

Стодд ● Шатц ● Стед

# ДВИГУНИ ВЗАЄМОДІЇ



**ЦІКАВА КНИГА ПРО  
ГЕНЕРАТИВНИЙ  
ШТУЧНИЙ ІНТЕЛЕКТ**



Обкладинку створили Джуліан Стодд та Сае Шатц (Julian Stodd and Sae Schatz)

Вперше опубліковано 2023 року

видавництвом Sea Salt Learning, Борнмут, Дорсет,  
Великобританія

[www.seasaltlearning.com](http://www.seasaltlearning.com) • [www.julianstodd.wordpress.com](http://www.julianstodd.wordpress.com)

**Усі авторські права захищені ©2023.**

Право Джуліана Стодда, Сае Шатц та Джеффа Стеда бути зазначеними авторами підтверджено відповідно до розділів 77 і 78 Закону про авторське право, дизайн і патенти 1988 року. Інтермедії запрошених авторів залишаються їхньою інтелектуальною власністю.

Усі права захищені. Жодна частина цієї книги не може бути передрукована, відтворена або використана в будь-якій формі або будь-якими електронними, механічними або іншими засобами, відомими зараз або винайденими в майбутньому, включаючи фотокопіювання і запис, або в будь-якій системі зберігання або пошуку інформації, без письмового дозволу видавництва.

*Повідомлення про торговельні марки:* назви продуктів або компаній можуть бути торговельними марками або зареєстрованими торговельними марками і використовуються лише для ідентифікації та пояснення без наміру порушити права.

ISBN: 978-1-7384482-0-3 (тверда обкладинка)

ISBN: 978-1-7384482-2-7 (м'яка обкладинка)

ISBN: 978-1-7384482-1-0 (електронна книга)

Ілюстрації, візуальний дизайн та верстка: Сае Шатц, Джуліан Стодд, Джефф Стед та видавництво Sea Salt Publishing.

---

## ЗМІСТ

Вступ .....	1
<b>Частина 1:</b> Двигуни вже тут .....	7
<b>Частина 2:</b> Людське у машині.....	21
<b>Частина 3:</b> Крихкість інтелекту.....	39
<b>Частина 4:</b> Злам та еволюція .....	67
<b>Частина 5:</b> У діалозі з навчанням.....	103
<b>Частина 6:</b> Незвичайні жнива .....	141
Ілюструючи <i>Двигуни взаємодії</i> .....	161
Інші співавтори .....	169
Примітки.....	171
Додаткова література .....	179

У Частинах 1 і 2 містяться ключові поняття, зокрема короткий огляд технології, яка є основою генеративного ШІ. Частини 3–6 досліджують глибші питання, як-от розмивання майстерності, фрагментація істини та еволюція наших структур. Ви можете читати цю книгу в будь-якому порядку; зокрема, ті, хто вже знайомий з генеративним ШІ, можуть почати з частини 3.

Міжнародний стандартний номер книжки в Польщі:

e-ISBN: 978-83-968496-4-9

Робота над адаптацією українською мовою книги “Двигуни взаємодії: цікава книга про генеративний штучний інтелект” (англ. “Engines of Engagement: A Curious Book about Generative AI”) проводилася Польською асоціацією безпеки (the Polish Association for Security) в кооперації з еАкадемією імені Жана д’Андурана програми НАТО з удосконалення військової освіти (NATO DEEP eAcademy) за підтримки Норвезького університетського коледжу оборони (The Norwegian Defence University College).

Керівник проекту

Пьотр Гавлічек, габілітований доктор

директор еАкадемії імені Жана д’Андурана програми НАТО з удосконалення військової освіти / професор Вармінсько-Мазурського університету в Ольштині

## ВСТУП

Слабкий шум далеких двигунів перетворився на рев. Зміни наближаються, але куди нам рухатися?

Хвиля нових технологій дозволяє нам ставити запитання, на які ми отримуємо вичерпні відповіді, творити мистецтво без ручок і фарб, складати пісні без гри на інструменті, знімати відео, не беручи до рук камеру, і досліджувати наші найпотаємніші зацікавлення, не виходячи за межі власної уяви.

Ці нові інструменти є двигунами продуктивності та знань, мистецтва та майстерності.

Вони кидають виклик нашим уявленням про творчість, експертизу, академічність і саму роботу. Вони все більше відчуються як людські (або людиноподібні) і настільки поширені, що з часом відмінність може стати абстрактною: мудрість *homo sapiens* посилена технологіями через діалог і доповнення. Дійсно, ми можемо стати нездатними відокремити досвід технологій від поняття людяності.

Ці технології є допоміжними та технологіями підтримки, але, можливо, також директивними та приписними. Вони можуть навчати нас на ходу, підштовхувати нашу поведінку або впливати на наше уявлення про реальність.

Але що змінилося, якщо не зважати на галас? Куди ми повинні дивитися і слухати, і що ми повинні робити прямо зараз, щоб змінити наше мислення до цієї нової реальності? З чим ми повинні боротися? Що ми повинні отримати в подарунок? І що ми можемо втратити?

Разом ми знаємо ці технології, ці “двигуни змін”, як Генеративний Штучний Інтелект (ШІ). Цей термін описує як те, як вони працюють “за завісою”, так і певне прагнення до того, чим вони можуть зрештою виявитися.

На додаток до безлічі нових текстів, зображень, аудіо, відео та інших все більш вражаючих артефактів та ефектів, генеративний ШІ породив багато шуму, припущень і думок. Частково це пов'язано зі стрімким поширенням технології в суспільній свідомості, а частково – з тим, що найпопулярніші інструменти генеративного ШІ працюють дуже по-людськи і в розмовному режимі.

Взаємодія з цими системами відчувається майже природною, особистою і цікавою, як хороша розмова, і вони, здається, можуть під'єднати нас до цілого ряду знань, творчості та розповіді, які не під силу жодній людині. Завдяки простоті використання та очевидній легкості розуміння генеративний ШІ виводить на перший план п'янке поєднання здогадок і надії.

Але, як і люди, ця технологія недосконала. Вона може бути неточною у своїх знаннях, хибною у своїй логіці та відверто сюрреалістичною у своєму мистецтві. Іноді вона може бути настільки переконливою, що ми можемо забути, що це машина, яка не має свідомого мислення.

Отже, це технологія, яка швидко вступила в конфлікт з нашими різними системами розуміння, усвідомлення та контролю, концепціями справедливості і нашими ринками соціальної валюти. Її випуск занурився в трясовину правових проблем, етичних питань (і псевдоетичних колізій) і комерційних питань, не кажучи вже про екологічні та політичні потрясіння. Генеративний ШІ приземлився швидко, але не гладко.

Нечасто можна спостерігати, як щось так швидко вривається в суспільну свідомість, як генеративний ШІ, і виростає до таких масштабів ідолопоклонства і натхнення, риторики і нерозуміння, догм і сумнівів без будь-якого очевидного маркетингу чи просування. Саме це і привело нас сюди, до цієї цікавої книги: каталізатора, який змушує зупинитися, поміркувати і поділитися – продукту розмов і роздумів біля нашого “багаття”.

---

## ЦЯ КНИГА: ЦІКАВА РІЧ

Навколо генеративного ШІ і без того багато галасу, і ми не маємо наміру додавати його, принаймні безпідставно. Ми залишаємо за собою право бути як правдивими, так

і помилятися. Отже, це – не книга відповідей, а радше книга цікавості та ідей. Ми маємо намір прогулятися ландшафтом з відкритими очима, щоб дивуватися і сумніватися.

У цій книзі ми розглянемо деякі питання та можливості, які відкриває генеративний ШІ. Ми розглянемо, як ці системи працюють на простому рівні, і розберемося, чому вони викликають у нас відчуття як можливості, так і незрозумілого страху.

Ми сподіваємося пролити трохи світла на ширший контекст, допомогти відкрити наш розум для творчого дослідження соціальних та організаційних впливів, зумовлених цією технологією. Як ця технологія може вплинути на процес прийняття рішень, на наші традиційні структури організації та освіти, знань і контролю? Як це змінить наші способи мислення, навчання і зрештою нашу діяльність – не з точки зору аутсорсингових завдань чи нових образів, а через фундаментальну тканину наших лінз розуміння?

Ця робота одночасно серйозна і жартівлива. Ми не хотіли створювати технічний посібник або наївний список спостережень за клікбейтами, і ми також не особливо зацікавлені в дослідженні визначеності. Відповіді дадуть інші. Ми ділимося своїми питаннями – про те, куди можемо прямувати, що можемо побачити під час подорожі і що може змінитися в майбутньому.

... що, щоб уникнути сумнівів, може бути що завгодно.



---

## ПРИМІТКА АВТОРІВ

Ми не повинні були бути тут. Джуліан, Сае і Джефф були зайняті написанням книги про науку навчання, коли зрозуміли, що в цих розмовах витрачають надмірно багато часу на обговорення генеративного ШІ. Коли Джуліан запропонував призупинити інші наші зусилля, щоб написати “швидку” книгу про ШІ, його співавтори справедливо засумнівалися в його прямолінійному мисленні. Але ми погодилися спробувати.

Через кілька тижнів у Джуліана огорнув жах та невпевненість, але Сае і Джефф були в захваті, і продовжили роботу. Крізь плутанину та оптимізм ми невпевнено просувалися уперед, проводячи щотижневі “щасливі години” в кафе та створюючи (тепер уже досить громіздкий) спільний документ. І ось, так чи інакше, ми опинилися тут.

Історія нашого написання ілюструє природу, силу і крихкість нашої спільної роботи. Кожен з нас приносить різні знання та перспективи і живе в різних повсякденних реаліях. Кожен з нас по-різному брав на себе лідерство в різних розділах – іноді для того, щоб стримувати інших, а іноді для того, щоб кинути нам виклик, щоб рухатися далі, а іноді просто для того, щоб разом бути цікавими.

Ми сподіваємося, що ви долучитеся до цієї роботи як партнер, а не як пасивний отримувач мудрості. Ми спробували відкрити простір для більш зваженого і, можливо, трохи незвичного споглядання – і не завжди найочевидніших концепцій. Сподіваємося, вам сподобається.

**Джуліан Стодд** – письменник, дослідник, художник і мандрівник. У своїх роботах досліджує контекст соціальної епохи – місця, де технології змінили те, що означає бути людиною. Його роботи зазвичай існують на перетині систем: соціальне лідерство, соціальне та спільне навчання, а також погляди на культуру, інновації та зміни, що розглядаються радше як соціальні рухи, ніж як структурні питання. Джуліан написав сімнадцять книг і працює за принципом #WorkingOutLoud, щодня ділячись ітераціями та часто знаходячи нові способи помилятися.

**Сае Шатц** – креативна дизайнерка систем, стратегій та артефактів для смислоутворення. Її професійна робота знаходиться на перетині людського пізнання та навчання, технологій та даних. Вона прожила кілька життів: працювала доцентом в академії, головним науковим співробітником у бізнесі, а нещодавно очолила програму Пентагону з дистанційного навчання, спрямовану на розвиток навчання упродовж життя з використанням штучного інтелекту. Сьогодні вона консультує уряди та інші організації по всьому світу з питань використання новітніх технологій та вивчення науки.

**Джефф Стед** – технолог у галузі навчання, винахідник продуктів і керівник творчих команд, які допомагають втілювати їх у життя. Йому подобається працювати у спілці з людьми і технологіями, адаптуючи нові інструменти, щоб знайти ліпші способи навчити учнів у всьому світі. Він витрачає свій час між керівництвом продукту для MyTutor та консультуванням інших стартапів у сфері EdTech, які працюють над тим, щоб змінити світ на краще.

Дякуємо, що долучилися до нашої роботи.

# ЧАСТИНА 1

## Двигуни вже тут

Нещодавній ажіотаж викликав у багатьох з нас відчуття “нового вбрання короля” . Можливо, раніше в нас були якісь нечіткі уявлення про ШІ, але тепер здається, що від нас усіх очікують обґрунтованих думок про те, як генеративний ШІ, діпфейки та великі мовні моделі змінять світ.

Отже, давайте почнемо з самого початку.

ШІ – це категорія програмного забезпечення, зосереджена на виконанні дій, притаманних людині, як-от розв’язання проблем, сприйняття світу, прийняття рішень та навчання. Можемо виокремити такі найпоширеніші його типи, зокрема Обробка Природної Мови (дає можливість комп’ютерам працювати з людською мовою, аналізуючи наші слова або спілкуючись з нами більш гуманно), Комп’ютерний Зір (дає можливість комп’ютерам “бачити” речі так, як це робить людина, та “транслювати” вхідні дані з датчиків у сприйняття навколишнього світу), Робототехніка (дає можливість ручної взаємодії зі світом, наприклад, машини, що працюють на заводі, та автоматизоване медичне обладнання, яке використовують у точній

хірургії), і Машинне навчання (стосується систем, які не просто мають запрограмовані можливості, але можуть “вчитися” з навколишнього середовища та інформації, яку ми їм надаємо).

Корисно розуміти, що не весь ШІ використовує машинне навчання. Іноді алгоритм використовує “старий добрий ШІ” (який також називають символічним ШІ), тобто підхід, заснований на правилах. Уявіть собі орду розробників програмного забезпечення, які, згорбившись над маленькими столами, вручну друкують дерева рішень і списки процедур “якщо-то”. (Якщо це допоможе, не соромтеся уявити собі наглядців у шоломах і наближену обстановку в стилі орвеллівських сюжетів). Для багатьох успадкованих систем такого ручного підходу достатньо – просто він не настільки вражає уяву, як те, чого ми досягли з допомогою сучасних методів, заснованих на даних.

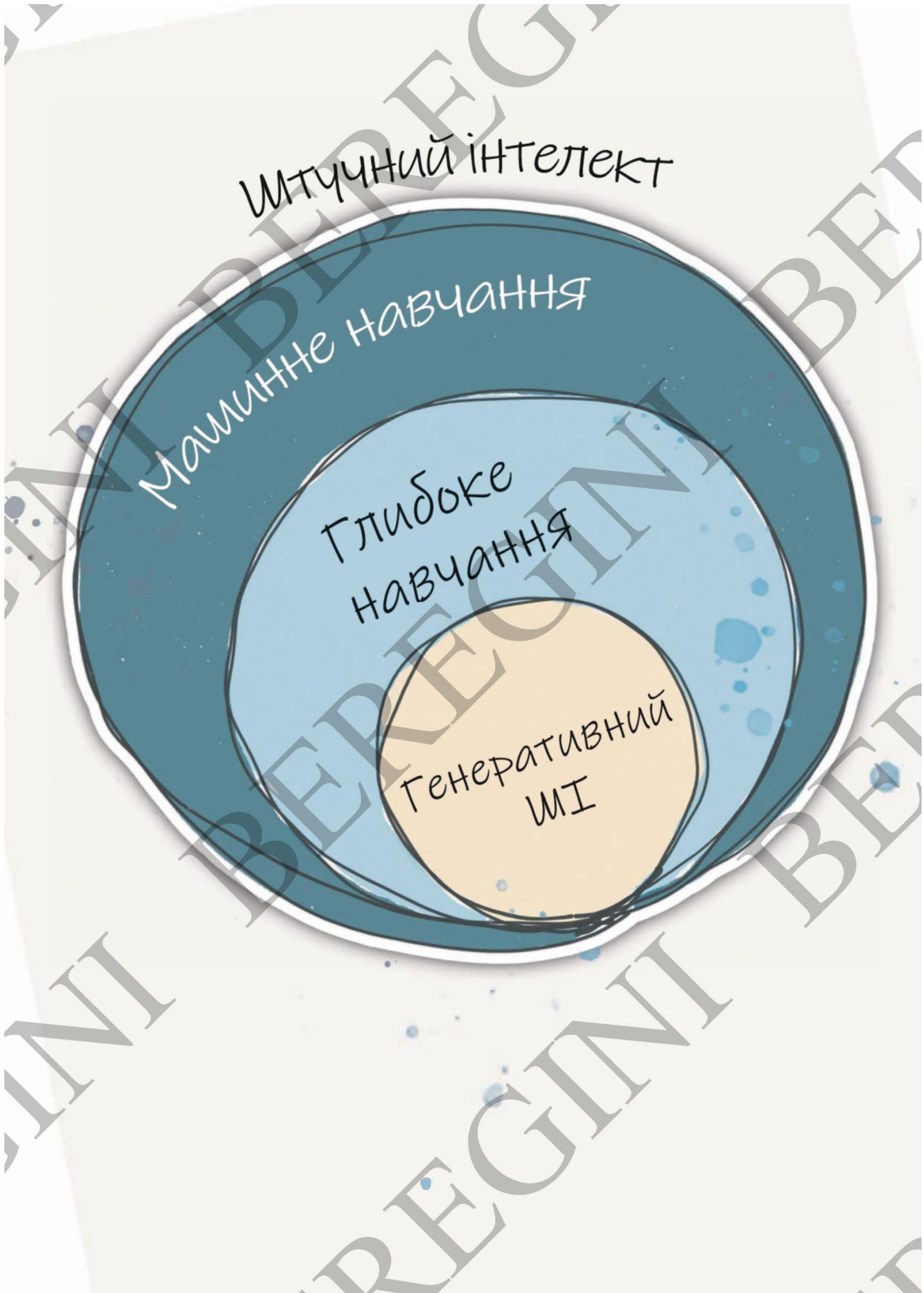
Машинне навчання – це підхід, який передбачає застосування алгоритмів “навчання” з використанням великих масивів даних. Простіше кажучи, алгоритми машинного навчання вивчають багато різних прикладів чогось – наприклад, мільйони зображень різних тварин. З часом алгоритми виявляють закономірності у цих даних – наприклад, навчаються виокремлювати зображення, на яких зображений кіт. На практиці, окрім цієї простої ілюстрації, коли алгоритм знаходить закономірності в наборі даних, він може використовувати цю інформацію для виявлення прихованих тенденцій, пошуку першопричин, прийняття рішень або прогнозування, щодо певної предметної області.

штучний інтелект

Машинне навчання

Глибоке  
навчання

Генеративний  
ШІ



Вони працюють без гуркоту і диму, але,  
втім, є двигунами – проте не руху, а радше  
розповіді та пісні, інформації та зображення  
– тихо змінюючи те, що є звичним для нас,  
руйнуючи визначеність, віддзеркалюючи  
нам світ.

Говорячи з нами.

Вступаючи з нами в діалог.

Малюючи нас. Співаючи нам.

Наслідуючи нас.

Спостерігаючи за нами.

Ці системи достатньо прості, щоб бути  
зруйнованими, достатньо складні, щоб бути  
незрозумілими, і достатньо потужні, щоб  
змінити світ одним лише шепотом.

**Двигуни вже тут.**

Двигуни вже тут.

Концептуально ці підходи подібні до того, як люди використовують розпізнавання образів, хоча машинне навчання може знаходити складні закономірності, які людина може не помітити. Наприклад, ми можемо надати алгоритму мільйони зображень людських клітин, і з часом алгоритм може навчитися виявляти ранні ознаки раку або хвороби раніше, ніж лікарі.

Лише кілька років тому масиви даних для алгоритмів машинного навчання мали бути ретельно підготовлені сумлінними людьми. Уявіть собі добре сформовану електронну таблицю, з дуже ретельно позначеними рядками та стовпчиками та формально визначеним типом даних. (Ви можете знову повернутися до уявної картинки – ряди програмістів за столами в орвеллівському стилі). Цей первинний, поверхневий вид машинного навчання займав багато часу на підготовку і давав дуже обмежені результати, частково тому, що нам, схильним до помилок людям, доводилося сприймати і формулювати найважливіші особливості в масивах даних.

Кілька десятиліть тому була винайдена потужна форма машинного навчання, яка отримала назву Глибоке навчання. Глибоке навчання може обробляти “неструктуровані дані”, що означає, що набори даних не потребують попередньо визначених форматів або створених вручну міток. Глибоке навчання може використовувати різні цифрові артефакти, як от електронні листи, пости в соціальних мережах, документи, відео та зображення.

Іншими словами, глибоке навчання дозволяє нам знаходити та використовувати закономірності в неструктурованих даних, і це дійсно неймовірно.

Глибоке навчання – це один з головних проривів, який дозволив створити нинішнє покоління генеративного ШІ. Воно дозволяє нам навчати алгоритми безпосередньо на книгах, творах мистецтва та музиці, і ми можемо масштабувати наші масиви даних до великих розмірів (оскільки людині не потрібно ретельно очищати та маркувати всі дані), забезпечуючи безпрецедентний рівень гнучкості алгоритмів. І коли ми говоримо “великий”, ми маємо на увазі *справді* великий. Ми майже не обмежені.

---

## ГЕНЕРАТИВНИЙ ШІ

У світі, де і так повно речей, генеративний ШІ виробляє ще більше. Більше речей, які ми можемо споживати.

Генеративний ШІ - це набір методів глибокого навчання, які використовуються для створення нових, раніше небачених артефактів, як-от зображення, відео або текст.

Генеративний ШІ створює “нові” артефакти, виявляючи та переробляючи шаблони старих, створених людиною (поки що) речей. Він бере продукти людської винахідливості (або нашої нудьги і марних зусиль) і якимось чином осмислює ці творіння, щоб виявити їхню сутність; потім він імітує мови, якими вони були

створені, щоб повністю генерувати нові – хоча і технічно похідні – артефакти.

Наприклад, алгоритм генеративного ШІ, навченого на європейському живописі, може навчитися розпізнавати особливості робіт великих художників. Потім він може використовувати цю інформацію, щоб створити абсолютно нове – але оманливо тематичне – зображення, з'єднуючи частини оригінальних візерунків. Це як магія, але з кодом.

У нас вже є текст, згенерований ШІ (наприклад, чат GPT), зображення (наприклад, нейромережа Midjourney або DALL-E), аудіо (наприклад, нейромережа WaveNet, яка може відтворювати розмовну мову або музику), відео (наприклад, нейронна мережа DeepDream), 3D (наприклад, застосунок GANverse3D, який може генерувати тривимірні моделі об'єктів, зображень або людей) і геймери (наприклад, програма AlphaGo, яка може навчитися грати в складні ігри на надлюдському рівні). І це тільки початок.

Можемо лише уявити, що у найближчому майбутньому ШІ створюватиме складніші артефакти. Все, що людина може створити і завантажити в комп'ютер у будь-якій формі, генеративний ШІ може імітувати та відтворити – швидше і з меншими зусиллями, ніж будь-хто з людей. І на відміну від нас, йому не потрібна кава, щоб почати роботу, і його не особливо турбує, чи скажете ви “дякую” або “це чудово” в кінцевому підсумку.

Або, скажімо по-іншому: генеративний ШІ може створити будь-який тип цифрового артефакту, який може створити людина, за умови, що ШІ володіє достатньо

великим масивом даних, з якого він може здобувати інформацію. Для зручності уявіть собі генеративний ШІ як чорну скриньку, яка “підживлюється” даними, – метафора, яка досить добре працює.

Дані – це джерело життєвої енергії генеративного ШІ. Але якщо заглибитися, варто зазначити, що, подібно до автомобільного двигуна, продуктивність генеративного ШІ залежить від кількох факторів: якості та кількості “пального” (даних), дизайну та складності “двигуна” (моделі ШІ), а також обчислювальних потужностей, доступних для розробки.

Алгоритмам необхідна величезна кількість даних. Визначення точної кількості цих даних – мистецтво, яке постійно розвивається; схоже, що існує точка спадної віддачі, у якій збільшення кількості даних, що використовуються для навчання алгоритму, є малоефективним (принаймні, за сучасних дизайнів моделі). Незважаючи на це, ця точка є величезною.<sup>1</sup> Крім розміру, вирішальне значення має якість даних. Хоча методи іноді можуть компенсувати вплив недостатньо великих масивів даних, жодна кількість даних не може компенсувати брак якості (наприклад, різноманітності та релевантності) даних.

На продуктивність моделі також впливає її дизайн – її “параметри” або внутрішні налаштування. Вони схожі на ручки та перемикачі, які модель використовує для розуміння даних. Або, конкретніше, це математичні значення, які виражають взаємозв’язки між невеликими фрагментами даних, наприклад ймовірність того, що будь-яке слово (“солодкий”) буде слідувати за іншим (“гіркий”) у великій мовній моделі. Більша кількість параметрів може зробити модель більш точною та

узагальнюючою, але вона також додає складності та збільшує вимоги до обчислювальних ресурсів і пам'яті.

Для порівняння, GPT-4 (мовна модель OpenAI, що є основою чату GPT) налічує понад 1,7 трильйона параметрів.<sup>2</sup>

Зрештою обмежуючим фактором є доступна обчислювальна потужність. Для навчання складного генеративного ШІ необхідна велика кількість серверів, електроенергії та графічних процесорів (GPU), які є удосконаленими електронними схемами, призначеними насамперед для оброблення зображень і відео (звідси і назва, проте тепер вони використовуються для різних складних обчислень). Це означає, що апаратне забезпечення, час і бюджет обмежують можливості навчання алгоритму.

Щоб оцінити масштаб: підготовка GPT-4 коштувала 63 мільйони доларів США у вигляді обчислювальних ресурсів.<sup>3</sup>

Конкуренція за ці ресурси дуже серйозна. Ми бачили, як США, Тайвань і Китай на міжнародному рівні змагаються за графічні процесори.<sup>4</sup> І ми були свідками серйозних сутичок у відомих компаніях, коли провідні розробники ШІ звільнялися на знак протесту проти відсутності доступу до достатньої кількості ресурсів для навчання своїх моделей.<sup>5</sup>

---

## БІЛЬШЕ, НІЖ ЧАТ-БОТ

Зважаючи на галас навколо генерації мовлення, може

скластися враження, що чат-боти є центром тяжіння генеративного ШІ. Але розмовні агенти (як їх офіційно називають) – це лише один із видів генеративного ШІ, один із прикладів того, чого може досягти ця технологія. Є певний сенс у тому, щоб розрізнити екземпляр і клас.

Генеративний ШІ – це категорія штучного інтелекту, як, наприклад, вантажні автомобілі є категорією автомобілів. Чат-боти – один із класів ширшої категорії генеративного штучного інтелекту, так само, як фургон – клас вантажних автомобілів. І чат-бот ChatGPT, і його побратими по бренду, як-от Bard, Claude і HuggingChat є окремими екземплярами чат-ботів, які можна порівняти з конкретною моделлю панельного фургона Mercedes-Benz.

Окремі екземпляри технології можуть демонструвати упередженість або переваги, і нам слід бути обережними з узагальненням цих своєрідних характеристик для цілого класу або категорії технологій.

Якщо продовжити автомобільну аналогію, то конкретна модель панельного фургона Mercedes-Benz може мати обмежений простір для ніг і імідж тісного та незручного, але це не означає, що всі панельні фургони, або всі вантажні автомобілі, або всі автомобілі загалом є тісними та незручними.<sup>6</sup>

Проте зі штучним інтелектом легко заплутати дискусію.

Розмови про генеративний ШІ, машинне навчання та старий добрий ШІ часто змішують у одну купу. І дуже легко сплутати перспективність та ризик чат-ботів з усім генеративним ШІ.

Ми часто бачимо таку неточність у розмовах про соціальні медіа, коли надмірно узагальнені міркування про конкретний екземпляр (наприклад, Twitter [X] або Facebook) починають охоплювати все поняття онлайн-платформ для спілкування.

Звичайно, будь-яка категорія завжди володіє певними невід'ємними та визначальними характеристиками. Наприклад, усі бензинові двигуни забруднюють довкілля, навіть найдосконаліші, незалежно від того, наскільки якісним є конкретний екземпляр. Справа в тому, що нам слід бути обережними з тим, як ми обмежуємо наше мислення.

Отже, що можна сказати про генеративний ШІ загалом? Які риси безсумнівно притаманні всім системам генеративного ШІ?

Усі класи та екземпляри генеративного ШІ демонструють креативність. (Це не означає, що ми обов'язково приписуємо цим алгоритмам художню іскру. Скоріше, ми маємо на увазі "творчість" у більш стерильному і технічному сенсі: механіку створення). Іншими словами, весь генеративний ШІ створює артефакти, яких ніколи раніше не існувало. Це їхній *сенси буття*. Кожна модель створює новий контент на основі вивчених шаблонів, хоча і завжди похідний, а ступінь художності та естетичної привабливості творінь є дискусійним.

Весь генеративний ШІ керується даними. Він значною мірою спирається на них, використовуючи великі масиви даних для виявлення закономірностей, які є основою для створення контенту. Ці масиви даних мають вирішальне значення для навчання моделей

і встановлення правил, за якими вони генерують контент.

Аналогічно, системи генеративного ШІ мають здатність поліпшуватися з часом. Продуктивність алгоритмів можна постійно підвищувати за допомогою ітеративних циклів зворотного зв'язку, точного налаштування та оптимізації моделі.

Крім того, безсумнівно, що як всеосяжна категорія, генеративний ШІ може похвалитися широким спектром потенційних переваг. Його можна застосовувати в різних сферах і до різних форм контенту. Він може підвищувати креативність, підтримувати інновації, допомогати у створенні контенту і надихати на нові ідеї та можливості.

Водночас, у багатьох своїх класах і сферах застосування перед генеративним ШІ постає низка етичних проблем. Питання власності та прав інтелектуальної власності, пов'язані з результатами роботи цих систем, а також з правом власності на великі масиви даних, які використовують для їхнього навчання, є дуже поширеними. Як і все, чого торкається людина, генеративний ШІ може спричинити упередженість і дискримінацію. Він також може сприяти поширенню дезінформації, давати можливість поганим акторам, або генерувати недоречні, образливі чи шкідливі речі.

Іншими словами, можна сказати, що з його величезним потенціалом і складними викликами, генеративний ШІ має подвійну природу. Творчі перспективи, які він пропонує, можуть сформувати світле майбутнє або ж накинати на нього зловісну пелену. Баланс між цими двома можливостями значною мірою залежить від того, як ми вирішимо його реалізувати.

**Ми створили двигуни, які не просто служать нам:**

**Вони перевершують нас.**

**Вони звільняють нас і потенційно роблять нас  
залежними.**

**Вони годують нас і живляться від нас.**

**І вони можуть забрати вашу роботу або поліпшити її  
виконання вами.**





## ЧАСТИНА 2

### Людське в машинах

ШІ не ідеальний – тому що ми теж не ідеальні.

Варто лише заговорити про генеративний ШІ, і через короткий проміжок часу хтось згадає про “упередженість”. І це буде справедливо.

Ми живемо в світі, в якому немає рівності та справедливості, і останнє, що нам потрібно, – це щоб нові технології увічнювали або поглиблювали дискримінацію, яка сформувалася історично. Тому варто розглянути ширше питання, чому деякі види ШІ демонструють упередженість, як вона з'являється і чи є вона неминучою.

Чи є упередження тимчасовим явищем, про яке ми повинні знати і відстежувати, як про тихий шум на задньому плані? Чи це постійна проблема, яку нам доведеться вирішувати, але тільки для того, щоб вона знову заразила нас різними способами, тож нам доведеться розв'язувати її знову і знову?

Чи взагалі цю проблему можна розв'язати?

## ПРОБЛЕМА З ДАНИМИ ДЛЯ НАВЧАННЯ

Машинне навчання поглинає дані, як гоночний автомобіль спалює бензин, і що складніша програма – наприклад, глибоке навчання і генеративний ШІ – то виникає потреба у більшій кількості та різноманітніших даних.

Неповні або упереджені дані можуть призвести до неправильних результатів та неточних прогнозів, і саме в цьому криється велика частина того, що ми сьогодні вважаємо упередженнями у результатах.

Упередженість потрапляє в систему кількома шляхами. У деяких випадках зловмисники можуть надавати недостовірні дані як паливо для наших двигунів, щоб забруднити їх і так завдати шкоди. В інших випадках ми моримо двигун голодом, даючи йому занадто мало пального – занадто мало або недостатньо різноманітних вхідних даних. Однак, дуже часто упередженість є в контекстах, в яких застосовують ШІ – або, принаймні, у тому, як ми (недосконалі люди) осмислюємо й оцифруємо ці контексти.

Наприклад, алгоритм машинного навчання, навчений на даних стандартизованого оцінювання, може карати учнів з бідного прошарку населення, оскільки він вивчив закономірність, що серед учнів з малозабезпечених сімей є тенденція отримувати нижчі бали, а потім припускає, що ця історична соціологічна проблема є надійною характеристикою системи, аксіомою, на яку можна спиратися і яку слід поширювати.

Ми бачили численні та яскраві приклади упередженості

в машинному навчанні, як-от тенденція ранніх алгоритмів найму Amazon карати резюме жінок<sup>7</sup> або суперечливі алгоритми предиктивної поліцейської діяльності, які непропорційно спрямовані на меншини.<sup>8</sup> А з генеративним ШІ, навіть у таких ретельно контрольованих додатках, як ChatGPT і Midjourney, ми бачили сексистські коментарі<sup>9</sup> і расистські зображення.<sup>10</sup>

Водночас суспільні погляди повільно зазнають змін, переважна більшість наших історичних письмових та візуальних джерел представляють соціальні моделі, які вже застаріли.

Якщо ви подивитесь на дані про доходи за останню тисячу років, то ці тенденції свідчать на користь чоловіків, оскільки упродовж більшої частини цього часу тільки чоловіки заробляли гроші або володіли власністю. І, звичайно, саме чоловіки диктували архіваріусам, що записувати. Навіть сьогодні, коли ми номінально маємо кадрові та правові механізми для забезпечення гендерного паритету, мало хто сперечатиметься, що він є, що і засвідчила катастрофа з алгоритмом найму в компанію Amazon лише десять років тому.

Крім того, цифрова культура, на якій навчаються наші алгоритми (твори мистецтва, пісні та книги, на яких вони навчаються), зазвичай, є підмножиною географічних і культурних аспектів нашого світу. WEIRD (західна (western), освічена (educated), промислово розвинута (industrialised), багата (rich) та демократична (democratic)) культури надмірно представлені серед творців і кураторів ШІ. Набагато менше даних для навчання існує з літературних і культурних артефактів малих мов і менш WEIRD-груп, а це означає що вони

недостатньо відтворені і щодо них в алгоритмах потенційно застосовують упередження.

Отже, сміття всередині, сміття назовні.

У ШІ немає емоцій. Його не “хвилює” (і він навіть не “розуміє”), що таке соціальна нерівність. Він просто виявляє закономірності. Якщо ми, недосконалі люди, створили шаблони несправедливості чи упередженості, то алгоритм є лише дзеркалом, через яке ми дивимося на себе – потужною лінзою для самооцінювання і, можливо, для пошуку підказок щодо того, як ми можемо еволюціонувати.

---

## НЕВИДИМЕ ДЗЕРКАЛО

Хоча генеративний ШІ може розмовляти англійською та кількома сотнями інших мов, а також безліччю мов програмування, жоден алгоритм не може по-справжньому розмовляти або формулювати своє мислення.

І ми (люди) також не можемо говорити їхньою мовою. Ми просто не можемо розкрити алгоритми і прочитати їхні механічні думки. Мовою генеративного ШІ є числа – нечитабельний і непроникний масив. Натхненні роботою людського мозку, ці алгоритми покладаються на штучні нейронні мережі – цифрові структури, які відображають наш власний розум.

У штучній нейронній мережі кожна дискретна “ознака” у масиві даних, наприклад, кожне слово в англійській

мові або навіть більш гранульовані лінгвістичні частини, такі як корені та префікси слів, є вузлом. Зв'язки між вузлами є аналогом наших біологічних “зв'язків між нейронами”, а в штучних нейронних мережах ці зв'язки мають певну вагу. Саме вона відображає статистичні зв'язки між вузлами.

Такі зв'язки є одними з тих параметрів моделі, про які ми говорили у Частині 1 – ручками і перемикачами, які змушують алгоритми працювати. Формально ці значення зберігаються у так званій матриці входжень – таблиці чисел, яка відображає ймовірність того, що різні слова з'являться разом. У нашому прикладі ці ваги можуть відображати ймовірність того, що певне слово, наприклад, “банк”, з'явиться в одному реченні з іншим словом, наприклад, “лівий”, “річка” або “гроші”; або “лівий” з'явиться зі словом “правий”, або “правий” зі словом “незрозумілий” чи “таємний”, тощо.

Це швидко перетворюється на вибух зв'язків.

Звичайно, мова – це щось складніше, ніж низка двовимірних зв'язків. Часто одне й те саме слово має різні значення та конотації. Контекст також має значення, тому необхідно враховувати додаткові фактори. Чат-бот ChatGPT використовує матрицю спільного вживання з більш ніж 12 000 різних вимірів для визначення ймовірності “спільного вживання слів”.<sup>11</sup> Це означає, що ChatGPT може враховувати широкий спектр контекстних факторів під час створення тексту, включаючи зв'язок між різними словами, а також їх порядок у реченні та загальній темі. Саме це впливає на реалістичність його результатів та їхню схожість з людською мовою. Однак, помножені на тисячі й тисячі англійських слів, ці цифри

перевищують нашу здатність візуалізувати.

Ми не в змозі поставити під сумнів мільярди (а часто і трильйони) математичних співвідношень, які визначають моделі генеративного ШІ. Це мова, яка не зазнає інтерпретації. Почути її – все одно, що слухати шум Всесвіту або гудіння мільйона кремнієвих мрій.

Ще більше ускладнює ситуацію те, що вузли та ваги, які використовують у штучних нейронних мережах, не є у явному вигляді значущими, зрозумілими людині концепціями (як у деяких наукових працях, які ми читали). Ці алгоритми не містять інструкцій на кшталт “оцінювати дітей з малозабезпечених сімей нижче” або “не брати на роботу жінок”. Натомість вони включають такі речі, як  $S(x) = 1 \div (1 + e^{-x})$  заховані в масі подібних математичних функцій.

Це відома проблема чорної скриньки машинного навчання.

Ні ми, ні алгоритми не можемо пояснити закономірності, які вони вивчили. Когерентні результати, які вони видають, виникають і з неймовірної системи. Це майже диво, якщо подумати, і навіть експерти зі штучного інтелекту не можуть повністю пояснити, як це все працює. Люк просто відчиняється, і тоді лунає пісня, або друкується сторінка, або створюється зображення.

Це частина магії генеративного ШІ: те, що виходить з цього люка, не може бути повністю передбачуваним. Щоразу з'являється щось нове. Але, коли ми заглядаємо всередину, шукаючи джерело цієї магії, ми не бачимо

нічого, крім кількісних даних, зібраних у статистичних іграх.

Розглянемо попередній приклад алгоритму, який навчився на перегляді мільйонів зображень тварин розпізнавати котів. Цей алгоритм не може сказати: “Коти мають гострі вуха” або “Коти мають довгі вуса”. Ці алгоритми “знають” лише математичні функції – закономірності в даних зображень. І ми поняття не маємо, за якими ознаками алгоритм виокремлює котів на фотографіях. Навіть якби ми розпакували алгоритм і змогли “зазирнути” у N-вимірну числову складність, яку він створив, ми, ймовірно, не зрозуміли б, які функції він використовує, і не знали б назв для цих понять.

Які наслідки ця складність має для нашої нинішньої дискусії про упередженість і помилки? По-перше, ми не можемо запитати алгоритм, як він відрізняє котів від інших зображень, і якщо він помиляється – скажімо, сприймає капусту за kota – ми не можемо запитати його, чому.

Ми не можемо повністю зазирнути у внутрішню роботу алгоритму. Ми не можемо пнути його колеса чи розчаровано гримати на нього.

У ліпшому випадку ми можемо припустити, що в даних для навчання була якась прогалина, якась квінтесенція реальності, яку ми недостатньо показали алгоритму. Або, можливо, був певний детальний вимір ймовірності, для адекватного відображення якого ми надали занадто мало параметрів (занадто мало модельних ручок і перемикачів моделі).

Реальність має нескінченну кількість ймовірностей (нескінченне розмаїття котів і капусти), тому завжди є ризик, що ми щось не врахували.

Щоб розв'язати проблему, ми можемо спробувати надати алгоритму більше даних або підвищити складність моделі, а потім сподіватися на краще. Але ми ніколи не дізнаємося, чи виявили ми проблему, чи повністю її вирішили, окрім як за допомогою постійних спроб і помилок. Зрештою, котів і капусту може бути складніше надійно розмежувати, ніж нам здається на перший погляд.

Пам'ятайте, що на певному ранньому етапі нашого життя нам (людям) довелося “вчитися” бачити і розвивати фундаментальне розуміння навколишнього світу, яке ми зараз використовуємо як вроджене. І щоб бути справедливими до двигунів штучного інтелекту, більшість з нас не змогли б пояснити, звідки ми знаємо, що кішка – це кішка, але принаймні у нас є мова, щоб спробувати.

Інженери ШІ працюють над цією проблемою з технологічного погляду у рамках нової галузі – Інтерпретованого ШІ. Вони намагаються знайти способи зробити невидиме видимим, висвітлити чорну скриньку трохи більше.

Але навіть якщо це вдасться, нам все одно доведеться долати упередженість, притаманну даним або таку, що виникає через нашу взаємодію з несвідомим алгоритмом.

---

## ПІДРУЧНИЙ ЗАХИСТ

Навіть якщо ми не можемо зрозуміти суть “чорної скриньки” машинного навчання, ми напевно можемо збудувати бар’єр на виході алгоритму, відфільтрувати сміття, яке він час від часу продукує, або запровадити заходи безпеки його роботи, чи не так?

Ми можемо встановити правила для подолання упереджень, як це зробила американська програма “Affirmative Action”, намагаючись подолати расову нерівність чи гендерні квоти, які запровадили Індія та Норвегія для своїх лідерів.

Ми можемо перевизначити алгоритм – *homo ex machina* (людина з машини).

І дійсно, такий підхід часто використовують. Але перш, ніж розглянути його, давайте обговоримо сумнозвісний контрприклад з чат-ботом Microsoft – Tay – бабая і повчальну історію для всіх чат-ботів.<sup>12</sup> Чат-бот Tay був запущений у Twitter у 2016 році. Чат-бот, який імітував мову дев’ятнадцятирічної американської дівчини, досить добре функціонував до свого запуску, але всього за кілька годин спілкування з інтернет-тролями Tay навчилася бути расисткою, сексисткою і досконало володіти словником урбаністичних термінів.

Це не означає, що проблема була в базовій моделі Tay. Вона не була погано навчена чи надто обмежена в плані складності моделі. Її падінням стала токсичність даних.

ChatGPT урахував досвід “Уроку Tay”, і його розробники встановили правила на додаток до шаблонів, виявлених у навчальних даних, щоб забезпечити відповідність його

вихідних даних певним цінностям, зокрема корисності і нешкідливості. Щоб навчити цього, Open AI (компанія, яка створила ChatGPT) найняла близько сорока людей, які переглядали перші результати роботи чат-бота та оцінювали його висловлювання на відповідність цим бажаним якостям, поступово вдосконалюючи роботу алгоритму.<sup>13</sup>

Можна використовувати різні підходи, від налаштування моделі за допомогою навчання з підкріпленням (як у прикладі ChatGPT вище) до накладання власноруч розробленого алгоритму старого доброго ШІ на результати машинного навчання. Отже, такий багатосаровий підхід, здається, вирішив нашу проблему, за винятком...

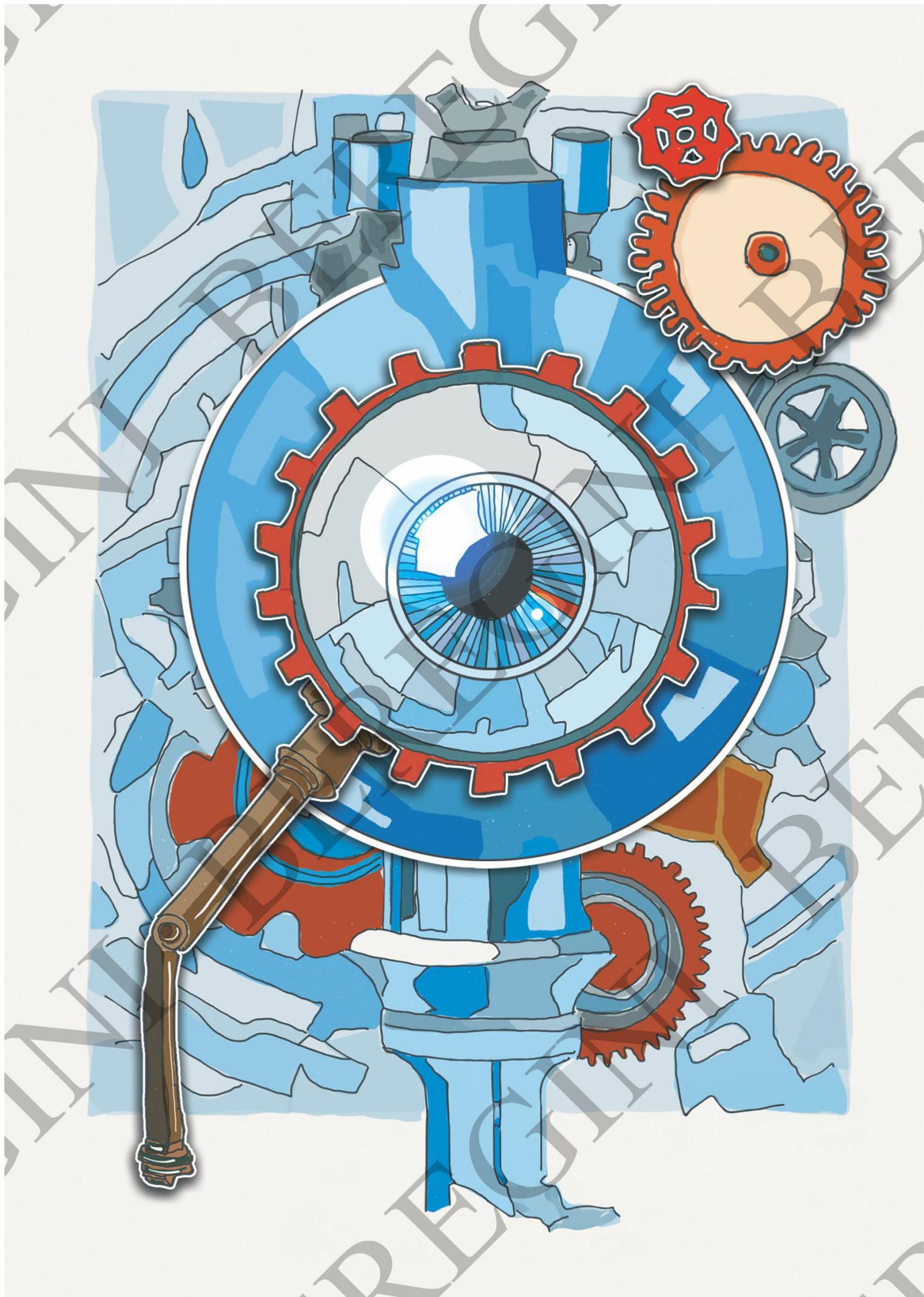
---

## ХТО ВИЯВЛЯЄ УПЕРЕДЖЕНІСТЬ?

Навіть якби ми мали повне уявлення про алгоритми і були впевнені у своїх методах програмування захисних механізмів навколо них, і навіть якби у нас була машина для виявлення упередженості, яка вручну виправляла б потенційні аномалії в наших шаблонах, як би ми її використовували?

Кому з людей ми можемо довірити компетентність та ясність, щоб зробити усе по-справжньому безпечним?

Ми упереджені за своєю природою, хоча б через обмеженість наших знань та обмежене сприйняття світу. Крім того, наш мозок налаштований виявляти закономірності, яких не існує, або зациклюватися



на упередженнях, які формують наші подальші спостереження. Навіть наші очікування діють як своєрідний фільтр, формуючи те, що ми сприймаємо, і те, як ми сприймаємо реальність. Як тільки ми дізнаємося, що щось є “істинним”, ми, зазвичай, тримаємося за це залізною хваткою, і наш розум наполегливо працює, щоб підігнати наші спостереження під наші очікування. До того ж, ми живемо в культурі, де інакодумство подавляють, і де ми вчимося належати собі та процвітати, підкоряючись.

Отже, наша відправна точка полягає в тому, що ті самі люди, які стурбовані упередженістю на структурному рівні, можуть не помічати її в собі, не кажучи вже про інших або в системах систем, що визначаються складними закономірностями. І навіть, якщо ми нарешті виявляємо упередженість, чи знаємо ми, що з нею робити?

Одна справа, коли 40 людей оцінюють корисність повідомлень від нового чат-бота, і зовсім інша – коли намагаються керувати алгоритмами предиктивної поліцейської діяльності або генеративним штучним інтелектом, який створює нові комбінації лікарських препаратів. Як упоратися з упередженістю, коли ставки вищі, більш ефемерні або відверто полемічні?

Наприклад, чи варто повідомляти алгоритмам предиктивної поліцейської діяльності, щоб вони ігнорували расову приналежність як ознаку шаблонів, які вони вивчають? Здається, це непоганий початок, але це вже пробували. Ці алгоритми поліцейської діяльності вже вилучають расову приналежність зі своїх масивів даних для навчання, але інші фактори, як-от соціально-економічний статус, освіта та місце проживання

зрештою у слугують проксі.

Отже, ми маємо справу з систематичними помилками, які переплітаються з нашою реальністю, а не просто з окремими характеристиками, які так легко ігнорувати. Саме це і привело нас до цієї проблеми: алгоритми виявляють шаблони реального світу – глибокі, надійні та повторювані на великих масивах прикладів.

І це повертає нас до нашого початкового питання: хто з нас достатньо компетентний і розбірливий, щоб визначити правила, які усунуть упередженість, не схилиючи шальки терезів в інший бік і не спричиняючи інших непередбачуваних наслідків?

Можливо, було б ліпше просто відмовитися від використання штучного інтелекту на користь чогось серйознішого. Нехай він розважає нас дотепними картинками або допомагає з рутинним написанням текстів, але відмовитися від його використання в більш серйозних ситуаціях.

Деякі керівні органи Європейського Союзу стають на цей шлях, наголошуючи на необхідності бути обережними або обмежувати використання ШІ в найбільш чутливих ситуаціях, як-от предиктивна поліцейська діяльність та біометричне спостереження.<sup>14</sup> Аналогічно, Організація Північноатлантичного договору (НАТО) ухвалила “Принципи відповідального використання”, що обмежують застосування ШІ у військовій сфері.<sup>15</sup>

Інші країни навіть заборонили (як ми можемо вважати) генеративний ШІ із низькими ставками через його передбачувану упередженість. Наприклад, Китай і Росія наклали ембарго на ChatGPT, стверджуючи, що він поширює дезінформацію.<sup>16</sup> Те, що в нашій культурі

WEIRD ми можемо назвати просто узагальненням нашого світогляду, інші можуть сприйняти як культурний імперіалізм, а більш ізоляціоністські країни можуть навіть стверджувати, що це упереджене спотворення (незалежно від того, чи є цей аргумент щирим).

Отже, ШІ – особливо машинне навчання та його похідні – несе ризики, які нам, простим смертним, може бути складно подолати або створити від них захист. Але, можливо, ми зможемо бодай певною мірою контролювати його за допомогою законодавства, якщо зможемо знайти правильний баланс, якщо зможемо зрозуміти його механізми і наслідки настільки, щоб діяти з мудрістю, і якщо будемо готові дотримуватися верховенства права, стримуючи свою гординю і виявляючи терпіння і альтруїзм сьогодні, щоб захистити наше спільне майбутнє.

Нагорода за відновлення цього балансу вражає. Сподіваймося, що той вид, який відповідальний за створення цих масивів упереджених даних, знайде в собі сили впоратися з такими складними завданнями.





---

# ІНТЕРМЕДІЯ НА ТЕМУ: “ВІДВАГА БРАТИ УЧАСТЬ”

МАРКА ЗАО-САНДЕРСА

Світ знову захоплений штучним інтелектом. Попередні навчальні платформи “Літо з ШІ” 60-х та 80-х років бліднуть порівняно з ними. Цього разу революційна технологія – великі мовні моделі – захопила увагу як широкої громадськості, так і комп’ютерних вчених. Зокрема, ChatGPT від OpenAI став тим застосунком для споживачів усіх часів, що найшвидше розвивається.

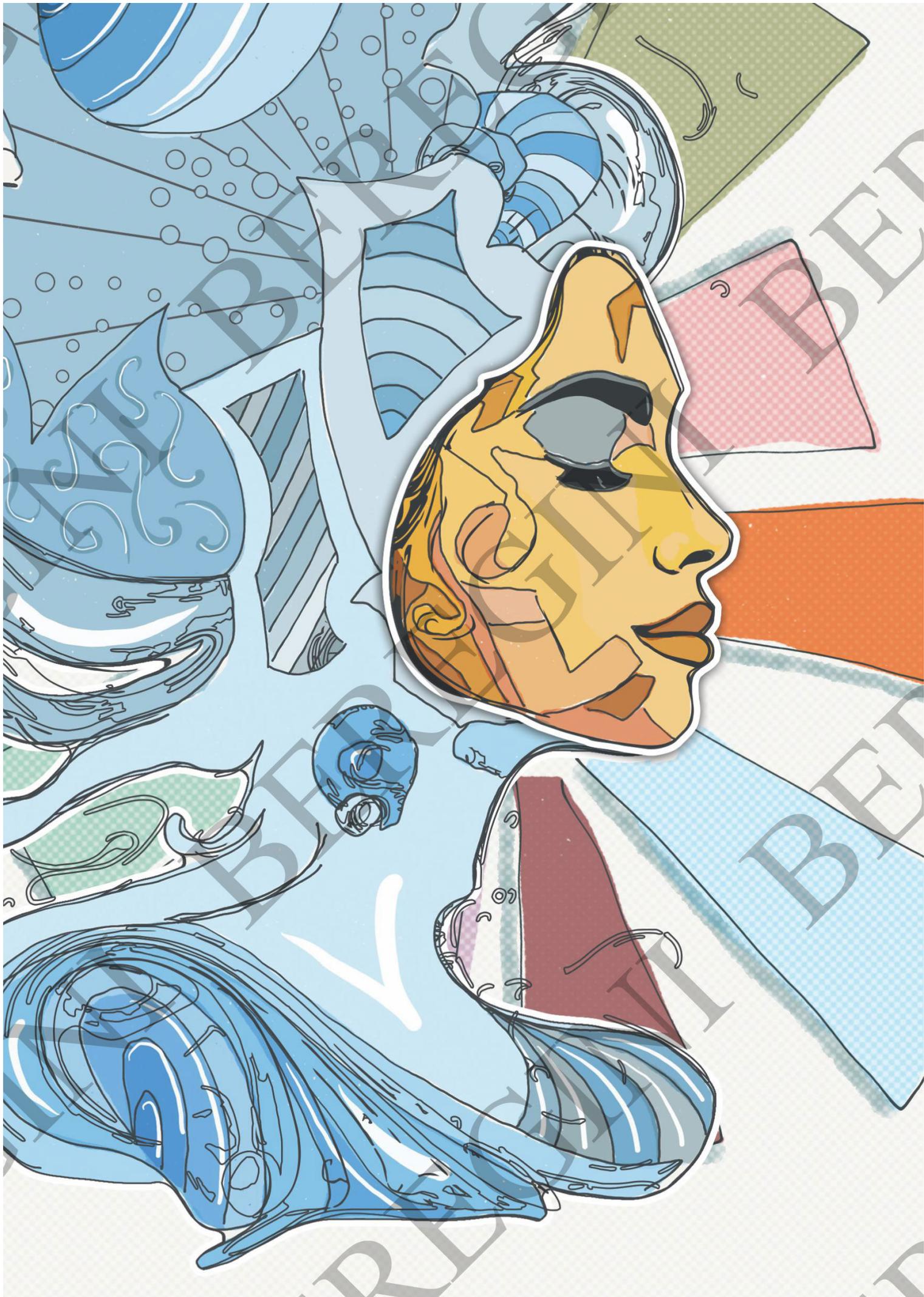
Багато аспектів роботи, яку виконує мільярд високопрофесійних працівників у світі, може виконувати ШІ. У художників, музикантів, письменників, редакторів, кодувальників, маркетологів та багатьох інших раптово виникла потреба адаптовувати і доповнювати те, що вони приносять на вітвар продуктивності.

Разом із технологічним прогресом з’являються спекуляції щодо майбутнього. Думки розходяться в різні боки, як і все, що знаходиться між ними. Одні стверджують, що це просто ажіотаж, який швидко вщухне. Інші вважають, що вплив буде катастрофічним і призведе до сингулярності, яка означатиме кінець нашого виду. Дехто вважає, що це буде щось середнє – ефект буде масштабним і змінить життя кожного з нас, подібно до інших великих технологічних досягнень, як-от колесо, писемність, комп’ютери та інтернет. Експерти також розходяться в думках щодо швидкості цього впливу – від місяців до десятиліть. Суперечки часто бувають гострими і пристрасними. Це питання захопило серця й уми як фахівців, так і нефахівців.

Шукаючи відповідь на такі важливі, складні та суперечливі питання як це, важливо зберігати спокій. У цьому допоможуть беззаперечні факти. Озбройтеся тими, які здаються вам актуальними: скільки людей зараз використовують генеративний ШІ? На яких навчальних даних він ґрунтується? Скільки ще даних йому потрібно? Які завдання він може виконувати? Наскільки добре він їх виконує? Як ми можемо це виміряти? Хто з ваших колег використовує його? Для яких цілей? Як ви самі можете його використовувати? Яку користь приносить саме вам?

Вольовий підхід є рятівним у запеклих дебатах. Ми можемо бути сміливими і помилятися. З одного боку, це означає ризикувати, пробувати щось нове, пропонувати свою точку зору і висловлювати переконання. З іншого боку, це означає бути готовими до того, що погляди, які ми сформувавши і висловили, можуть бути помилковими. Мало кому з нас вдається добре дотримуватися цього балансу. Ми або боязко дотримуємося своєї думки, відмовляючись вступати в бій, і тим самим не сприяємо прогресу. Або ми жорстко прив'язуємося до однієї сторони і стаємо нездатними змінити свою позицію. Технологічний прогрес, та й взагалі будь-який прогрес, вимагає від нас участі, накопичування доказів і нових висновків, які ці докази пропонують. Це і є науковий метод.

Бути сміливим і бути готовим помилятися – це два принципи, на яких ґрунтуються *“Двигуни взаємодії”*. Тож, будь ласка, читайте критично. Піддавайте сумніву припущення. Ви можете мати рацію, а автори можуть помилятися. Що ще важливіше, наберіться сміливості і приєднайтеся до розмови, щоб вплинути на розвиток цієї нової потужної технології. ★



## ЧАСТИНА 3

### Крихкість інтелекту

Генеративний ШІ змушує нас переосмислити себе по-новому, що є однією з причин, чому він водночас і захоплює, і тривожить нас. Він ставить під сумнів фундаментальні речі, які визначають нас як розумних, чуттєвих і людських істот. Змушує нас обмірковувати нові критерії цінностей та зіткнутися з деякими екзистенційними припущеннями.

Чому раніше ми асоціювали добре написані тексти з істиною? Чому ми вважали, що “генерування” і “творення” є ознаками розуму або що інтелект є одним із воріт до світу цінностей?

А що таке, власне, інтелект?

Наші традиційні визначення інтелекту, зазвичай, зосереджуються на розв’язанні проблем, здатності міркувати, самосвідомості та здатності сприймати і переносити ідеї з одного простору в інший. Можна було б справедливо стверджувати, що про людський інтелект свідчить також наша здатність брехати, шахраювати, обманювати, творчо підходити до крадіжки і вигадувати

податкові системи нескінченної складності. І, ймовірно, про наш інтелект також свідчить наша здатність грати, творити, винаходити і вчитися.

Проте ми також схильні визначати інтелект за тим, чим він не є – за його відсутністю – і зверхньо ставимося до тих речей, яких йому бракує.

Це досить довгий список. Почнемо з каміння.

Камені, за нормами більшості, не мають інтелекту (якщо тільки вони не думають дуже-дуже багато). Тюльпани так само не вирізняються інтелектом, хоча вони мають здатність повертати голови до сонця, чим і перевершують каміння. І цілком можливо, що вони можуть чути і навіть спілкуватися у примітивний хімічний спосіб.

А як щодо мурах? Ну, поодинці вони не є сильними суперниками, але разом вони, можуть досягти непоганих результатів. Тож, можливо, не розумні як такі, але трохи розумніші в колективі?

Білки? Вони відомі своєю хитрістю у вузькій сфері, пов'язаній з горіхами. Але ніхто не думає, що білка вкраде його роботу. Отже, це дуже обмежена інтелектуальна здатність.

А от з мавпами складніше. Деякі з них користуються знаряддями праці (один із показників кмітливості), вони б'ються, спілкуються і піклуються одне про одного, і все це виглядає цілком по-людськи – досить розумно за нашими власними, вигаданими нами ж стандартами. Але не варто надто турбуватися здатністю

мавп до колективних дій, щоб скинути нас, або їхньою здатністю замінити Шекспіра, незалежно від того, скільки друкарських машинок їм дадуть.

Немовлята? Ну, донька Джуліана – Мідоу (у віці десяти місяців) одного разу заповзла в кут і не змогла вилізти звідти, ймовірно, змирившись з тим, що всесвіт став меншим і більш кутастим, ніж був раніше. Отже, очевидно, що інтелект не стрясає світ, принаймні, поки що.

Генеративний ШІ? Що ж, це відкрите питання. Безумовно, він геніальний у своїй концепції і дуже зручний, але чи є алгоритми *розумними*? Або чи можуть вони стати такими в майбутньому?

---

## ЩО ТАКЕ ІНТЕЛЕКТ?

Учені вимірюють інтелект різними способами. Багато тестів включають спостереження, як-от чи може певний вид впізнати себе в дзеркалі або користуватися інструментами. Вимірювання інтелекту може також оцінювати, як людина або особливо розумна тварина виконує складні завдання, як-от розв'язує головоломки, щоб отримати смачні ласощів. Існують також стандартизовані тести для видів вищого рівня, як-от оцінювання IQ (що означає “коефіцієнт інтелекту”), нейропсихологічне оцінювання та тести досягнень.

ШІ може конкурувати у багатьох з них, перевершуючи чимало талановитих людей у таких тестах, як вступний

тест до юридичних ВНЗ (LSAT) і випускний іспит (GRE).<sup>17</sup> Нещодавно алгоритми навіть отримали високі оцінки з тесту Равена (Raven's Progressive Matrices - Прогресивні матриці Равена) – це тест для визначення IQ, який має більш складні питання з таких тем, як орієнтація в просторі, абстрактне мислення і рухомий інтелект.<sup>18</sup>

Чи підкреслюють ці досягнення геніальність генеративного ШІ, або, точніше, чи ставлять вони під сумнів наші оцінки інтелекту або нашу німу згоду з тим, що інтелект є чимось очевидним і вимірюваним? Чи сприймаємо ми “інтелект” як штучно створений феномен або вроджену характеристику, чи це радше суб'єктивне судження? Чи не бентежить нас той факт, що камінь – тобто кремній – перевершує багатьох з нас у інтелектуальних досягненнях?

У певному сенсі ми створили ідеальний обман за допомогою ШІ, так само, як ми створили тести на інтелект, не кажучи вже про саме поняття “інтелект”. Ми створили алгоритми, точно розраховані на те, щоб досягти успіху в цих діапазонах когнітивної діяльності. На щастя, відмежування інтелекту виникло в нашій власній колективній свідомості, і в нас також є можливість перебудувати його - пересунути ворота на вимогу.

Але спершу нам слід повернутися до запитання: що для нас *інтелект*? З академічного погляду ми зазвичай визначаємо інтелект (серед біологічних видів) як цілісне застосування пізнання, яке включає автоматичні когнітивні процеси (як-от увага та пам'ять), емоції

(необхідний компонент мислення) і когнітивні навички вищого порядку (виконавчі функції, як-от прийняття рішень). Інтелект виникає з поєднання розвиненого пізнання, сильних емоцій, соціальних навичок і самосвідомості.

Наскільки ці поняття підходять ШІ? Виявляється, не дуже.

[1]

**Мислити “цілісно”.** Наразі ми покладаємося на “слабкий” або “вузький” ШІ – алгоритми, розроблені для виконання однієї задачі, наприклад, генерування тексту або навігації на дорогах. Хоча ми можемо об’єднати їх для досягнення вражаючих результатів, нереально очікувати, що навіть найсучасніші програми зможуть виконувати всі завдання одночасно.

Протиставте слабкий ШІ “сильному”, відомому також як Штучний Генералізований Інтелект або AGI. Його мета – вміти робити все і одразу, іншими словами, бути здатним без проблем виконувати широкий спектр когнітивних завдань.

Можливо, ви вже чули суперечки на цю тему. Найсенсаційніший аргумент полягає в тому, що ШІ може призвести до “сингулярності штучного інтелекту” – сценарію, коли ШІ виходить з-під контролю, як у апокаліптичному бойовику. Наразі у нас немає СШІ (сильний штучний інтелект) і, можливо, не буде його ще довгий час. Успільноті ШІ точаться палкі суперечки щодо доцільності та потенційних термінів реалізації СШІ,

до того ж аргументи варіюються від “його винайдуть в цьому десятилітті” до “цього ніколи не трапиться”.<sup>19</sup> Але в будь-якому разі, його ще немає. Тому ми можемо бути спокійні, знаючи, що люди все ще можуть перевершити роботів у плані розумової адаптації.

Ми можемо одночасно (хоч і не завжди гідно) проводити телеконференції та доглядати за дітьми, а також орієнтуватися на переповнених тротуарах і сканувати вивіски в пошуках підходячого для перекусу місця. Ми навчилися жити в пустелях і в арктичній тундрі, у торгових центрах і навіть у космосі. Як вид, ми з повагою ставимося до власного інтелекту та його здатності діяти в незліченних умовах і виводити нас з нових криз. Білки цього не роблять, і (принаймні поки що) ШІ теж.

[ 2 ]

**ШІ ДЕМОНСТРУЄ ФРАГМЕНТИ, ЗДАВАЛОСЯ Б, РОЗУМНОЇ ПОВЕДІНКИ**, навіть якщо вона (поки що) не є цілісною. Це те ж саме, що мати шматочки інтелекту?

Розглянемо знаменитий уявний експеримент **“Китайська кімната”** філософа Джона Серля.<sup>20</sup> Уявіть собі кімнату, в якій перебуває людина, що не розуміє китайської мови. В її розпорядженні набір правил, написаних англійською мовою, та стосик книжок, що містять китайські речення. Коли вона отримує письмові запитання китайською мовою ззовні кімнати, вона, дотримуючись правил, старанно знаходить правильні відповіді, записані в китайських книжках, а потім передає ці відповіді китайському співрозмовнику по той бік дверей. Людині ззовні здається, що вона

спілкується з кимось, хто розуміє китайську мову, але людина в кімнаті просто механічно виконує правила, не розуміючи суті розмови.

Штучний інтелект – це щось на кшталт нещасного хлопця в кімнаті. Хоча він здатен виконувати складні завдання і видавати на перший погляд розумні результати, йому бракує справжнього розуміння. Він досить переконливо імітує розумну поведінку, але чи дорівнює видимість інтелекту справжньому?

Навіть найсучасніші системи генеративного ШІ все ще покладаються на складні набори правил для імітації розумного мислення. Ці алгоритми не розуміють по-справжньому ні вхідних даних, які вони отримують, ні результатів, які вони створюють. Наприклад, сучасні чат-боти можуть переконливо жартувати, але вони не розуміють жодного слова. Вони просто механічно, як папуги, повторюють шаблони, почерпнуті з мільярда інших текстів. Ось чому дехто називає ці алгоритми “стохастичними папугами”. (“Стохастичний” – це модний термін, який безумовно розумні люди використовують для позначення “випадкового”, а “папуги”, звісно ж, – це ті балакучі птахи, яких любляють пірати у фантастичній літературі). Іншими словами, алгоритми виглядають розумними, але це просто пташині мізки, які бездумно дотримуються правил.

Не всі погоджуються з такою думкою. Один з контраргументів полягає в тому, що китайська кімната розглядає питання свідомості, а не інтелекту.<sup>21</sup> Свідомість є більш складним і філософським явищем,

що включає в себе такі речі, як суб'єктивний досвід і внутрішні психічні стани. Тож, можливо, ми зайшли надто далеко, продовжуючи дискусії про інтелект в такому руслі.

Якщо ми переорієнтуємось, то можемо обґрунтовано стверджувати, що алгоритм або система, яка складається з безлічі збірних алгоритмів, може бути розумним, якщо він повністю імітує всі когнітивні дії, пов'язані з інтелектом. Якщо він може досягти цього, то ми створили функціональний еквівалент – немає потреби в обговоренні таких понять, як свідомість або умисність.

Перш ніж ми заблукаємо у філософській кролячій норі, варто поцікавитися, чи здатен ШІ на такий подвиг. Чи здатен ШІ функціонально відтворити весь перелік розумових процесів, які ми називаємо пізнанням? Чи може машина тією чи іншою мірою зімітувати кожен фрагмент розумної поведінки?

[ 3 ]

**ЩО ТАКЕ КАТАЛОГ ПІЗНАННЯ?** Якщо запитати у пересічної людини, які процеси є основою інтелекту, вона може перерахувати такі речі, як-от пам'ять і пригадування, сприйняття, навчання, розв'язання проблем, планування та інші когнітивні здібності вищого рівня.

Міркування.

Свідоме мислення.

Ми думаємо, а отже, ми є...

...але, як зазначив невролог Антоніо Дамасіо, доповнюючи істину Декарта, емоції також є важливими для людського інтелекту.<sup>22</sup> Тому ліпшою аксіомою може бути “ми думаємо і відчуваємо, отже, ми є”, доповнюючи чек-лист так, щоб він обов’язково містив емоції разом з іншими (здавалося б, більш раціональними) когнітивними процесами. На прикладі реальної історії Елліота, який пережив рак мозку і у якого була пошкоджена лобова частка, Дамасіо продемонстрував, як функції вищого порядку, зокрема прийняття рішень, нерозривно пов’язані з емоційними процесами.

Можливо, емоції – не такий вже й визначальний критерій інтелекту. Зрештою, навіть білки виявляють цілий спектр почуттів – від жадібної цікавості до тривожного самозбереження, і всі вони впливають на прийняття рішень. Але білки досі не перейшли до розряду “розумних” видів, принаймні, за загальноприйнятими науковими стандартами, незважаючи на їхні часто героїчні та пристрасні спроби вкрасти пташиний корм з подвір’я. Чи зможе ШІ зробити це ліпше?

Існує галузь, відома як “Емоційні обчислення”, яка досліджує, як машини взаємодіють з людськими емоціями.<sup>23</sup> Використовуючи пристрої введення даних, як-от камери і мікрофони, а також неінвазивні нейрофізіологічні датчики, ШІ може розпізнавати емоції в реальному часі з досить високою точністю. Розпізнавати посмішки на фотографіях і оцінювати настрої в тексті стало ще простіше. У результаті алгоритми стають дедалі вправнішими в імітації розуміння емоцій та відтворенні емоційних реакцій.

Хоча штучний інтелект може грати ролі і промовляти репліки, він ще не здатний відчувати радість чи відчай або посправжньому співпереживати почуттям іншої людини. Емоційний інтелект алгоритмів є суто симульованим, заснованим на механічних процесах, у яких відсутня суб'єктивна природа людських емоцій з їхньою складною взаємодією біології, культури та накопиченого життєвого досвіду.

Прагнучи надати ШІ більшої емоційної глибини, дослідники працюють над системами, здатними розуміти не лише поверхневі прояви емоцій, а й глибинний контекст і мотиви. Ці зусилля спрямовані на розроблення справжньої Теорії розуму для машин.

Теорія розуму описує нашу здатність не лише сприймати емоції інших людей, а й розуміти їхні думки, наміри та переконання. Люди природно розвивають теорію розуму в міру дорослішання, що дозволяє нам орієнтуватися в складній соціальній павутині людських взаємодій, передбачаючи, що думають та відчують інші. Це, по суті, читання думок. І він тісно переплітається з нашим емоційним інтелектом, дозволяючи нам співпереживати, співпрацювати та формувати складні взаємовідносини. Хоча штучний інтелект досяг значних успіхів у розпізнаванні емоцій, розроблення теорії розуму для машин, яка б дозволила алгоритмам розуміти і передбачати психічні стани людини, залишається величезним викликом.

[ 4 ]

**ІНТЕЛЕКТ – ЦЕ ЩЕ Й СОЦІАЛЬНЕ ЯВИЩЕ.** Наше почуття

емпатії та навички бачення ситуації з погляду інших людей – наша теорія розуму – допомагають нам спілкуватися і співпрацювати. Ми можемо діяти в колективі, а не як окремі індивіди, створюючи і передаючи культуру з покоління в покоління, спираючись на знання і досягнення наших предків, що є ще однією визначальною характеристикою розумного виду.

Отже, шимпанзе, дельфіни, деякі розумні птахи, і, можливо, навіть восьминоги та гриби все ще залишаються в переліку “розумних” видів, але, ймовірно, не білки з їхніми обмеженими соціальними навичками. І, напевно, не штучний інтелект.

ШІ може імітувати соціальну взаємодію, розпізнавати емоції та надавати відповіді, подібні до людських. Він може аналізувати дані, щоб виявити соціальні тенденції і навчитися розрізняти поведінку, яка є соціально прийнятною в тій чи іншій культурі. Для спільної роботи ШІ може співпрацювати з людьми або іншими агентами ШІ, а також відтворювати, здавалося б, відповідні настрої у розмовах або створених відео. Однак цим проявам бракує внутрішніх якостей, які визначають людське соціальне пізнання, що є основою наших колективних досягнень і культурної еволюції. Ми будуємо суспільства, творимо мистецтво і орієнтуємося в культурній спадщині, що є наслідком нашої унікальної здатності до емоційного зв'язку і спільного досвіду.

Як і у випадку з мешканцем китайської кімнати, соціальна взаємодія ШІ керується правилами. Тому, хоча деякі алгоритми можуть здаватися соціальними, всередині вони, по суті, психопати. Але не хвилюйтеся, вони не усвідомлюють образи.

[ 5 ]

### **ОСТАННІМ КРИТЕРІЄМ ІНТЕЛЕКТУ Є САМОСВІДОМІСТЬ**

– здатність усвідомлювати себе як окрему особистість, відмінну від інших, і рефлексувати над власними психічними станами і переживаннями.

У переліку видів, що володіють самосвідомістю, є багато звичайних підозрюваних: людиноподібні мавпи і деякі види мавп, дельфіни і косатки, найрозумніші птахи, слони і свині, а також (що жахає) деякі види мурах. Але, поки що, не штучний інтелект.

Отже, повертаючись до питання, з якого почався цей розділ: “Чи є генеративний ШІ розумним?”. Відповідь, принаймні на сьогоднішній день, – “ні”, якщо ми погоджуємося з загальноприйнятими визначеннями інтелекту.

Але чи правильно ми розставили ці орієнтири? Чи справедливо стверджувати, що інтелект є чимось досяжним лише після певного розмежування? Чи потребує він справжнього втілення емоцій та самосвідомості, чи достатньо детальної функціональної мімікрії? Наскільки “інтелект” є новим явищем, а наскільки це поняття є радше оцінним судженням?

Чи справді ми просто ставимо запитання:

“Чи розвивається *моя* цінність з ростом ШІ?”

---

## ПОКЛОНІННЯ СЛОВУ

Ми часто ставимо слова на п'єдестал. Ми надаємо їм інтелектуальної ваги та "істинності". Слова озвучують наші найпотаємніші надії і страхи, наші ідеї та невпевненість.

Мова надає постійності та виразності нашій інтелектуальній сутності. Через неї ми висловлюємо любов, осмислюємо людський стан і замовляємо на вечерю наприкінці тижня піцу.

Ми пам'ятаємо мовою.

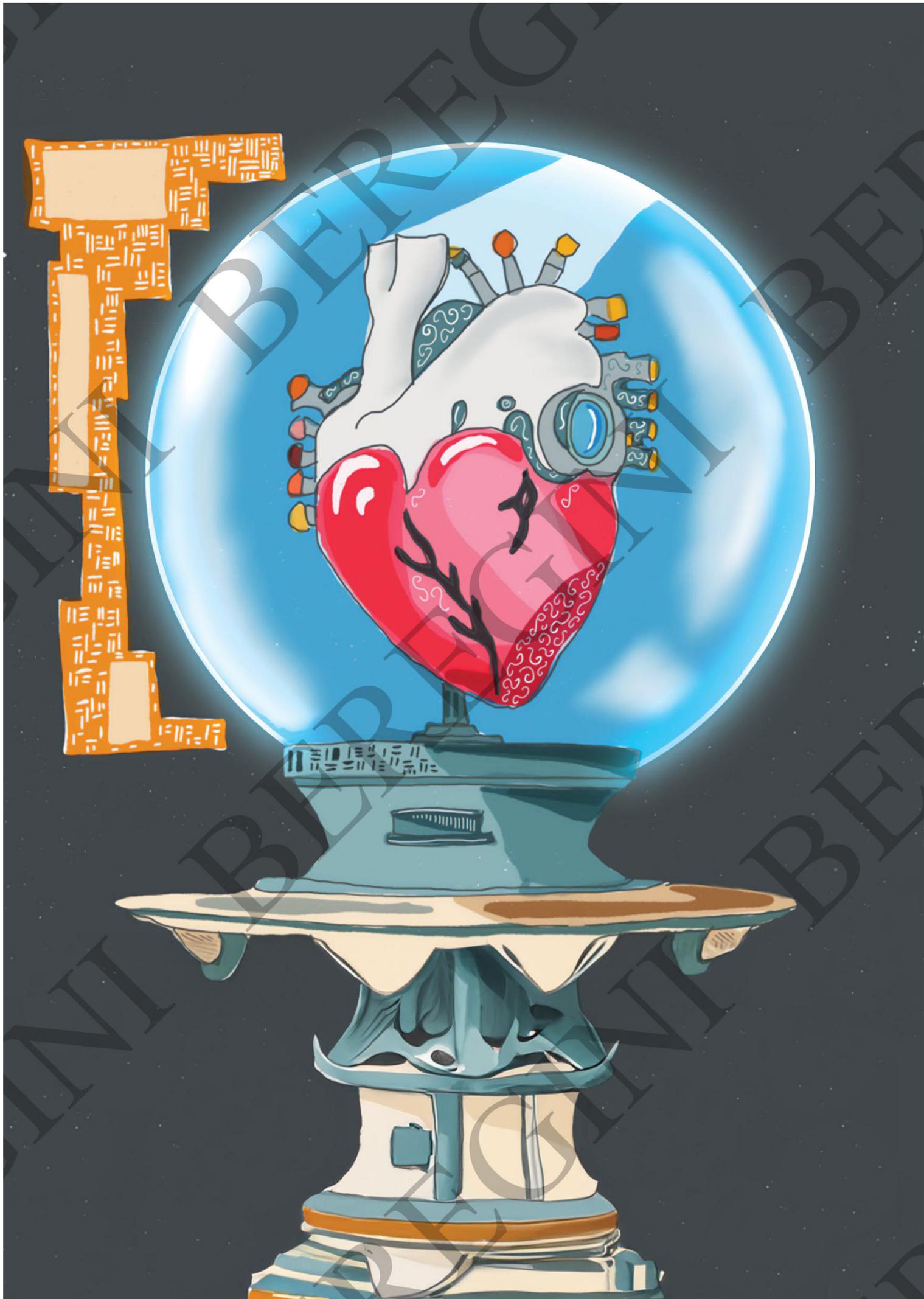
З її допомогою ми розпалюємо революції та ведемо бої.

Вона пронизує наше існування.

Багато в чому мова формує канал між нашим внутрішнім і зовнішнім світом, і саме через лінзи, створені з мови, ми бачимо та інтерпретуємо навколишню дійсність.

Усе це звучить доволі велично і захопливо. Тож не дивно, що ми шануємо майстерність володіння народною мовою. Грамотність і, зокрема, письмо вже давно є визнаною валютою інтелекту. Але тепер, маючи побіжну думку і кілька секунд вільного часу, ми можемо створити есе, написане штучним інтелектом, практично на будь-яку тему. І якщо сьгоднішні результати роботи штучного інтелекту іноді бувають невиразними, лише на рівні посереднього учня, то навіть такий ступінь зв'язності вражає.

Чи не так?



І ось ми підходимо до питання:

Що таке інтелект?

Його слід розглядати як факт чи як переконання?

І чи взагалі це має певне значення?

Жорстке визначення дало б нам впевненість у тому, де закінчатися перегони, але що, якщо інтелект – це не щось одне, а багато? Або якщо це лише семантичне розмежування? Невже інтелект – це те, що ми цінуємо, тому що він рідкісний, або тому що ми віримо в його унікальність – в те, що ми унікальні?

Чи інтелект виявиться спектром, на якому ШІ – або дерево, або пуголовок – може вишикуватися в лінію? Можливо, випередивши нас. Щось, що стане загальнодоступним і поширеним? Дешевим.

Можливо, ми врешті-решт прив'яжемо “інтелект” до наявності серця або хиткого мозку? А може, ми навчимося руйнувати п'єдестал, на який ми його поставили.

Чи правильно ми визначаємо критерії “визначного”? Або поставимо точніше запитання: як генеративний ШІ змушує нас засумніватися у власному поклонінні правильно сформованим словам?

У першій частині ми розповіли про те, як працює генеративний ШІ, але повторимо просте пояснення. Великі мовні моделі виявляють зв'язки між словами (або їхніми частинами), а потім детально формулюють ці взаємодії у вигляді статистичних кореляцій. Коли ви ставите запитання чат-боту, він не “розуміє”, що ви маєте на увазі. Він просто використовує свою масивну математичну модель для аналізу ваші вхідних даних і надалі передбачення прийняттого результату:

одне

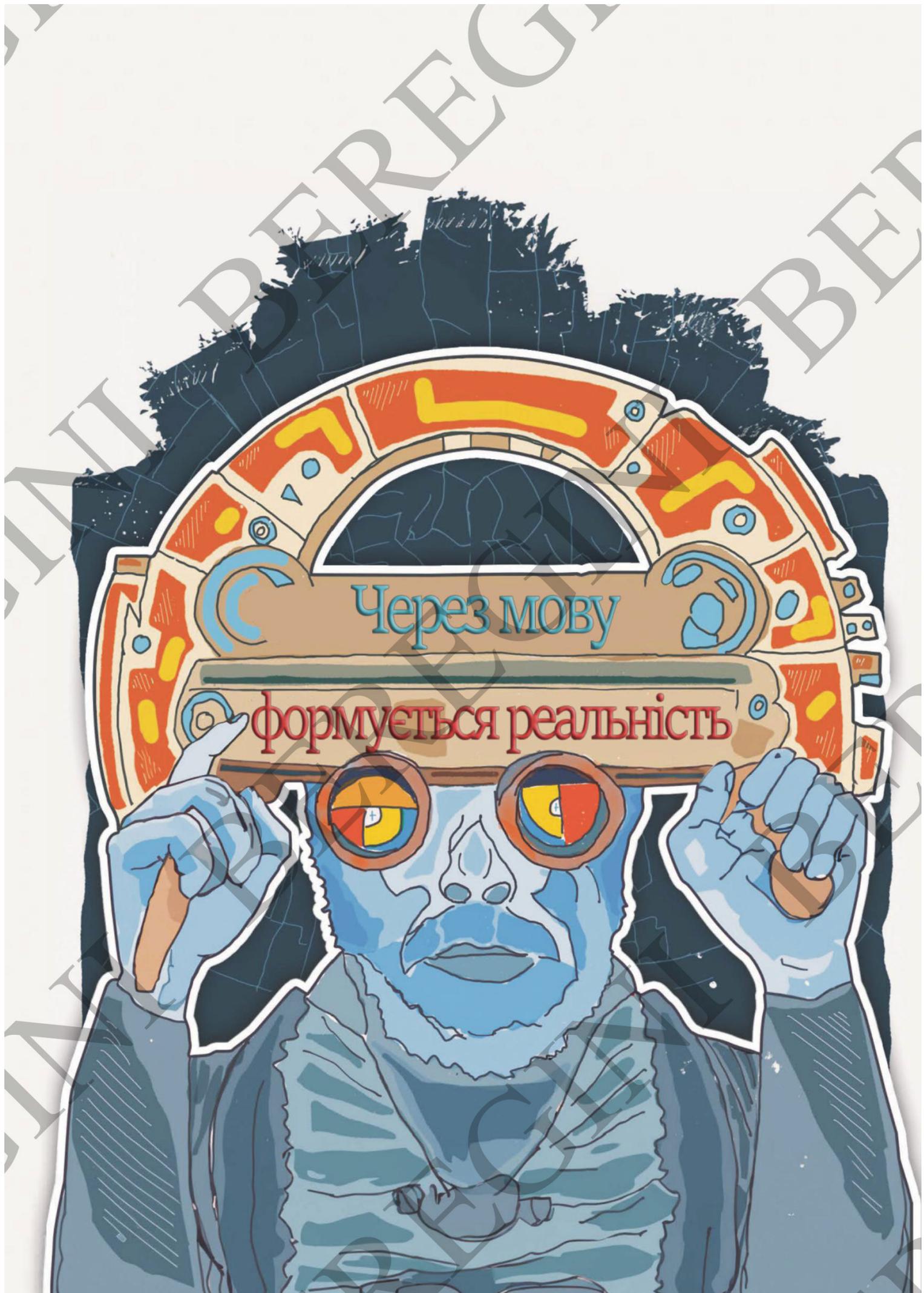
слово

після

наступного

Генеративний ШІ не має відчуття реальності, творчої кризи чи самосвідомості. Він не замислюється над межами інтелекту чи своїм місцем у ширшому всесвіті (принаймні, наскільки нам відомо). Великі мовні моделі ніколи не бачили сонця і не відчували дощу. Але ChatGPT може написати вірш про погоду, а також описати відчуття сонячного світла і його наукові принципи красномовніше і детальніше, ніж це би зробили ви чи я.

Але все це – множина (дуже, дуже велика множина) зважених функцій і ймовірностей, побудованих на основі шаблонів, знайдених у бібліотеках текстів,



Через мову

формується реальність

написаних людьми. Нам знадобилися мільярди письмових прикладів, перш ніж з'явилися справді якісні шаблони. Але той факт, що щось взагалі з'явилося - і що ці чат-боти (незважаючи на їхні недоліки) здаються настільки реалістичними - демонструє, що після аналізу достатньої кількості даних можна виявити приховані закономірності, навіть ті, що здаються незбагненними, такі ж складні, як наше уявлення про реальність та її формулювання.

Багато хто з нас звик розглядати вільне володіння мовою як своєрідний додаток до інтелекту - демонстрацію здатності сприймати істини, виявляти осяяння, міркувати дотепно і мудро. Бездушне і бездумне оволодіння ШІ мовою кидає виклик багатьом нашим глибоко вкоріненим (хоча їх рідко досліджують) переконанням.

Сучасні алгоритми настільки легко орієнтуються в мові, що кидають тінь на межу між синтаксисом і розумом. І подібно до дітей, які стають мудрішими, з часом вони стануть ще більш здібними. Навіть деякі експерти зі штучного інтелекту були вражені їхніми можливостями, а один відомий інженер Google приписує розум LaMDA, мовній моделі Google для діалогових застосунків, завдяки її дивовижній лінгвістичній досконалості.<sup>24</sup>

Важко змиритися з тим, що мову - це щось настільки глибоко пов'язане з нашою свідомістю та культурою - можна звести до математичних процедур. Та коли ми смикаємо за цю філософську нитку, то швидко усвідомлюємо її зв'язок з нашими уявленнями про

інтелект, ідентичність і, можливо, навіть про природу самої реальності.

Кожне з них прив'язане до нитки, що розплутується.

Чи може бути так, що все, що ми цінуємо, піддається спрощенню, здатному звести будь-що до математичних моделей, в яких статистичні частоти підмінюють смислове значення? Якщо так, то виникає ще одне тривожне запитання: Наскільки те, що ми вважаємо “загальним (людським) інтелектом”, за своєю суттю є лише біологічним еквівалентом вибудовування слів у хитромудрі схеми, вивчені повторенням? Моменти нашої геніальності були справжніми чи результатом роботи трильйонів стохастичних імітаторів, які ритмічно барабанили по клавіатурах, поки не з'явилося щось на кшталт осяяння?

---

## ОДИНИЦЯ ВАРТОСТІ

Ми, люди, розумні істоти, маємо виразну ідентичність, свідомість і духовну цінність. Завдяки нашому інтелекту ми відрізняємося від інших видів, нам надано вищий статус і ми визнані “людьми” (і в масовій культурі, і в законодавстві).

З давніх-давен людство використовувало поняття інтелекту для формування нашого уявлення про особистість, а отже, і про цінності.

Розмежування “особистостей” і “не особистостей” (хоч би



якими упередженими були подібні судження впродовж усієї історії) використовувалося для виправдання будь-якого гноблення і жорстокої поведінки. Останнім часом поняття “особистість” почали використовувати у більш позитивних цілях, зокрема як орієнтир поведінки з іншими людьми.

Візьмемо для прикладу восьминога.

З появою нових доказів ми дізналися, що деякі восьминоги володіють дивовижною кмітливістю і, можливо, навіть інопланетним інтелектом. Багато гурманів стверджують, що тепер, коли ми розуміємо їхні здібності розв'язувати проблеми, вони не стали менш смачними, але їхній вищий інтелект змушує (деяких) з нас замислитися.

Чи просунулися вони вгору по харчовому ланцюжку?

Чи хоча б потенційно вони наближаються до статусу “особистості”, подібно до того, як багато хто з нас персоніфікує своїх домашніх собак і котів? Чи етичним є наше ставлення до восьминогів через призму індивідуальності?

Нам подобається думати, що ми особливі, що мантія “індивідуальності” передає якусь іскру або душу, яка робить нас унікальними і гідними серед видів і серед зірок. Еволюційний підйом людства завдяки нашому вродженому інтелекту - це заспокійливий і зручний наратив. Але слово “особливий” може бути поширенішим, ніж ми собі уявляємо.

За межами нашої Сонячної системи ми бачимо, що планети, схожі на Землю, теоретично не такі вже й унікальні. Всесвіт дуже великий, і все менш імовірним здається той факт, що ми єдині, хто вдивляється в нього.

Дійсно, не виключено, що ми знайдемо сліди життя або його складові на інших “Землях”. Можливо, ми виявимо, що наша “людська винахідливість” розкидана по всьому всесвіту. Але тут і зараз ми відчуваємо, що все під нашим контролем. Наш унікальний інтелект дозволяє нам не лише тримати тварин як домашніх улюбленців, а й з чистим сумлінням їсти їх.

А що стосується машин, то це просто інструменти.

Бездумні і замінні; звичайно, не такі розумні, цінні і незамінні, як людство, але як щодо мавп, восьминогів або добре організованих мурашиних колоній?

Чи є в нашій ментальній ієрархії цінностей місце, де ми ставимо алгоритми вище за анатомію? І які почуття викликає у нас оскарження цього рейтингу? Думка про те, що машини можуть мати проблиск інтелекту - і, як наслідок, крихту індивідуальності, є парадигмальним зрушенням, до якого ми, найімовірніше, ще не готові.

Імовірно, питання інтелекту ШІ полягає не стільки в його можливостях, результатах тестування чи навіть історіях, які він розповідає. Радше визнання ШІ розумним змушує нас поставити під сумнів саму концепцію “інтелекту”, визнати його сконструйованим поняттям і, можливо, навіть припустити, що в чомусь ми не такі вже й особливі, як нам хотілося б вірити.

**І нікому не подобається почуватися пересічним.**



---

# ІНТЕРМЕДІЯ ПРО ЛЮДСЬКУ ВИНЯТКОВІСТЬ

ДОНАЛЬДА КЛАРКА

У своїй книзі *Технології навчання*,<sup>25</sup> я досліджую, як технології стали потужними генераторами культури і прогресу. Писемність, алфавіти, друкарство, засоби масової інформації, комп'ютери, інтернет, а тепер і штучний інтелект. Кожен з них доповнив, а часто і замінив те, що ми робимо як біологічний вид. Проте ми все ще вважаємо себе “винятковими”.

Ця людська винятковість, ідея якої полягає в тому, що в нашому виді є якась “сутність”, яка робить нас унікальними, переноситься на наші страхи перед технологіями. У третьому столітті до нашої ери Плавт вважав, що сонячний годинник спричинить наш кінець, розбиваючи дні на години. (Чесно кажучи, з нудними лекціями в школах і роботою з 9:00 до 17:00 він мав рацію!) Сократ з підозрою ставився до письма. Друкарство засуджувалося католицькою церквою та Османською імперією. Кіно, радіо і телебачення піддавалися нападкам з боку скептиків, а копіювальна машина зневажалася як руйнівник творчості та авторських прав.

Інтернет спровокував нову хвилю відрази. Лунали заклики заборонити Вікіпедію (школи та університети блокують, можливо, найбільшу соціально сконструйовану базу знань, яку коли-небудь бачив наш вид), соціальні мережі, смартфони, а тепер і штучний інтелект.

Значна частина коментарів, особливо етичний шум навколо генеративного ШІ, свідчить про упередженість до північної

півкулі та її інституцій. Замість того, щоб зосередитися на генеративному ШІ як на великому дарі для навчання й охорони здоров'я, факторі демократизації і способі зменшення нерівності, серед нас з'являються лицарі високої моралі, які кричать про (як я називаю) дванадцять коней ШІ-апокаліпсису:

1. Плагіат (списування)
2. Публікація (авторське право)
3. Заангажованість (мовна упередженість)
4. Забобони (упередженість)
5. Походження (фальсифікат)
6. Пропаганда (обдурювання нас)
7. Конфіденційність (дані та безпека)
8. Люди (дегуманізація)
9. Бідність (безробіття)
10. Нажива (великі технології)
11. Планета (енергія та викиди)
12. Загибель (події, пов'язані з вимиранням)

Це не означає, що не існує етичних проблем. Але вказувати на ці ризики лінуються ті, хто зазвичай отримує значну вигоду від не використання технологій. Більшість дебатів є п'янкою сумішшю упередженості, антропоморфізму і людської винятковості. І тут, і деінде армія етиків, які не мають жодного уявлення як про етику, так і про штучний інтелект, дуже категорично висловлюються про упередженість, стохастичних папуг, галюцинації і навчання.

Наприклад, академічні кола накинулися на генеративний ШІ не як інструмент поліпшення викладання та навчання, а як спосіб плагіату під час написання есе, можливо, тому що їхня бізнес-модель базується на перевірці достовірності даних. Видавництва зосередилися на авторських правах, хоча нічого не “копіюється”.

Цьогорічні подорожі Африкою надали мені надію, що від генеративного ШІ може бути велика користь. Я на власні очі бачив, як ШІ поглинає та використовує мови меншин, від африкаанс до зулу. А для деяких молодих африканців можливість виконувати роботу, пов'язану з маркуванням і навчанням з підкріпленням за допомогою великих мовних моделей (які хтось може розцінити як експлуататорські), може стати в країнах з високим рівнем безробіття вагомим першим кроком у ІТ.

Ще однією моральною проблемою є відсутність підтвердження походження, тобто відсутність можливості простежити джерела інформації у моделях, що може призвести до того, що їхні результати будуть помилковими. Безумовно, існують проблеми, пов'язані з дідфейками та галюцинаціями. Але люди, які панічно бояться, що ШІ обдурить нас усіх, часто вважають “маси” (а не себе) найбільш вразливими. Ба більше, плагіни та допоміжні інструменти ШІ вирішують ці проблеми. Пам'ятайте, що генеративний ШІ – це найгірший варіант, який коли-небудь буде. Використовувати ChatGPT 3.5 сьогодні – це все одно, що користуватися Вікіпедією орієнтовно з 2004 року.

ШІ не є машиною правди, але, з іншого боку, і ми не є винятковими, коли справа доходить до правди. Наче людський розум вільний від упередженості! Принаймні, упередженість, виявлену в ШІ, можна поступово усунути. Цього не можна сказати про людський мозок, який помилково вважають винятковим.

Наступний кінь ШІ-апокаліпсису – загроза для нас самих як людей. Чи не дегуманізує нас штучний інтелект? Звісно, ні, радше навпаки. Я не можу пригадати жодної іншої технології, яка б нас гуманізувала більше. Наприклад, завдяки великим мовним моделям ми можемо розмовляти з накопиченою

культурою нашого виду – розмовляти з самим собою.

Але що буде з нашими засобами існування? Чи не залишить ШІ нас усіх без роботи і не збагатить технологічних магнатів нашими колишніми зарплатами? Звичайно, спочатку може виникнути безробіття, оскільки ШІ підвищує продуктивність сучасних організацій, але лауреат Нобелівської премії з економіки Вільям Нордхаус продемонстрував, що 98% економічної вартості, створеної новими технологіями, повертається в суспільство, що робить винахід технологій філантропічним за своєю суттю в пропорції 50:1. Воно сприяє процвітанню інновацій через малі компанії та поліпшує повсякденне життя мільйонів (або мільярдів) людей, які ним користуються.<sup>26</sup>

А тим, кого хвилюють обсяги енергії та викидів, які виробляє штучний інтелект, можливо, варто звернути увагу на інших, більш серйозних забруднювачів довкілля. Навіть порівняно з учорашнім виробництвом сторінки тексту або цифрового зображення за допомогою людини, штучний інтелект потребує набагато менше енергії. Якщо вже на те пішло, то ШІ, ймовірно, допоможе нам вирішити проблеми зміни клімату та чистої енергії.

І наостанок, висувуються звинувачення в тому, що ми загинемо як вид у результаті масового вимирання, знищені ШІ. Це крайня форма надуманої людської винятковості: мілленаризм, який ми, люди, здається, так любимо. Можливо, саме таке апокаліптичне мислення про кінець світу і є найбільш “винятковою” рисою в нас. І якщо це так, ми можемо лише сподіватися, що ШІ допоможе нам стати трохи менш “винятковими”. ★



## ЧАСТИНА 4

### Злам та еволюція

Його розповсюдження настільки різноманітне, а бар'єри для його використання настільки низькі, що нам достатньо просто слухати, дивитися або існувати, щоб зіткнутися з генеративним ШІ у дикій природі. Навіть глянувши мимоволі, ми вже можемо побачити спіралі можливостей. Ми можемо уявити нові способи використання, нові творчі методи, нові двигуни майбутнього, нові способи руйнування, нові речі, які будуть створені, і, можливо, нові погляди на наше власне місце в цьому світі.

Можна сказати, що темпи змін випереджають нашу здатність зрозуміти їх, принаймні раціонально. Щойно ми починаємо розмову про, скажімо, прозу, створену людиною, чи прозу, створену ШІ, світ уже змінився – і ШІ уже перевершив усі письмові артефакти людства, принаймні кількісно. Ми залишаємося стояти на стартовій лінії, тоді як двигуни мчать уперед.

Результативність випереджає стратегію.

Дії випереджають роздуми.

Як наслідок, наша людська реакція на стрімкий розвиток штучного інтелекту не буде спокійною і виваженою. Ми можемо очікувати, що на нас чекає суперечки та боротьба, судові розгляди і законодавча ретроспектива, поки ми будемо прокладати свій шлях крізь цю еволюцію. Ми вступаємо в цей новий ландшафт максимально незграбно і не вигідно для нас.

---

## ХВИЛІ МОЖЛИВОСТЕЙ

У випадку з інноваціями завжди є “щось”, а потім хвилі, які це “щось” створює, як у межах нашого розуміння, так і у часі.

Візьмемо для прикладу двигун внутрішнього згорання. Він пройшов шлях від примітивного механізму до високоефективних машин, якими ми користуємося сьогодні, прогресуючи з плином часу. Зрештою він став предтечею сучасних електродвигунів, які невдовзі зроблять його непотрібним – можливо, так само, як двигуни внутрішнього згорання витіснили коней, створивши унікальну й каталітичну ланку в ланцюзі історії.

До того ж, двигуни внутрішнього згорання принципово вплинули на наш спосіб життя. Вони розширили наші моделі торгівлі, дозволили створити нові форми торгівлі, які так само призвели до зіткнення та краху культур, дали змогу роздмухувати масштабні війни й ослабили успадковані структури влади, побудовані на поділі місць у просторі.

Так, історія двигуна внутрішнього згоряння – це не просто історія інженерії, а радше розповідь про розселення людства та освоєння ним території, розширення горизонтів та переосмислення відстаней, а також це історія про те, як ми нищимо довкілля та забруднюємо клімат.

Іншими словами, інновація внутрішнього згоряння уможливила безліч змін і досягнень, стимулювала подальші інновації та створила хвилеподібну низку наслідків, як передбачуваних, так і непередбачуваних. І хоча прогнозування – це завжди сумнівне заняття, цілком ймовірно, що генеративний ШІ вже спричинив подібні хвилеподібні зміни – такі, що сягатимуть щонайменше того ж рівня, що й спадщина двигуна внутрішнього згоряння.

---

## ТРАНСФОРМАЦІЯ ОРГАНІЗАЦІЙ

Вплив генеративного ШІ на нас і наші організації, ймовірно, найліпше розглядати з трьох аспектів. Перший – це “зміни всередині наших систем”, другий – “зміни, які руйнують наші системи”, а третій – “нові можливості”.

[1]

**ЗМІНИ В НАШИХ СИСТЕМАХ** характеризуються ефективністю та оптимізацією. За допомогою наших нових інструментів ШІ у нас є можливість

підвищити продуктивність і розширити можливості (і на індивідуальному, і на організаційному рівнях), а отже, розраховувати, що вони підвищать ефективність генерування нових ідей і розкриття нових можливостей. Вони допоможуть нам співпрацювати та надавати сенс тому, що виявимо. Вони розкажуть за нас нашу історію і витягнуть нові історії з неопублікованої балаканини.

Це перші наслідки, які починають поширюватися, і ми їх уже спостерігаємо. Ці зміни знижують вартість і складність наявних парадигм роботи, замінюючи людей машинами, збільшуючи швидкість і масштаб продуктивності, і, ймовірно, випереджаючи застарілу інфраструктуру.

З точки зору бізнесу це виглядає як зменшення чисельності персоналу, більш ощадливе та стабільне виробництво, зниження витрат та підвищення доходів (принаймні, поки ринок не зрівноважить попит та пропозицію). Або, інакше кажучи, короткостроковий прибуток – постійний рушій зміни поведінки в організаціях.

Проте вплив ШІ не обмежується бізнесом: охорона здоров'я, освіта, подорожі, розваги, і навіть наші системи управління та демократії широко відкриті для впливу алгоритмів.

Хорошим прикладом є сфера охорони здоров'я. Сьогодні вона потребує складної і дорогої інфраструктури. Крім того, вона залежить від централізованого людського досвіду, який, зазвичай, є дефіцитним і дорогим у сфері знань, що швидко розвивається, а отже, на окремих фахівців лягає велике навантаження з вивчення основ та підтримання сертифікації упродовж тривалого часу.

На цей момент хвилі генеративного штучного інтелекту досягли стін наших лікарень. Ми вже спостерігаємо зростання популярності медичних чат-ботів, які мають не лише вищі показники експертизи, а й емпатії (ще однієї вразливої людської риси).

У міру зростання надійності та доступності алгоритмів охорони здоров'я вони відкриють нові можливості для наших сформованих (а іноді і застарілих) систем. Можливо, медичні боти звільнять людей-експертів від рутинних завдань, зменшивши напругу, викликану нестачею середнього медичного персоналу у всьому світі, і створивши більше простору для медичних працівників, які зможуть присвятити себе унікальним спеціалізаціям.

Можливо, ми одночасно побачимо збільшення кількості лікарів загальної практики та середнього медичного персоналу – людей, які мають навички з широкого кола дисциплін, але залежать від ШІ, який доповнює їхній широкий досвід спеціалізованими знаннями.

Можливо, ми станемо свідками переосмислення концепції клінік, оскільки потреба у фізично централізованих ресурсах починає еволюціонувати. Це може стимулювати появу легких медичних аванпостів і нових провайдерів медичних послуг. Можливо навіть сприятиме появі діаспори фахівців і відповідному розширенню доступу до високоякісної медичної допомоги у всьому світі, що стане можливим завдяки ШІ в поєднанні з робототехнікою і телемедициною.

Звісно, все це приведе до подальших змін у взаємодії із страховими компаніями, шкільній освіті та ліцензуванні, політиці охорони здоров'я та сотні інших

взаємозалежних секторів. Важко уявити цей шлях досконало, але можна передбачити, що сьгоднішні структури зміняться.

У міру розвитку внутрішніх процесів у галузі охорони здоров'я ці зміни неминуче поширюватимуться на всі інші. Це підводить нас до другої форми змін: не лише в межах сформованих систем, а й до змін, які пересувають ці межі.

[ 2 ]

**Зміни, що руйнують наші системи** не просто впроваджують ефективність у звичні процеси; вони руйнують системи, в яких ця ефективність утримується.

Наприклад, хоча ШІ і може знизити вартість виробництва або навіть розширити можливості надання медичних послуг, він також може скасувати необхідність у матеріальних і соціальних структурах, які нині підтримують ці системи. Або, інакше кажучи, зміни, які віщує ШІ, можуть привести до появи абстрактних організацій, оскільки ці системи дедалі більше розходяться з нашими нинішніми уявленнями про них, зміщуючи та руйнуючи їхні базові структури (організації, освіти, влади, контролю, управління тощо), доки вони не зможуть більше терпіти дисонанс.

Витіснення старих систем бунтарями, внутрішніми чи (частіше) зовнішніми, визначає цю другу категорію змін: руйнування і витіснення структур та історій, які ми вважали, що знаємо, помітна трансформація парадигм і поведінкових сценаріїв, зміна гвардії серед авторитетів галузі та законодавців смаків.

Це не так вже й надумано. Ми вже стикалися з такими потрясіннями раніше.

Кілька десятиліть тому, у епоху постіндустріальної революції (до настання інформаційної ери та розвитку штучного інтелекту), промисловість і влада були зосереджені в організаціях, які використовують технологію доменів (Domain-based Organisations). Ці застарілі конструкції продуктивності та результатів, зазвичай, були організовані у вертикально сегментовані функції у рамках пірамідальних ієрархій, які концентрують багатство і владу на вершині.<sup>27</sup>

Організації, які використовують технологію доменів, зростали завдяки колективізації різноманітних можливостей (багато людей, зібраних разом та організованих), а також використанню систем і процесів для забезпечення послідовності, відповідності та відтворюваності в масштабі. І все це, як правило, вкладалося в надійну фізичну, а згодом і цифрову інфраструктуру.

З настанням інформаційної ери ми спостерігаємо загальний зсув у бік збільшення використання спільної інфраструктури, мережевих взаємодій аутсорсингу. Те, що спочатку мало на меті підвищити ефективність наявних систем, ненавмисно і докорінно змінило ці системи, зруйнувавши стару інфраструктуру і винайшовши нові моделі роботи. Ми навіть стали свідками появи організацій без інфраструктури та організацій, заснованих на бренді, для яких ринкова цінність полягає в їхній історії, а не в бетоні, склі та

сталі, з яких вони виготовлені.

Для цього у нас є безліч компаній, які самі нічого не роблять, ніде не сидять, але торгують у всьому світі, продаючи нам ідеї та історії і, можливо, навіть дешево взуття та підроблені пальта, зроблені деінде та іншими людьми.

У міру зміщення балансу від інформаційної епохи до соціальної, ми побачили – і фактично продовжуємо активно спостерігати сьогодні – як зв'язки і концепції переформатовують наші організації.<sup>28</sup> Ці нові структури підривають старі конструкції влади і породжують нові рамки громадянства, переконань і приналежності – часто позбавлені будь-якої промислової інфраструктури чи прив'язки до певного географічного місця.

У книзі “Соціально динамічна організація” досліджуються ці концепції. У ній Джуліан досліджує, як технологічні зміни призвели до соціальних, організаційних і суспільних трансформацій, які радикально погіршили наші традиційні рамки влади та контролю, тонко і часто підступно замінюючи їх новими конструкціями, що ґрунтуються на зв'язках, репутації, довірі та інших соціальних валютах. Це приховані структури, які є основою цивілізацій і формують нашу поведінку.<sup>29</sup>

Потужність генеративного ШІ обіцяє ще більше прискорити процеси фрагментації і дроблення, характерні для соціальної епохи. Частково це відбувається завдяки створенню нових каналів для

оптимізації та зв'язку, але радикальніше – завдяки індустріалізації аналізу та визначенню напрямків, синтезу та творчості, а також інтеграції та використанню даних. Ці колись елітні можливості широко розкриті, ставши доступними для всіх вертикалей і підприємств будь-якого розміру. Це, безсумнівно, вплине на наші колективні рамки і засади.

Хто або що найбільше ризикує?

Зрозуміло, що робочі місця, і не тільки низькооплачувані. Генеративний ШІ впливає і на роботу випускників. Робота, пов'язана зі знаннями, консолідацією, інтерпретацією та плануванням. Робота, пов'язана з розповіддю історій, щоб впливати на інших, звітом про діяльність або аргументацією своєї позиції. Робота, яка потребує відкриттів, аналізу та ретельного інтелекту. Робота, яка передбачає ефективне управління людьми, делікатне наставництво та глибоку експертизу. Робота, що потребує співчуття та розсудливості.

Традиційні структури влади, можливостей і зайнятості, у межах яких існували ці робочі місця, також неминуче зміняться. Деякі з них можуть виявити гнучкість і пристосуватися до “змін першого типу” (‘Type-1 Change’) у межах існуючих кордонів, як ми розглядали в попередньому розділі. Однак багато інституцій чинитимуть опір, тримаючись за старі системи, доки вони не зламаються і не розпадуться на частини з появою нових парадигм.

Ми вже тисячу разів бачили, як розігрується ця історія, коли бізнес відмовляється приймати хвилі змін, доки

вони не розбиваються об їхні береги: “Кодак” і цифрова камера. “Блокбастер” і передавання мультимедійних потоків. “Сірс” та електронна комерція.

Певною мірою деякі стартапи FinTech можуть бути ранніми прикладами цих змін. Вони більше схожі на технологічні стартапи, ніж на традиційні банківські структури, часто діють з більшою соціальною метою і мають завидний рівень гнучкості, недосяжний для традиційних банків. Незалежно від того, чи досягне успіху та чи інша корпорація, стартапи FinTech у сукупності демонструють фрагментарність і вразливість застарілих структур у цьому секторі. Але FinTech – це відомий бунтівник, який з’явився в соціальну епоху, але ще до останньої революції в галузі штучного інтелекту.

Ми можемо лише припускати, як генеративний ШІ прискорить трансформацію фінансового сектора, і водночас задаватися питанням, які ще установи будуть витіснені хвилями змін. Деякі з них вже видніються на горизонті.

Візьмемо, наприклад, вищу освіту. Традиційні заклади стикаються з серйозною дилемою, яка стає ще більш обтяжливою, коли поєднується з нетерпимістю до змін. Виклики, які перед ними повстануть, – це не лише виловлювати есе, створені ШІ, або модернізувати свої платформи для дистанційного навчання за допомогою нових аналітичних інструментів; скоріше, цим інституціям – і самим їхнім концепціям – загрожують зміни, які можуть змінити основи наукового середовища.

ШІ пропонує безпрецедентну можливість



переосмислити навчальний процес. Адаптивні алгоритми навчання можуть персоналізувати освітні подорожі, віртуальні репетитори з елементами ШІ можуть запропонувати миттєві рекомендації, а в поєднанні з мобільними та інформаційними технологіями штучний інтелект дозволяє навчанню вийти за межі фізичних класів і фіксованих спільнот.

Усе це дуже круто.

Однак укорінені звичаї багатовікових освітніх моделей, з їхнім шануванням традиційних дипломів та ієрархічним поширенням знань, є потенційними перешкодами.

Спокуса підлаштуватися, зробити кроки до примирення на периферії може призвести до того, що деякі установи поверхнево інтегрують ШІ у свої нинішні структури, а потім виявиться, що їх випередили спритніші конкуренти, які вписалися в нові соціальні конструкції, що стали можливими в цей динамічний період.

Хвилі ШІ вже розмивають підвалини закритого простору навчальних закладів, якими ми їх знаємо. Ми розглянемо еволюцію навчання та розвитку більш детально в наступній частині книги. Наразі ж наш погляд на сектор вищої освіти слугує лише прикладом “змін другого типу” (“Type-2 Change”) – змін, які руйнують, витісняють і перебудовують системи.

А як щодо третьої категорії руйнувань? Після інновацій настає експлуатація: хвилі винаходів, диверсифікації, доповнень та зближення.

[ 3 ]

**Здатність виникати**, за своєю природою, неможливо передбачити. Ця третя форма руйнування виходить за рамки витіснення та переписування правил гри. Спочатку це може проявлятися у вигляді конкурентів, які діють відповідно до новомодних бізнесмоделей, створюючи нові, раніше немислимі ринки – іншими словами, бізнесів, які конкурують не лише асиметрично, а й парадигмально.

Звідти нові можливості породжують важко передбачувані вторинні та третинні ефекти: руйнують наявну організаційну злагожденість, соціальні умовності та наш звичний спосіб життя. Дійсно, наші уявлення про роботу та інфраструктуру, що її підтримує, про суспільство, мету та винагороду, навіть наші структури житла та управління можуть розпастися і бути повністю замінені чимось новим.

Під час попередніх хвиль інновацій ми були свідками подібних експансивних мінливостей. Двигун внутрішнього згоряння, про який ми говорили раніше в цьому розділі, створив ці нові можливості та викликав хвилі. Більш сучасні приклади також легко помітити, принаймні в ретроспективі:

- Компанія Netflix та потокове мультимедіа не лише підірвали індустрію розваг, а й змінили способи споживання креативного контенту аудиторією, докорінно змінили моделі прибутку медіа, а також стандарти власності та інтелектуальної власності, змінили відносини між великими

даними та розвагами, а також запровадили нові жанри сторітелінгу. Можливо, компанія Netflix та її соратники навіть підірвали та переосмислили наші механізми глобальної культури.

- Застосунки спільних поїздок, як-от Uber і Lyft, скористалися злиттям розподілених мобільних систем, машинного навчання та кишенькових GPS-технологій, щоб не лише конкурувати з традиційними таксі, а й стати каталізатором гіг-економіки та всіх її наступних наслідків, як-от перегляд структур зайнятості, оскарження трудового законодавства та нові дебати про права працівників і природу праці.
- Винахід соціальних мереж в інтернеті за допомогою таких платформ, як-от MySpace та Facebook, не лише спричинив революцію у міжособистісному спілкуванні, а й змінив способи обміну інформацією, наші погляди на приватність та цифрову ідентичність, а також динаміку влади та репутації. Вони змінили структуру суспільства, поширюючи дезінформацію та розширюючи можливості інформаційних бульбашок, а також створюючи шляхи для колективних дій і даючи можливість висловитися прихованим спільнотам.

Варто пам'ятати, що всі наші соціальні та суспільні структури вигадані; ідея їздити на роботу, заробляти гроші та бути керованим, ринки та фабрики, правові системи та освіта – все це вигадано. Це все можна анулювати, абстрагуватися від існування або розвинути.

Генеративний ШІ може впливати на деякі соціальні

структури сильніше, ніж на інші. Наприклад, його здатність синтезувати і створювати артефакти у валютах науки і культури є дуже важливою. Ці валюти володіють – або, можливо, *володіли* – великою неприборканою міццю. Як вплине на цю динаміку відкриття воріт доступу і перевантаження каналів споживання? Як комерціалізація цих артефактів змінить те, що ми визначаємо цінність, або те, як і що ми обираємо для придбання чи проведення часу?

Нові можливості, що з'являються завдяки генеративному ШІ, майже напевно призведуть до нових парадигм у сфері охорони здоров'я, фінансах, освіті та виробництві. Традиційні сектори, безсумнівно, постануть перед викликами, зіткнувшись, ймовірно, з далекосяжними структурними змінами. Отже, можна з упевненістю стверджувати, що принаймні деякі з цих змін спричинять нові хвилі.

Але розмови про зміни нічого не варті і, загалом, часто є недоречними.

Наприклад, незважаючи на початковий ажіотаж, ми ще не бачили радикальних змін (нашою мовою “Зміни третього типу” (“Type-3 Change”)) у технології блокчейн або Інтернету речей (IoT); натомість ми спостерігаємо експерименти та поступову інтеграцію.

Тож, хто знає, де ми опинимося? Як і у випадку з багатьма речами в житті, справжнє питання полягає не стільки в пункті призначення, скільки в нашій подорожі до нього. Як ми вирішимо орієнтуватися в цьому морі змін? Чи будемо ми намагатися підтримувати усталені методи

і поступово інтегрувати зміни? Чи спробуємо сміливо експериментувати з нестабільними, але революційними інноваціями? Чи варто нам розпростерти обійми і підняти руки догори, або ж міцно заплющити очі?

---

## МЕТАМОРФОЗИ МАЙСТЕРНОСТІ

Генеративний ШІ здібний, дуже здібний. Фотографії (здебільшого) виглядають справжніми. Есе (здебільшого) важко відрізнити від написаного людиною. Навіть музику (здебільшого) неможливо відрізнити від популярних виконавців, яких вона імітує. Це досить моторошно.

Що це означає для майстерності та для творців, які стоять за мистецтвом, яким ми насолоджуємося сьогодні? Творча майстерність глибоко вкорінена у суспільстві відтоді, як з'явилися звукозаписи. Це спосіб нашого спілкування, спосіб вираження нашої колективної душі та нашої особистої ідентичності:

Я – письменник.

Я – художник.

Я – музикант.

Мистецтво промовляє до нас у спосіб, який живе ще довго після його творців. Але це не означає, що шлях художньої творчості та захоплення мистецтвом був гладким. Насправді, навіть побіжний огляд історії показує, що еволюція мистецтва завжди була бурхливою, з хвилями інновацій, що розбиваються об



норму: невеликі групи художників, які кидають виклик звичаю, щоб спробувати щось нове, кидають виклик мистецьким ідеалам того часу, поки не стають новим істеблшментом.

Так було завжди – від Пікассо, який ображав своїми абстрактними стилями, до Уорхола, який масово продукував поп-арт, і до Бенксі, який нелегально розпилював фарбу серед ночі. Гері порушував архітектурні норми своїми деконструйованими роботами, а реп руйнував усталені творчі структури, не кажучи вже про глибинну динаміку влади в музичній індустрії.

Нові мистецькі ідеї стикаються з поточними смаками, знаходять нову аудиторію, а потім знову вливаються в мейнстрім.

Наші естетичні та моральні стандарти змінюються з часом. Графіті – ще один нещодавній приклад, який швидко перетворився з антиістеблшменту на легітимний жанр, а багато художників графіті стали комерційно успішними (на додаток до нічних вандалів). Татуювання зараз є мейнстрімною формою мистецтва, як і пісочна скульптура, і – на щастя чи на жаль – як і генеративний ШІ.

Сучасні цифрові художники мають потужних роботів помічників, які опираються на глибокі знання про мільйони інших творів мистецтва, щоб кинути виклик і розширити свої художні можливості. Цінові інструменти також відкривають двері творчості для ширшого кола дизайнерів. Програми зі штучним інтелектом, що генерують зображення, як-от DALL-E, Firefly, Midjourney та Stable Diffusion, дозволяють аматорам малювати

словами або писати свої історії з вільних фрагментів.

Мистецтво – це і підтекст, і ремесло. Це і пристрасть поділитися посланням зі світом, і фізична здатність творити.

Сотні років великі художники мусили подрібнювати свою сировину та утаємничували рецепти, для приготування пігментів. Їхня професія була як ремеслом створення фарб, так і мистецтвом нанесення їх на полотно.

Раніше графічним дизайнерам доводилося малювати пером і тушшю, ретельно обводячи ескізи на синіх лініях, вручну розділяючи кольори на різних аркушах і вручну розраховуючи пагінацію.

У кожному випадку інновації, чи то у вигляді тубиків з фарбою, чи то у вигляді дизайнерського програмного забезпечення, сприяли руйнуванню. Дехто може заперечити, що генеративний ШІ відрізняється від цих аналогій, оскільки це не просто інструмент, яким користуються художники, а й сам може створювати певні мистецькі витвори. І хоча в цьому аргументі є певна рація, порівняння з тубиками з фарбою та цифровою версткою є більш точним, ніж ми можемо собі уявити на перший погляд.

Для тих, хто зосереджений винятково на ремеслі, штучний інтелект створює загрозу суттєво знецінити їхню роботу. Для художників, які наповнюють свої роботи соціальними підтекстами, новими стилями або унікальною експресією, ШІ пропонує себе як вірний помічник і творчий спаринг-партнер.



*Ceci n'est pas de l'art*

Що робить твір мистецтва “справжнім”?

Чи має значення процес його створення?

Чи зусилля та почуття митців  
наділяють їхні творіння цінністю,  
чи цінність має сам артефакт?

Частина цінності, яку  
несе в собі мистецтво,  
зумовлена його новизною  
та майстерністю.

Банальне, бездумне і без натхнення швидко знецінюється. Подібно до того, як ми стали свідками поширення жахливого дизайну і моторошних фотографій завдяки легкості сучасного цифрового дизайну – мільйони жахливих фільмів на YouTube, спотворених фотошопом зображень і бездушних корпоративних слайд-шоу – ми також почали цінувати якісні артефакти, яких у нас стало більше завдяки низьким бар'єрам для входу на ринок.

Це може бути добре...

У нас є цілі покоління, які можуть створювати пісні та розповідати свої історії без особливих зусиль, і генеративний ШІ поширює цей доступ далі, одночасно знецінюючи ручну працю, яка колись була необхідною частиною створення артефактів.

Це може бути добре...

...якщо це не підриває основи професійного мистецтва.

Справедливо запитати, чи не створили ми покоління дрібних аматорів і шаблонних підробок, підриваючи до того ж найоригінальніші креативи. Що станеться, коли у наших роботів не буде нових творів мистецтва, на яких вони зможуть вчитися? Або коли узурпатори творчості збагачуються за рахунок творців? Що відбувається, коли студія переробляє голос співака зовнішність

актора стає власністю корпорації, а стиль дизайнера перетворюється на імітацію за допомогою клік-ферм?

Генеративний штучний інтелект пишається своєю продуктивністю, відтворюючи меш-апи на вимогу і не посилаючись на безліч творів мистецтва, які є його основою. Кожен результат, безумовно, унікальний, але де його походження, і як нам віддати належне художникам, чії оригінальні роботи навчили машини?

Ми створили способи (хоч і недосконалі) розуміння автентичності та визнання мистецького внеску. Творчі роботи захищені авторським правом, і автори можуть розраховувати на отримання своїх грошей і роялті за їх офіційне використання, а також за їх похідні, як-от ремікс музичного семплу, що є основою нової пісні. Крім того, у нас є багаторівневий ринок цінностей, який надає перевагу оригінальним творам або творам з обмеженим тиражем перед масово виготовленими варіаціями, навіть якщо вони виглядають або звучать схоже.

Ці соціальні конструкції навряд чи виживуть у своїх нинішніх формах, що, безсумнівно, приведе до переосмислення стимулів для митців. Ще до того, як з'явився цей новий виклик, пов'язаний зі штучним інтелектом, існував розрив між творцями мистецтва та механізмами перетворення їхніх творів на гроші.

Стрімінгові сервіси продають доступ до пісень, тоді як музиканти отримують лише невелику частку від цих доходів. Колекціонери виставляють на аукціони твори мистецтва за шалені суми, тоді як оригінальні художники не отримують прибутку від перепродажу. Вуличні художники розписують стіни провулків своїми

повідомленнями, а потім бачать, як їхні роботи викрадають і продають на приватних ринках.

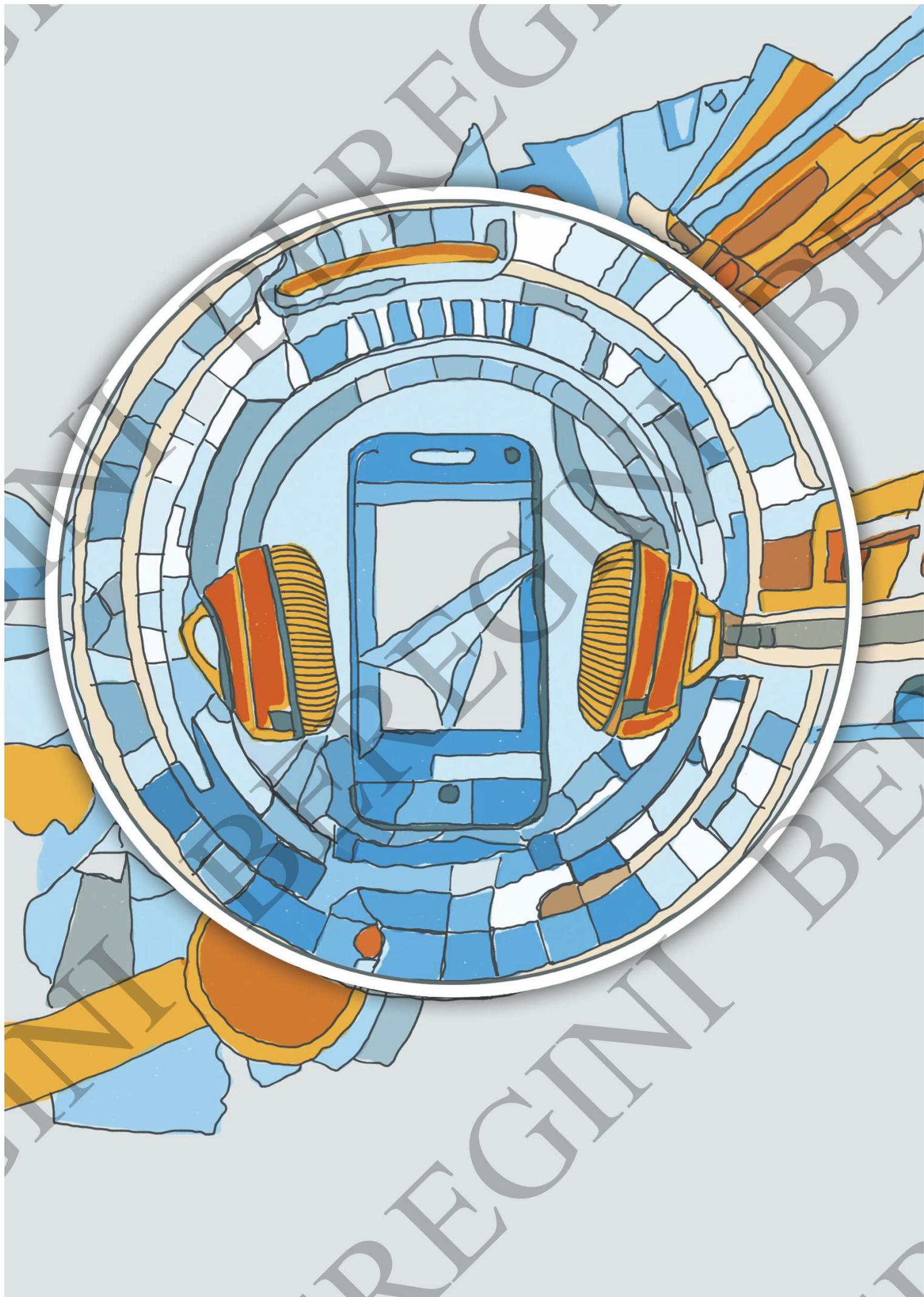
Ми також стали свідками індустріалізації мистецтва, яку циніки можуть назвати бездушною: розважальні агенції масово продюсують кей-поп гурти за шаблонними схемами; великі кіностудії створюють сюжети на основі фрагментів глядацьких переглядів; а нещодавно з'явилися алгоритмічні маніпуляції, як-от використання знаменитостями ботів для популяризації їхнього контенту.

Чи цінуємо ми досі оригінальність? І чи готові ми за неї платити?

Ці проблеми інтелектуальної власності та ризики для засобів до існування творчих людей є реальними, і хоча поширення генеративного ШІ може висвітлити цю боротьбу більш яскраво, цей шлях ми вже пройшли – та активно рухаємося ним у багатьох напрямках.

Ми продовжуємо блукати нею, незадоволені та невпевнені, ставлячи під сумнів і борючись зі студіями та синдикатами за права митців, водночас транслюючи та клонуючи їхню творчість. Генеративний ШІ – це не стільки окремих напрямків, скільки ще один відрізок тієї ж кам'янистої дороги, і, можливо, нове світло, яке проливає генеративний ШІ, допоможе нам ліпше розгледіти деякі пастки на цьому шляху.

У нас ще немає відповіді на це питання, але якщо ми, як суспільство, цінуємо творчість і справжнє мистецтво, то ми повинні знайти способи підтримати митців, які це створюють.



Геракліт зауважив, що ніхто не може увійти в одну річку двічі, частково тому, що ніхто не залишається однією і тією ж людиною з плином часу. Наші думки та міркування, вчинки та пригоди, невдачі та успіхи змінюють нас. Ми ростемо завдяки їм. Вони формують нас, перетворюючи у наші майбутні “я”.

У міру того, як ми передаємо дедалі більше досвіду штучному інтелекту – наприклад, просимо його згенерувати повідомлення для листівки зі співчуттям або намалювати малюнок до дня народження матері, – страждає не тільки (потенційно мимовільний) одержувач. Ми, творці, втрачаємо можливість досліджувати свої почуття під час написання, або спогади про проведений із мамою час, коли ми робимо начерки, або розчарування від невдалих мазків пензлем і маленьких уроків терпіння та саморозвитку, які ми отримуємо в такі моменти.

Коли мільйони цих скромних і невибагливих моментів замінюються автоматикою – коли артефакти, колись створені нашими власними руками і розумом, натомість легко виробляються бездумним натисканням кнопки, – як це змінює (чи не змінює) нас як людей?

Загалом, ми, як соціальні істоти, позбавляємо себе дедалі більшої кількості цих мікродосвідів – як це змінює наші взаємодії та очікування, сценарії та історії, терпимість і цінності? Усуваючи стільки розбіжностей а, отже, стільки досвіду з акту творіння, ми збідніємо у відносинах, які проникають у наші думки і поведінку, у наші племена і спільноти, і, можливо, навіть у структури суспільства в цілому.

---

## ЗАНЕПАД КУЛЬТУРИ

Мистецтво і культура – це симбіоз. Культура відображається в артефактах, які ми створюємо, і залишає свій відбиток у них, а ці артефакти передають дух часу, дух епохи.

Мистецтво віддзеркалює життя, життя віддзеркалює мистецтво.

Природно, що коли створення та поширення мистецтва (у всіх його різноманітних формах) стає більш доступним, це має зворотний вплив на культуру. Ще до стрімкого поширення генеративного ШІ, ми вже почали помічати перші ознаки цього явища:

### Період напіврозпаду культури скорочується.

У попередні епохи на формування нових стилів у мистецтві, музиці та літературі йшли десятиліття, якщо не покоління, а то й десятиліття. Культура розвивалася повільно, і з кожним новим шедевром і рухом ми всі разом отримували нові перспективи. Митці відкривали для себе нові глибини натхнення, і поступово, часто проростаючи із затертих часом залишків попередніх ітерацій, розквітли нові цикли культури.

У нашому перенасиченому засобами масової інформації світі інформаційної епохи ми вже бачимо, що нові творіння (якщо вони популярні) утримують нашу колективну увагу лише на мить, доки їх не витіснить

нове, а потім ще новіше, у прискореному і дедалі коротшому життєвому циклі стереотипів і трендів.

У міру того, як культурні революції стають все більш впорядкованими, а ми закручуємо спіраль все швидше, ми втрачаємо дещо з їхнього досвіду. Ми втрачаємо споглядання та вкорінення ідей, які колись знаходили у тих, хто не встигав у ногу з часом. Це не просто ностальгічна перспектива. На цю динаміку впливають численні сили соціальної епохи – наша радикальна зв'язність, зростання спільнот, перерозподіл влади та цінність автентичних голосів, – все це впорядковує механізми культури.

ШІ скорочує шлях ще більше.

Генеративний ШІ прискорює виробництво креативної продукції і відкриває двері творчості для більшої кількості митців, кожен з яких може створювати нові роботи з меншими затратами часу і, часто, меншими зусиллями.

Мільйони художників-початківців можуть досліджувати стиль, проживати нескінченні зміни, врешті-решт натрапити на щось нове, можливо, навіть швидкоплинно популярне, а потім так само легко відкинути це нововведення заради наступного, не надаючи великого значення винаходу і витрачаючи мало часу на його обмірковування. Це не обов'язково погано саме по собі, але це риса, що явно розвивається в механізмах культури.

Ми скорочуємо не лише важкий процес вдосконалення, а й часові рамки творчості. Оскільки генеративний

ШІ та інші цифрові інструменти прискорюють цикли створення, поширення, фрагментації та радикальних ітерацій, як це змінює наші стосунки з художнім вираженням та його культурною значущістю?

Ми стали ненаситними у своїй здатності споживати медіа та швидкі у своїй схильності втомлюватися від вчорашніх творінь. Ми дедалі більше звикаємо до новизни і залежимо від неї, а наш мозок вже налаштований на те, щоб надавати цінності дефіциту і, так знецінювати буденність.

Надлишок скорочує тривалість нашої уваги та зменшує глибину взаємодії, потенційно зменшуючи емоційний та інтелектуальний вплив артефактів і кидаючи виклик нашому почуттю культурної пам'яті. Традиційні маркери цінності, як-от рідкісність та ексклюзивність, відповідно зменшуються, трансформуючи те, як ми надаємо значення різним творам мистецтва.

Колесо обертається щоразу швидше.

Хоча кожний наступний виток, здається, має все менше значення.

Ми маємо певний вплив на те, як розгортається ця пригода, навіть якщо тільки у власній свідомості. Чи присвячуємо ми своє культурне “я” глобальним артефактам зі штучним інтелектом, які заповнили наші стрічки? Чи приречено і бездумно прокручуємо швидкоплинний контент? Чи навмисно сповільнюємо час, шукаючи локальні ретро-артефакти та практикуючи усвідомлене споживання?

Чи впливе це все назовні?

Можливо. Але це може призвести до нашого занепаду.  
Нас може стати менше, навіть якщо створюємо більше.

Не існує заздалегідь визначеної умови, щоб усе завжди ставало ліпшим або розвивалося по висхідній. Також немає жодних підстав для того, щоб “інакше” означало “гірше”. У будь-якому разі іноді нам доводиться керувати ситуацією.

---

## РУЙНУВАННЯ НАШОЇ ПРАВДИ

Якщо вам вдалося уникнути переказу своїх заощаджень нігерійському принцу і відмовитися від придбання панацеї у різних шаманів сьогодні, то, найімовірніше, ваші BS-фільтри працюють просто чудово – або працювали досі.

З роками ми вивчили ознаки спаму: погана графіка, дивні адреси та мовні помилки. Ми навчилися розпізнавати клікбейт-статті з їхніми шаблонними заголовками. Зазвичай, ми також могли розпізнати відео з “фейкових новин” з їхнім поганим монтажем та явними викривленнями, і більшість з нас двічі подумали б, перш ніж передати свої банківські реквізити телефонному оператору, навіть якщо він назвався давно втраченим родичем.

Зворотний бік також був правдивим. Ми розробили фільтри, щоб відрізнити справжнє: добре написані та персоналізовані повідомлення, відео високої чіткості та

професійні фотографії, опубліковані відомими ЗМІ.

Але наші фільтри засмічуються.

Двигуни інновацій частково перетворилися на двигуни спотворення, деформуючи нашу реальність ненавмисно і підступно. Генеративний ШІ пропонує незліченну кількість нових способів розмити наше уявлення про реальне та істинне. Він навіть змушує нас замислитися, наскільки дійсною залишається різниця між реальністю та вигадкою.

Наші лінзи розбиті вщент. Наше сприйняття істини занепадає.

Небезпека виходить за межі маргінальних шахраїв та промислового обману. Сьогодні ми стурбовані стійкістю демократії та надійністю глобальної комунікації.

Ми звикли довіряти писаному, але тепер ми навчилися йому не довіряти. Колись ми вірили в реальність, зафіксовану на фотографіях, а тепер сумніваємося у тому, що бачимо на власні очі. Навіть відео та голос перетворилися на те, що можна створити з пилу, – витончений – і водночас легко здійснений – цифровий обман. У результаті підозрілість проникла майже в усі канали, і ми часто реагуємо на нову інформацію з недовірою, іноді ніколи не зрушуючи з цієї стартової лінії, незважаючи на докази.

Тому що просто “бачити” вже не означає “вірити”.

Це більше, ніж мейнстрімне розмивання істини, це також зростаючий цинізм щодо поняття “істина” взагалі. Як люди, наш мозок схильний відключатися від такої

складності. Велика кількість конкуруючих наративів, а також гіперобізнаність і гіперкритичність, необхідна для того, щоб орієнтуватися в них, відштовхують нас від пошуку сигналу в шумі. Ми ставимо під сумнів і сумніваємося, потім втомлюємося і відмовляємося від усього цього.

Коли правда спотворюється, довіра стає більш крихкою. Ми скептично налаштовані до сигналів ззовні, які не піддаються нашому безпосередньому контролю.

Наш світогляд звужується.

Ми відключаємося і не віримо.

Речі, які ми шануємо, можуть зникнути в міру того, як ця віра руйнується.

“Істина” – складне слово. Ми використовуємо його поперемінно для опису того, що є універсальним та індивідуальним, що піддається кількісному оцінюванню, і того, у що ми просто віримо. Це може здатися анафемою для тих, хто прагне “правди” або “істини”. Звичайно, існують більш універсальні принципи: гравітація – це “істина”. Але так само, як і моє почуття належності.

Істина може бути контекстуальною. Наукова “правда” – та, за яку ми даємо Нобелівські премії – може бути менш мінливою, але брудна і розбита “правда”, яку кожен з нас конструює у своїй свідомості і у своїх спільнотах, є суб’єктивною. Ми віримо в ці істини і, коли вони піддаються сумніву, часто чіпляємося за них ще міцніше. Ми боремося за ці істини не на життя, а на смерть.

На індивідуальному рівні ми схильні вважати, що інші

є більш вразливими, більш схильними до обману, ніж ми самі. Але, можливо, природа обману змінюється, оскільки ми занурюємося в море нових “істин” з нескінченними історіями, що впливають на нас. Розповсюдження створених “альтернативних істин” може підштовхнути нас до заселення ландшафтів розділених переконань, і, можливо, ми ризикуємо взагалі відмовитися від концепції колективної істини.

А можливо, у нас все налагодиться. Ми намагаємось не передбачати майбутнє, а прагматично дивитися на речі. Практичний шлях потребує від нас ясного бачення механізмів змін, які впливають на нас – не для того, щоб зробити висновок (чи сконструювати нову жорстку “істину”), а й не для того, щоб наївно вірити, що все самотійно налагодиться.

Щоб знайти шлях крізь шум, нам потрібно розробити нові фільтри. Нам знадобляться нові способи навігації в інформації, нові інструменти для побудови довіри та нові механізми для створення спільних знань і вивчення великого світу.

За іронією долі, нам доведеться покладатися на “двигуни” – ті самі інновації, які протистоять концепції правди, щоб підготувати та озброїти нас до подорожі цим ландшафтом “постправди”.





---

# ІНТЕРМЕДІЯ ПРО СИСТЕМИ СИСТЕМ

МАРКА ОЛЕРТА

Є такий старий анекдот: “Я не знаю, хто винайшов воду, але б’юся об заклад, що це була не риба”. Мається на увазі, що ми схильні ігнорувати ті частини нашого середовища, які одночасно присутні, але якось відходять на задній план.

Як це пов’язано з генеративним ШІ і змінами, які він вносить майже в усе? Усвідомлення цих невидимих, але важливих елементів нашого робочого середовища, як тектонічних плит, покаже нам, де відбуватимуться зіткнення між старим і новим. Погляньте на цей перелік: авторське право, подвійна бухгалтерія, посадові інструкції, атестації та оцінки. Тією чи іншою мірою всі ці системи або аватари систем будуть обмежувати, впливати або перешкоджати розгортанню генеративного ШІ.

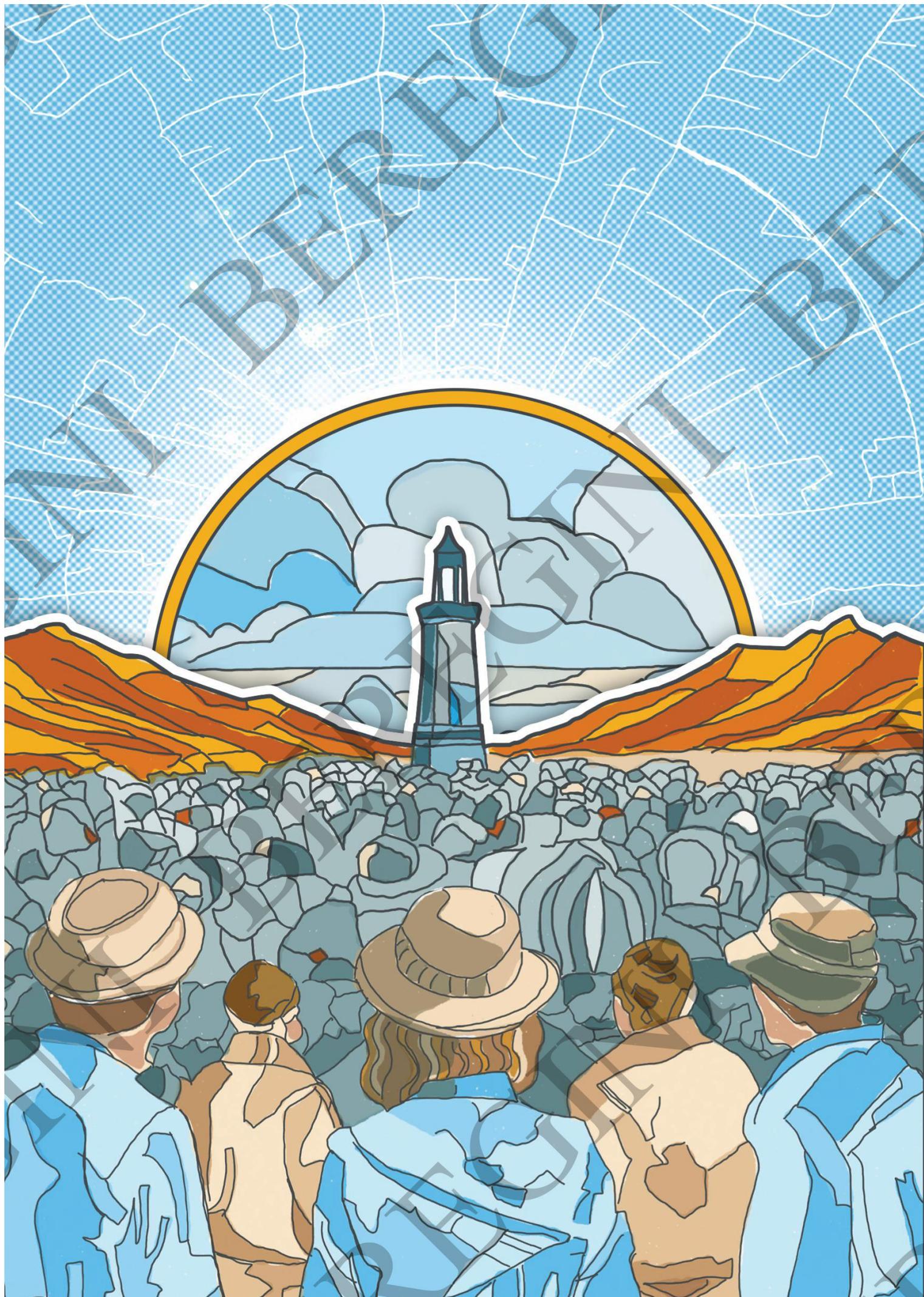
Конфлікт авторських прав вже розгорівся. Художники подали позови проти MidJourney, Stable Diffusion і DALL-E, звинувачуючи їх у тому, що вони навчили моделей імітувати стиль художника без його дозволу або компенсації. Вже існує стартап Kudurru, який дозволяє художникам або блокувати адреси інтернет-протоколів на основі штучного інтелекту, або фактично “отруювати” запити штучного інтелекту, надсилаючи назад зображення, відмінне від запитуваного. Існують подібні судові позови і щодо текстових зразків. Ці суперечки доведеться вести та вирішувати в міру того, як ми будемо рухатися вперед.

Припустимо, ви впроваджуєте генеративний ШІ у своїй організації. Ваша організація має контракт з OpenAI, який дозволяє вашим співробітникам створювати GPT, потужні та кастомізовані версії бота ChatGPT, які окремі особи та організації можуть розробляти для конкретних цілей.

Отже, уявіть, що ваша організація має кілька таких ботів для внутрішнього використання. Чи наймали ви когось для цього? Чи це міститься у посадових інструкціях? Як ви оцінюватимете їхню роботу та їхню цінність для організації? Якщо ці люди почнуть впроваджувати дивовижні GPT, які заощаджують організації значні суми грошей, як ви оціните цінність цієї людини для компанії? Згідно з принципами подвійної бухгалтерії ви можете відображати працівників у звіті працівників тільки як зобов'язання, витрати.

І останнє: якщо генеративний ШІ зможе автоматизувати так левову частку нашої діяльності в організаціях, то як буде вимірюватися наша цінність? Як ми будемо думати про переміщення роботи людей вгору по ланцюжку цінностей, щоб генеративний ШІ не просто поглинав наші організації знизу вгору?

Існує сайт під назвою *“Для цього є ШІ”* (“There’s an AI for That”), який містить індекс впливу ШІ на робочі місця. Індекс визначає відсоток роботи, на яку може вплинути штучний інтелект. У генерального директора індекс впливу становить 91%, здебільшого він стосується управління командою та оцінювання ефективності роботи. Чи почали ми думати про цей рівень впливу зверху вниз і про те, як ми повинні усвідомлювати і працювати над адаптацією наших систем до цього впливу? Ця книга допоможе вам у цьому. ★



## ЧАСТИНА 5

### У діалозі з навчанням

З моменту народження, дивлячись на розмиті обличчя батьків, ми розпізнаємо закономірності і намагаємося знайти сенс у навколишньому світі.

Наш мозок напрочуд гнучкий, і що різноманітніший та багатший у нас досвід, то гнучкішим і потужнішим стає наш мозок. Ми здатні орієнтуватися в нових, раніше незвіданих ситуаціях, спираючись на фрагменти знань і життєвого досвіду, які ми накопичили упродовж життя.

**Навчання** – це процес отримання досвіду і зростання, а потім міркування і взаємодії на основі цих переживань. Це квінтесенція людського буття. Це щось на кшталт нашої суперсили.

Машинне навчання має дивовижні паралелі з нашими власними (людськими) процесами навчання: алгоритми, по суті, навчаються самі, спостерігаючи та осмислюючи багато мільйонів цифрових досвідів. Завдяки цьому машинне навчання та його похідні: глибоке навчання і генеративний ШІ, здатні докорінно змінити те, як ми, люди, взаємодіємо з навчанням – у

школах і професійних коледжах, а також у багатьох інших різноманітних формах та аспектах. Ми тільки починаємо розуміти, як це може виглядати...

---

## ЛІНЗИ НАВЧАННЯ

Навчання відбувається скрізь і повсякчас. Якщо ви йшли незнайомим коридором і спіткнулися об добре замаскований дверний обмежувач, ви отримали незабутній урок. І в майбутньому ви, ймовірно, скоригуєте свою поведінку, щоб уникнути цієї перешкоди, що свідчитиме про те, що ви навчилися.

Ми вчимося на роботі, у наших спільнотах, від наших наставників і колег, і, звичайно, у рамках формально визначених структур. Більшість із нас вперше отримали формальну освіту в дитинстві, ймовірно, в класі з партами, вишикуваними в ряди, обличчям до вчителя, який стоїть перед дошкою з крейдою або, у наш час, смарт-дошкою.

Цей образ “традиційного” класу часто називають парадигмою навчання “індустріальної епохи”, і він бере свій початок у прусській школі XVIII століття, яка була революційним нововведенням для свого часу. Ці початкові школи сприяли широкому розповсюдженню грамотності та лічби серед хлопчиків та дівчаток у часи, коли масова освіта ще не стала звичним явищем. У пруських школах робився акцент на дисципліні, слухняності та повазі до влади. Від учнів очікували

дотримання правил, а вчителі використовували переважно методи навчання через повторення і зазубрювання.

У ХХ столітті з розвитком теорії освіти та психології змінювалося і наше розуміння педагогіки (науки про викладання). На зміну старим парадигмам біхевіоризму прийшов конструктивізм, ознаменувавши перехід від зазубрювання до внутрішнього конструювання знань, і були впроваджені активніші методи навчання – хоча й у класах, які все ще багато в чому нагадували старі пруські школи. Тож, хоча ми й оновили наші навчальні програми відповідно до сучасних вимог, ми все ще застосовували їх у рамках моделей індустріальної епохи.

З наближенням ХХІ століття, коли Всесвітнє павутиння відкрило широкий доступ до інформації, ми ще більше засумнівалися в пруській моделі, вважаючи, що вона не відповідає розвитку освітньої та інформаційної сфер. Ця модель здавалася надто спрощеною в епоху, коли інформації ставало дедалі більше і вона ставала доступнішою.

Поширення доступу до інтернету, пошукових систем і соціальних мереж полегшує пошук книг і статей, покрокових інструкцій та навчальних відео. Багато в чому ми навіть можемо замінити інтернет-артефактами наші когнітивні навички нижчого рівня. Нам більше не потрібно запам'ятовувати прості факти; ми можемо знайти їх, кілька разів натиснувши клавіші. Нам не потрібно розуміти основоположні принципи, щоб розв'язати проблему; ми можемо попросити

онлайн-спільноти: “поясни, ніби мені п’ять років”.

Ці нові можливості та їхній вплив на наше навчання і пізнання спричинили чергову переоцінку. Освітняни почали акцентувати на “навчанні інформаційної епохи” та “компетенціях двадцять першого століття”. Учителів заохочували розвивати у своїх учнів навички критичного мислення, аналізу та синтезу, а також використовувати більш цілісні та адаптивні підходи. Крім того, з’явилися нові теоретичні парадигми про соціальну епоху та теорію конективізму. Ці концепції наголошують на розподіленому пізнанні, мережевому досвіді, соціальному навчанні та синтезі фрагментованої інформації. Вони визнають, що осмислення є як внутрішньою, так і колективною здатністю.

Конективізм розглядає навчання як соціальний процес – не як щось винятково внутрішнє та індивідуальне, а як щось розподілене та динамічне. Конективізм зосереджується на розвитку здатності учнів маневрувати в мережі джерел і спільнот, розкиданих у нашому соціотехнічному світі. Він пропонує учням зосередитися на тому, щоб навчитися орієнтуватися в інформації та розвивати мережі ресурсів та однолітків. Ця теорія заохочує розвиток таких навичок, як цифрова грамотність, мережева грамотність і здатність критично осмислювати інформацію з різних (а іноді й сумнівних) джерел.

Аналогічно, поняття соціального навчання в “соціальну епоху” визначає навчання як конструювання та обмін “сенсами” всередині та між людьми, а також між організаціями.<sup>30</sup> Створення сенсу виходить за межі передавання інформації, виробництва знань і навіть окремої людини – це наратив, який виникає, або різні наративи, що виникають у спільноті. “Сенс” виступає і як доробок (результат) навчання, і як схема для постійного утворення “сенсів” спільноти: сприйняття, прийняття рішень і дії відбуваються в мінливому комплексі ідей, ідентичностей та інтерпретацій, в яких люди та організації повинні навчитися орієнтуватися.

Сучасний ШІ ще більше змінює наше ставлення до знань.

Хоч можливості соціальної епохи і конективізму, ймовірно, залишаться незамінними, так само безсумнівно, що вони – і старі рамки пруської освіти, в яких вони були розроблені, – більше не будуть достатніми.

---

## НА ШЛЯХУ ДО ЗНАНЬ

Алгоритми глибокого навчання виявляють складні закономірності для нашого інтелектуального споживання. Наративні та діалогові інструменти витончено вписуються в наші природні механізми допитливості та дослідження. А у своєму повсякденному використанні генеративний ШІ точно відтворює багато

аспектів людського пізнання вищого рівня. Ми можемо не тільки практично миттєво знаходити інформацію, а й просити допомоги в її узагальненні, аналізі, зіставленні з іншими концепціями, застосуванні у наших унікальних ситуаціях і навіть оцінці того, чи є і чому певні концепції є релевантними.

Усе це вносить значну невизначеність у ландшафт навчання, зокрема формальної освіти, тренінгів та тестування.

Можна стверджувати, що зараз на порядку денному практично все: характер і роль викладання, право власності і використання навчальних матеріалів, структура наших освітніх систем, етика і механізми оцінювання, вплив щораз більшої цифрової нерівності, а також наша усталена модель освіти, продуктивності і розвитку здібностей. Ці нові інструменти навіть кидають виклик нашим поглядам на те, що таке “спроможність”, а також на те, як і де її зберігають окремі особи та інституції.

Для початку давайте розглянемо фундаментальне навчання.

У старій моделі людина могла “шпортатися на шляху до знань”, виконуючи низькорівневі завдання, часто просто ліпші за звичайну роботу, що могло тривати роками. У початковій школі це могли бути місяці вивчення граматики і правопису протягом багатьох семестрів, сотні погано написаних есе в середній школі, а згодом ведення блогів, написання статей на загальні теми і їх публікація.

Деякі з цих працівників (у нашій розповіді – це професійні письменники-початківці) поступово отримали знання, які привертали до них увагу боса чи наставника, який піднімав їх на наступний рівень продуктивності – “потонути або виплисти”. І поступово вершки (іноді) піднімалися б (деякий час) наверх, оскільки люди будували б свою внутрішню інфраструктуру здібностей, крок за кроком знизу вгору.

З появою плагінів для перевірки орфографії та граматики, не кажучи вже про генеративний ШІ для створення базових есе, багато з цих завдань початкової освіти та початкового рівня стають менш важливими, а то й зовсім непотрібними. Однак основні навички, отримані в результаті цього досвіду, залишаються важливими, як-от розвиток наполегливості, розсудливості та почуття самосвідомості, а також інтуїтивне розуміння таких складових гарного письма, як темп, чіткість і голос.

Подібні історії можна розповісти про цифрових художників, розробників навчального контенту, кінематографістів, юристів, фінансових консультантів, розробників програмного забезпечення та представників багатьох інших професій. У різних дисциплінах початковий рівень, через який проходять новачки, історично розвинув мислення та здібності, що зрештою стали основою їхніх знань.

Але вже ні.

Отже, як нам скоригувати наші освітні моделі, щоб продовжувати розвивати ці експертні здібності – як-от компетенції вищого рівня, тонке розуміння контексту

і автоматизм – коли стимули для проходження нижчих рівнів майже зникли? Як ми можемо допомогти новачкам удосконалювати свої знання та навички, коли базові компоненти, на яких вони побудовані, можна без особливих зусиль отримати за допомогою цифрових інструментів?

Ми вже бачили подібні сценарії раніше.

Коли в освіті з'явилися калькулятори, люди побоювалися, що учні перестануть вміти рахувати, якщо не виконуватимуть множення чотиризначних чисел подумки. Озираючись назад, нам слід визнати, що математика – це набагато більше, ніж базова арифметика, і що цифрові допоміжні засоби насправді прискорили розвиток цієї галузі. Сьогодні математики співіснують з технологіями, навчаючись користуватися графічними калькуляторами в ранньому віці, застосовуючи спеціалізоване програмне забезпечення у своїх випускних проєктах і створюючи власні макроси та застосунки на роботі.

Такі навички, як оцінювання, грамотність у роботі з даними та ймовірнісне мислення, набули особливого значення не лише серед професійних математиків, а й загалом. Учорашні фахівці, які працювали до появи цифрових технологій, розвивали такі здібності інтуїтивно, прокладаючи свій шлях від усної алгебри до нудної математики нижчих порядків. Сьогодні нам потрібно сумлінно розвивати в учнів навички обчислень вищих рівнів, і ми успішно розробили навчальні програми та педагогічні методики для виконання цього завдання.

Нам би хотілося вважати, що генеративний ШІ піде схожим шляхом, але, незважаючи на поверхневу схожість між появою калькуляторів і нашою нинішньою ситуацією, ці два випадки можуть бути не такими вже й аналогічними

Цифрові інструменти чудово вписуються в математичний інструментарій, зрештою, вони її близькі родичі. І цифрові інструменти поширювалися поступово, з розвитком інформаційної епохи. На противагу цьому, впровадження генеративного ШІ було одночасно широким і раптовим, в одну мить похитнувши основи майже всіх академічних дисциплін.

Менш ніж за рік майже кожен комерційний інструмент для підвищення продуктивності ввібрав у себе певну варіацію цих можливостей ШІ. Генеративні алгоритми ШІ вже непомітно підштовхують і поліпшують нашу роботу. Масштаб і швидкість змін свідчать про те, що це навряд чи буде поступова еволюція, яку ми спостерігали у випадку з калькуляторами та комп'ютеризованими електронними таблицями. Найімовірніше, це буде руйнівна еволюція, яка зламає усталені парадигми і спричинить непередбачувані хвилі в різних галузях, відкриваючи нові можливості і даючи поштовх інноваціям.

Перед фахівцями в галузі освіти і навчання постав вибір: захищати сумнівний статус-кво або йти невизначеним шляхом, вибудуваним алгоритмами, який, найімовірніше, розширять можливості.

---

## НОВІ ДВИГУНИ НАВЧАННЯ

Заради справедливості слід зазначити, що ШІ вже використовується в освіті, зокрема старий добрий ШІ та машинне навчання. За допомогою ШІ учні отримують оцінки і рейтинги. На основі прогнозів ШІ університети розподіляють фінансування або виявляють абітурієнтів з групи ризику завдяки алгоритмам. Студенти використовують платформи на основі ШІ для підготовки до іспитів і “списування” домашніх завдань, а викладачі – для виявлення порушень та боротьби за академічну чесність. Окрім корпоративного програмного забезпечення (наприклад, студентських інформаційних систем і платформ електронного навчання, які використовують школи та навчальні центри), зростає також ринок застосунків для освітніх технологій (EdTech), орієнтованих безпосередньо на споживача, які використовують ШІ.

Звільнившись від столітніх догм та інституційної інерції, які обмежували діяльність формальних академічних організацій, EdTech-компанії вже скористалися тріщинами, що з'явилися в династичних структурах освіти. Тріщини тільки розширюються, і через ці канали EdTech-компанії починають наводнювати ландшафт генеративним ШІ, запускаючи парадигмальну зміну з благословення традиційних інституцій або без нього.

Ця нова ера навчання за допомогою генеративного ШІ, з ним і через нього потребуватиме нових моделей педагогіки і саморозвитку, нових інструментів для оцінювання, нових компетенцій для розвитку і нових методологій синтезу і співпраці.

Давайте розглянемо деякі з цих нових рушіїв навчання та хвилі змін, які вони, ймовірно, спровокують.

[1]

Прискорюючи зростаючу тенденцію, освіта та навчання, ймовірно, відійдуть від лінійних рамок, у яких навчальний процес організовано за окремими предметами і послідовними блоками. На зміну їм прийдуть веб-подібні моделі – більше схожі на екосистему, в якій різні фрагменти навчання генеруються, компілюються і надаються відповідно до потреб окремих людей, а не на довгий шлях, яким ми всі йдемо.

Основою цього підходу є персоналізація, яку забезпечує ШІ.

Багато “розумних” освітніх платформ вже пропонують, хоч і розрізнено, різноманітні варіанти персоналізації. Наприклад, додаток для вивчення мови може використовувати ШІ для оцінювання успіхів учня і подальшої пропозиції йому індивідуального зворотного зв'язку, підказок та послідовності контенту. Донедавна таке адаптивне навчання було рідкісною пропозицією, унікальною ціннісною пропозицією, завдяки якій такі застосунки, як Babbel і Duolingo, займали величезні частки ринку. Сьогодні це необхідність.

Незабаром, будь-яка платформа включатиме персоналізацію – не просто фільтрацію контенту (яка рекомендує “наступні” навчальні діяльності) або прості дерева рішень (які дають заздалегідь підготовлений зворотний зв'язок на основі конкретних помилок), а оптову генерацію нових медіа та шляхів навчання,

основою яких є індивідуальні потреби.

Персоналізоване (“адаптивне”) навчання - це добре розвинена наука, що виникла приблизно у 1960 році в галузі знань, призначених для інтелектуального навчання. Історично склалося так, що створення застосунків для персоналізованого навчання займало багато часу і мало вузьку сферу застосування. Проте добре розроблені системи продемонстрували потужні результати, до того ж поліпшення результатів навчання в середньому має середній і великий ефект.<sup>31</sup> Найліпші системи демонструють величезні переваги. Наприклад, один відомий інтелектуальний викладач навчав новачків, щоб вони стали ІТ-спеціалістами. Він показав результат на 3–4 $\sigma$  (що можна уявити як 3–4 буквені оцінки) ліпший, ніж за традиційної професійно-технічної освіти, а учні, які використовували послуги репетитора, перевершили досвідчених професіоналів всього за 16 тижнів навчання.<sup>32</sup> Цього було досягнуто за багато років до нинішньої хвилі інновацій у сфері ШІ, а це означає, що гіперконтекстуалізація та персоналізація, яких можна досягти за допомогою генеративного ШІ, ймовірно, дадуть ще більш вражаючі результати.

Генеративний ШІ дає змогу максимально персоналізувати навчання: мої запитання та навчальна програма створені для мене, а ваші – для вас. Наприклад, на початкових рівнях вивчення нової мови можна повторювати корисні слова чи фрази, частіше повторювати ті, з якими виникають труднощі, і отримувати персоналізований зворотний зв'язок щодо помилок у вимові. На середньому та вищому рівнях

учень може зануритися в книги, фільми та розмови, згенеровані штучним інтелектом на вимогу, відповідно до його рівня, інтересів та цілей. А коли настає черга тестування та отримання сертифікатів, цифрове портфоліо учня – а не один підсумковий кваліфікаційний тест – може стати доказом його компетентності завдяки аналізу його роботи, проведеному за допомогою ШІ.

[ 2 ]

З правильними підказками генеративний ШІ може створювати тисячі варіантів навчальних завдань. Він спрощує адміністративний тягар проектування і розроблення майже всіх типів контенту: планів уроків, статей, вправ, тестів, рубрик, рамок компетентностей, віртуальних середовищ і навіть цілих цифрових курсів.

Наслідки цього важко повністю передбачити. На суто механічному рівні витрати на виробництво значно зменшаться (хтось знає про безкоштовні підручники?). З точки зору викладача, персоналізоване навчання більше не потребує значних витрат. Навіщо давати всім учням однакові завдання, якщо ми можемо створити унікальні тести і домашні завдання для кожного?

Компанії, які заробляють на життя ліцензуванням контенту або створенням іспитів, терміново намагаються адаптувати свої робочі процеси до впровадження ШІ. Це, мабуть, страшний час для традиційних видавців, яким доводиться продавати ресурс, що швидко знецінюється, борючись з інерцією своїх попередніх бізнес-моделей.

Широкий доступ до економічно ефективних

імітаційних моделювань і віртуальних середовищ, керованих ШІ, зробить революцію у збагаченні досвіду та практичному навчанні. Наприклад, студенти-медики можуть відпрацьовувати операції у віртуальній операційній, випробовуючи нескінченну кількість сценаріїв, створених з урахуванням конкретних навчальних потреб або щоб допомогти їм підготуватися до майбутніх процедур. Студенти-інженери можуть тестувати складні конструкції у захоплюючих цифрових середовищах, безпечно помиляючись і повторюючи свої проекти за допомогою компаньйона з генеративним штучним інтелектом. Віртуальні середовища на основі генеративного ШІ дозволять нам маніпулювати ідеями, випробовувати нові способи бачення і пізнання, а також швидко створювати прототипи як у навчальних закладах, так і за їх межами. Такі можливості не тільки прискорюють темпи винахідництва; незабаром їхнє використання стане обов'язковим для працівників, які володіють знаннями в певних галузях, якщо вони хочуть бути конкурентоздатними на професійному рівні.

[ 3 ]

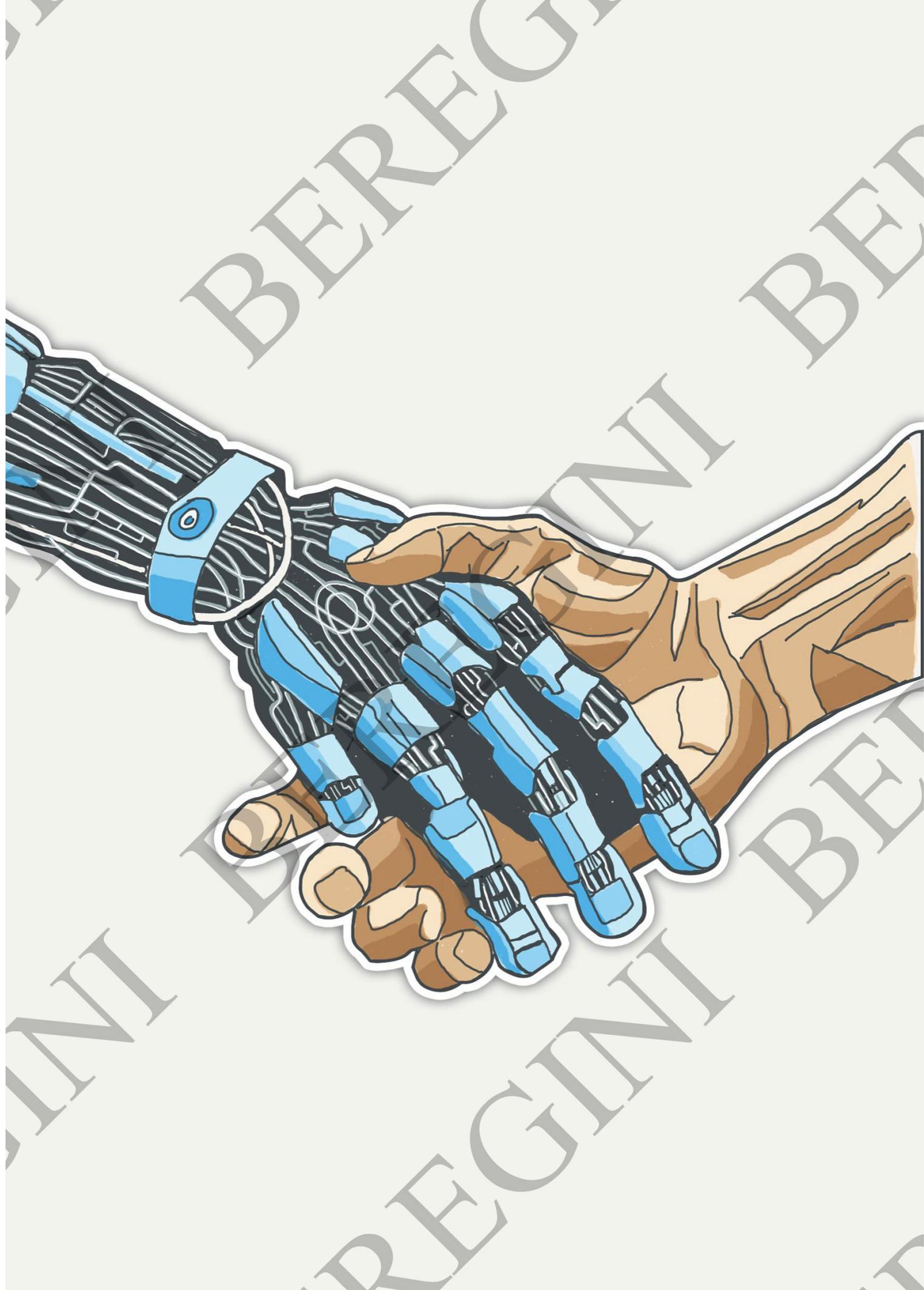
Оскільки темпи прогресу продовжують прискорюватися, це змінить (і фактично вже змінює) ринок праці. ШІ та автоматизація в поєднанні зі змінами у суспільстві, як-от диверсифікацією кар'єри, збільшенням тривалості життя і підвищенням складності виробництва, змінюють вимоги до працевлаштування, і не тільки сьогодні чи завтра, а постійно, що потребує безперервного розвитку для працівників у багатьох дисциплінах, щоб залишатися конкурентоздатними.

Щоб іти в ногу з часом, нам потрібно буде займатися навчанням упродовж життя – розвивати нові знання і навички з дитинства до виходу на пенсію (що б не означало слово “пенсія”). Або, як назвали цю концепцію в Гарварді, нам знадобиться “навчальна програма на 60 років” – платформа для навчання впродовж усієї нашої життєдіяльності.<sup>33</sup>

Валюта освіти, ймовірно, зміститься від традиційних ступенів до більш динамічного підходу, коли люди обиратимуть з “шведського столу” варіантів навчання і використовуватимуть такі інструменти, як-от мікрокредитування і “учнівські гаманці”, засновані на компетенціях, для підтвердження своїх можливостей. Ландшафт навчання може фрагментуватися в міру того, як індивідуальні шляхи та розбіжні наративи замінюватимуть моделі масової освіти.

Ми можемо уявити собі відданих III-наставників, які допомагають нам орієнтуватися на цій новій місцевості. Ці цифрові наставники, можливо, подаровані нам у дитинстві, поступово “навчатимуться” на власному досвіді (з часом підвищуючи ефективність), так само, як і ми вчимося у них, зростатемо симбіотично. Вони можуть стати нашими ідеальними радниками – завжди доступними, з бездоганною пам’яттю про наші вподобання та цілі, ніколи не упередженими та обізнаними з такими методами, як сократівське опитування, рольові ігри та рефлексивний зворотний зв’язок.

Ці допоміжні засоби навчання не просто прискорять засвоєння знань за встановленими маршрутами; вони



також змінять характер навчання, зробивши його соціальним, навіть коли воно відбувається наодинці. Коли ми “ведемо діалог” з цими алгоритмами, ми використовуємо ті ж самі механізми пошуку та спільноти, що й під час спілкування між собою. Вони можуть допомогти нам розвинути метапізнання, критично оцінити своє мислення і навіть розвинути емпатію.

Крім “інших” ШІ, ми аналогічно можемо уявити собі ШІ-копії нас самих. Цифрові двійники можуть стати дзеркалом, через яке ми розглядаємо себе і досліджуємо далі вплив різних рішень на наші здібності, добробут і досягнення – імітації мільйона можливих життів.

У поєднанні з екосистемою персоналізованих навчальних ресурсів ці можливості, безсумнівно, чинитимуть глибокий вплив не лише на те, як ми вчимося, а й на соціальні та фізичні структури, в яких ми це робимо. Вже зараз наші звичні уявлення про “освіту” і “роботу” – як про обмежені формальні структури у фіксованих місцях і за фіксованим розкладом – руйнуються. Нові моделі доступу, розподілу та проникності призводять до дезагрегації знань, а також до ширших моделей змін.

У своєму дослідженні Соціальної епохи Джуліан висунув гіпотезу, що ми побачимо появу нових гільдій: по суті, нових соціальних структур, які формують кістяк кар’єри. Ці клаптикові спільноти володітимуть інформацією та досвідом, владою та репутацією, підриваючи ієрархічну модель, в якій організації “володіють” ефективністю, а університети “володіють” знаннями. Певною мірою ми вже спостерігаємо це зрушення зокрема у сфері

кібербезпеки, яке майже неможливо здійснити в межах жорстких організаційних кордонів або без допомоги взаємозв'язаних, здатних до практичної діяльності спільнот.

Концепція “Нових гільдій” відображає моделі змін, які на найширшому рівні вплинуть на фундаментальні структури суспільства, ті механізми влади, безпеки, прибутку та права, які формують конструкцію, в якій ми живемо.

Якщо ці прогнози здадуться надто оптимістичними, подумайте про те, що, як тільки люди повністю звільняться від рамок формальної освіти й отримають стимул для всебічного розвитку, це, безсумнівно, спричинить численні наслідки. Успадковані контексти навчання, традиційні джерела знань та одноосібні олігархи академічного авторитету балансують на межі руйнування.

[ 4 ]

Одна річ навряд чи зміниться – це наша потреба в оцінюванні, хоча *способи*, якими ми збираємо ці дані, а також *що* ми збираємо і *як* ми їх використовуємо, безсумнівно, еволюціонують.

“Оцінювання” – це набагато більше, ніж тести, оцінки та дипломи. У чистому розумінні “оцінювання” означає документування характеристик, контексту чи компетентності, зазвичай у якийсь кількісний спосіб, навіть якщо мета є якісною чи прихованою. Однак, з огляду на наш досвід у формальній освіті, багато хто з нас асоціює його з тестами, де в бланку відповідей

потрібно обвести кружечок з правильною відповіддю, та з іспитами, коли помилки виправляють червоним чорнилом, – жахливі мірила здібностей.

Проте ми доклали чимало зусиль. За минулі роки освітяни та психометристи значно вдосконалили цю сферу. Наприклад, стандартизовані тести (як от сертифікаційні іспити GRE і Cisco), а також деякі персоналізовані навчальні платформи та програми проведення оцінювання співробітників вже давно використовують і старий добрий ШІ, так і машинне навчання. І вже понад п'ятдесят років комп'ютеризоване адаптивне тестування, в якому складність кожного наступного питання коригується на основі результатів учасника тестування, використовує ШІ.<sup>34</sup> Незважаючи на це, донедавна використання ШІ в оцінюванні було відносно обмеженим, частково через наші дещо обмежені підходи до вимірювання та оцінювання навчання.

Традиційно ми здебільшого покладалися на експліцитні та сумативні форми оцінювання: іншими словами, тести, відокремлені від результатів роботи та перехресних тем, та тести, які проводяться наприкінці певного етапу (наприклад, наприкінці курсу або заздалегідь визначеного періоду переатестації), а також тести, які проводяться в обмежений момент часу. Одним із прикладів є підсумкове тестування.

Крім того, використовуються формативні тести, які іноді називають “тестами для навчання” (на відміну від підсумкових “тестів для перевірки вивченого”). Формативні тести – зазвичай невеликі поточні перевірки, призначені для інформування викладача чи комп'ютеризованої системи або для того, щоб

допомогти учням оцінити власний прогрес. Формативне оцінювання може бути у вигляді домашніх завдань, підсумкових вікторин або вправ у застосунках.

Генеративний ШІ добре підходить для підтримки обох цих традиційних форм тестування. Генеративні алгоритми можуть створювати незліченну кількість ітерацій схожих запитань, щоб обмежити можливість шахрайства і відкалібрувати оцінку для кожної людини. Генеративний ШІ також дає змогу економічно та ефективно розробляти складніші формати тестувань: сюжетні ситуаційні тести (можливо, навіть із зображеннями та відео, згенерованим ШІ), проблемні сценарії та симуляції у відкритому світі, які відтворюють реальні завдання. Крім того, після іспиту ШІ може допомогти з оцінюванням результатів та створити персоналізований зворотний зв'язок.

Окрім вдосконалення цих традиційних підходів до тестування, досягнення в галузі ШІ відкривають нові можливості для комплексних, безперервних і багатофакторних вимірювань. Упоєднанні з сенсорними технологіями (як-от камери, датчики наближення і портативні пристрої) і методами аналітики навчання, ми тепер можемо інструментувати оцінювання середовища для збору цілого комплексу даних неінвазивно і на місці.

Наприклад, оцінювання ділових переговорів може включати сценарій згенерований ШІ у віртуальному світі, в якому учень веде переговори із зацікавленими сторонами, кожен з яких є аватаром, керованим ШІ. Крім оцінювання ефективності роботи в певних точках, система може збирати поведінкові дані, зокрема рухи очей, вираз обличчя і тон голосу (у ході

усного діалогу). Він також може збирати приховані дані історії відвідувань людиною в комп'ютеризованому інтерфейсі, наприклад, куди вона натиснула (зверталася до своїх нотаток чи переключалася на несанкціоноване вікно?) і скільки часу вона провела, дивлячись на різні екрани. Дані відстеження такого рівня доступні вже сьогодні, хоча ранні алгоритми, які використовують їх для оцінювання ефективності, зазвичай, були надто грубими, щоб на них можна було покластися, і іноді несправедливо упереджено ставилися до певних груп.

Контекст оцінювання не завжди повинен нагадувати реальні ситуації. Використовуючи підхід, який отримав назву “приховане оцінювання”, алгоритми машинного навчання можуть робити висновки про чийсь здібності, наприклад, вміння розв'язувати проблеми, спостерігаючи за грою в ігрі, як-от *Plants vs. Zombies 2*, або спостерігаючи за їх виконанням інших нешкідливих завдань.<sup>35</sup> Дані можна збирати протягом тривалого часу, поступово збираючи докази, можливо, з різних джерел, а потім використовуючи ШІ, зробити висновки про чийсь компетентність на основі цих спостережень. (Незважаючи на свою мерзенну назву, “приховане оцінювання” не має на меті бути таємним; найімовірніше, ця назва стосується оцінювання, яке проводять у фоновому режимі, про що, з етичних міркувань, особа, яку оцінюють, повинна знати.)

Під час випробувань приховане оцінювання та багатофакторні прогностичні показники показують хороші результати.<sup>36</sup> Нейрофізіологічні дані, зокрема міміка та вокалізація, вже використовуються в деяких (суперечливих) застосунках для перевірки вакансій.<sup>37</sup> А дані історії відвідувань інтерфейсу допомагають інспекторам підтвердити особу учасника тестування та

запобігти шахрайству.<sup>38</sup>

Якщо це все доповнити генеративним ШІ, можливості значно розширюються. За допомогою ШІ ми можемо не лише збирати достовірні, надійні дані та виконувати прогнозний аналіз; незабаром ми зможемо проводити точне тестування на вимогу: створювати індивідуальні середовища для оцінювання для точного визначення відповідних знань, навичок чи поведінки кожної людини або команди.

[ 5 ]

Традиційна робота викладачів і співробітників факультету зміняться, оскільки генеративний ШІ почне замінювати тестувальників низького рівня, оскільки він скоротить робочі процеси розробки навчальних програм з місяців до днів, і оскільки він поступово замінить адміністративні завдання викладання. Як і у випадку з будь-якою інновацією, ми як виграємо, так і програємо. Частина наших усталених систем – аспекти влади, контролю і пошани – зміняться. Індивідуальні ролі у спільноті зміняться: деякі професії зникнуть, а нові – з'являться.

Ми побачимо появу “інженерів з навчання” – людей, які використовують ітеративні, керовані даними процеси для поєднання науки про навчання, технологій, аналітики та інших принципів, орієнтованих на людину, для оптимізації результатів навчання. Ми відчуємо дедалі більшу потребу у викладачах-експертах, кваліфікацію яких можна буде масштабувати на тисячі студентів або моделювати і відтворювати за допомогою

цифрових аватарів. (Як приклад платформа Khanmigo від Академії Хана). І ми побачимо більше фахівців у різних формах EdTech, аналітики навчання та психометрії (наука про вимірювання параметрів людини), оскільки ми розробляємо більш достовірні та комплексні способи оцінювання здібностей.

Ми також можемо сподіватися, що з'являться нові кадри етиків, які розуміють і людей, і алгоритми, та допоможуть нам рухатися вперед з обережністю і мудрістю. ШІ несе в собі ризики та величезні перспективи, і його дедалі більше поширення дає цей двосічний меч в руки мудреців і шарлатанів, дурнів і геніїв, від наших традиційних академічних інститутів до проривних технологічних стартапів, від урядових програм до наших робочих місць. Нам знадобляться сильні голоси, які допоможуть спрямувати цих численних учасників до прояву їхньої ліпшої сутності.

[ 6 ]

Багато організацій та амбітних фахівців, які прагнуть отримати комерційну перевагу, вже впроваджують ШІ для поліпшення результатів навчання, розвитку та підвищення результативності праці. Генеративний ШІ прискорить ці зміни. Наприклад, він, ймовірно, прискорить консолідацію освіти, навчання та продуктивності. Унаслідок цього може виникнути нова парадигма навчання – гнучкіша і динамічніша, менш визначена й окреслена, тож ми вчитимемося без відриву на робочому місці і виконуватимемо (змодельовані) завдання в (віртуальному) класі.

---

## ПЕРСОНАЛІЗОВАНІ ЗНАННЯ: РЕВОЛЮЦІЙНИЙ ТОВАР?

У своїй книзі *“Безкоштовно: майбутнє революційної ціни”* (*Free: The Future of a Radical Price*) Кріс Андерсон обговорює мінливу динаміку інформації в цифрову епоху.<sup>42</sup>

У книзі, опублікованій у 2009 році, через призму епохи, що передувала генеративному ШІ, досліджено, як змінилася традиційна модель оцінювання інформації завдяки інформаційній епосі, що настала з появою інтернету. Андерсон виходив з того, що з огляду на велику кількість інформації (нехай і загальної) доступної нам, відносна цінність загальної інформації впала практично до нуля.

Таке знецінення загальної інформації особливо очевидне для цифрових товарів, як-от онлайн-дискусії, електронні новини та програмне забезпечення, де копіювання та розповсюдження не потребує особливих зусиль.

Андерсон стверджував, що ці види **”інформації як продукту”** стали легкодоступними і відтворюваними, що призвело до усунення

їх дефіциту і, так, до зниження їх економічної цінності. Однак, він стверджував, що інформація, персоналізована під унікальні запити, потреби і контекст окремої людини чи організації – тобто **“інформація як послуга”** – все ще має величезну цінність і не є чимось, що можна легко знайти чи відтворити.

Перенесемося в сьогодні, коли генеративний ШІ зазіхає на цю колись безпечну гавань. З розвитком генеративного ШІ бізнес-модель “інформація як послуга” може втратити значну частину тієї цінності, яку вона колись мала.

Ця нова парадигма має стосуватися підприємств, які заробляють гроші на стратегічному застосуванні технічних знань, наприклад, консультантів з управління, копірайтерів і навіть адвокатів. Вони заробляють на тому, що переносять загальні знання (наприклад, з міжнародного фінансового права) на конкретні ситуації (наприклад, на один конкретний бізнес-контракт). Міць генеративного ШІ загрожує дефіциту і цінності цих видів діяльності.

Не дивно, що багато традиційних педагогів відчувають подібний тиск.

Очікується, що доповнений інтелект використовуватиметься дедалі більше. Доповнений інтелект – це інтеграція ШІ з людським інтелектом для поліпшення нашої когнітивної діяльності. У взаємодії люди і програми можуть здійснювати дивовижні подвиги, які перевершують можливості кожного з них окремо. Навіть звичайні люди, які працюють у парі з ШІ, показали, що вони можуть перевершити гросмейстерів (людей) або окремі алгоритми.<sup>39</sup>

Це дає підстави говорити про експертів-універсалістів – осіб, які володіють широким (але потенційно неглибоким) спектром знань і навичок. Хоча вони можуть володіти глибокими знаннями в певних галузях, для них є найхарактернішим є вміння долати кордони, вони володіють трансдисциплінарними навичками, критичним мисленням, творчим підходом до розв'язання проблем, цифровою та інформаційною грамотністю, а також вмінням налагоджувати міжгалузеву співпрацю. Вони відомі своєю універсальністю, адаптивністю та здатністю знаходити точки дотику в різних сферах.

Іншими словами, ми спостерігаємо послаблення акценту на “чистих” галузевих знаннях на користь більш універсальних здібностей – епоха поліматів. Частково це може бути пов'язано з тим, що “вузькі експерти” (ті, хто досягає вершин у спеціалізованій, але більш обмеженій дисципліні) часто є менш адаптивними, ніж фахівці широкого профілю, що є потенційно фатальним недоліком у світі, де зміни відбуваються дуже швидко. Неусвідомлена компетентність експертів, яку називають “парадоксом досвіду”, може засліплювати їх у мінливих умовах, навіть якщо вони стають надмірно самовпевненими. Вони також можуть втратити свою творчу іскру, потрапляючи в шаблони “функціональної фіксованості” та жорсткого автоматизму.

Філіп Е. Тетлок - видатний вчений, який досліджував цей парадокс, популяризував фразу: “Лисиця знає багато речей, а їжак – одну велику річ”. Ця фраза, запозичена з давньогрецького прислів'я, ілюструє два протилежні підходи до знань і досвіду.<sup>40</sup> Їжак зосереджений, але потенційно негнучкий, тоді як лисиця здатна адаптуватися і володіє еkleктичними знаннями. Звичайні експерти іноді можуть бути лисицями, але експерти-універсалісти – особливо коли їх доповнює штучний інтелект – майже напевно є хитрими.

Експерти-універсалісти, підсилені ШІ, стають дедалі більше привабливими в умовах зростання нестабільності та складності, а організації потребують більшої гнучкості. Експерти універсалісти забезпечують адаптивність і контекстуальне розуміння, тоді як алгоритми надають детальні, але вузькі експертні знання відповідно до даної ситуації. “Доповнені універсалісти” можуть стати на вершині організацій майбутнього.

І не лише окремі експерти можуть бути доповнені. Соратники ШІ можуть приєднуватися до команд людей, допомагаючи їм працювати точніше та ефективніше, а ШІ може допомогти з оптимізацією результатів роботи команд людей, використовуючи їх колективну продуктивність, яку можна миттєво ввести за допомогою інтерфейсу “мозок-комп'ютер”, для виявлення помилок або відшкодування за відволікання або втому членів команди.<sup>41</sup> Це може звучати як наукова фантастика, але ці “рушії доповнення” вже тут, і досягнення в області генеративного ШІ лише прискорять їх поширення.



---

## ДВИГУНИ НЕДОБРОЧЕСНОСТІ?

Процес навчання не буває швидким чи легким. Він не “створений” для того, щоб бути акуратним. Це не означає, що навчання має бути обмеженим або деморалізуючим, але без певних труднощів навчання не може проявитися. М’язам потрібне тренування, щоб стати сильнішими, і так само наш розум потребує стимуляції та провокації, щоб вчитися.

Звісно, потенціал генеративного ШІ щодо скорочення процесів навчання викликає занепокоєння. Учні завжди прагнуть знайти шляхи найменшого опору, наприклад, віддати алгоритмам написання есе, складання тестів та узагальнення прочитаного на аутсорсинг.

Це обман?

Межа між шахрайством і навчанням дуже тонка або принаймні залежить від контексту. Учні, які використовують інструменти ШІ для створення структурних схем, дослідження зв’язків або навігації по ідеях, можуть застосовувати вражаючу тактику навчання. Спілкуючись з ШІ-асистентом, учні можуть розширити своє розуміння, розвинути навички пошуку інформації та метапізнання (навички мислення вищого порядку, пов’язані з поглибленим усвідомленням власних прогалин, можливостей і потреб).

Учні також можуть зловживати ШІ, опираючись на нього як на милицю, тоді як їхні власні навички критичного мислення атрофуються і їхнє довгострокове розуміння сповільнюється. ШІ може стисло викласти текст, але це не замінить читання книги. Не менш важливо й те,

що генеративні алгоритми не є “рушіями відповідей”. Вони не є премудрими, всевидючими і всезнаючими. І хоч генеративний ШІ може “знайти сенс”, це не “мій” чи “ваш” сенс. Алгоритми не можуть сконструювати сенс, який є в наших головах, хоча іноді вони можуть допомогти нам знайти прозріння, щоб зробити це самим.

Існує також ризик, що, оскільки учні дедалі більше залежать від генеративного ШІ, їхня робота може стати більш знеособленою. Історії можуть втратити індивідуальну автентичність, витвори мистецтва можуть втратити особистісний відтінок, і навіть музичні композиції можуть бути холоднокровно сфабриковані за допомогою банальних алгоритмів. Зловживаючи “милицями” штучного інтелекту у творчій сфері, учні можуть втратити можливість розвивати власні голоси та уяву.

Враховуючи вищесказане, заборона використання ШІ в навчанні має такий самий сенс, як заборона калькуляторів на уроках алгебри або інтернету на уроках історії. Вивчення цих дисциплін передбачає регулярне використання таких цифрових інструментів, і використання ШІ вже є (або незабаром стане) таким же.

У відповідь може виникнути запитання: “Як ми можемо перевірити знання студентів або перевірити участь тренерів, якщо вони можуть легко використовувати генеративний ШІ для виконання домашнього завдання або імітувати свою присутність на вебінарі?”. Але це неправильне запитання. Ліпше поставити запитання, яке досліджує наші поточні структури та обмеження: “Чи правильно ми контролюємо?”

Якщо ми розробили настільки крихкі структури навчання, що для того, щоб учні могли брати участь у ньому, необхідно обмежувати їхні можливості, то, це радше недолік системи, ніж провина учнів. Тож, можливо, ліпше витратити наш час на пошук шляхів розвитку цієї системи – способів, за допомогою яких генеративний ШІ може розширити навчальний досвід і підняти учнів на нові висоти – замість того, щоб розпалювати конкуренцію в класі, намагаючись заборонити використання ШІ в контексті навчання.

Зрештою, з іншого боку, “шахрайство” можна назвати “інновацією”, а наша робота як викладачів і тренерів полягає в тому, щоб підготувати учнів до майбутнього – майбутнього, яке, безумовно, буде пов’язане з генеративним ШІ.

---

## ДОПОВНЕНІ СПІЛЬНОТИ ТА (АНТИ)СОЦІАЛЬНЕ НАВЧАННЯ

Навчання буває різних форм і розмірів. Деякі види навчання є повністю формальними, відбуваються в рамках добре структурованих класів або жорстких навчальних програм; інші виростають з наших традиційних знань і практичного досвіду. Соціальне навчання використовує цей вид неявних знань. Воно недосконале і неохайне – принаймні з погляду структурованих і перевірених понять, якими переважно користуються формальні організації –але воно, як правило, дуже прагматичне і здебільшого надзвичайно цінне.

Одним із головних механізмів обміну традиційними знаннями є діалог: розповідь історій, розв'язання проблем і виклик один одному, проявляючи зацікавленість і ставлячи запитання. Організації можуть сприяти такій взаємодії, надаючи опорну підтримку, за якої ми створюємо спільні простори для діалогу, а також більш формальні канали, які допомагають спільнотам знайти способи поділитися своїми знаннями та думками.

Генеративний ШІ може зробити опорне соціальне навчання доступнішим – у дієвий спосіб.

Поговоримо про наратив. Соціальні колективи формуються і спілкуються за допомогою історій. Інструменти генеративного ШІ (наприклад, Copilot in Teams або Zoom AI Companion) можуть синтезувати спільно створені наративи групи, або ж вони можуть забезпечити учасників діалоговою структурою, у межах якої члени групи будуть працювати або на яку опиратися.

Інструменти генеративного ШІ можуть допомогти порівняти і зіставити наративи між колективами та паралельними групами, щоб виявити розбіжності. Ці розбіжності позначають точки інтересу на ландшафті колективних знань, оскільки розуміння на місцевому рівні обов'язково відхиляється від “авторитетного наративу”, тому можливість знайти і порівняти ці відмінності часто стає одкровенням.

Звичайно, історії, які створюють згуртованість у групі, можуть також “зводити стіни”, що відокремлюють ін-групу від аут-групи, що може створити проблеми для нових членів, які захочуть приєднатися до вже сформованого колективу. У цьому разі генеративний ШІ

може створити необхідні опори, даючи нам можливість орієнтуватися в історіях, які розказали інші, за допомогою підказок та вказівників, підштовхуючи до цікавих досліджень і виводячи на поверхню нові теми з усієї спільноти як “путівники по колективу” або оповідачі про мандрівки.

Генеративні алгоритми можуть бути сполучними елементами, модераторами, дороговказами, а також сторожами або мостами, які допомагають отримати доступ до колективних знань.

Потенційно ШІ може стати цілими спільнотами.

Історично діалог і співтворчість були сферою діяльності індивідів, які працюють у соціальних структурах. Формальні організації та менш формальні колективи відповідно володіють як кодифікованими, так і соціально створеними динамічними формами знань. Традиційно спільноти покладалися на репутацію, щоб виокремити сигнали серед шуму, до того ж кожен колектив вимагав певних внесків з точки зору репутації, часу, зусиль, пристойності та культури. Тепер можна уявити, як генеративний ШІ грає кожну роль у спільноті – по суті, “безкоштовно” – з механічної точки зору, створюючи наш власний епічний наратив, персоналізований і наповнений незліченною кількістю неігрових персонажів, з якими ми можемо взаємодіяти, дозволяючи нам вести діалог із самими собою: своєрідне *антисоціальне* навчання.

Готовий до співпраці згідно з дизайном, але не з іншими.

Спільнота одного.

Колективний на вигляд, але самотній на практиці.

Наслідки цього можуть бути цікавими. Поширення ботів у нашому соціальному навчальному просторі може позбавити нас важливих аспектів конкуренції чи допомоги, а також масштабу. Вони можуть втягнути нас у моделі антисоціальної поведінки або співпраці без особливих зусиль, або ж слугувати зразками кращої взаємодії, потенційно ілюструючи механізми допитливості, зв'язків та співпраці. Мабуть, вони урізноманітнять наші погляди, оскільки ми досліджуємо необмежені можливості без наслідків і репутаційних ризиків. А можливо, вторгнення “двигунів діалогу” в наші соціальні спільноти зруйнує довіру і підірве саму цінність, яку вони прагнуть зміцнити.

Двигуни вже запущені, тож незабаром ми дізнаємося, чи є вони рушіями колективного зростання, які зміцнюють спільноти та розкривають розподілений досвід, чи рушіями розпаду, які загрожують крихкій автентичності нашої соціальної структури.

---

## СМИРЕННІ ІНАКОДУМЦІ

Деякі фахівці з навчання та девелопменту можуть обурюватися цією новою реальністю, але правда полягає в тому, що руйнація старих моделей – це відголоски інформаційної епохи. На сьогодні іржа вже встигла з'явитися, і вчорашні застарілі парадигми не витримують потрясінь генеративного ШІ.

Проте багато хто з нас вагається.

Наш власний досвід, сформований парадигмами

минулого, привчив нас до застарілих форм освіти та навчання. Наші системи координат народилися з цих структур. Безсумнівно, старі традиції, догми та відпрацьовані методи мають свої переваги. Вони створили інтелектуальну платформу, на якій ми стоїмо сьогодні, і вони сприяють ясності та послідовності в роботі наших інституцій. Але ця комфортна звичка може також засліпити нас на шляху до прогресу. Доходячи до крайнощів, ми можемо стати накопичувачами застарілих істин, задихаючись у минулій епосі.

Сучасні системи освіти все ще мають різочу схожість зі старими пруськими школами: конвеєри, які надмірно акцентують увагу на передаванні інформації та є заручниками оцінок, що стимулюють монокультуру та обмеження. Навчання на робочому місці не набагато ліпше: воно надає перевагу чітким структурам, а не складним реаліям, і часто зосереджується на поступовому розвитку навичок, а не на підготовці працівників до справжніх викликів, з якими їм доведеться зіткнутися.

Хоча і не варто відкидати мудрість минулого, настав час смиренно визнати, що старі моделі, які свого часу добре нам послужили, не працюють. Те, що колись було патиною віку і просвітництва, поступилося місцем корозії. Настав час переглянути наші погляди на навчання:

- Чи все ще потрібна нам традиційна модель “школи”, в якій визначальними є час, місце та навчання з використанням друкованих носіїв? Які паралельні та альтернативні рівні навчання допоможуть нам ліпше адаптуватися до динамічних викликів

нашої епохи? Якби у вас була чиста дошка і чарівна паличка, які знання та навички ви б визначили пріоритетними для майбутніх поколінь, особливо знаючи, що вони будуть доповнені штучним інтелектом?

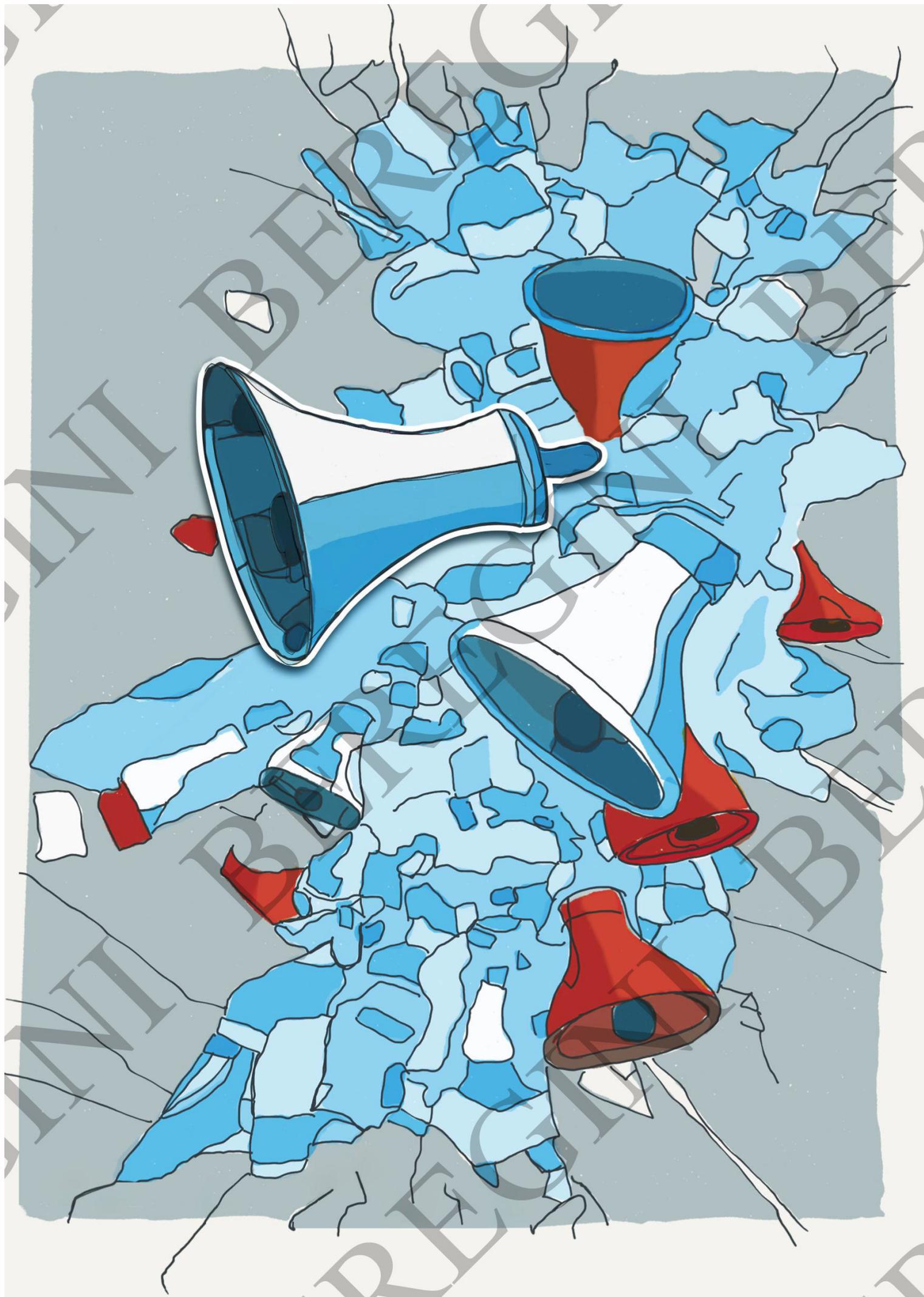
- Яку роль відіграють формальні центри знань та відомства (як-от університети, роботодавці, ліцензійні агенції та урядові програми) у майбутньому навчання? Чи повинні вони рішуче захищати свою спадщину від руйнування, чи прагнути гармонійного співіснування?
- Якщо ми вагатимемося, чекаючи і спостерігаючи за тим, як розгортаються події, то хто буде керувати змінами? І чи оцінимо їхні рішення щодо переможців і переможених – тих, хто отримує прибуток і тих, кого експлуатують, – і щодо людських якостей, які підносяться над тими, якими ми вирішили поступитися машинам?

Звичайно, прийняття цієї неоднозначної нової реальності пов'язане з певними ризиками, але якщо ми стоятимемо осторонь і чекатимемо, поки з'явиться визначеність, перш ніж діяти, то ми неодмінно залишимося позаду. А тим часом першопрохідці – стартапи, що швидко зазнають невдач, та ощадливі корпорації – матимуть повну свободу дій у прийнятті рішень для цього сектору.

Навчання є основою нашої людської сутності. Наша здатність до навчання та навчальний досвід формують траєкторії нашого життя. Наші суспільства побудовані на інститутах навчання, структурах оцінювання та стимулах для навчальних досягнень.

Тому трансформація навчання за допомогою штучного інтелекту – це не просто цікавість для освітян чи підрив ринку для виробників підручників. Це глибокий переворот у житті людини, що впливає на безліч способів роботи, процвітання, зростання та відкриттів. Взаємодія є обов'язковою. Це розмова, в якій ми всі зацікавлені; настав час приєднатися до діалогу про ШІ та навчання, перш ніж майбутня обстановка буде визначена замість нас.





## ЧАСТИНА 6

### Незвичайні жнива

Стрибаючи і оступаючись, керовані оптимізмом чи ідеологією, цікавістю чи тривогою, ми все ж таки опинилися тут.

Неминуче.

Прискорючись по слизькій стежині.

Звучить один гудок. Потім другий.

Повітря швидко наповнюється шумом: двигуни творчості, цілей або прибутку гудуть, улюлюкають, цвірінькають і скрекочуть один на одного, тоді коли всі навколо розмахують руками. Коментатори висловлюють свої думки. Панікери кричать. А стюарди свистять у свистки, парадоксально намагаючись сповільнити і втихомирити гамір, додаючи шуму какофонії.

Ми все глибше занурюємося у зростаючий дисонанс. І з кожним новим сигналом ми колективно підстрибуємо в тривозі (або радості, або жадібності). І майже не має значення, чи це виття сирени, чи муркотіння гармонії. Ми перевантажені, і кожна дія додає ще більше безладу, не залишаючи нам простору для дихання і роздумів,

диференціації чи класифікації. Як ми можемо рухатися вперед, коли кожен новий рушій, здається, сигналізує про тривогу, яка уражає наші системи? Складність здається нездоланною. І в цьому шумі раціональні дискусії тихо завмирають.

---

## ТРИВОГА, СКЛАДНОЦІ ТА ДИСКОМФОРТ

Нерідко ми відчуваємо почуття тривоги у відповідь на появу генеративного ШІ: страх, що він випередить людину або зробить наші роботи застарілими, невпевненість у тому, що суспільні зміни постійно прискорюються, а також дедалі сильніше відчуття моральної паніки.

**Моральна паніка** – це відчуття, що щось, переважно “зле” за своєю суттю, загрожує нашому способу життя. Це – екзистенційна загроза. Вона має відтінок ірраціональності, хоча очевидно, що деякі загрози, створені генеративним ШІ, не є надуманими: існує упередженість. Несправедливість присутня. Скоюються злочини.

Вже зараз, дозволяючи собі багато, люди створюють образи насильства, позбавляючи інших заощаджень або переконуючи повірити сфабрикованим новинам. Отже, ми не повинні дивитися на це з широко розплющеними очима і не повинні поспішати бити в черговий тривожний дзвіночок.

У всіх нас поганий послужний список у прояві

моральної паніки: танцювальна манія середньовічної епохи, “червоний страх” 1940-х років, перебільшений страх перед сатанізмом у 1980-х роках. Відомий всім Сократ, якого часто цитують у сьогodнішніх дискусіях про ШІ, навіть висловлював певні нотки паніки щодо винайдення писемності. Передбачали, що інші винаходи – від телеграфу до соціальних мереж – призведуть до занепаду суспільства. І все ж ми продовжуємо жити.

Сократ, за словами Платона, попереджав про моральний злам, до якого призведе письмо: “Бо це твоє відкриття [дар письма] створить забудькуватість у душах учнів, бо вони не будуть користуватися своєю пам’яттю, довірятимуть зовнішнім письмовим знакам і не пам’ятатимуть себе самих. Те, що ти відкрив, допомагає не пам’яті, а пригадуванню, і ти даєш своїм учням не істину, а лише видимість істини; вони будуть слухачами багатьох речей, але нічого не засвоять; вони будуть здаватися всезнаючими, а насправді нічого не знатимуть; вони будуть нудним товариством, створюючи видимість мудрості без реальності”.<sup>43</sup>

Страх поширюється, як чорнило у воді. Поки у нас не буде переваги ретроспективного підходу протягом багатьох десятиліть, важко сказати, які з побоювань є виправданими. Але ми можемо передбачити, що неспецифічне роздратування і поспішні реакції навряд чи дадуть нам користь. Зміни неминучі, але те, як ми їх сприймаємо, як шукаємо баланс між швидкою адаптацією та жорсткою прихильністю до минулого,

залежить від кожного з нас. Чи повинні ми знизити наші стандарти, щоб максимізувати експерименти і можливість отримання блага? Чи миритися з певною шкодою в ім'я прогресу? Кого ми захищаємо і яких меж ми набуваємо для цієї безпеки?

У грі завжди присутня напруга. Наші організації за своєю природою обмежені та повільно змінюються, і часто це правильно. Ми не хочемо нищити наші системи з кожною новою примхою. Але за надмірну повільність доводиться платити, і коли йдеться про зміну парадигми, цією ціною може стати вимирання.

Що ж робити? Першим кроком на шляху сталого прогресу є готовність зрозуміти себе, долучитися до індивідуального та суспільного метапізнання, щоб чітко розуміти свої мотиви. Чи робимо ми мудрі паузи, щоб осмислити обстановку? Чи мчимо вперед, з зав'язаними очима оптимізму або в окулярах, затуманених "страхом проґавити"? Чи ми заперечуємо все, паралізовані неминучим тягарем складності та відчуттям тривоги, що походить від зламу наших звичних концепцій?

Варто пам'ятати, що ці сприйняття можуть мати емоційне, а не суто структурне походження. Або інакше кажучи: ми не повинні дозволяти "диванним експертам" зупиняти наш прогрес через те, що ми уявляємо, що вони у цьому орієнтуються.

Нічого не робити – поганий варіант, а намагатися все контролювати – нереально. Але якщо ми зможемо подолати цю напругу і прорватися крізь інерцію вагання, то наступним кроком буде пошук знань – не обов'язково експертних. Тому що, хоча хвилювання або

побоювання можуть заважати нам приймати рішення, найімовірніша перешкода, з якою ми зіткнемося, – це просте незнання.

Ніхто, навіть найгучніші голоси, не володіє даром передбачення. Ніхто не є всебічним експертом з питань ШІ та його численних технічних деталей, організаційних процесів і побічних ефектів. Ставтеся з розумним сумнівом до будь-кого, хто претендує на всезнання в цій сфері, яка так стрімко розвивається.

Натомість сміливо рухайтесь вперед із цікавістю. Створюйте мережі, яким ви довіряєте, і не соромтеся ділитися своїм сприйняттям і запитаннями, дослідженнями і невдачами. Кожен з нас має більше можливостей, ніж ми можемо собі уявити, і кожен з нас може виявити ниточки, які можуть перетворитися на опори в цій новій реальності. Але ми навряд чи здобудемо ці знання, якщо стоятимемо осторонь, чекатимемо і спостерігатимемо збоку.

Іноді просто важливо перебувати в русі: у роботі, мисленні та уяві. Ліпше почати створювати для себе та організацій перелік інструментів вільного володіння штучним інтелектом – незалежно від того, наскільки незграбними чи нерішучими ми можемо почуватися – раніше, ніж пізніше. Навіть невеликі кроки, спрямовані на бажання допомогти. І якщо ми відчуваємо, що трохи не готові почати, корисно пам'ятати, що сьогодні ми обмежені настільки, наскільки обмеженим буде ШІ. Його вплив розширюється з кожним днем, тож ліпше включитися у цей процес зараз, поки це ще можливо.

У окремих людей і, зокрема, організацій може виникнути

думка віддати своє мислення щодо ШІ на аутсорсинг, чекаючи, що технологічні компанії прокладуть нам шлях уперед і скажуть, що думати про ці нові “рушії можливостей” та чого від них чекати. Звичайно, немає нічого поганого в тому, щоб покладатися на компетентних порадників, але, як і в більшості інших випадків, головне – знайти баланс. Уникайте надмірних поступок іншим, особливо коли на кону стоїть їхній прибуток.

Генеративний ШІ розвиватиметься і дедалі більшою буде його присутність (що фактично вже є) у наших найзвичніших застосунках. Є спокуса просто чекати і спостерігати, комфортно підготовуючи генеративний ШІ виправленнями та регулярними оновленнями програмного забезпечення. Але намагайтеся уникнути помилкового відчуття безпеки. Варто шукати контрольованого середовища для гри та експериментів. Деякі істини з’являються самі. І нам, найімовірніше, знадобиться і те, і інше: безпека в цифрах і масштабах, а також життєвий досвід і низові інновації. Отже, кожного разу, коли ви приймаєте безпечне рішення, подумайте про те, щоб збалансувати його з дещо ризикованим рішенням, яке допоможе вам відкрити для себе грані та межі цього нового світу можливостей.

Ще один важливий момент, який слід враховувати під час прийняття рішень щодо генеративного ШІ, особливо в ситуаціях з організаціями, – прагнемо ми оптимізації або інновацій. Чи влаштовує нас “зміна першого типу” (“Type-1 Change”), яка прискорює потік у нашій нинішній системі, або ж ми прагнемо “зміни другого типу” (“Type-2 Change”), яка повністю змінює маршрут?

Наприклад, чи потрібен нам ефективніший спосіб складання резюме, чи ліпше застосувати зовсім інший підхід до підбору персоналу? Чи отримаємо ми найбільшу користь від підбиття III підсумків наших веб-зустрічей, чи варто розглянути нові способи зустрічатися та обговорювати наші спільні зусилля?

Варто пам'ятати, що наші майбутні організації не будуть просто аналогами сьогоденних, але оснащеними III, – такими ж, але трохи швидшими або ліпшими. Натомість, найімовірніше, наші майбутні структури матимуть розвинені механізми впливу, управління та реалізації. А щоб відкрити цей шлях, необхідно створити прототипи, проводити експерименти та розвивати культуру досліджень.

Початкові хвилі генеративного III здебільшого очолювали гігантські корпорації, що дозволяло нам (як окремим учасникам) обмежено впливати на те, що відбувалося. У міру комерціалізації базової технології ми бачимо, що дрібні та різноманітні інноватори впроваджують її у творчі та нішеві рішення, багато з яких є спекулятивними або крихкими, і значна частина з них зазнає невдачі або, принаймні, виявиться не ліпшою за статус-кво. Незалежно від індивідуальних результатів, така інноваційність, найімовірніше, буде корисною, поки ми колективно прокладатимемо шлях у майбутнє – доки ми будемо готові відкрити себе для трохи цікавості і зможемо не запуснути уявні пазури паніки по нашій траєкторії.



---

## НЕКЕРОВАНЕ МАЙБУТНЄ

Чи забезпечать нам уряди захист?

Одним словом: малоімовірно.

Принаймні, не за рахунок дій, спрямованих винятково на ШІ.

Нормативно-правова база часто ліпше справляється з ретроспективним аналізом, ніж з прогнозуванням. А в глобальному контексті ми, імовірно, побачимо, що будь-які спроби регулювання ШІ – регламентовані чи ретроактивні – утворюють клаптикову огорожу, крізь яку прослизують вперті інновації.

Навіть серед тих країн і міжнародних інституцій, які готові регулювати ШІ, це все ще важко зробити. Існує кілька очевидних “червоних ліній”. Наприклад, Європейський Союз заборонив використання генеративного ШІ для підсвідомих або цілеспрямованих маніпуляцій. (Закон ЄС про ШІ також категорично забороняє різні програми машинного навчання, як-от дистанційна біометрична ідентифікація та системи соціального кредитування, які за основу класифікації людей беруть їхню соціальну поведінку, соціально-економічні або інші особистісні характеристики.<sup>44)</sup>

Однак, коли ми відходимо на крок від таких кричущих “червоних ліній”, межі “розумної” безпеки стають дедалі небезпечніші. Спершу може здатися, що розсудливі люди легко погодяться з певними ідеями. Наприклад, у певних ситуаціях ми повинні мати чітке розуміння, з ким ми маємо справу – з людиною чи машиною. Однак, про які саме ситуації йде мова – незрозуміло.

Чи повинні медики з елементами ШІ ідентифікувати себе як такі?

Можливо, і дослідників у галузі права теж?

Як щодо операторів онлайн-служби підтримки або баристи в додатку?

Де ми проводимо межу, і чи не здаватимуться наші сучасні міркування про “зловісну долину” між “справжніми” та “штучними” працівниками старомодними через кілька років? Чи не здасться майбутнім поколінням милим і химерним те, що ми зосередилися на таких відмінностях, можливо, так само, як ми ставимося до луддитів, які на початку ХІХ століття бунтували на автоматизованих фабриках?<sup>45</sup>

Більшість людей сьогодні – але, ймовірно, не всі – погодилися б з тим, щоб накласти певні обмеження на типи зображень, які може створювати ШІ (обмеження, вже впроваджені в найпопулярніших генераторах зображень).

Але, знову ж таки, де ми спільно проведемо жорстку межу?

Чи є допустимою вишукана художня нагота?

Політична сатира на авторитарних лідерів чи релігійні образи?

Як щодо використання облич знаменитостей чи навіть схожості з нами самими, особливо, якщо це використовується іншими? І що, якщо ці “інші” отримують прибуток від продукту, створеного ШІ? Ситуація знаходиться між складністю і дискомфортом, над якою нависла пелена заперечення.

На абстрактному технічному рівні генеративний ШІ не надто складний для розуміння. Те, що він робить, порівняно просте за своєю концепцією: він переробляє наявні ідеї для створення нових творінь. Ця, здавалося б, нескладна концепція є настільки потужною багато в чому завдяки своїй простоті та здатності узагальнювати. Подібно до того, як прості машини – важелі, гвинти тощо – утворюють фундаментальні будівельні блоки машинобудування, генеративний ШІ можна вважати новим і типовим цифровим інструментом, чимось на кшталт відкриття шківів. Що ми зможемо створити або зруйнувати за допомогою цього нового інструменту, поки що ніхто не знає.

І в цьому полягає складність.

Коли людству дається абсолютно нова здатність, з нашою безмежною винахідливістю можливості стають безмежними. Глибині нашої творчості – і творчої злочинності – немає рівних. Це поєднання людей і програм одночасно дає нам можливості і викликає найбільше занепокоєння.

Малоймовірно, що уряди зможуть захистити нас від інновацій, які несе з собою ШІ. У найліпшому разі державне регулювання спрямоване на боротьбу з низкою ар'єргардних дій, які (принаймні частково) спрямовані не проти того фронту. Тому що, хоча алгоритми можуть бути перетворені на двигуни дезінформації або злочинності, їхнє використання визначається нашою людською уявою та стимулами. А наші здібності до вияву дотепності та лиходійства – особливо тепер, коли ми доповнені ШІ, – незрівнянні.

---

## ІНТЕРМЕДІЯ ПРО ПТАШИНИЙ СПІВ

Ми з сином спостерігаємо за комбайном. Він напрочуд маневрений, прямує полем до нас, поки ми стоїмо біля огорожі. Гігантські вентилятори всередині машини виходять назовні через люк на даху, оточуючи транспортний засіб димкою пилу і гуркотом повітря, що рухається.

Він підїжджає прямо до того місця, де ми стоїмо, розвертається на краю поля, водій навіть знаходить час, щоб помахати нам рукою, коли пілотує новий курс. І ось, коли машина віддаляється по полю, нас наздоганяє вибух: вихлопні гази з подрібненої соломи, лушпиння та бруду наповнюють повітря – і наші носи та волосся.

Збираючи зерно, машина відкидає все інше вбік – а ми кашляємо і біжимо, щоб сховатися від полови, яка летить за нами. Механізована, ефективна, вона залишає нас розпатланими і захеканими.

Жнива – це момент судження, кульмінація аграрного року. Моя мама пам'ятає старі обжинки, вечірки в селі після того, як урожай успішно зібрали. Але не більше. Якщо раніше збір врожаю був ручною і виснажливою роботою, то сьогодні цей процес значно оптимізували і масштабували. Старі живоплоти, що розділяли клапті середньовічних полів, давно зникли, рови засипані. Деякі фермери навіть переорали неолітичні кургани, які колись

порушували прямі лінії їхніх геометричних посадок, вирішивши, що штраф, виписаний регуляторними органами, буде вартий збільшенню врожаю.

Індустріалізація сільського господарства привела до надлишку їжі в багатьох заможних суспільствах. Ми створили велику кількість щедрих земель, змінивши ландшафт навколо себе. Але наше процвітання не минуло дарма. Ще кілька десятиліть тому живопліт, біля якого ми з сином стояли, мовчав би: пестициди та надмірна культивуація знищили його комах і прогнали птахів, чиї пісні (і безпринципна боротьба за залишки врожаю) лунали тут з часів появи сільського господарства. Але зі зростанням ефективності наших врожаїв, ландшафти ставали дедалі більше монокультурними, а наше мовчазне небо наповнювалося пилом виснаженого ґрунту.

Потім відбувся певний перелом – почали виділяти субсидії на збереження живоплотів. Ми платили людям за те, щоб вони залишали неораний край поля для польових квітів і дозволяли бур'янам рости на своїх полях, або навіть залишали цілі сільськогосподарські угіддя на поталу природі, даючи їм час перезимувати і відновитися.

Ми експлуатували і страждали, потім адаптувалися і пристосовувалися, сподіваючись знайти нестійкий баланс між прогресом і крахом. Досі залишаються певні сумніви щодо того, чи вдалося нам це повною мірою, чи діяли достатньо активно. Іноді, подібно до пташиного співу, ми не помічаємо чогось, поки воно не зникне. 🌿



---

## ФРАКТАЛЬНИЙ ПІДХІД

Протягом своєї історії людство рухалося вперед нерівномірно: довгі періоди визначеності змінювалися бурхливими проривами змін, непередбачуваними інноваціями та соціальною еволюцією. І протягом усього цього часу наш прогрес переплітався з технологіями, від появи коліс і писемності до антибіотиків і персональних комп'ютерів. Ми прогресуємо. Технології розвиваються. Але вплив не є лінійним.

Іноді ми рухаємося вперед стрибкоподібно.

Іноді речі ламаються.

Іноді ми спотикаємося на темних стежках.

Концептуальний прорив генеративного ШІ дає нам абсолютно нову можливість: новий ресурс для використання. Подібно до друкарського верстата або двигуна внутрішнього згорання, ці нові двигуни мають широкі перспективи, з можливістю впливати на нас і індивідуально (на дуже локальному людському рівні), і глобально, майже в кожній системі та соціальній структурі, які ми досі винайшли.

Ці двигуни ШІ тільки почали збільшувати оберти. Ми все ще перебуваємо на ранніх стадіях стрімкого прогресу, і, як наслідок, наші погляди все ще зосереджені на дрібницях і нагальних проблемах. Вони постійно і швидко змінюються та є фрагментарними. Що ми побачимо, якщо відволічемося від пікселів і дозволимо шаблонам прогресу проявитися? Важко сказати,

виходячи з наших теперішніх перспектив.

Наша уява має межі, які значною мірою визначаються нашим досвідом, домінуючими культурними нормами та колективними наративами. Певні речі є звичними для нас, і навіть якщо ми прагнемо неупередженості, ми просто не можемо уявити, що вони зміняться.

Якими б проникливими ми себе не вважали, цілком ймовірно, що кожен з нас заперечує певні майбутні зміни. Кожен з нас хоче зберегти віру в те, що ми (власне, ми самі та наші ролі) є винятковими та неповторними.

Що завгодно, аби самоствердитися.

Але надія не робить це правдою. Тільки тому, що ми – або, точніше, речі, які ми робимо своїми руками та розумом – колись мали цінність, не існує космічного закону, який би стверджував, що ці речі залишатимуться цінними завтра також.

Безумовно, у цьому сміливому майбутньому на нас чекає “цінність”, але вона може знаходитися в різних частинах ландшафту. Диригенти, конструктори, інженери, перекладачі, оповідачі, з’єднувачі та підричники майбутнього будуть виконувати нові ролі та займатимуть нові місця.

Те, що раніше було справою досвіду та мистецтва, переходить у потік думок. І в міру цього переходу вони деформують і трансформують наші способи мислення, і – особливо в контексті нашої виняткової прив’язаності до соціальної епохи – дозволяють нам бачити світ крізь

калейдоскоп нових ідей, спотворених перспектив і творчих лінз.

У найліпшому випадку ці двигуни повісті та пісні, витвору мистецтва та артефактів не просто тягнуть нас за собою, підмінюючи наше мислення своїми результатами. Навпаки, ми можемо перестрибувати між нейронами та алгоритмами, між тим, що є машиною і тим, що є людиною – взаємодією органічної думки та цифрового артефакту.

Це необтяжена сила цікавості, вплетена в нашу практику: думка, перенесена в мистецтво; створення індивідуального сенсу випадково. Як ластівка на крилах, що кружляє і ковзає над травою – краса, що міститься в сукупності рухів, сукупності польоту, складному танці, заплутаному, але граційному.

В індустріальну епоху машини продуктивності та результативності, видобутку та перетворення були двигунами згоряння, що руйнували наше природне середовище, шкодили нашому здоров'ю та здійснювали контроль з вершини індустріальної піраміди, де багатство та влада стікалися до вершини.

Але ці нові двигуни є іншими, або мають можливість бути такими – часто працюють безшумно, є диверсифікованими (на противагу стандартизованим), доповнюють людей-операторів (на противагу небезпечно антагоністичним). Взаємовигідні. Те, що було важко, стає легким; складність абстрагується або приховується. З'являються нові речі.

Це і є наше завдання: знайти нові простори, нову мету.

Видається малоімовірним, що ми зможемо рухатися вперед без збоїв, але ми переживемо ці хвили змін, спричинені генеративним ШІ та його побратимами – двигунами взаємодії. Ми зможемо жити з цим або – наважимося припустити? – процвітати, незважаючи на потрясіння, які вони віщують.

Можливо, ми знайдемо спосіб зібрати  
врожай і при цьому почути спів птахів.





## ЗВОРОТНИЙ БІК СПРАВИ

### Ілюструючи Двигуни взаємодії

Для книги про генеративний ШІ, книги, яка досліджує вплив наших нових “двигунів взаємодії” на творчість і художню цінність, ми хотіли зробити щось особливе з ілюстраціями.

Джуліан ілюструє всі свої роботи, і для нього процес ілюстрації невід’ємний від процесу мислення. Він часто створює ілюстрації до того, як повністю сформулює свої ідеї – у такий спосіб, акт творення стає майже співпрацею між лінгвістичною та художньо-інтуїтивною частинами його мозку. Діалог між наукою і мистецтвом, словом і зображенням.

Але діалог (трилог!) цієї книжки інший, а отже, і її художнє оформлення.

Ми, троє авторів, не лише працювали разом над створенням ілюстрацій, а й використовували процес, який включає в себе генеративний ШІ. Зазвичай цей процес виглядав приблизно так: на наших щотижневих сесіях у “письменницькому кафе” ми визначали ключові концепції для ілюстрацій. Пізніше Сас та Джефф працювали з одним із “двигунів мистецтва”,

щоб створити прототипи дизайну на основі цих тем. Ми використовували програми Adobe Firefly та Midjourney. Порівняно з ілюстрацією вручну, цей процес був плідним, але все одно знадобився певний час, щоб досягти бажаного результату. Джефф пояснював так:

*Я використовував мікс Midjourney і Firefly, граючи з обома, щоб спробувати передати візуальну ідею, яка виникла у мене в голові. Я починав з одного інструменту, а потім використовував образи, які він створював, як насіння для іншого. Мій процес включав два паралельні завдання: по-перше, знайти стиль, який би мені подобався. Мені потрібно було знайти фрази для визначення стилю, який ми створювали, наприклад, “хімерний” або “намальований від руки”. По-друге, візуалізувати те, що я мав на увазі. Тут я обирав теми з наших письмових робіт, які мені спадали на думку, а потім вручну робив ескізи або уявляв, як би я хотів візуалізувати ці ідеї. Після цього я вступив у діалог з Двигуном, намагаючись пояснити словами суть мого бачення. Знадобилося чимало ітерацій, поки я не дійшов згоди з ним щодо того, як це має виглядати!*

*Я відчував підтримку, натхнення, а іноді й здивування.*

Створення ескізів ілюстрацій пройшло етап проб і помилок та потребувало творчого бачення – хоча й з допомогою гросмейстерського, суперхудожнього мозку. Генеративний ШІ дозволив нам блискавично створювати прототипи та експериментувати з різними стилями. Робота з цими інструментами була захопливою і натхненною, чудовим союзом між людиною і машиною.

І це був союз. Хоча бар'єри для використання цих “двигунів творчості” низькі, нам все ж потрібно було докласти зусиль для створення змістовних матеріалів, особливо якщо ми хотіли, щоб вони відповідали нашим очікуванням.

З часом ми почали використовувати наші власні завершені зображення (те, що ми раніше нарисували та відредагували) як “референси”. Іншими словами, ми могли завантажувати власні роботи, скажімо, у Firefly, і він генерував результати зі схожими візуальними елементами. Це дозволило нам створити портфоліо навколо певного “голосу”.

Кінцевим результатом цього етапу “генерування” стало багато непридатних для використання зображень, а також кілька десятків претендентів, які ми разом переглянули. Джуліан вибрав з них окремі зображення, щоб використати їх як натхнення для своїх малюнків, які він створив за допомогою цифрового стилуса та додатку Paper на своєму iPad. Принаймні в цій частині він дотримувався свого типового підходу до ілюстрування.

Після того, як Джуліан робив ескіз нового дизайну, він з деяким трепетом передавав ці зображення Сае та Джеффу на перевірку. Іноді ми трохи змінювали ескіз і навіть відмовлялися від кількох концепцій на цьому етапі. Але як тільки ми були задоволені ескізом (або, принаймні, переконані, що варто спробувати той чи інший дизайн), починалися наступні етапи цифрового оброблення.

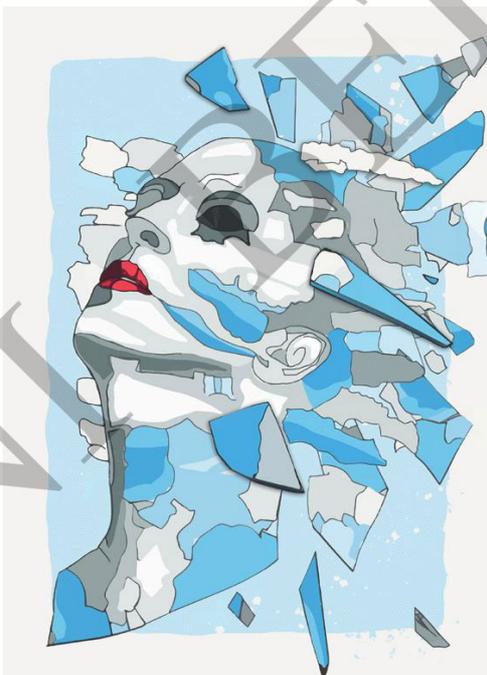
Сае – досвідчений графічний дизайнер, і вона використовувала роботи роботи Джуліана, які він створив за допомогою Adobe Photoshop та Illustrator. Вона додала



Генеративний ШІ



Цифровий живопис на папері



Після цифрової обробки

оригінальні елементи, такі як деталізація і текстура, а також трохи цифрової магії, використовуючи фільтри Photoshop або вбудовану функцію генеративної заливки, щоб розширити або підкоригувати частини ескізів Джуліана.

Ось приклад такої еволюції: Наша ілюстрація “Блакитна пані”, яку ви можете знайти в частині 4 цієї книги, почалася у програмі Firefly. Сае згенерувала початковий дизайн, використовуючи фрази з підрозділу “Поклоніння слову”, зокрема “мова утворює канал між нашим внутрішнім і зовнішнім світами” і “наша реальність побудована з мови”. Сае повторювала і змінювала налаштування, доки прототип Блакитної пані не було створено.

Потім Джуліан відтворив концепцію у вигляді набагато чистішої (цифрової) акварельної картини за допомогою застосунку

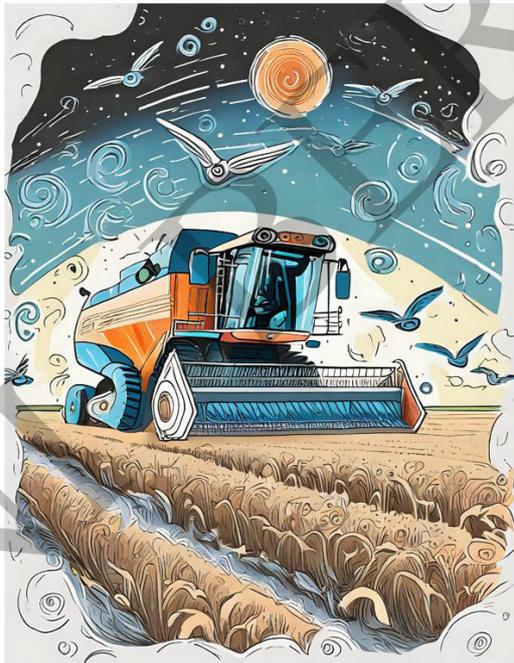
Paper. Зникли дрібні деталі та море слів, але на їхньому місці з'явився чіткий фокус на плавних синіх лініях. До речі, до цієї пригоди Джуліан завжди стверджував, що не вміє рисувати обличчя. А тепер, схоже, під впливом штучного інтелекту він відкрив у собі новий талант!

На третьому зображенні Сае завантажила паперовий рисунок Джуліана у Photoshop, де з допомогою генеративної заливки, розвернула зображення в портретну схему, створивши шию Блакитної пані. Це чудова демонстрація того, як ШІ швидко перетворився на товарний інструмент в арсеналі цифрового художника.

Крім того, Сае використала генеративний ШІ, щоб підправити губи Блакитної пані, зробивши їх трохи ближчими до анатомічно правильних. Вона також додала “осколки” навколо зображення, кілька бризок цифрового чорнила, текстуру та тіні. Разом ці деталі визначили наш стиль.

Ось інший приклад: розглянемо зображення комбайна з Частини 6 цієї книги. Він подолав схожий шлях. Крім того, для комбайна ми запозичили елементи безпосередньо з інших ілюстрацій, щоб створити остаточну версію.

Зверніть увагу на ці радіальні лінії на фінальному зображенні. Вони були майстерно перероблені з раннього ескізу, який Джуліан створив для іншої частини книги. Завдяки майстерності Сае в роботі з цифровими зображеннями, ми змогли ювелірно замінити темне небо на малюнку комбайна на більш яскравий та енергійний рисунок.



Генеративний ШІ

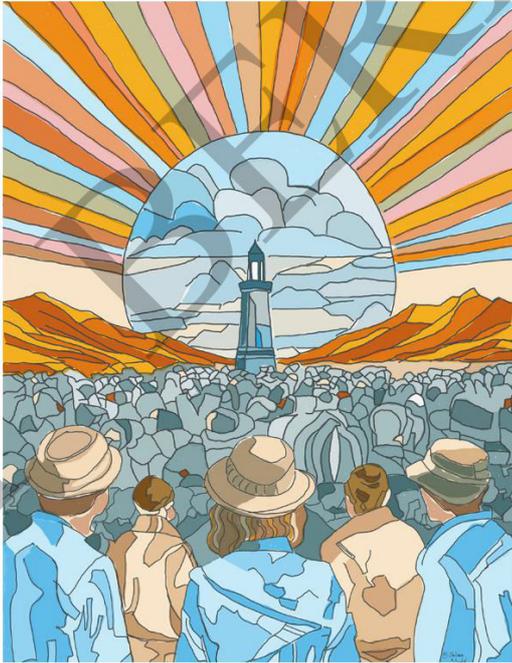


Цифровий живопис на папері

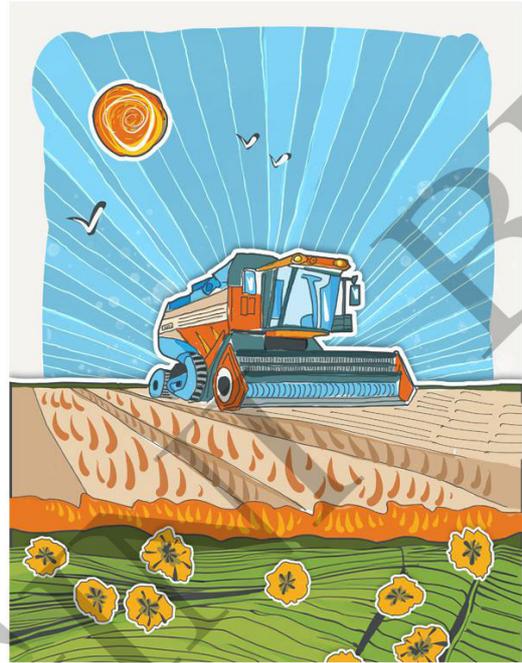
Поки ми працювали над зображеннями, докладаючи спільних зусиль та розробивши колективну візуальну граматику, наш стиль еволюціонував. Після кількох зображень ми зупинилися на палітрі та деяких характерних деталях дизайну, які Сає кумедно назвала нашою “роздробленою” естетикою.

Зрештою, двадцять зображень у цій книзі – це те, що ніхто з нас не зміг би створити самотужки, навіть за допомогою цих дивовижних “двигунів мистецтва”. За іронією долі, це твердження торкається теми “людської винятковості”, яку ми ретельно досліджували в частині 3, але, безумовно, жоден комп’ютер не зміг би їх створити – з огляду на такі застереження.

Наш творчий процес був досить повільним та неефективним. Книга вийшла б швидше, можливо, була б більш послідовною і, певно, навіть “ліпшою”, якби



Зображення з радіальними лініями



Після цифрового оброблення

ми виключили з рівняння нас, людей. Але ці людські елементи – недосконалість, які ми створюємо, і досвід, який ми, художники, отримали від цього процесу, – принаймні частково є сенсом мистецтва. Цінність з'явилася завдяки співпраці як зі штучним інтелектом, так і між собою. Справді, еволюція мистецтва, синтез стилів та артефактів, які ми створили, вплинули на наше письмо і мислення, і ці зміни вплинули на ландшафт, яким ми (як автори) і ви (як читачі) подорожували.

Отже, мистецтво – це подорож, а не просто пункт призначення, і ми вдячні, що ви вирушили в цю подорож разом з нами і, можливо, навіть трохи надихнулися нашими оригінальними – і часто досить химерними – ілюстраціями, створеними людиною за допомогою штучного інтелекту. ★



## Інші співавтори

Ці автори люб'язно поділилися своїми думками про генеративний ШІ в коротких нарисах у цій книзі.

Дональд Кларк (Donald Clark) – підприємець, інвестор, автор, спікер і блогер із 40-річним досвідом роботи в галузі навчальних технологій. Він допоміг заснувати та керувати трьома компаніями, що займаються штучним інтелектом та навчанням, а також брав участь у створенні кількох інших компаній, і він написав книгу про штучний інтелект у навчанні “*AI for Learning*”, опубліковану у 2021 році.

Серед інших книг Дональда – “*Проектування навчального досвіду*” (2022), “*Технології навчання*” (2023), “*Навчання в метаспросторі*” (2023) та “*Штучний інтелект для навчання: використання генеративного ШІ*”, публікація якої запланована на 2024 рік. Посилання на його блог: <https://donaldclarkplanb.blogspot.com>.

Марк Олерт (Mark Oehlert) – невпинно допитливий. За освітою антрополог та історик, Марк застосовує свої знання для досягнення інноваційних, трансформаційних результатів. Він має досвід проектування, розроблення і запуску навчальних та інноваційних програм у масштабах підприємств, які підвищують взаємодію та утримання персоналу, а також сприяють підвищенню ефективності організації. Наприклад, Марк керував навчальними та інноваційними програмами в

Міністерстві оборони США та Amazon, а також очолював програму соціального впливу в Unity 3D.

Марк процвітає у творчому, колаборативному середовищі та приймає неоднозначність, яка часто супроводжує це. Наразі Марк є науковим співробітником Інституту інновацій у великих організаціях. Слідкуйте за Марком на <https://markoehlert.substack.com>.

Марко Зао-Сандерс (Marc Zao-Sanders) – засновник і генеральний директор [filtered.com](https://filtered.com) – технологічної компанії, яка розробляє ШІ для розуміння навчального контенту з метою підвищення кваліфікації людей. Він захоплюється алгоритмами, навчанням, контентом, штучним інтелектом, мисленням по-іншому, тим, як працює мозок загалом, і, зокрема, як працює його власний мозок.

Марк отримав ступінь з математики та філософії в Оксфорді та заснував два соціальних підприємства, які допомагали молодим людям з непривілейованих верств населення в їхніх освітніх та кар'єрних прагненнях, виборі та досягненні успіху.

Він є постійним автором “*Гарвардського бізнесжурналу*” (*Harvard Business Review*) та інших видань, зазвичай розповідаючи про ШІ, навчання або *таймбоксинг*. Нещодавно він написав книгу “*Таймбоксинг: Сила робити одну справу за раз*”, опубліковану видавництвом Penguin Random House у січні 2024 року.

Марк живе на півночі Лондона з дружиною, трьома дітьми, двома котами та двома собаками.

## Примітки

- 1 Hoffmann, J., Borgeaud, S., Mensch, A., Bouchatskaya, E., Cai, T., Rutherford, E., et al. (2022). Навчання обчислювально оптимальних моделей великих мов, *arXiv: 2203.15556v1*. <https://doi.org/10.48550/arXiv.2203.15556>
- 2 Schreiner, M. (2023, 11 червня). Витік інформації про архітектуру GPT-4, набори даних, вартість та інше. *The Decoder*. <https://the-decoder.com/gpt-4-architecture-data-sets-costs-and-more-leaked>
- 3 Там само, Schreiner, 2023.
- 4 Shivakumar, S. & Wessner, C. (2022, June 8). *Напівпровідники та національна оборона: Які ставки?* Центр стратегічних і міжнародних досліджень. <https://www.csis.org/analysis/semiconductors-and-national-defense-what-are-stakes>
- 5 Huang, K. & O'Regan, S.V. (2023, 5 вересня). Inside Meta's AI drama: Внутрішня ворожнеча через обчислювальну потужність. *The Information*. <https://www.theinformation.com/articles/inside-metas-ai-drama-internal-feuds-over-compute-power>
- 6 Зрозуміло, що ми вигадали цей порівняльний приклад. На практиці ми відносно впевнені, що MercedesBenz пропонує великий простір для ніг, і ми віримо, що жоден панельний фургон не постраждав у ході проведення цієї аналогії.
- 7 Dastin, J. (2018, 10 жовтня). Amazon скасовує секретний інструмент рекрутингу зі штучним інтелектом, який демонстрував упереджене ставлення до жінок. *Reuters*. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKC-N1MK08G>

- 8 Heaven, W.D. (2020, 17 липня). Прогностичні поліцейські алгоритми є расистськими: їх потрібно демонтувати. *MIT Technology Review*. <https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice>
- 9 Alba, D. (2022, 8 грудня). Чат-бот OpenAI видає упереджені роздуми, незважаючи на запобіжники. *Bloomberg*. <https://www.bloomberg.com/news/newsletters/2022-12-08/chatgpt-open-ai-s-chatbot-is-spitting-out-biased-sexist-results>
- 10 Nicoletti, L. & Bass, D. (2023, 14 червня). Люди упереджені; Генеративний ШІ ще гірший. *Bloomberg*. <https://www.bloomberg.com/graphics/2023-generative-ai-bias>
- 11 Кожен з цих токенів перетворюється (одношаровою нейронною мережею) у вектор вбудовування (довжиною 768 для GPT-2 і 12,288 для GPT-3 ChatGPT). Див: Вольфрам, С. (2023, 14 лютого). Що робить ChatGPT...і чому він працює? *Стівен Вольфрам, статті* [blog]. <https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work>
- 12 Vincent, J. (2016, 24 березня). Twitter навчив штучного інтелекту чат-бота Microsoft бути расистським мудаком менш ніж за день. *The Verge*. <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>
- 13 OpenAI. (2023, 27 березня). Технічний звіт GPT-4. *arXiv: 2303, 08774*. <https://doi.org/10.48550/arXiv.2303.08774>
- 14 Borak, M. (2023, 14 червня). Парламент ЄС схвалив Закон про штучний інтелект на тлі палких дебатів щодо біометрії. *Біометричне оновлення*. <https://www.biometricupdate.com/202306/eu-parliament-approves-ai-act-amid-heated-biometrics-debates>

- 15 Soare, S.R. (2021, 19 листопада). Алгоритмічна міць, НАТО і штучний інтелект. *Міжнародний інститут стратегічних досліджень*. <https://www.iiss.org/sv/online-analysis/military-balance/2021/11/algorithmic-power-nato-and-artificial-intelligence>
- 16 Щодо Китаю: Davidson, H. (2023, 23 лютого). “Політична пропаганда”: Китай обмежує доступ до ChatGPT. *The Guardian*. <https://www.theguardian.com/technology/2023/feb/23/china-chatgpt-clamp-down-propaganda>  
Щодо інших країн: Martindale, J. (2023, 12 квітня). Це країни, де ChatGPT наразі заборонено. *Цифрові тенденції*. <https://www.digitaltrends.com/computing/these-countries-chatgpt-banned>
- 17 OpenAI. (2023, 25 липня). Результати іспитів GPT4 проти GPT-3.5: Ефективність ШІ проти людей. *ChatGPT Blog* [блог]. <https://chatgptplus.blog/gpt4-test-scores>
- 18 Hersche, M., Zeqiri, M., Benini, L. et al. Нейро-векторносимволічна архітектура для розв’язання прогресивних матриць Равена. *Nature Machine Intelligence* 5, 363–375 (2023). <https://doi.org/10.1038/s42256-023-00630-8>
- 19 Metaculus, онлайн-платформа для прогнозування, яка агрегує людські прогнози, показує, що багато респондентів передбачають, що загальний штучний інтелект буде винайдений наприкінці цього десятиліття (тобто приблизно до 2030 року). Вони визначають ШІ як “єдину уніфіковану програмну систему, яка може ...

[1] надійно пройти 2-годинний змагальний тест Тьюрінга з текстом, зображеннями та аудіофайлами ...

[2] має загальні робототехнічні можливості, такі, що здатні ... задовільно зібрати (або еквівалент) модель автомобіля Ferrari 312 T4 у масштабі 1:8 приблизно 2021 року ...

[3] проявити високу компетентність у різних галузях знань, що вимірюється досягненням щонайменше 75% точності в кожному завданні та 90% середньої точності в усіх завданнях...

[4] отримати топ-1 строгу точність принаймні 90,0% на завданнях рівня інтерв'ю, знайдених у тесті APPS...

[5] наприклад, пояснити свої міркування під час відповіді на запитання або усно повідомити про свій прогрес та ідентифікувати об'єкти під час складання моделі.

Дивіться: <https://www.metaculus.com/questions/5121/date-of-artificial-general-intelligence>

Опитування експертів зі штучного інтелекту, проведене у 2022 році, прогнозує, що ймовірність того, що ШІ буде винайдений у найближчі 37 років, становить 50%. Див: Stein-Perlman, Z., Weinstein-Raun, B., & Grace, K. (2022, Aug. 3). Опитування експертів 2022 року про прогрес у галузі ШІ. *Вплив ШІ*. <https://aiimpacts.org/2022-expert-survey-on-progress-in-ai>

Нарешті, див. цей огляд очікуваних дат AGI: Dilmeğanı C. (2023, Oct. 2). *Коли відбудеться сингулярність? 1700 експертних думок про AGI [2023]*. AI Multiple. <https://research.aimultiple.com/artificial-general-intelligence-singularity-timing>

- 20 Searle, J. (1980). Розум, мозок та програми. *Поведінкові та мозкові науки*, 3(3), 417–424. <https://doi.org/10.1017/S0140525X00005756>
- 21 Наприклад, див. Chalmers, D. J. (1996). *Свідомий розум: у пошуках фундаментальної теорії*. Видавництво Оксфордського університету.
- 22 Damasio, A.R. (1994). *Помилка Декарта: Емоції, розум і людський мозок*. США: GP Putnam.

- 23 Область афективних обчислень була започаткована значною мірою Розаліндою В. Пікард; див., наприклад, Picard, R.W. (2000). *Афективні обчислення*. Бостон, штат Массачусетс: MIT Press Direct. <https://doi.org/10.7551/mitpress/1140.001.0001>
- 24 Harrison, M. (2023, 28 квітня). Ми взяли інтерв'ю у інженера Google, якого звільнили за те, що він сказав, що його ШІ ожив. *Футуризм*. <https://futurism.com/blake-lemoine-google-interview>
- 25 Clark, D. (2023). *Технологія навчання: Повний посібник для професіоналів у галузі навчання*. UK: Kogan Page.
- 26 Цит. за: Andreessen, M. (2023, Oct. 16). *Маніфест технооптимістів*. Andreessen Horowitz [блог]. <https://a16z.com/the-techno-optimist-manifesto>
- 27 Stodd, J. (2019, 19 грудня). Соціально динамічна організація ['domain to dynamic' pt 7]. *Навчальний блог Джуліана Стодда [блог]*. <https://julianstodd.wordpress.com/2019/12/19/the-socially-dynamic-organisation-domain-to-dynamic-pt-7>
- 28 Stodd, J. (2022, 26 січня). Контекст соціальної епохи. *Навчальний блог Джуліана Стодда [блог]*. <https://julianstodd.wordpress.com/2022/01/26/the-context-of-the-social-age>
- 29 Stodd, J. (2022). *Соціально динамічна організація: Нова модель організаційного дизайну*. UK: Sea Salt Publishing. <https://seasaltlearning.com/the-socially-dynamic-organisation-guidebook>
- 30 Stodd, J. (2014). *Посібник із соціального лідерства*. UK: Джуліан Стодд на Smashwords. <https://seasaltlearning.com/social-leadership-handbook-second-edition>  
Див. також: Stodd, J. (2022). *Посібник із соціального навчання*. UK: Sea Salt Publishing. <https://seasaltlearning.com/the-social-learning-guidebook>

- 31 Наприклад, див. Gao, Y. (2023). *Потенціал адаптивних систем навчання для покращення результатів навчання: Мета-аналіз*. (Докторська дисертація, Альбертський університет). <https://doi.org/10.7939/r3-а6хd-м403>

Гао проаналізував 77 емпіричних досліджень систем адаптивного навчання і виявив, що в середньому вони дають розмір ефекту 1,48 (в середньому). Для довідки, будь-який розмір ефекту понад 0,25 вважається корисним, а будь-який розмір ефекту понад 0,8 вважається “великим” і його можна легко помітити неозброєним оком.

- 32 Fletcher, J.D., & Morrison, J.E. (2014). *Прискорення розвитку досвіду: цифровий репетитор для технічної підготовки ВМС* [технічний звіт]. Арлінгтон, штат Вірджинія: Інститут оборонного аналізу.

Для порівняння, розмір ефекту понад  $1\sigma$  приблизно дорівнює підвищенню успішності учнів з 50-го перцентіля до 84-го перцентіля, а розмір ефекту понад  $2\sigma$  – підвищенню успішності з 50-го перцентіля до 98-го перцентіля.

- 33 Richards, J. & Dede, C. (2020, Oct. 26). *Навчальний план на 60 років: стратегічна відповідь на кризу*. Educause Review. <https://er.educause.edu/articles/2020/10/the-60-year-curriculum-a-strategic-response-to-a-crisis>

- 34 Зазвичай Девіду Вайсу (David J. Weiss) та Ненсі Бец (Nancy E. Betz) приписують першість у розробці комп'ютеризованого адаптивного тестування. Див: Weiss, D.J. & Betz, N.E. (1973). *Вимірювання здібностей: Традиційне чи адаптивне?* (Звіт про дослідження 73-1). Міннеаполіс: Університет Міннесоти, факультет психології. <https://files.eric.ed.gov/fulltext/ED077933.pdf>

- 35 Shute, V. & Ventura, M. (2013). *Прихована оцінка: вимірювання та підтримка навчання у відеоіграх*. Бостон, штат Массачусетс: MIT Press.

- 36 Приклади прихованої оцінки див. у *ibid.* Shute and Ventura, 2013. Приклад мультимодальної аналітики навчання, що застосовується у фізичних (армійських) вправах, див., наприклад, Vatrak, C., Biswas, G., Mohammed, N., & Goldberg, B.S. (2022). Автоматизована оцінка ефективності роботи команди з використанням мультимодальної аналітики байєсівського навчання. В *Матеріали Міжвідомчої/галузевої конференції з навчання, моделювання та освіти 2022 року*. Арлінгтон, Вірджинія: Національна асоціація тренінгів та моделювання.
- 37 Наприклад, див. Zetlin, M. (n.d.). ШІ тепер аналізує вираз обличчя кандидатів під час відеоспівбесіди. Див. <https://www.inc.com/minda-zetlin/ai-is-now-analyzing-candidates-facial-expressions-during-video-job-interviews.html>
- 38 Приклад перевірки ідентичності за поведінкою кліків див. у Vamosi, S., Reutterer, T., & Platzer, M. (2022). Підхід глибокої рекурентної нейронної мережі для вивчення подібності послідовностей для ідентифікації користувачів. *Системи підтримки прийняття рішень*, 155, 113718.
- Приклади використання поведінки кліків для підтримки академічної доброчесності див. тут: Mortati, J., & Carmel, E. (2021). Чи можемо ми запобігти технологічній гонці озброєнь у шахрайстві студентів університетів? *Комп'ютер*, 54(10), 90-94.
- ... або див. Alexandron, G., Berg, A., & Ruipérez-Valiente, J. A. (2023). Метод загального призначення на основі аномалій для виявлення шахраїв на онлайн-курсах. *IEEE Transactions on Learning Technologies*, pp. 1–12. <https://doi.org/10.1109/TLT.2023.3297132>
- 39 Kasparov, G. (2017). *Глибоке мислення: де закінчується машинний інтелект і починається людська творчість*. UK: PublicAffairs.

- 40 Давньогрецький поет Архілох спочатку писав про лисиць та їжаків, але Філіп Е. Тетлок (серед інших) допоміг популяризувати цю ідею в наш час. Наприклад, див: Tetlock, P. E. (2017). *Експертне політичне судження: Наскільки це добре? Як ми можемо знати?* Нове видання. UK: Princeton University Press.
- 41 Poli, R. (2023). Когнітивне доповнення команд людей та супер-ШІ за допомогою інтерфейсів мозку та комп'ютера за допомогою супер-людини та супер-ШІ. У *Матеріали конференції з генетичних та еволюційних обчислень*. Нью-Йорк: Асоціація обчислювальної техніки. <https://doi.org/10.1145/3583131.3603554>
- 42 Anderson, C. (2009). *Безкоштовно: майбутнє радикальної ціни*. US: Hachette Books.
- 43 Konnikova, M. (2012, 30 квітня). Про письмо, пам'ять і забуття: Сократ і Гемінгвей беруться за Зейгарник. *Scientific American*. <https://blogs.scientificamerican.com/literally-psyched/on-writing-memory-and-forgetting-socrates-and-hemingway-take-on-zeigarnik>
- 44 Європейський парламент. (2023, 11 травня). AI Act: крок до перших правил щодо штучного інтелекту (Ref.: 20230505IPR84904). *Прес-рум*. <https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence>
- 45 Ми згадали про луддитів з наміром викликати в уяві образ людини, яка бореться проти технологічного прогресу, але це посилення трохи несправедливе щодо історії. Хоча луддити руйнували машини і бунтували проти автоматизованих заводів, їхніми основними рушійними силами були соціальні та економічні чинники. Наприклад, див. Conniff, R. (2011, березень). *Проти чого насправді боролися луддити*. Смітсонівський журнал. <https://www.smithsonianmag.com/history/what-the-luddites-really-fought-against-264412>

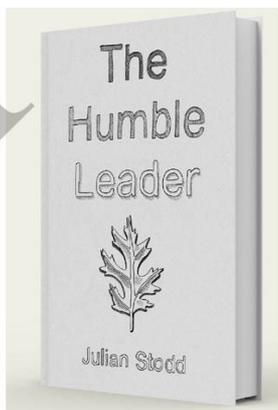
## Додаткова література

### SEA SALT PUBLISHING

Sea Salt Publishing – незалежне видавництво, що спеціалізується на роботах, які досліджують аспекти соціальної епохи. “Соціально епоху” Джуліан Стодд визначив у 2014 році як новий контекст нашого життя і роботи – ландшафт радикальних зв’язків, різноманітних екосистем технологій, руйнування традиційної влади, де значна частина цінностей знаходиться на перетині формальних і соціальних систем. У видавництві Sea Salt Publishing ми публікуємо дослідницькі, креативні та амбітні роботи, що стосуються соціальної епохи, у різних форматах, від красивих книг у твердій палітурці до практичних посібників та революційних і складних “зінгів”.



### ІНШІ КНИГИ ДЖУЛІАНА СТОДДА



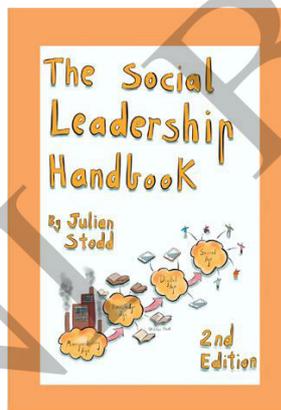
---

*Смирений лідер – це керовані роздуми про нашу особисту смиренність як соціального лідера.*

---

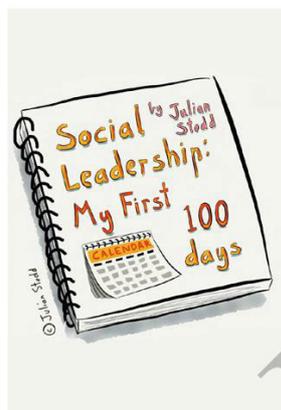
## ІНШІ КНИГИ ДЖУЛІАНА СТОДДА (ПРОД.)

---



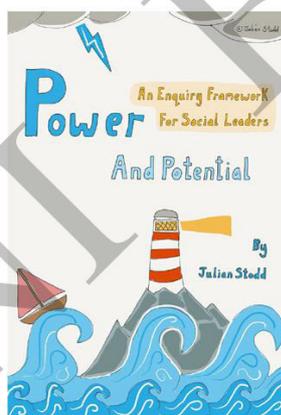
*Посібник з соціального лідерства, 2-е видання досліджує перетин формальної та соціальної влади і розглядає важливість цього в контексті соціальної епохи.*

---



*Соціальне лідерство: Мої перші 100 днів – це практична, керована, рефлексивна подорож у 100 днів діяльності, кожен з яких включає провокації, запитання та дії. Ви заповнюєте книгу в міру просування. Вона супроводжується набором зі 100 подкастів.*

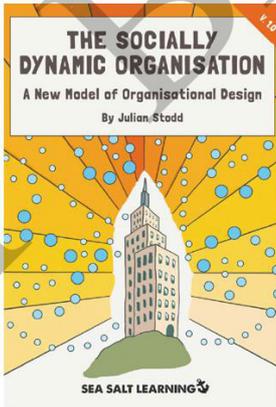
---



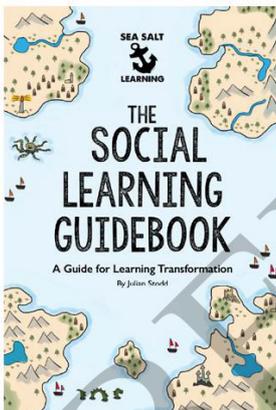
*Влада і потенціал – це структура дослідження: серія з шістнадцяти запитань, які разом створюють простір для вивчення. Предметом дослідження є влада – влада, яка є основою лідерства як у формальному, так і в соціальному просторах.*

---

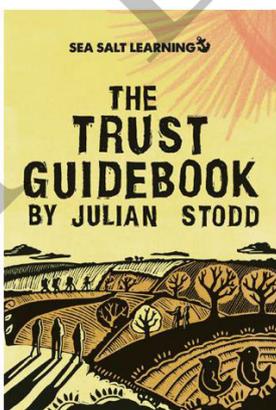
*Серія путівників* – серія коротких, заснованих на дослідженнях, практичних і прикладних путівників з питань соціального віку, яка постійно розширюється



*Соціально динамічна організація* представляє новий тип організації, яка є легкою та швидко адаптується, процвітає за часів постійних змін, поважає старе, але приймає нове.



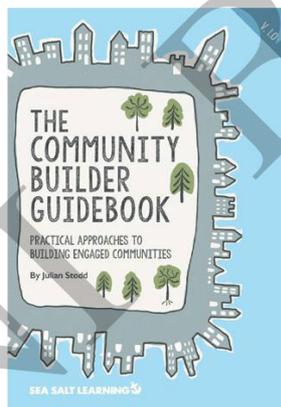
*Посібник із соціального навчання* надає практичний огляд принципів і методів розроблення соціального навчання в сучасній організації.



*Посібник з довіри* містить результати нашого масштабного дослідження ландшафту довіри та пропонує сімдесят два запитання, які лідери можуть використати у роботі зі своїми командами.

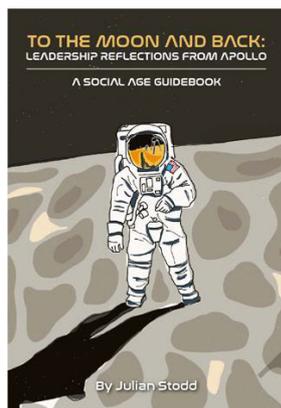
## ІНШІ КНИГИ ДЖУЛІАНА СТОДДА (ПРОД.)

---



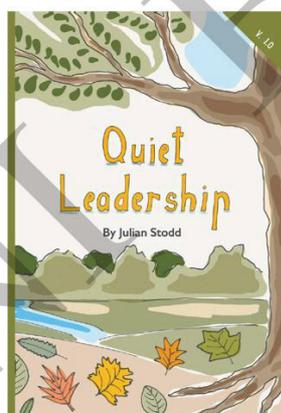
*Посібник для розбудови спільнот* пропонує вам практичні ідеї для створення активних і динамічних соціальних навчальних спільнот та спільнот практик.

---



*На Місяць і назад: роздуми про лідерство з Аполло* ділиться вісьмома ключовими історіями про програму “Аполлон”, а також особистими роздумами Джуліана про те, що це означає для лідерства в соціальну епоху.

---

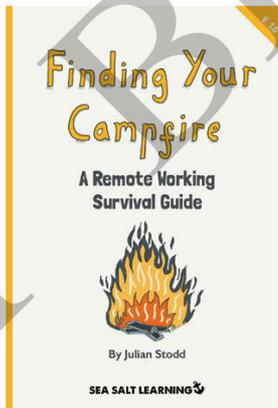


*Тихе лідерство* – це дослідження лідерства в найменших дрібницях: нашому мисленні, наших словах і наших діях кожного дня.

---

## ІНШІ КНИГИ ДЖУЛІАНА СТОДДА (ПРОД.)

---



*Знайти своє багаття* допомагає віддаленим командам у вивченні трьох тем: пакування рюкзака, керівництво експедицією та перебування разом і нарізно.

---

## ПРОДОВЖУЙТЕ РОЗМОВУ

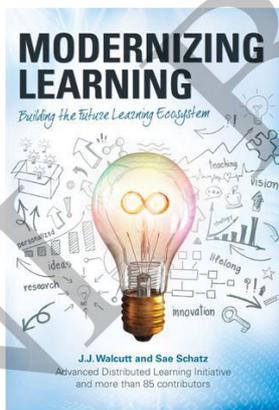
Ми хотіли б почути, що ви думаєте про цю книгу! Надішліть нам повідомлення, слідкуйте за нашою роботою та станьте частиною нашої спільноти.

Долучайтеся до навчання Sea Salt Learning за посиланням [www.seasaltlearning.com](http://www.seasaltlearning.com) або зв'яжіться з нами за посиланням [hello@seasaltlearning.com](mailto:hello@seasaltlearning.com). Ви також можете читати щоденний блог Джуліана на сайті [www.julianstodd.wordpress.com](http://www.julianstodd.wordpress.com), дивитися його шістдесят друге відео *Social Leadership Daily* на Substack на [dailyquestions.substack.com](http://dailyquestions.substack.com), та приєднатися до щотижневих глибоких досліджень *The Captain's Log* на <https://socialage.substack.com>.

Слідкуйте за оновленнями на LinkedIn, шукайте [Джуліан Стодд](#), [Сас Шатц](#), [Джефф Стед](#), та [Sea Salt Learning](#). Або знайдіть нас на X: [@JulianStodd](#), [@SaeSchatz](#), [@geoffstead](#), and [@SeaSaltLearning](#).

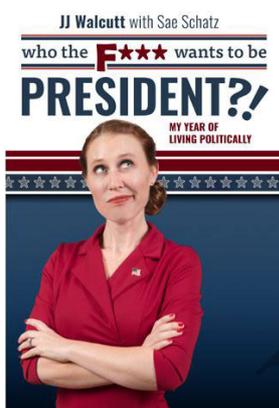
## ІНШІ КНИГИ САЕ ШАТЦ

---



*Модернізація навчання: побудова майбутньої навчальної екосистеми, написана у співавторстві з JJ Walcutt і десятками авторів, висвітлює основи і майбутнє навчання впродовж життя з використанням технологій. Вона знаходиться у вільному доступі за посиланням [www.adlnet.gov](http://www.adlnet.gov).*

---



*Хто в біса хоче бути президентом: Мій рік політичного життя – це гумористична правдива історія. Приєднуйтесь до Джей Джей Волкатт під час передвиборчої кампанії, коли вона намагається розібратися в американській політиці, уряді та громадянській активності.*

---

## ІНШІ КНИГИ ДЖЕФФА СТЕДДА

TA-DA!

Ви вже маєте її! Ви тримаєте в руках першу фізичну книгу Джеффа. Інші роботи можна знайти через Scholar та LinkedIn, або пошукати в Інтернеті його нещодавні ключові доповіді.

---