

Кейт О'Ніл

# BIG DATA

ЗБРОЯ МАТЕМАТИЧНОГО ЗНИЩЕННЯ

*Як великі дані збільшують нерівність і загрожують демократії*

*З англійської переклала Ольга Калініна*

Київ  
BOOKCHEF  
2020

# ЗМІСТ

- 7** ВІДГУКИ ПРО КНИГУ «ЗБРОЯ МАТЕМАТИЧНОГО  
ЗНИЩЕННЯ»
- 15** ПОДЯКА
- 17** ВСТУП
- 
- 35** Розділ 1  
**ЧАСТИНИ БОМБИ**  
*Що таке модель?*
- 
- 58** Розділ 2  
**КОНТУЗІЯ**  
*Моя подорож до втрати ілюзій*
- 
- 82** Розділ 3  
**ГОНИТВА ОЗБРОЄНЬ**  
*Ідучи до коледжу*
- 
- 109** Розділ 4  
**ПРОПАГАНДА МАШИН**  
*Онлайн-реклама*
- 
- 130** Розділ 5  
**ЦИВІЛЬНІ ЖЕРТВИ**  
*Справедливість в Ері Великих Даних*

- 159** Розділ 6  
**НЕПРИДАТНИЙ СЛУЖИТИ**  
*У пошуках роботи*
- 182** Розділ 7  
**ОБЛИВАЮЧИСЬ ПОТОМ**  
*Нумо до роботи!*
- 207** Розділ 8  
**НЕПРЯМИЙ ЗБИТОК**  
*Кредит на споживання*
- 232** Розділ 9  
**НЕБЕЗПЕЧНА ЗОНА**  
*Застрахуватися*
- 257** Розділ 10  
**ЦІЛЬОВА АУДИТОРІЯ**  
*Суспільне життя*
- 282** ВИСНОВОК
- 309** ЛІТЕРАТУРА

## **Відгуки про книгу «ЗБРОЯ МАТЕМАТИЧНОГО ЗНИЩЕННЯ»**

«У книзі О'Ніл окреслено страхітливий огляд того, як алгоритми дедалі більше регулюють життя людей... Знання авторки про силу та ризики математичних моделей у поєднанні з даром за аналогією робить її однією із найцінніших спостерігачів постійного збройного нарощування Великих Даних... Вона віртуозно працює, пояснюючи всепроникність і ризики алгоритмів, які регулюють наше життя».

*New York Times Book Review*

«“Зброя математичного знищення” — це історія Великих Даних Кремнієвої долини, про яку не розкажуть прихильники Кремнієвої долини... Вона, по суті, розкриває недоліки того, як використовують інформацію для оцінки всього — від кредитоспроможності до тактики політики... Ґрунтовне читання для кожного, хто схильний вірити, що дані не брешуть».

*Reuters*

«Прониклива та хвилююча».

*New York Review of Books*

«“Зброя математичного знищення” — це нагальна критика того, як невинно зловживають математикою в кожному з аспектів нашого буття».

*Boston Globe*

«Неймовірна й дуже захоплива книжка»

*Ювал Ной Харарі, автор бестселера 2016 року*

*«Людина розумна»<sup>1</sup>*

«Просто фантастична книжка... О'Ніл робить переконливий крок до того, щоби ми зрозуміли, як далеко можуть зайти алгоритми».

*The Atlantic*

«Важливе нагадування усім: Великі Дані, як і люди, залишаються добрими, допоки володіють собою».

*Wired*

«Якщо ви підозрювали, що щось не так відбувається з нашою довірою до даних, і вам завжди не вистачало математичних навичок зрозуміти їхню природу, тоді ця книжка для вас».

*Salon*

«Кейт О'Ніл — це один із найсильніших голосів про те, як потрібно грамотно обмежувати діяльність

---

1 Харарі Ю. Н. Людина розумна. Історія людства від минулого до майбутнього / пер. з англ. Я. Лебеденка. — Харків: КСД, 2016. — 544 с.

алгоритмів, яким ми дозволяємо впливати на всі сфери нашого життя».

*Scientific American*

«Доволі зрозуміло та захопливо... лаконічно та прозоро... “Зброя математичного знищення” — це роман Ептона Синклера “Джунглі” сьогодення».

*Марк Ван Холлебеке, автор блогу Data and Society: Points*

«Неймовірна книжка. Попри всю технічну складність цього явища, книжка “Зброя математичного знищення” чітко веде читачів скрізь ці непрості системи математичного моделювання».

*National Post*

«Ця книжка є мудрою, жорстокою і водночас надзвичайно потрібною».

*Джордан Елленберг, професор математики в університеті Вісконсин-Медисон, автор книги «Як ніколи не помилятися»*

«О'Ніл стала справжньою [володаркою] у світі Великих Даних».

*Time*

«О'Ніл говорить про те, як наші постійні використання перевантажених ботів використовують дані для укріплення дискримінації та несправедливості демократичного вибору. Якщо ви працюєте з даними чи продукуєте їх у режимі онлайн, то маєте прочитати цю книжку».

*ArsTechnica*

«Ця книжка має не менш серйозний посил, ніж роман Синклера “Джунглі” чи книжка еко-захисниці Рейчел Карсон “Тиха весна”. Як і ці книжки, вона має змінити курс американського суспільства».

*Aspen Times*

«Доступно, переконливо — і найважливіше, про що більшість не говорять — і її дуже легко читати».

*Inside Higher Ed*

«Найнебезпечнішими алгоритмами є також найтаємніші. Ось чому... дослідження в книжці Кейт О'Ніл показують, як це важливо; і вона говорить, де можна їх знайти».

*The Guardian*

«Розважальна і водночас серйозна книжка, яка дасть читачам засоби для проходження ідеологічного туману сучасності, серед якого заховалися всі небезпеки революції в епісі Великих Даних».

*In These Times*

«Бачення світу О'Ніл керується алгоритмами, пронизане темним гумором і озлобленням — як сучасний Доктор Стрейнджлав або “Пастка-22”<sup>1</sup>. Це повчальне, тривожне і глибоко важливе значення».

*Стівен Строгац, професор університету  
Корнелла, автор книги «Свято невідомого»*

1 Геллер Дж. Пастка-22 / пер. з англ. О. Фешовець. — Вавилонська бібліотека, 2016. — 464 с.

«Нічні жахи технофобів... На відміну від деяких інших останніх книг про збір даних книга О'Ніл не історична; вона пропонує більше варіант холодного душу, коли проводить читачів через шляхи індустрії "великих даних", що сприяє розвитку соціальних недуг».

*Publishers Weekly*

«Це маленький Філіп К. Дік, маленький Орвелл і маленький Кафка в портреті могутніх бюрократій, що передають контроль над найінтимнішими рішеннями наших життів до гіперповноважених комп'ютерних моделей, пронизаних всіма невирішеними атавістичними людськими недоліками».

*Paris Review*

«Таємничість бореться з ясністю, а плутанина — із розумінням — ця книга може допомогти нам змінити це, допоки ще не стало надто пізно».

*Астра Тейлор, авторка книги «Громадянська платформа»*

«Аналіз О'Ніл чудовий, її письмо привабливе, а висновки — невтішні».

*Дана Бойд, засновниця некомерційної організації Data & Society і авторка книги «Це складно»*

«Книга "Зброя математичного знищення" є фантастичним, відвертим закликом до зброї. Вона визнає, що моделі нікуди не зникають: як інструмент для ідентифікації людини у скруті, вони чудові. Та як інструмент для покарання і позбавлення громадянського права, вони є нічними жахами. Книга Кейт



О'Ніл є, певно, важливою, бо вона вірить у науку про дані. Це життєво важливий короткий курс, де саме ми маємо досліджувати системи навколо нас і вимагати від них ліпшого».

*Корі Докторов, автор «Меншого брата»  
та співредактор у «Voing Voing»*

«Багато алгоритмів працюють до знемоги до невідповідності влади й упередження. Якщо ви не хочете, щоб ці алгоритми стали вашими господарями, прочитайте книгу “Зброя математичного знищення” Кейт О'Ніл, щоб зруйнувати останні паростки тиранії зарозумілого панівного класу».

*Ральф Нейдер, автор «Небезпечний  
на будь-якій швидкості»*

«Наступного разу, коли ви почуєте чиїсь надмірні некритичні сентименти про дива Великих Даних, покажіть їм книгу “Зброя математичного знищення”. Це буде корисно».

*Фелікс Селмон, Fusion*

«Від отримання роботи до пошуків одного з членів подружжя, алгоритми прогнозування тихо формують і контролюють наші долі. Кейт О'Ніл бере нас у подорож між обуренням і дивом із прозовою розповіддю, що змушує відчувати, ніби це всього лише проста бесіда. Проте вона важлива. Ми маємо зважати на технології».

*Лінда Тірадо, авторка «Рука біля рота:  
жити в перевантаженій Америці»*

## ПОДЯКА

Дякую моему чоловікові й дітям  
за їхню неймовірну підтримку.

Дякую також Джону Джонсону, Стівену Вольдману,  
Макі Інаді, Беккі Джаффе, Аарону Абрамсу, Джулії Стіл,  
Карен Бернс, Метту ЛаМантіа, Марті Пун, Лізі Редкліфф,  
Луїсу Даніелю та Меліссі Білскі.

Наостанок дякую людям, без яких ця книга не існувала би:

Лаурі Страусфелд, Аманді Кук, Еммі Беррі,  
Джордану Елленбергу, Стівену Бейкеру, Джей Мендел,  
Семові Кенсон-Бенанаву й Ерні Девісу.



## ВСТУП

Коли я була маленькою дівчинкою, то часто спостерігала за міським рухом крізь вікно машини й вивчала цифри на номерних знаках. Спрощувала кожен із них до найменших значень — простих чисел, які лишень можна полічити. 45 — це три на три і на п'ять. Це називають факторингом<sup>1</sup>, і це було моїм улюбленим видом розслідувань. Як математичний ботанік я була особливо заінтригована простими числами.

Моя любов до математики, зрештою, стала моєю пристрастю. Я приєдналася до математичного табору у чотирнадцять, постійно приходячи додому з кубиком Рубика, який притискала до себе. Математика ставала для мене притулком від цього божевільного

---

1 Декомпозиція об'єкта (приміром, числа) у добуток інших об'єктів або множників, які після перемноження дадуть вихідний об'єкт. — Тут і далі — прим. перекл., якщо не зазначено іншого.

світу. Вона просувалася чимраз далі, її царина знань розросталася безжалісно, доказ за доказом. І мені ще було куди рухатись. Я була найліпшою в коледжі з математики та продовжила своє навчання на доктора. Моїм напрямком стала алгебраїчна теорія чисел — сфера, яка своїм корінням сягала у факторинг, яким я займалася ще дитиною. І якось несподівано отримала місце професора в Бернард-коледжі, де кафедра математики була поєднана з Колумбійським університетом.

А потім я зробила неймовірне. Полишила свою роботу й пішла працювати помічником до Девіда Елліота Шоу, керівника гедж-фонду<sup>1</sup>. Покидаючи коледж через фінансові труднощі, я перенесла математику з абстрактної теорії до практики. Обчислення, які ми здійснювали з цифрами, переходили в трильйони доларів, які переходили з одного рахунку на інший. Спочатку я була в захваті та вражена роботою в новій лабораторії, у лабораторії світової економіки. Та восени 2008 року, коли я пробула в цьому місці понад рік, усе пішло шкереберть.

Ця криза доволі чітко показала, що математика, як мій прихисток, не тільки глибоко проникнула в проблеми світу, а й також підживлювала багато з них. Криза будівництва, колапс багатьох фінансових інституцій, зростання безробіття — усе це підтримували й заохочували за допомогою розмаху магії математичних формул. Ба більше, завдяки надзусиллям, які я безмежно обожнювала, математику поєднували

---

1 Приватний інвестиційний фонд із використанням об'єднаних коштів, який займається керуванням капіталом, щоб максимально примножити кошти з чітко обумовленим або плаваючим ризиком.

з технологією, щоби сіяти хаос і нещастя, тим самим збільшуючи ефективність і розмах систем, які я зараз визнавала помилковими.

Якби в нас були світлі голови, то ми зробили б крок назад до цієї точки, щоби визначити, як зловживали математикою і як ми могли перешкодити схожій катастрофі в майбутньому. Натомість, на початку кризи, нові математичні техніки були популярними як ніколи, тож їх починали використовувати у багатьох царинах. Вони штампували в режимі 24 години на добу та 7 днів на тиждень скрізь петабайти<sup>1</sup> інформації, багато з якої було вилучено з соціальних мереж чи комерційних веб-сайтів. І дедалі частіше ці техніки фокусували не на світових фінансових ринках, а на людських істотах, на нас. Математики й статистики досліджували наші прагнення, пересування і зусилля, які ми витрачали. Вони передбачали нашу доброчесність і обраховували наш потенціал як студентів, працівників, коханців і злочинців.

Це була Економіка Великих Даних, яка тоді передбачала нечуваний успіх. Комп'ютерна програма могла проглянути тисячі резюме чи заяв на отримання кредиту за секунду чи дві й розсортувати їх відповідно до списків із зазначенням найперспективніших клієнтів у топі рейтингу. Це не тільки економило час, а й рекламувалося як чесний і об'єктивний підхід. Зрештою, вона не включала упереджених людей, які поралися зі стосами паперів, а лише машинну обробку сухих цифр. Десь 2010 року чи близького того математику сприймали цілком належно та як ніколи

1 Одиниця виміру інформації, яка дорівнює 1015 байт, або ж 1024 терабайтам.

раніше в людських відносинах, і громадськість загалом вітала це.

Та я вбачала в цьому проблему. Застосунки, які створили на основі математики і які продукували економіку даних, виникли з огляду на вибір ненадійних людських істот. Деякі з таких рішень робили, безсумнівно, з найліпшими помислами. Попри це багато з таких варіантів зашифрували людську упереженість, непорозуміння і забобонність у програмне забезпечення, яке почало поступово керувати нашим життям. Наче боги, ці математичні моделі були невизначеними, їхня робота була невидимою для всіх, але були й утаємниченіші жерці у цій справі — математики й розробники. Їхні вердикти, навіть якщо вони помилкові чи шкідливі, були безапеляційними й беззаперечними. І вони зазвичай карали бідних і пригнічених у нашому суспільстві, продовжуючи робити багатих ще багатшими.

Саме тоді я й придумала назву для таких шкідливих типів моделей — Зброя Математичного Знищення, або коротко — ЗМЗ. І покажу це на одному прикладі, який викриє всі її руйнівні характеристики протягом книжки.

Як часто трапляється, цей випадок почався з мети, яка гідна похвали. В 2007 році новий мер Вашингтона, округ Колумбія, Едріан Фенті, був налаштованим на те, щоби змінити відсталі школи міста. Перед ним постало геть непросте завдання: на той час лише один з двох [1] учнів старшої школи завершив своє навчання після 9-го класу, і тільки 8 % випускників [2] мали хороші оцінки з математики. Тому Фенті найняв на роботу освітню реформаторку на ім'я Мішель Рі, щоби

вона обійняла нову посаду — керівниця за наглядом вашингтонських шкіл.

Суть теорії полягала в тому, що учні недостатньо вчилися через те, що їхні вчителі не виконували своїх обов'язків. Тому 2009 року пані Рі впровадила план, щоби відсіяти тих викладачів, які працювали погано. Це був тренд у проблемних освітніх округах в межах округу Колумбія, і в цю перспективну систему оцінювання почали вкладати чудову ідею під гаслом «Оцінуй учителя!». Усе просто: варто позбутися найгірших викладачів і замінити їх найліпшими — тими, які приносять більшу користь. Мовою вчених з даних, це «оптимізує» шкільну систему і, ймовірно, гарантувало би поліпшення результатів для дітей. За винятком «поганих учителів», звичайно; ну хто з цим сперечався би? Пані Рі розробила інструмент для оцінки вчителів [3] під назвою ІМРАСТ (або ж ВПЛИВ), і до кінця 2009–2010 навчального року округ звільнив усіх викладачів [4], чії оцінки в рейтингу становили менше 2 %. До кінця наступного року ще 5 %, або 206 вчителів [5], виперли з роботи.

Сара Висоцкі, учителька п'ятих класів [6], начебто не мала причин для хвилювання. Вона працювала в середній школі МакФарленда лише два роки й уже мала прекрасні рекомендації від директора та батьків її учнів. Одні оцінювали її [6] надзвичайну увагу до дітей; інші називали її «однією з найліпших викладачок, з якою колись спілкувалися».

Перед кінцем 2010–2011 шкільного навчального року Висоцкі отримала низькі оцінки [6] за системою перевірки від ІМРАСТ. Її проблема полягала в тому, що нова система проведення оцінювання, більше відома як модель додаткового оцінювання, яка спря-

мована на оцінку ефективності викладання математики й мовних навичок. Той показник, який отримали з алгоритму обчислювання, показував половину її загальної оцінки [6], і він переважував позитивні відгуки від шкільної адміністрації й громади. Тож округу нічого не лишилося як звільнити вчительку, яка опинилася серед 205 викладачів, чії оцінки за системою IMPACT були нижчі за мінімально допустиме значення.

Здавалося, що ця історія нагадувала радше полювання на відьом чи на зведення рахунків. Та насправді в цьому й полягала логіка підходу навчального округу. Їхні керівники, зрештою, могли бути друзями з найгіршими викладачами. Вони могли захоплюватися їхнім стилем викладання чи способом донести інформацію. Погані вчителі, здавалося, могли бути хорошими. Тож Вашингтон, як і багато інших шкільних систем, хотів мінімізувати ці людські недоліки та приділити більше уваги оцінкам, які постали на важких результатах: досягнення оцінюються в математиці та читанні. Цифри «усе чітко покажуть», обіцяли чиновники округу. Мовляв, оцінки будуть справедливими.

Звісно, Сара Висоцкі вважала ці значення дуже нечесними й хотіла дізнатися, як так вийшло. «Я не думала, що їх розумів бодай хтось», — пізніше сказала вона мені. Як міг учитель узагалі отримати такі низькі оцінки? І що це за модель додаткового оцінювання?

Зрозуміло, що вона проглянула систему, і все виявилось доволі заплутаним. Округ найняв консультантів [6] з Організації дослідження Математики, що базувалася в Принстоні для реалі-



зації програми оцінювання. Завданням цієї установи було здійснити обрахунки навчального прогресу учнів в окрузі, а потім полічити, як їхні переваги й недоліки можна приписати до їхніх викладачів. Очевидно, що це було досить нелегко. Дослідники знали, що багато змінних від учнівських соціокультурних знань минулого до впливів вивчення людей із обмеженими можливостями можуть позначитися на оцінках учнів. Алгоритми також враховували й кишенькові витрати дітей на різні потреби, що насправді теж стали однією з найбільших проблем при здійсненні аналізу.

Насправді намагання мінімалізувати чинники людської поведінки, продуктивності чи потенціалу до алгоритмів були непростим завданням. Щоби зрозуміти, з чим консультаційна організація Математика мала справу, достатньо поглянути на десятирічну дівчинку, яка мешкає в бідних районах на південному сході Вашингтона, округ Колумбія. Наприкінці одного навчального року вона склала звичайний тест для п'ятикласниці. Потім життя було ключем, як зазвичай. У неї могли бути проблеми з родиною чи з грошима. Певне, вона могла переїхати з одного будинку до іншого, чи непокоїтися через те, що в її старшого брата були проблеми з законом. Може й так, що вона була нещасливою від своєї ваги чи зазнавала цькування у школі. У будь-якому випадку наступного року вона складе наступний стандартний тест, який буде розраховано вже на шестикласників.

Якщо порівнювати результати тестів, то вони будуть достатньо очікуваними, або й зростуть. Та якщо її оцінки стануть нижчими, то буде легко полічити різницю між її навчанням і активністю успішних учнів.