ДОБРОВІЛЬНИЙ КОДЕКС ПОВЕДІНКИ

З ЕТИЧНОГО ТА ВІДПОВІДАЛЬНОГО ВИКОРИСТАННЯ **ШТУЧНОГО ІНТЕЛЕКТУ**

16.12.2024

ПОЯСНЮВАЛЬНА ЗАПИСКА

Системи штучного інтелекту (далі — ШІ) стрімко розвиваються, проникаючи в усі сфери нашого життя — від чатботів служби підтримки клієнтів — до складних систем аналізу даних, прийняття рішень та навіть рекрутингу. Широке впровадження систем ШІ робить використання (впровадження) етичних принципів, зазначених у цьому Добровільному кодексі поведінки з етичного та відповідального використання штучного інтелекту (далі — Кодекс) нижче, вкрай важливим для компаній. Дотримуючись етичного використання принципів, компанії-розробники можуть запобігти поширенню упереджень, дискримінації та ненавмисній шкоді. Етичне використання принципів, зміцнює довіру та підзвітність, сприяючи таким перевагам для суспільства, як підвищення ефективності процесів прийняття рішень, покращення захисту конфіденційності, стійкому та інклюзивному прогресу в суспільно важливих секторах, включаючи охорону здоров'я, освіту та державне управління.

ДЛЯ КОГО?

Цей Кодекс розроблений для:

- Компаній, які розробляють або використовують системи ШІ з широким спектром застосування: генеративні системи, корпоративні додатки для управління знаннями, інструменти обслуговування клієнтів та інші подібні рішення;

- Розробників, дослідників, юристів та менеджерів, які працюють з цими системами;

- Організацій громадянського суспільства, які беруть участь у регулюванні та впровадженні ШІ.

НАВІЩО?

Метою Кодексу є дотримання представниками ШІ-бізнесу прав людини, охоплених нижче, щоб сприяти формуванню культури саморегулювання на території України у сфері ШІ для забезпечення етичного та відповідального використання ШІ.

Кодекс описує низку заходів, які компанії, що є підписантами Кодексу (далі — Компанії-підписанти) повинні вжити, щоб:

- Визначити, уникнути або пом'якшити потенційні ризики у сфері дотримання прав людини, пов'язані з їхніми системами ШІ та забезпечити етичне використання ШІ. Кодекс прагне гарантувати, що системи ШІ не порушують права людини.

- Сприяти розвитку культури саморегулювання у сфері ШІ в Україні. Цей Кодекс спрямований на створення середовища, де розробники систем ШІ беруть на себе відповідальність за етичний розвиток та використання їхніх систем.

- Забезпечити узгодженість і прозорість. Кодекс забезпечує уніфікований підхід до взаємодії між різними організаціями та компаніями, що працюють у сфері ШІ. Ця узгодженість є важливою для прозорості, дотримання нормативних вимог і зміцнення довіри між зацікавленими сторонами.

- Підвищувати обізнаність персоналу. Кодекс заохочує підвищувати обізнаність персоналу Компаній-підписантів стосовно систем ШІ, пропонуючи інструменти для прийняття обґрунтованих рішень під час роботи з системами ШІ з метою дотримання відповідних етичних принципів та прав людини.

ДОБРОВІЛЬНИЙ КОДЕКС ПОВЕДІНКИ З ЕТИЧНОГО ТА ВІДПОВІДАЛЬНОГО ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ

8 Принципів етичного та відповідального використання, розробки та управління системами штучного інтелекту

ПРЕДМЕТ ТА СФЕРА ЗАСТОСУВАННЯ

Усвідомлюючи, що ШІ несе в собі ризики для прав та свобод людини та суспільства загалом, які потребують ретельного управління, та прагнучи до відповідального та етичного розвитку й використання ШІ, Компанії-підписанти добровільно беруть на себе зобов'язання застосовувати ШІ у спосіб, що відповідає найвищим етичним стандартам, та сприяти розвитку, впровадженню відповідальних практик використання ШІ. У межах допустимих для Компаній-підписантів (враховуючи внутрішні політики, положення NDA, угоди про комерційну таємницю та інші договори й зобов'язання Компаній-підписантів) зобов'язуються впроваджувати принципи, викладені в цьому Кодексі.

Цей Кодекс не охоплює питань, що стосуються розробки чи використання систем ШІ в секторі національної безпеки та оборони.

Ступінь та рівень впровадження кожного нижчепереліченого принципу повинні відповідати особливостям відповідної системи та рівню потенційного ризику системи ШІ, визначених OECD.

ТЕРМІНИ ТА ВИЗНАЧЕННЯ

Вхідні дані означає дані, надані або безпосередньо отримані системою ШІ, на основі яких система виробляє вихідні дані.

Життєвий цикл включає планування та проєктування системи; збір та обробку даних відповідно до норм чинного та міжнародного законодавства; побудову системи (систем) ШІ та/або адаптація наявної системи (систем) ШІ для конкретних цілей; тестування, оцінку, верифікацію та валідацію; розгортання та випуск на ринок; оперування і моніторинг; а також списання, припинення використання або знищення системи ШІ.

Навчальні дані означає дані, що використовуються для навчання системи ШІ шляхом налаштування її параметрів, що підлягають навчанню.

Ризик означає поєднання ймовірності настання негативних наслідків та серйозності цих наслідків.

Користувач означає фізичну особу, яка взаємодіє з сервісом, застосунком, платформою чи іншим інструментом, що функціонує повністю або частково на основі системи ШІ, незалежно від мети чи способу такої взаємодії.

Персонал Компанії-підписанта означає сукупність працівників та підрядників, залучених оператором до різних етапів роботи з системами ШІ. Це охоплює такі види діяльності, як розробка, тестування, розгортання, впровадження, технічне обслуговування, а також поширення та моніторинг систем ШІ.

Система штучного інтелекту (система ШІ) означає машинну систему, яка розроблена для роботи з різним рівнем автономності та може демонструвати адаптивність після її впровадження, і яка для явних або неявних цілей робить висновок на основі вхідних даних, які вона отримує, як генерувати вихідні дані, такі як прогнози, контент, рекомендації або рішення, які можуть впливати на фізичне або віртуальне середовище.

Чинне законодавство означає сукупність нормативно-правових актів, що ухвалюються органами державної влади України, а також міжнародних угод, конвенцій та інших міжнародно-правових документів, які були ратифіковані Україною та мають юридичну силу на території України.

Терміни володілець персональних даних, розпорядник персональних даних, персональні дані, обробка персональних даних, обробка персональних даних, яка становить особливий ризик для прав і свобод суб'єктів персональних даних, суб'єкт персональних даних використовуються у значенні, визначеному в Законі України від 01.06.2010 № 2297-VI "Про захист персональних даних".

ПЕРШИЙ ПРИНЦИП: РИЗИКО-ОРІЄНТОВАНИЙ ПІДХІД (RISK-BASED APPROACH)

Компанії-підписанти повинні розуміти свою відповідальність щодо систем ШІ, які вони розробляють або якими керують, і в разі необхідності вживати належних технічних та організаційних заходів, визначених на їхній розсуд, для управління ризиками та безпечного і відповідального використання таких систем на всіх етапах їх життєвого циклу.

Дотримання принципу може передбачати, але не обмежується таким:

- Визначення та оцінку потенційного ризику та впливу ШІ на права людини на різних етапах Життєвого циклу системи ШІ.
- Розробка та впровадження організаційних і технічних заходів для управління ризиками, пропорційних характеру діяльності та ризикам системи ШІ, з урахуванням її повного життєвого циклу.
- Розробку і впровадження політик, процедур і навчання персоналу, спрямованих на забезпечення етичного та відповідального використання ШІ.
- Оновлення систем ШІ для усунення безпекових вразливостей та покращення їх продуктивності.
- Розробку та впровадження кризових протоколів, спрямованих на своєчасне та належне реагування на виявлені ризики, їх усунення чи зменшення їхніх негативних наслідків.
- Поширення інформації та найкращих практик управління ризиками з іншими Компаніями-підписантами з урахуванням необхідності збереження конфіденційності.
- Застосування різноманітних методів тестування та заходів для оцінки, усунення та зменшення безпекових ризиків ШІ та дискримінаційних упереджень в системах ШІ перед випуском на ринок.
- Надання персоналу відповідальному за роботу з системами ШІ подальших версій інструкцій щодо належного використання системи ШІ, включаючи інформацію про заходи, вжиті для усунення та мінімізації ризиків.

ДРУГИЙ ПРИНЦИП: БЕЗПЕКА ТА НАДІЙНІСТЬ (SAFETY AND ROBUSTNESS)

Компанії-підписанти повинні вживати належних заходів для гарантування безпеки користувачів та надійності систем ШІ протягом усього життєвого циклу таких систем.

Дотримання принципу може передбачати, але не обмежується таким:

- Створення процедур для повідомлення про безпекові інциденти та документування зареєстрованих інцидентів у роботі систем ШІ, а також подальше виправлення безпекових прогалин.

- Застосування різноманітних методів тестування на різних завданнях і в різних контекстах перед впровадженням (англ. deployment), з метою оцінки продуктивності та забезпечення надійності системи ШІ. Компанії-підписанти мають право самостійно визначати найбільш відповідні методи тестування для своїх систем ШІ з урахуванням потенційних ризиків.

- Застосування належних на розсуд Компанії-підписанта заходів кібербезпеки для захисту системи від атак, включаючи шифрування даних та контроль доступу.

- Проведення навчання та внутрішнього тестування щодо цифрової безпеки та кібербезпеки для персоналу та залучених осіб з урахуванням специфічних потреб Компанії-підписанта та рівня потенційних ризиків системи ШІ.

ТРЕТІЙ ПРИНЦИП: ПРИВАТНІСТЬ ТА ЗАХИСТ ДАНИХ (PRIVACY AND DATA PROTECTION)

Компанії-підписанти мають проєктувати, впроваджувати та здійснювати обробку персональних даних у межах життєвого циклу системи ШІ з дотриманням норм чинного законодавства про захист персональних даних.

Дотримання принципу може передбачати, але не обмежується таким:

- Повідомлення суб’єктів персональних даних про джерела збирання, передання персональних даних до третіх країн, мету їхньої обробки, місцезнаходження Компанії-підписанта як володільця чи розпорядника персональних даних.

- Повідомлення суб’єктів персональних даних про механізм автоматичної обробки персональних даних, зокрема про автоматизоване прийняття рішення, яке має правові наслідки для особи, логіку та критерії його прийняття та можливість перегляду таких рішень людиною.

- Визначення наявності та документування правової підстави на обробку персональних даних в системах ШІ, зокрема включаючи фіксування виняткових випадків, коли обробка персональних даних становить особливий ризик для прав і свобод суб’єктів персональних даних.

- Впровадження належних організаційних та технічних заходів захисту даних для запобігання порушенням захисту даних, зокрема запобігання випадковому або незаконному знищенню, втраті, зміні, несанкціонованому розкриттю або доступу до персональних даних. Компанії-підписанти самостійно визначають необхідні заходи, пропорційні ризикам, такі як політики інформаційної безпеки, шифрування даних, контролі доступу та регулярне резервне копіювання.

- Залучення лише тих розпорядників, які надають достатні гарантії щодо вжиття необхідних технічних і організаційних заходів у спосіб, що дозволяє забезпечити виконання вимог чинного законодавства та гарантувати захист прав суб'єкта персональних даних.

- Дотримання принципу мінімізації даних, тобто збирання й обробка лише тих персональних даних, які необхідні для досягнення конкретної мети обробки в системах ШІ.

ТРЕТІЙ ПРИНЦИП: ПРИВАТНІСТЬ ТА ЗАХИСТ ДАНИХ (PRIVACY AND DATA PROTECTION)

- Дотримання принципу обмеження строків зберігання даних, щоб обробка персональних даних забезпечувала ідентифікацію суб'єктів даних не довше, ніж це необхідно для досягнення мети обробки в системах ШІ або мети обробки, визначеної чинним законодавством.

- Забезпечення дотримання принципу точності та достовірності даних і, за необхідності, зважаючи на мету обробки таких даних, здійснювати перегляд баз даних, щоб інформація, яка міститься в них, зберігалася в достовірній та актуальній формі.

- Надання користувачам доступу до їхніх персональних даних та можливості їхнього виправлення або видалення відповідно до вимог чинного законодавства.

- Наявність процедур для проведення оцінки ризиків впливу на приватність (DPIA/PIA).

- Впровадження принципів проектованої приватності (Privacy by design) та технологій підвищеної приватності (Privacy enhanced techniques).

- Використання методів псевдонімізації для мінімізації ризиків можливості ідентифікувати суб'єкта персональних даних, якщо це необхідно, на розсуд Компанії-підписанта та пропорційно ризикам.

- Наявність процедур видалення даних або використання методів анонімізації даних після спливу строку їх обробки та зберігання з урахуванням вимог чинного законодавства.

- Наявність структурного підрозділу або відповідальної особи, що організовує роботу, пов'язану із захистом персональних даних при їх обробці за умови, що вимоги чинного законодавства зобов'язують Компанію-підписанта як володільця персональних даних призначити таку особу або підрозділ.

- Впровадження механізмів для швидкого виявлення, реагування та вирішення інцидентів, пов'язаних із витоком або порушенням захисту персональних даних Компанії-підписанта.

- Проведення періодичного навчання та внутрішнього тестування персоналу Компанії-підписанта з питань відповідальної роботи з персональними даними в системах ШІ.

ЧЕТВЕРТИЙ ПРИНЦИП: СПРАВЕДЛИВІСТЬ І РІВНІСТЬ (EQUITY AND FAIRNESS)

Компанії-підписанти повинні створювати системи ШІ таким чином, щоб уникнути порушень прав людини, дискримінації, упереджень на всіх етапах життєвого циклу систем ШІ, а також забезпечити репрезентативність даних, що використовуються як навчальні дані.

Дотримання принципу може передбачати, але не обмежується таким:

- Створення структурного підрозділу чи призначення відповідальної особи, які здійснюють нагляд за системою ШІ, включаючи регулярну оцінку впливу системи ШІ на права людини та дотримання принципів, закладених у Кодексі.

- Використання репрезентативних наборів даних для навчання систем ШІ. Репрезентативні набори даних повинні відображати різноманітність населення (зокрема, гендерну, вікову, за соціальним статусом, за місцем проживання тощо), щоб уникнути упереджень та дискримінаційних результатів у роботі системи.

- Утримання від використання широко заборонених практик для систем ШІ, наприклад: здійснення соціального ранжування (social scoring), маніпулювання поведінкою людей та розпізнавання облич у режимі реального часу, якщо інше не буде встановлено відповідно до чинного законодавства.

- Управління наборами даних, що використовуються для навчання систем ШІ, включаючи перевірку наборів даних для виявлення та усунення упереджень та дискримінаційних положень, за можливості, а також забезпечення їх безпечного зберігання.

П’ЯТИЙ ПРИНЦИП: ПРОЗОРІСТЬ (TRANSPARENCY)

Компанії-підписанти повинні забезпечувати прозорість функціонування систем ШІ та процесів прийняття рішень, надавати чітку інформацію про можливості та обмеження систем ШІ, а також про джерела даних, які використовувались для їх навчання.

Дотримання принципу може передбачати, але не обмежується таким:

- Оприлюднення достатньої та зрозумілої для користувача інформації про можливості та обмеження системи ШІ, включаючи загальний опис використовуваних алгоритмів та типів навчальних даних залежно від ризику системи ШІ та обмежень можливостей систем. Таке інформування має надавати можливість споживачам приймати обґрунтовані рішення.

- Використання та за доцільності розроблення методів маркування. Наприклад, чітке маркування або повідомлення, що інформує користувачів про те, що вони взаємодіють саме з ШІ-системою, а не з людиною або використання повідомлень, які вказують на обмеження системи ШІ, щоб попередити користувача про те, що результат може мати похибки або не застосовуватися у певних випадках.

- Оприлюднення узагальнених описів типів навчальних даних, без розкриття конфіденційної інформації, які використовуються для розробки системи, а також заходів, вжитих для виявлення, усунення та мінімізації ризиків та інформації про проведене тестування. За можливості надання публічності наборам даних, на яких навчається система ШІ.

- Належне та зрозуміле інформування користувачів про факт взаємодії з системою ШІ.

ШОСТИЙ ПРИНЦИП: КОНТРОЛЬ З БОКУ ЛЮДИНИ І МОНІТОРИНГ (HUMAN OVERSIGHT AND MONITORING)

Компанії-підписанти повинні передбачати можливість перегляду і коригування рішень систем ШІ людиною та право користувача відповідної системи оскаржити таке рішення у випадках, коли це рішення має юридичні або інші значні наслідки для користувача, або має значний вплив на демократію, верховенство права та права людини; забезпечення суспільного порядку, суспільної безпеки та здоров'я.

Дотримання принципу може передбачати, але не обмежується таким:

- Розробка чітких ролей і сфер відповідальності для структурних підрозділів чи осіб, які здійснюють нагляд за системою ШІ.
- Впровадження процедур моніторингу для відстеження продуктивності систем ШІ та виявлення потенційних проблем, таких як моніторинг точності, неупередженості та справедливості результатів роботи системи.
- Визначення структури підзвітності в Компанії-підписанта. Встановлення чітких механізмів звітності, щоб персонал Компанії-підписанта міг висловлювати застереження, повідомляти відповідальних менеджерів Компанії-підписанта про неетичне використання системи ШІ та звертатися за вказівками до таких менеджерів.
- Включення положень щодо періодичного перегляду та оцінки систем ШІ у внутрішні політики Компаній-підписантів задля виявлення, усунення та зменшення недоліків та ризиків, що можуть мати вплив на дотримання принципів етичного та відповідального використання ШІ. Періодичність перегляду та спосіб оцінювання визначається самостійно Компаніями-підписантами відповідно до ризиків та мети системи ШІ.
- Включення зобов'язань персоналу Компанії-підписанта щодо етичної розробки та відповідального використання систем ШІ в їхні посадові інструкції, з урахуванням внутрішніх політик Компанії-підписанта.
- Створення механізмів та/або порталів для отримання відгуків, скарг та звернень від користувачів систем ШІ та зацікавлених сторін, що уможливлюють комунікацію з особою, яка здійснює нагляд за системою ШІ та отримання зворотного зв’язку.

СЬОМИЙ ПРИНЦИП: ЛІДЕРСТВО ТА РОЗВИТОК ПОІНФОРМОВАНОСТІ (AWARENESS BUILDING AND LEADERSHIP)

Компанії-підписати повинні підтримувати ініціативи щодо розробки та впровадження інноваційних рішень, нових стандартів та методик, а також здійснювати заходи з підвищення обізнаності користувачів у сфері етичного та відповідального ШІ на розсуд та відповідно до можливостей кожної Компанії-підписанта.

Дотримання принципу може передбачати, але не обмежується таким:

- Долучення до розробки та впровадження галузевих стандартів для забезпечення етичного та відповідального використання ШІ.
- Інвестування в дослідження та розробку безпечного, надійного та справедливого ШІ, а саме підтримку академічних досліджень, створення власних дослідницьких лабораторій та співпрацю з іншими організаціями, зокрема громадянським суспільством.
- Долучення до розробки та впровадження освітніх програм для громадськості щодо безпечного та етичного використання ШІ та обміну найкращими практиками.
- Долучення до співпраці із зацікавленими сторонами, університетами та дослідницькими центрами для підтримки досліджень з розробки стандартів етичного та відповідального використання ШІ.
- Співпрацю з профільними організаціями для розробки законодавства про інтелектуальну власність (ІВ), яке відповідає наявним потребам та особливостям використання ШІ.
- Спільно з іншими підписантами Кодексу сприяти розробці секторальних рекомендацій (implementing guidelines) етичного та відповідального використання ШІ.

ВОСЬМИЙ ПРИНЦИП: ІНТЕЛЕКТУАЛЬНА ВЛАСНІСТЬ (INTELLECTUAL PROPERTY)

Компанії-підписанти зобов’язуються дотримуватися норм чинного законодавства у сфері захисту прав інтелектуальної власності (ІВ) у розробці та використанні ШІ, включаючи: уникнення порушення авторського права на комп'ютерні програми, тренувальні дані, які використовуються для навчання систем ШІ; визначення власності на контент, згенерований системами ШІ або з їх використанням, та забезпечення справедливого розподілу вигод від використання ШІ.

Дотримання принципу може передбачати, але не обмежується таким:

- Своєчасну адаптацію систем ШІ до вимог національного законодавства та міжнародних зобов’язань України у сфері ІВ.
- Сприяння доступу до необхідних інструментів та процедур для сторін із законними інтересами, включаючи правовласників, для здійснення та захисту своїх прав ІВ. Наприклад, проведення консультацій щодо ймовірного порушення права ІВ та пошук можливих рішень або впровадження систем повідомлення та реагування на ймовірні порушення прав ІВ.

ЗОБОВ’ЯЗАННЯ НА ЕТАПІ ПОЧАТКОВОЇ ІМПЛЕМЕНТАЦІЇ КОДЕКСУ

Протягом шести місяців від дати підписання Кодексу Компанії-підписанти докладатимуть максимальних зусиль для підготовки до впровадження процесу імплементації принципів Кодексу та звітування з метою повного дотримання положень Кодексу. Дотриманням положень Кодексу вважається імплементація протягом звітного періоду мінімум 1 прикладу з принципів, передбачених цим Кодексом та включення опису його імплементації до звіту. Звітний період розпочнеться через 6 місяців після підписання Кодексу. Одночасно Компанії-підписанти беруть на себе зобов’язання протягом 6 місяців з дати підписання Кодексу активно сприяти створенню саморегулівної організації, яка забезпечить підтримку, моніторинг та спільний розвиток практик дотримання Кодексу.

КІНЦЕВІ ПОЛОЖЕННЯ

Підписанням цього Добровільного кодексу поведінки з етичного та відповідального використання штучного інтелекту (ШІ) Компанії-підписанти **добровільно беруть на себе зобов'язання в межах допустимих для Компаній-підписантів:**

- Дотримання принципів та положень, викладених у цьому Кодексі.
- Впровадження та підтримки систем та процесів для забезпечення дотримання цих принципів та положень.
- Перегляду та оновлення своїх політик та практик щодо ШІ, щоб забезпечити їх відповідність цьому Кодексу та чинному законодавству.
- Публікування серед Компаній-підписантів щонайменше 1 раз у рік звітів про свою діяльність щодо дотримання цього Кодексу.
- Публікування загальнодоступних щорічних матеріалів про свою діяльність щодо імплементації положень Кодексу (наприклад, блоги про ШІ тощо).
- Компанії-підписанти беруть на себе зобов’язання активно сприяти створенню саморегулівної організації, яка забезпечить підтримку, моніторинг та спільний розвиток практик дотримання Кодексу.

Звіти повинні серед іншого містити:

- Опис заходів та інструментів, впроваджених Компанією-підписантом та спрямованих на виконання принципів та положень цього Кодексу.
- Інформацію про будь-які проблеми чи виклики, з якими Компанія-підписант зіткнулася у зв'язку з дотриманням цього Кодексу.

- Плани Компанії-підписанта щодо подальшої діяльності на виконання цього Кодексу.

ВАЖЛИВО:

Цей Кодекс жодним чином не звужує (вимоги) норми національного законодавства та міжнародного права. Цей Кодекс може час від часу оновлюватися відповідно до технологічного та законодавчого прогресу.

ПІДПИСАНТИ КОДЕКСУ

Grammarly
Мар'яна Романишин, Area Tech Lead for Computational Linguistics

MacPaw
Володимир Кубицький, Head of AI

LetsData
Андрій Кусий, CEO & Co-founder

DroneUA
Сергій Маленький, Director

WINSTARS.AI
Дмитро Софина, CEO R&D Center

Gametree.me
Богдана Сидоренко, CEO

YouScan.io
Олексій Орап, Founder & Chief Growth Officer

EVE.calls
Олексій Скрипка, CEO & Founder

Valtech
Дмитро Козловський, Head of Legal & Partnership

SoftServe
Олег Денис, Co-founder, Executive Vice President, Audit

ТОВ "УКЛОН УКРАЇНА"
Сергій Гришков, CEO

Preply
Богдан Кревський, Head of Finance

BlueBird
Богдан Станкевич, Co-founder

ЛУН
Денис Суділковський, CBBO