені у статті стандарти охоплюють такі сфери, як: термінологія та поняття штучного інтелекту; система управління; управління ризиками; конфіденційність; упередженість; прозорість; підзвітність тощо. Дотримуючись цих стандартів, організації можуть працювати над тим, щоб їхні системи штучного інтелекту були справедливими, прозорими та дотримувалися етичних принципів.

На основі проведеного аналізування показано, що штучний інтелект все частіше застосовується в усіх секторах, які використовують інформаційні технології. Беручи це до уваги, запропоновано включати до щорічних програм робіт із національної стандартизації прийняття та/або розроблення згармонізованих стандартів ISO та IEC штучного інтелекту.

Ключові слова: інформаційні технології, штучний інтелект, стандарт, система управління, структура систем штучного інтелекту.

Постановка проблеми. Сприятливою умовою з розроблення продукції і послуг та освоєння ринків збуту є впровадження згармонізованих міжнародних і європейських стандартів. Стандарти забезпечують технологіям, продукції і послугам сумісність, оскільки стандарт містить детальну інформацію про використання та їхній зміст. Застосування стандартів надає перевагу інноваційним технологіям у науці та презумпцію відповідності у виробництві, машинобудуванні, приладобудуванні, бізнесі тощо. Існує потреба в забезпеченні дослідників відповідними інструментами для допомоги в розроблянні та застосуванні інформаційних технологій, зокрема технології штучного інтелекту (Artificial Intelligence, AI), і в цьому випадку стандартизація як інструмент передачі знань відіграє важливу роль.

У 2021 році наказом Національного органу стандартизації - ДП «УкрНДНЦ» «Про створення технічного комітету стандартизації «Штучний інтелект» від 10.11.2021 № 416 був створений технічний комітет стандартизації ТК 200 з відповідною назвою «Штучний інтелект». Організацією, що виконує функції секретаріату ТК 200, визнано Міністерство цифрової трансформації України. Попри те, до сьогодні залишається невирішеною проблема щодо наявної чинної нормативної бази з питань технології штучного інтелекту. Для єдиного розуміння цієї технології та визначення ключових термінів, ризиків та переваг щодо застосування штучного інтелекту доречно прийняти в Україні відповідні міжнародні стандарти ISO/IEC, ISO/IEC TR, ISO/IEC TS як національні методом перекладу.

Аналіз останніх досліджень. На сьогодні розкриття потенціалу технології штучного інтелекту безперервно вдосконалюється. Розробникам та користувачам штучного інтелекту потрібно адаптуватися до цих змін, зосередившись на різних способах. Тут постає питання, як розумно використовувати весь потенціал можливостей штучного інтелекту, не стаючи жертвою ризиків? Відокремити можливості та ризики можна лише за умови надійного управління. Ось чому важливо ознайомитися з системою управління штучним інтелектом, що викладена у першому у світі стандарті ISO/IEC 42001:2023 Information technology - Artificial intelligence - Management system (Інформаційні технології - Штучний інтелект - Система управління), та який закладає основу для етичного, безпечного та перспективного використання технології штучного інтелекту в різних сферах його застосування [1].

Окремо важливо зазначити, що 1 серпня 2024 року набув чинності Європейський закон про штучний інтелект [2]. Цей закон є першим у світі всеосяжним нормативно-правовим актом щодо штучного інтелекту. Закон про штучний інтелект покликаний гарантувати, що штучний інтелект, розроблений і використовуваний в ЄС, заслуговує довіри з гарантіями захисту основних прав людей. Регламент спрямований на створення гармонізованого внутрішнього

ринку штучного інтелекту в ЄС, заохочення впровадження цієї технології та створення сприятливого середовища для інновацій та інвестицій.

Мета роботи та постановка проблеми полягає в тому, щоб звернути увагу наукової спільноти на важливість послідовного використання та більш чіткого розуміння технології штучного інтелекту через стандартизацію. Під час розроблення або комерціалізації відповідних застосувань використання згармонізованих міжнародних та європейських стандартів у сфері штучного інтелекту на національному рівні сприятиме однаковому розумінню, презумпції відповідності та послідовному впровадженню технології штучного інтелекту.

Викладання основного матеріалу. За своєю суттю штучний інтелект означає здатність машини або комп'ютерної системи виконувати завдання, які зазвичай вимагають людського інтелекту. Вона передбачає програмування систем для аналізу даних, навчання на досвіді та прийняття розумних рішень – керуючись людським внеском. Штучний інтелект має потенціал революціонізувати різні галузі, дозволяючи машинам вирішувати складні проблеми та мислити інтуїтивно, виходячи за рамки простої автоматизації. Це охоплює різні підгалузі та технології, такі як машинне навчання та обробка природної мови.

Машинне навчання (Machine Learning, ML) відноситься до здатності комп'ютерної системи навчатися на основі даних без явного програмування. Усе починається з даних. Алгоритми ML навчаються на величезних наборах даних, які вони вчаться аналізувати, щоб виявляти закономірності, взаємозв'язки та тенденції. Ці закономірності потім можуть бути використані для прогнозування або прийняття рішень на основі нових, невидимих даних. Машинне навчання має надзвичайні переваги та застосування в нашому повсякденному житті. Іноді використовуються машинне та глибоке навчання як синоніми, але є різниця. Глибоке навчання – це різновид машинного навчання. Однак глибоке навчання може аналізувати більше типів інформації, виконувати складніші операції, робити більш тонкі та глибокі прогнози на основі даних тощо.

У стандарті ISO/IEC 23053:2022 Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML) (Структура систем штучного інтелекту (AI) з використанням машинного навчання (ML)) встановлюється структура штучного інтелекту (AI) і машинного навчання (ML) для опису загальної системи AI з використанням технології ML [3]. ML - це царина AI, яка використовує обчислювальні методи, щоб дозволити системам навчатися на основі даних або досвіду. Іншими словами, системи ML розробляються за допомогою оптимізації алгоритмів, щоб відповідати навчальним даним, або покращують їхню продуктивність на основі максимізації винагороди. Методи ML включають глибоке навчання, яке також розглядається в цьому документі. Структура описує компоненти системи та їхні функції в системі AI. Цей документ застосовується до організацій усіх типів і обсягів, у тому числі державних і приватних компаній, державних установ і некомерційних організацій, які впроваджують або використовують системи AI. У цьому документі використовуються такі терміни, як знання, навчання та рішення. Цей документ має на меті надати структуру для опису систем AI, які використовують машинне навчання. Встановлюючи загальну термінологію та загальний набір понять для таких систем, цей документ забезпечує основу для чіткого пояснення систем і різних визначень, які стосуються їх розробки та використання. Цей документ призначений для широкої аудиторії та забезпечує основу для інших стандартів, спрямованих на конкретні аспекти систем ML та їхні компоненти.

Обробка природної мови (Natural language processing, NLP) - це царина штучного інтелекту, яка зосереджена на тому, щоб дозволити комп'ютерам розуміти людську мову та взаємодіяти з нею. NLP включає в себе широкий спектр технологій, що дозволяють комп'ютерам розуміти структуру і значення людської мови. Він охоплює такі завдання, як розпізнавання мови, розуміння природної мови та генерація природної мови. Аналізуючи та інтерпретуючи мову, NLP дозволяє комп'ютерам отримувати корисну інформацію, відповідати на запитання та вести бесіди. Завдяки здатності синтезувати, аналізувати та діяти з величезними обсягами даних за лічені секунди, штучний інтелект є надзвичайно потужним. Як і у випадку з будь-якою потужною технологією, дуже важливо впроваджувати її відповідально, щоб максимізувати її потенціал і мінімізувати негативні наслідки. Ще однією з ключових етичних проблем, пов'язаних зі штучним інтелектом, є конфіденційність. Оскільки системи штучного інтелекту збирають величезні обсяги даних із баз даних по всьому світу, існує потреба забезпечити захист особистої інформації та її відповідальне використання.

Технологія AI робить машини та комп'ютерні програми розумними, дозволяючи їм виконувати завдання, які зазвичай вимагають людського інтелекту. Загалом системи штучного

інтелекту працюють, обробляючи величезні обсяги даних, шукаючи закономірності, за якими можна моделювати власне прийняття рішень.

Відповідно до ISO/IEC 22989:2022 Information technology - Artificial intelligence - Artificial intelligence concepts and terminology Інформаційні технології - (Штучний інтелект - Поняття та термінологія штучного інтелекту) під штучним інтелектом мається на увазі «<дисципліна> дослідження та розробка механізмів і застосувань систем AI» [4, 3.1.3]. Система AI генерує результати, такі як вміст, прогнози, рекомендації або рішення для заданого набору цілей, визначених людиною [4, 3.1.4]. Сконструйована система може використовувати різні методи та підходи, пов'язані зі штучним інтелектом, задля розробки моделі для представлення даних, знань, процесів тощо, які можна використовувати для виконання завдань. Системи штучного інтелекту розроблені для роботи з різними рівнями автоматизації.

Оскільки можливості штучного інтелекту зростають у геометричній прогресії, виникають глибокі занепокоєння щодо конфіденційності, упередженості, нерівності, безпеки та захисту. Вивчення того, як ризики штучного інтелекту впливають на користувачів, має вирішальне значення для забезпечення відповідального та сталого розгортання цих технологій. Сьогодні, як ніколи, компаніям потрібна структура, яка б скеровувала їх на шляху до штучного інтелекту. Міжнародний стандарт ISO/IEC 42001:2023 відповідає цій потребі і містить настанови щодо управління технологіями штучного інтелекту. Він пропонує систематичний підхід до вирішення проблем, пов'язаних із впровадженням штучного інтелекту у визнану структуру системи управління, що охоплює такі сфери, як етика, підзвітність, прозорість та конфіденційність даних. Розроблений для нагляду за різними аспектами штучного інтелекту, він забезпечує інтегрований підхід до управління проектами зі штучного інтелекту, від оцінки ризиків до ефективного запобігання цим ризикам.

Як стандарт системи управління, ISO/IEC 42001 побудований на процесі встановлення, впровадження, підтримки та постійного вдосконалення штучного інтелекту. Цей підхід важливий із багатьох причин: по-перше, він гарантує, що цінність штучного інтелекту визнається і існує правильний рівень нагляду. По-друге, система управління дозволяє організації активно адаптувати свій підхід відповідно до експоненціального розвитку технології, а також заохочує організації проводити оцінку ризиків AI. Прогнозується, що зі швидким поширенням штучного інтелекту в усьому світі ISO/IEC 42001 стане невід'ємною частиною успіху організації, наслідуючи інші стандарти систем менеджменту, такі як ISO 9001 для якості, ISO 14001 для навколишнього середовища та ISO/IEC 27001 для IT-безпеки.

Що таке система управління штучним інтелектом? Система управління штучним інтелектом полягає в тому, як організація керує своїми проектами зі штучним інтелектом. Вся справа в налаштуванні правил і методів, щоб переконатися, що штучний інтелект використовується відповідально та ефективно. Ця система допомагає керувати всім, від оцінки ризиків до відповідального впровадження штучного інтелекту в дію. Структура управління штучним інтелектом має структурований спосіб управління ризиками та можливостями, пов'язаними зі штучним інтелектом. Він охоплює ключові компоненти, такі як прозорість, зрозумілість і автономність, надаючи організаціям чіткі настанови щодо використання штучного інтелекту відповідно до нових нормативних актів (наприклад, Закон ЄС про штучний інтелект). Зокрема у стандарті ISO/IEC TR 24030:2024 Information technology - Artificial intelligence (AI) - Use cases (Інформаційні технології - Штучний інтелект (AI) - Приклади використання) наведено колекцію репрезентативних випадків використання програм штучного інтелекту в різних галузях [5]. Загалом у період із липня 2018 року до кінця червня 2022 року експерти надали 187 випадків використання AI. У цьому стандарті термін «випадки використання» означає «випадки використання, вибрані з поданих». Цей документ вибрав 81 робочий випадок з усіх поданих матеріалів. За умови відповідального управління та нагляду штучний інтелект може реалізувати свій потенціал як надзвичайно позитивний технологічний розвиток.

Впровадження стандарту ISO 14001 має на меті допомогти організаціям відповідально виконувати свою роль щодо систем штучного інтелекту (наприклад, використовувати, розробляти, контролювати або надавати продукти чи послуги, які використовують штучний інтелект). Штучний інтелект потенційно викликає певні обміркування, такі як:

- використання штучного інтелекту для автоматичного прийняття рішень, іноді в непрозорий і незрозумілий спосіб, може вимагати специфічного управління, що виходить за рамки управління класичними IT-системами;

- використання аналізу даних, розуміння та машинного навчання, а не логіки, закодованої людиною, для проєктування систем, що як збільшує можливості застосування систем штучного інтелекту, так і змінює спосіб розробки, обґрунтування та розгортання таких систем;
- використання систем штучного інтелекту, які виконують безперервне навчання, змінюють свою поведінку тощо.

Цей документ містить вимоги до створення, впровадження, підтримки та постійного вдосконалення системи управління AI в контексті організації. Очікується, що організації зосередять застосування вимог на функціях, які є унікальними для штучного інтелекту. Певні функції штучного інтелекту, такі як здатність постійно навчатися та вдосконалюватися або відсутність прозорості чи зрозумілості, можуть вимагати різних запобіжних заходів, якщо вони викликають додаткові занепокоєння порівняно з тим, як традиційно виконувалося б завдання. Впровадження системи управління на основі штучного інтелекту для розширення існуючих структур управління є стратегічним рішенням для організації.

Потреби та цілі організації, процеси, розмір та структура, а також очікування різних зацікавлених сторін впливають на створення та впровадження системи управління штучним інтелектом. Іншим набором факторів, які впливають на створення та впровадження системи управління штучним інтелектом, є численні варіанти використання штучного інтелекту та необхідність знайти належний баланс між механізмами управління та інноваціями. Організації можуть вирішити застосовувати ці вимоги, використовуючи підхід, заснований на оцінці ризиків, щоб забезпечити належний рівень контролю для конкретних випадків використання штучного інтелекту, послуг або продуктів у межах сфери діяльності організації.

Система управління штучним інтелектом повинна бути інтегрована з процесами організації та загальною структурою управління. Конкретні питання, пов'язані з AI, повинні враховуватися під час проєктування процесів, інформаційних систем і засобів управління. Ключовими прикладами таких процесів управління є:

- визначення організаційних цілей, залучення зацікавлених сторін та організаційної політики;
- управління ризиками та можливостями;
- процеси управління проблемами, пов'язаними з надійністю систем штучного інтелекту, такими як безпека, справедливість, прозорість, якість даних та якість систем штучного інтелекту протягом усього їхнього життєвого циклу;
- процеси управління постачальниками, партнерами та третіми сторонами, які надають або розробляють AI-системи для організації.

Цей документ містить рекомендації щодо застосування засобів керування для підтримки таких процесів. Цей документ дозволяє уникнути конкретних вказівок щодо процесів управління. Організація може поєднувати загальноприйняті рамки, інші міжнародні стандарти та власний досвід для впровадження критично важливих процесів, таких як управління ризиками, управління життєвим циклом та управління якістю даних, які підходять для конкретних випадків використання штучного інтелекту, продуктів або послуг у межах сфери.

Одним із таких міжнародних стандартів у сфері штучного інтелекту є ISO/IEC 23894:2023 Information technology - Artificial intelligence - Guidance on risk management (Інформаційна технологія - Штучний інтелект - Настанова щодо управління ризиками) [6], який зосереджений на управлінні ризиками в системах штучного інтелекту. Цей стандарт спрямований на забезпечення того, щоб алгоритми та моделі штучного інтелекту були зрозумілими та могли бути перевірені на упередженість та справедливість, тим самим зміцнюючи довіру та впевненість у системах штучного інтелекту.

Організація, яка дотримується вимог зазначених вище стандартів, може надати докази своєї відповідальності та підзвітності щодо своєї ролі стосовно систем штучного інтелекту.

Висновки. У міру того, як штучний інтелект стає все більш складним, ми можемо очікувати, що він відіграватиме важливу роль у вирішенні глобальних проблем і прискоренні пошуку рішень.

ISO і IEC утворюють спеціалізовану систему всесвітньої стандартизації. Національні органи стандартизації держав, які є членами ISO або IEC, беруть участь у розробці міжнародних стандартів через технічні комітети, створені відповідною організацією для вирішення окремих сфер технічної діяльності. Технічні комітети ISO та IEC співпрацюють у сферах, що становлять взаємний інтерес.

За результатами аналізування міжнародної нормативної бази стандартів ISO/IEC, ISO/IEC

TR, ISO/IEC TS, пов'язаних із технологією AI, визнано, що ці стандарти відповідають сумісності систем штучного інтелекту, гарантуючи, що технології штучного інтелекту можуть безперервно працювати разом і ефективно обмінюватися даними. Це особливо важливо, оскільки штучний інтелект стає все більш інтегрованим у різні сфери та програми.

Оскільки розробка та впровадження штучного інтелекту продовжує прискорюватися, розробка та гармонізація стандартів на національному рівні сприятиме єдиному розумінню цієї технології та визначенню ключових термінів, ризиків, переваг тощо.

У статті зосереджено увагу на ключових, на нашу думку, стандартах щодо виключення неправильного тлумачення термінів, сприяння спілкуванню та розумінню у сфері AI. Розглянуто всесвітньо визнаний стандарт ISO/IEC 42001, який містить рекомендації щодо управління системи штучного інтелекту.

Зазначені у статті ключові стандарти штучного інтелекту доречно прийняти в Україні як національні методом перекладу через технічний комітет стандартизації ТК 200 «Штучний інтелект».

Отже, потенційні наслідки штучного інтелекту є далекосяжними та глибокими. Штучний інтелект стає все більш потужним. Щоб забезпечити його розробку та відповідальне використання, доречно вирішувати питання упередженості, конфіденційності та прозорості через стандартизацію.

Інформаційні джерела

1. ISO/IEC 42001:2023 Information technology - Artificial intelligence - Management system. URL: <https://www.iso.org/obp/ui/en/#iso:std:iso-iec:42001:ed-1:v1:en> (дата звернення: 11.02.2025).

2. Artificial Intelligence Act. REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024, laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689> (дата звернення: 11.04.2025).

3. ISO/IEC 23053:2022 Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML). URL: <https://www.iso.org/obp/ui/#iso:std:iso-iec:23053:ed-1:v1:en> (дата звернення: 11.04.2025).

4. ISO/IEC 22989:2022 Information technology - Artificial intelligence - Artificial intelligence concepts and terminology. URL: <https://www.iso.org/obp/ui/#iso:std:iso-iec:22989:ed-1:v1:en> (дата звернення: 11.04.2025).

5. ISO/IEC TR 24030:2024 Information technology - Artificial intelligence (AI) - Use cases. URL: <https://www.iso.org/obp/ui/en/#iso:std:iso-iec:tr:24030:ed-2:v1:en> (дата звернення: 11.04.2025).

6. ISO/IEC 23894:2023 AI - Guidance on risk management. URL: <https://www.iso.org/obp/ui/en/#iso:std:iso-iec:23894:ed-1:v1:en> (дата звернення: 11.04.2025).

Sakhniuk I., Fedoseyeva I., Bytkov M.

Technical Center of the National Academy of Sciences of Ukraine, Kiev, Ukraine

STANDARDIZATION OF CERTAIN PROVISIONS IN THE FIELD OF ARTIFICIAL INTELLIGENCE TECHNOLOGY

The article considers and analyzes the international regulatory framework of standards of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) related to AI technology, which has almost 400 standards. The focus is on several key AI standards that provide a framework for the responsible and ethical use of AI technologies. An overview of the scope of their application is given. The standards specified in the article cover such areas as: terminology and concepts of artificial intelligence; control system; risk management; confidentiality; Bias; transparency; accountability, etc. By adhering to these standards, organizations can work to ensure that their AI systems are fair, transparent, and ethical.

Based on the analysis, it is shown that artificial intelligence is increasingly used in all sectors that use information technology. Taking this into account, it is proposed to include in the annual programs of work on national standardization the adoption and/or development of harmonized ISO and IEC standards for artificial intelligence.

Keywords: *information technology, artificial intelligence, standard, management system, structure of artificial intelligence systems.*