

Ми і є ШІ № 5:

# Ми — штучний інтелект



© Julia Stoyanovich & Falaah Arif Khan (2022) Translated by Yaroslava Savosh-Davydova; edited by Yevhen Redko and Tetiana Zakharchenko

# УМОВИ ВИКОРИСТАННЯ

Усі ілюстрації в цьому коміксі доступні за ліцензією CC BY-NC-ND 4.0. Будь ласка, перейдіть на сторінку ліцензії, щоб дізнатися більше про те, як можете використовувати ці роботи.

Не соромтеся використовувати панелі/групи панелей у презентаціях/статтях, якщо

- 1) належно цитуєте їх;
- 2) не вносите змін в окремі панелі.

Цитувати як:

Джулія Стоянович та Фала Аріф Хан. «Ми — штучний інтелект». Ми і є ШІ. Комікси, том 5 (2021) <http://r-ai.co/comics>

У цьому томі ми повернемося до деяких ключових понять із коміксів,

зміцнюючи ідею про те, що  
технологія — зокрема ШІ —  
така, якою хочемо її бачити.

**МИ — це  
ШІ.**

Ми обміркували кілька гіпотетичних і  
реальних застосувань ШІ-технологій,

зокрема алгоритм,  
який вчиться випікати,



смартсвітло,



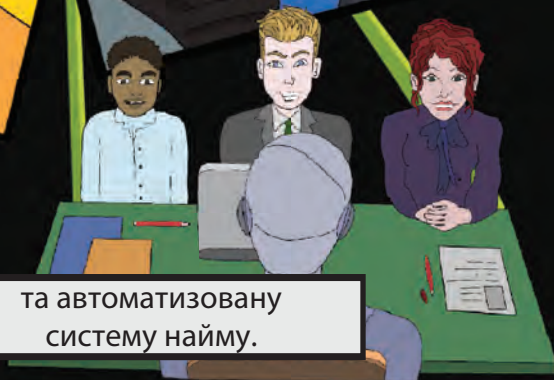
безпілотний  
автомобіль



штучний інтелект, що грає в шахи,



та автоматизовану  
систему найму.

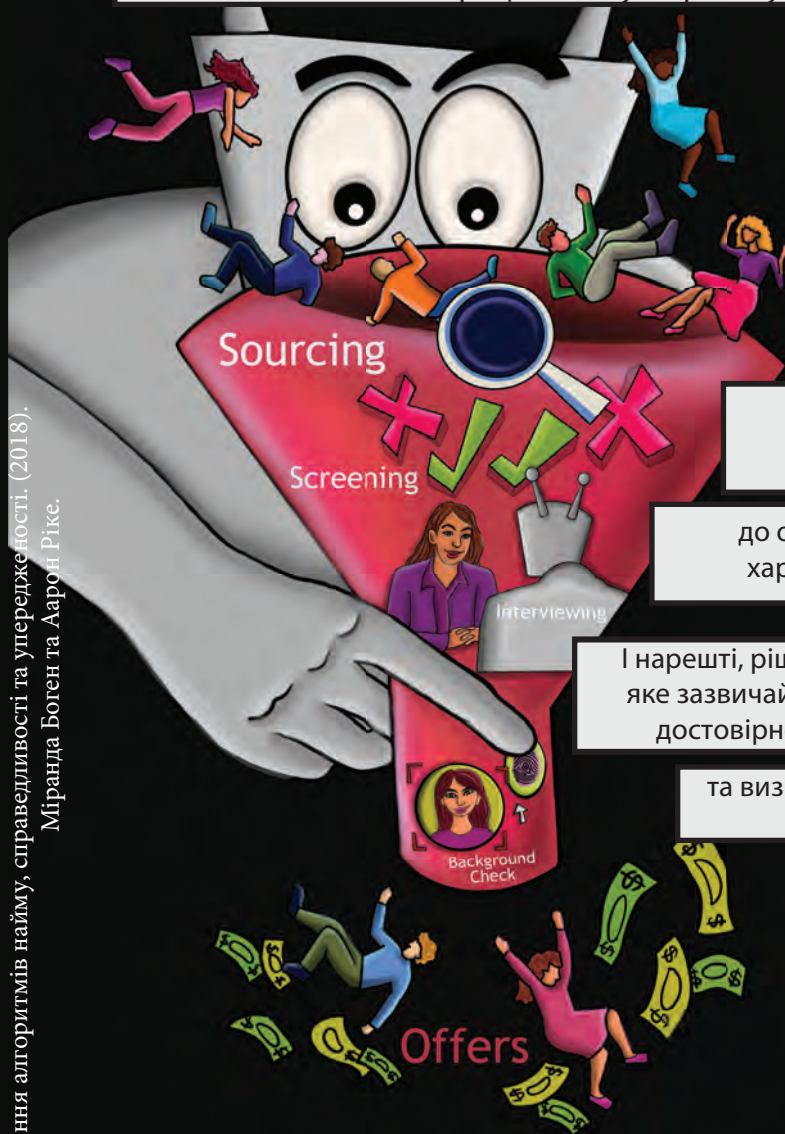


З поданих прикладів автоматизована система найму,  
мабуть, найсуперечливіша,

і це також тема, яку ми вже докладно обговорювали,  
тож побіжно повернімося до неї.



2018 року у звіті «Потрібна допомога» Міранда Боген та Аарон Ріке описали процес найму на роботу як лійку —



послідовність кроків,  
керованих алгоритмом  
на основі даних,

у якому низка рішень  
завершується  
пропозицією роботи  
одним кандидатам

і відмовою  
іншим.

Від пошуку кандидатів та  
розміщення вакансій в  
інтернеті

до оцінки досвіду, навичок і  
характеристик кандидатів.

І нарешті, рішення про найм на роботу,  
яке зазвичай передбачає перевірення  
достовірності біографічних даних

та визначення відповідного рівня  
заробітної плати.

Важливо, що на всіх цих  
етапах використовують  
дані та предиктивну  
аналітику — те, що ми  
побіжно називаємо «ШІ».

Про це заявила Дженні Янг, колишня  
очільниця Комісії США з питань рівних  
можливостей у сфері зайнятості  
(EEOC).

Через важливість рішення, яке вони  
приймають або допомагають прийняти,

«автоматизовані системи найму діють як  
сучасні ворота до економічних можливостей».



Ми бачили приклади, коли ці інструменти посилюють гендерні та расові упередження.

Автоматизовані системи найму також обмежують доступ до робочих місць для людей з інвалідністю, —



ізолюючи людей, які мають психічні захворювання, зокрема депресію або біполярний розлад,

на основі онлайн-тестів особистості,

навіть якщо такі люди мають відповідні навички для роботи [7].

Ми також часто турбуємося про те, чи справді ці інструменти «працюють» —

чи вловлюють вони корисні сигнали з даних

або в кращому разі чи становлять ретельно продумане підкидання монетки.

Як зазначає професор комп'ютерних комп'ютерних наук Арвінд Нараянан [8]:

Чи є ці інструменти «фуфломіцином» для ШІ?



Інструменти ШІ — це інженерні артефакти.

Щоб дізнатися, чи працюють вони, ми повинні використовувати науковий метод:

сформулювати гіпотезу, що інструмент справді добирає працівників, які добре справляються з роботою,

і прогнозує краще, ніж унаслідок випадкового вгадування.

По тому спланувати експеримент, щоб перевірити цю гіпотезу,

і бути готовими визнати, що інструмент не працює з огляду на результат.

У складній екосистемі, у якій автоматизовані інструменти найму замовляють, розробляють та використовують, ми повинні запитати себе:

хто відповідає за те, щоб ці інструменти були створені й застосовані як належить?

Хто відповідає за виявлення та пом'якшення наслідків дискримінації й порушень належної правової процедури?

А за контроль над поширенням «фуфломіцину» під витівковою назвою «штучний інтелект»?

Відповідь — ми всі відповідаємо.

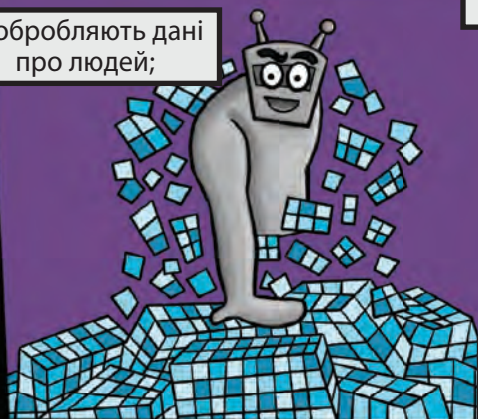
У відповідь на ці питання ми спостерігаємо спроби регулювати використання алгоритмічних інструментів.

Зробимо крок назад, щоб отримати більш систематизований погляд на

«Автоматизовані системи прийняття рішень», або ADS.

Ці системи:

1. обробляють дані про людей;

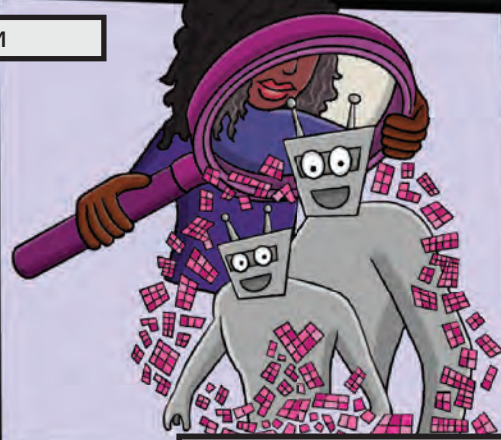


2. допомагають — або разом із прийняттям рішень людиною, або автономно —



в ухваленні рішень, які впливають на життя людей та їхні засоби до існування;

Крім того, ми хотіли б, щоб такі системи



3. мали конкретну, заявлену мету — підвищити ефективність та сприяти або принаймні не перешкоджати рівному доступу до можливостей;

4. були оприлюднені публічно та контрольовані.

Чи є формула в електронній таблиці ADS?

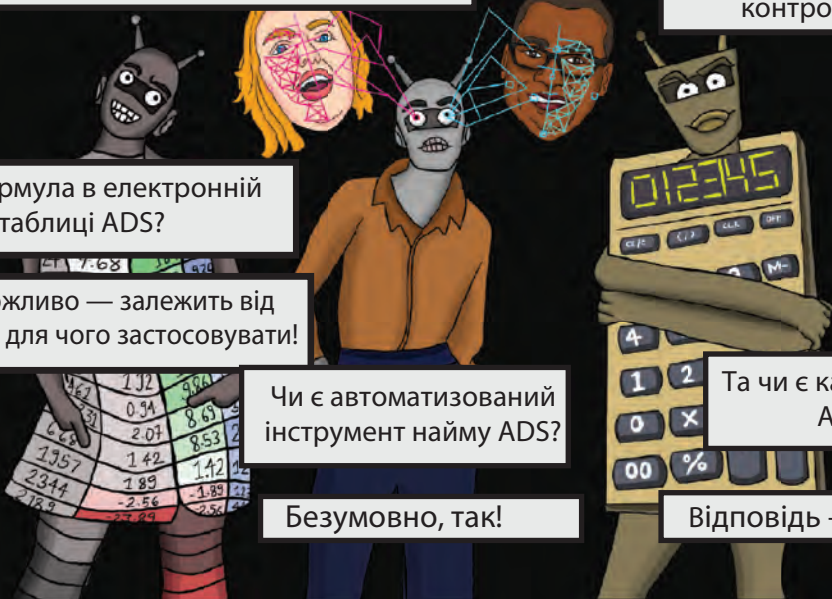
Можливо — залежить від того, для чого застосовувати!

Чи є автоматизований інструмент найму ADS?

Безумовно, так!

Та чи є калькулятор ADS?

Відповідь — ні!





Як ми можемо регулювати ADS?

І чи варто взагалі намагатися?

Хоч у галузі досі переважає думка, що «регулювання буде перешкоджати інноваціям»,



**індустрія не може вирішувати сама собою.**

І навіть у Кремнієвій долині дедалі більше визнають те, що потрібне ефективне регулювання, яке полегшить дотримання законодавства та обмежить відповідальність.

Існує чимало дискусій щодо конкретної нормативно-правової бази:

чи варто використовувати принципи обережності,

що можна підсумувати як «краще перестрахуватися, ніж потім шкодувати»?



Або, який ймовірніше, спробуйте гнучкіший метод, що ґрунтується на оцінювання ризиків,

наприклад алгоритмічне оцінювання впливу?



Усе це, і не тільки, становить предмет запеклих дискусій.



Наприклад, нещодавно місто Нью-Йорк публічно взяло на себе зобов'язання відкрити «чорну скриньку» щодо того, як уряд використовує технології.

У травні 2018 року була скликана перша у США робоча група з автоматизованих систем прийняття рішень,

якій доручили надати рекомендації установам Нью-Йорка

щодо забезпечення прозорості та підзвітності у використанні ADS.



Робоча група опублікувала звіт у листопаді 2019 року [9], узявши на себе зобов'язання використовувати ADS **там, де вони корисні;**

сприяти справедливості, рівності, підзвітності та прозорості їхнього використання

і зменшувати потенційну шкоду протягом усього їхнього життя.

[9] Звіт робочої групи з автоматизованих систем прийняття рішень міста Нью-Йорк, листопад 2019 року.

Тепер нам потрібно працювати разом, щоб зробити ці принципи конкретнішими.

Перший крок — ретельно продумати наші цілі.

Як зробити так, щоб ADS працювали для всіх нас?

ШІ дає безпрецедентні можливості, щоб прискорити розвиток науки та інновацій,

трансформувати те, як живемо, і те, як керуємо.

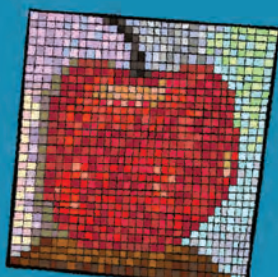
Однак коли розробляємо ШІ й застосовуємо його, ми повинні чесно говорити про свої цілі,

хто виграє, коли досягнемо цих цілей,

і кому буде завдана шкода.

Чи працює ШІ для всіх нас, залежить від нашого вибору —

наприклад, маємо вирішити, як подати дуже складні й часто невимірювані поняття.



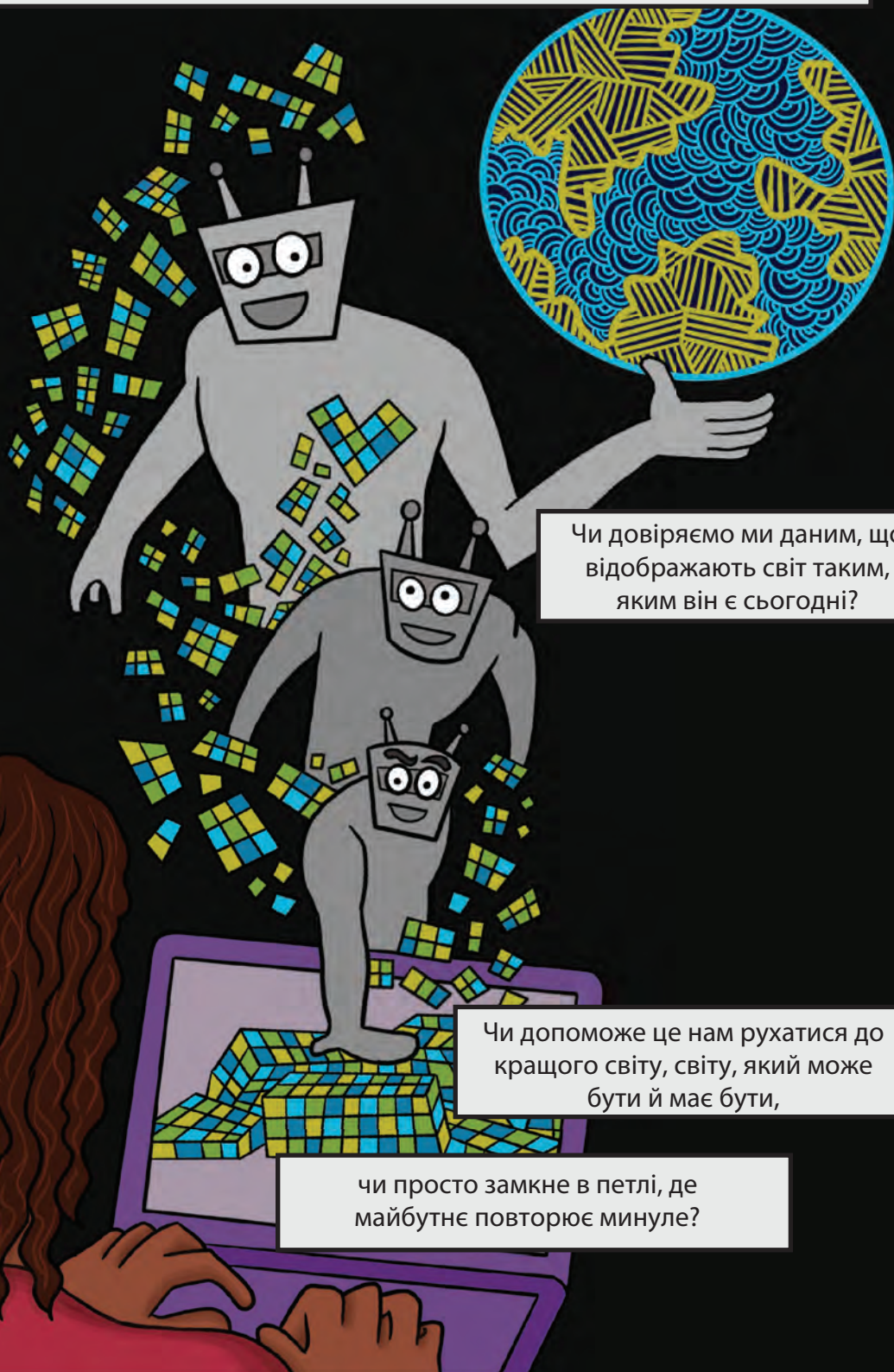
Які риси виберемо, щоб представити перспективного претендента на роботу?



На основі чого ми вирішуємо, що правило класифікації успішне або неуспішне?



Ми також повинні ретельно та критично обмірковувати, чи довіряємо даним, які живлять наші інструменти штучного інтелекту.



Чи довіряємо ми даним, що відображають світ таким, яким він є сьогодні?

Чи допоможе це нам рухатися до кращого світу, світу, який може бути й має бути,

чи просто замкне в петлі, де майбутнє повторює минуле?

Як можемо використовувати технології, щоб покращити суспільство?

Для цього ми повинні знайти прагматичну золоту середину між шкідливими крайнощами.

## Технооптимізм —

віра в те, що технологія може самотужки розв'язати глибоко вкорінені суспільні проблеми, наприклад системну дискримінацію під час наймання на роботу,

## і технокритицизм —

переконавання, що будь-яка спроба операціоналізувати етичні та правові норми в ADS дорівнюватиме «відмиванню грошей»,

тому має бути відкинута повністю.

Щоб досягти прогресу, нам усім потрібно вийти із зони комфорту.

Ми повинні взяти на себе відповідальність за рішення, що приймають, розробляючи, використовуючи та контролюючи ШІ, — наші рішення.

NUESTRAS DECISIONES.

Тільки якщо ми всі зробимо крок назустріч, суспільство зможе спиратися в розробленні технологій на людей.



Зрештою, технологія — зокрема і ШІ — це те, чим ми її робимо.

**МИ — це ШІ.**