Київський столичний університет імені Бориса Грінченка

Володимир Соколов, Павло Складанний

ОСНОВИ НАУКОВИХ ДОСЛІДЖЕНЬ

Навчальний посібник

для здобувачів другого (магістерського) та третього
(освітньо-наукового) рівнів вищої освіти
галузі знань F/12 «Інформаційні технології»

Київ — 2026

УДК 001.891(075.8)
ББК 72
С594

Рекомендовано до видання кафедрою інформаційної та кібернетичної безпеки імені професора Володимира Бурячка факультету інформаційних технологій та математики Київського столичного університету імені Бориса Грінченка

(протокол № 12 від 9 вересня 2026 р.)

Рецензенти:

Опірський Іван Романович, доктор технічних наук, професор, Національний університет «Львівська політехніка»;

Смірнов Олексій Анатолійович, доктор технічних наук, професор, Центральноукраїнський національний технічний університет.

Соколов В. Ю., Складанний П. М. Основи наукових досліджень : навчальний посібник для здобувачів другого (магістерського) та третього (освітньо-наукового) рівнів вищої освіти. Київ, КСУБГ, 2026, 251 с. — 10 розділів, додатки.

Навчальний посібник системно викладає методологічні, нормативні, технологічні та етичні основи наукової діяльності в галузі інформаційних технологій. Розглянуто наукове знання й методологію науки, організацію та нормативно-правове регулювання наукової діяльності й підготовки докторів філософії в Україні, логіку розгортання дослідження від проблемної ситуації до задуму й дослідницької програми, систему загальнонаукових і конкретно-наукових методів (зокрема моделювання та імітаційний експеримент, експертні й кваліметричні процедури), спеціальні дослідницькі методології інформаційних технологій та кібербезпеки (планування експерименту, контрольований експеримент, дослідження випадку, опитування, бенчмаркінг, коректне оцінювання моделей машинного навчання, етика кібербезпекових досліджень, генеративний штучний інтелект як предмет і як інструмент дослідження), технології пошуку й систематизації наукової інформації та проведення систематичного огляду літератури, наукометрію й сучасні підходи до оцінювання досліджень, статистичне опрацювання експериментальних даних, відтворюваність, керування дослідницькими даними та практики відкритої науки, структуру наукової статті за схемою IMRaD і типи наукових публікацій, академічну доброчесність і дослідницьку етику, зокрема за Законом України «Про академічну доброчесність».

Кожен розділ завершується висновками, питаннями для самоперевірки та практичними завданнями дослідницького характеру. Посібник призначено для здобувачів другого (магістерського) і третього (освітньо-наукового) рівнів вищої освіти галузі знань F/12 «Інформаційні технології», а також буде корисним науковим і науково-педагогічним працівникам, керівникам наукових робіт, гарантам програм і молодим дослідникам суміжних галузей.

Зміст

Основні терміни та скорочення

Для англомовних скорочень наведено розгорнуту назву, а в маррівських лапках — її український відповідник; у круглих дужках подано пояснення, коли назва сама по собі не розкриває змісту.

ДІ	довірчий інтервал
ДСТУ	національний стандарт України
ЄКТС	Європейська кредитна трансферно-накопичувальна система
ЄС	Європейський Союз
ЗВО	заклад вищої освіти
ІТ	інформаційні технології
МОН	Міністерство освіти і науки України
НАЗЯВО	Національне агентство із забезпечення якості вищої освіти
НБУВ	Національна бібліотека України імені В. І. Вернадського
НДР	науково-дослідна робота
НРАТ	Національний репозитарій академічних текстів
НФДУ	Національний фонд досліджень України
ОНП	освітньо-наукова програма
ШІ	штучний інтелект
ACM	Association for Computing Machinery ‘асоціація обчислювальної техніки’ (найбільше фахове товариство в галузі інформатики)
AHP	Analytic Hierarchy Process ‘метод аналізу ієрархій’
ALLEA	All European Academies ‘усі європейські академії’ (федерація європейських академій наук)
ANOVA	Analysis of Variance ‘дисперсійний аналіз’
APC	Article Processing Charge ‘плата за опрацювання статті’
AUC	Area Under the Curve ‘площа під кривою’
CC	Creative Commons ‘творчі спільноти’ (сімейство відкритих ліцензій)

CDIO	Conceive — Design — Implement — Operate ‘задумати — спроєктувати — реалізувати — застосувати’ (рамка інженерної освіти)
CoARA	Coalition for Advancing Research Assessment ‘коаліція за поступ в оцінюванні досліджень’
COPE	Committee on Publication Ethics ‘комітет з етики публікацій’
CORE	Computing Research and Education Association of Australasia ‘асоціація комп’ютерних досліджень і освіти Австралазії’ (провідний рейтинг наукових конференцій з інформатики)
CRediT	Contributor Roles Taxonomy ‘таксономія ролей учасників’
CVD	Coordinated Vulnerability Disclosure ‘координоване розкриття вразливостей’
DMP	Data Management Plan ‘план керування даними’
DOAJ	Directory of Open Access Journals ‘каталог журналів відкритого доступу’
DOI	Digital Object Identifier ‘цифровий ідентифікатор об’єкта’
DORA	San Francisco Declaration on Research Assessment ‘Сан-Франциська декларація про оцінювання досліджень’
DSR	Design Science Research ‘проєктувальне дослідження’ (парадигма здобуття знання побудовою й оцінюванням артефакту)
DSRM	Design Science Research Methodology ‘методологія проєктувального дослідження’ (шестикроковий процес за Пефферсом)
FAIR	Findable, Accessible, Interoperable, Reusable ‘знаходні, доступні, сумісні, придатні до повторного використання’ (принципи керування дослідницькими даними)
FDR	False Discovery Rate ‘частота хибних відкриттів’
FFP	Fabrication, Falsification, Plagiarism ‘фабрикація, фальсифікація, плагіат’ (три діяння грубої недоброчесності)
FWCI	Field-Weighted Citation Impact ‘нормований за галуззю вплив цитувань’
FWER	Family-Wise Error Rate ‘сімейна помилка’ (імовірність хоча б однієї помилки першого роду в сім’ї гіпотез)
GDPR	General Data Protection Regulation ‘Загальний регламент про захист даних’

ICMJE	International Committee of Medical Journal Editors ‘міжнародний комітет редакторів медичних журналів’
IEEE	Institute of Electrical and Electronics Engineers ‘інститут інженерів з електротехніки та електроніки’
IMRaD	Introduction, Methods, Results and Discussion ‘вступ, методи, результати й обговорення’ (канонічна структура емпіричної статті)
ISO	International Organization for Standardization ‘міжнародна організація зі стандартизації’
JCR	Journal Citation Reports ‘звіти про цитування журналів’
JIF	Journal Impact Factor ‘імпакт-фактор журналу’
MCC	Matthews Correlation Coefficient ‘коефіцієнт кореляції Меттьюза’
OECD	Organisation for Economic Co-operation and Development ‘Організація економічного співробітництва та розвитку’
ORCID	Open Researcher and Contributor ID ‘відкритий ідентифікатор дослідника та учасника’
OSF	Open Science Framework ‘платформа відкритої науки’
OUCI	Open Ukrainian Citation Index ‘відкритий український індекс цитувань’
PICOC	Population, Intervention, Comparison, Outcome, Context ‘сукупність, втручання, порівняння, результат, контекст’ (шаблон дослідницького питання)
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses ‘рекомендовані пункти звітування для систематичних оглядів і метааналізів’
QRP	Questionable Research Practices ‘сумнівні дослідницькі практики’
ROC	Receiver Operating Characteristic ‘робоча характеристика приймача’ (крива залежності чутливості від частки хибнопозитивних)
ROR	Research Organization Registry ‘реєстр дослідницьких організацій’
SJR	SCImago Journal Rank ‘ранг журналу за SCImago’
SNIP	Source Normalized Impact per Paper ‘нормований за джерелом вплив на статтю’
SoK	Systematization of Knowledge ‘систематизація знань’ (тип статті, що впорядковує розрізнені результати галузі)

TRL	Technology Readiness Level ‘рівень готовності технології’
UNESCO	United Nations Educational, Scientific and Cultural Organization ‘Організація Об’єднаних Націй з питань освіти, науки і культури’

Вступ

Здобувач, який приходить в аспірантуру з інформаційних технологій, зазвичай уміє багато: проєктувати системи, писати код, розгортати інфраструктуру, читати специфікації. Чого він майже напевно не вміє — це перетворювати зроблене на *наукове знання*. Різниця не в складності роботи й не в обсязі витраченого часу. Інженерний результат вважається досягнутим, коли система працює; науковий — лише тоді, коли сформульовано твердження, вказано умови, за яких воно було б хибним, показано процедуру, що дає змогу іншому дослідникові це перевірити, і чесно окреслено межі, поза якими твердження не діє. Саме цей перехід — від «я зробив систему, і вона працює» до «я встановив, що за таких-то умов виконується таке-то співвідношення, і ось як це перевірити» — і є предметом цього посібника.

Потреба в такому переході стала гострішою, ніж була ще десять років тому. Обсяг наукової продукції зростає швидше, ніж спроможність спільноти її перевіряти; частка публікацій із доступними даними й кодом в інформатиці лишається низькою; криза відтворюваності зачепила й машинне навчання, і кібербезпеку. Водночас з'явився інструмент, який здатен за секунди породити текст, зовні невідрізненний від наукового, разом із переліком джерел, яких не існує. У цих умовах методологічна підготовка перестає бути формальною дисципліною освітньої складової й перетворюється на практичну навичку самозахисту: дослідник має вміти відрізняти доказ від його імітації — у чужих роботах і насамперед у власних.

Посібник призначено для здобувачів другого (магістерського) та третього (освітньо-наукового) рівнів вищої освіти галузі знань F/12 «Інформаційні технології». Він не є ані курсом філософії науки, ані збіркою нормативних витягів, ані підручником зі статистики, хоча містить елементи всіх трьох. Його логіка — логіка *маршруту*: від з'ясування, що взагалі робить твердження науковим, до вміння подати власний результат так, щоб він витримав перевірку.

Структура посібника

Десять розділів утворюють чотири змістові блоки.

Розділи 1–2 — контекст. Перший розділ з'ясовує, чим наукове знання відрізняється від технічного й буденного, як працюють критерії демаркації, чому твердження «наш алгоритм кращий» без указаних умов спростування не є науковим, і в чому полягає специфіка інформатики з її трьома традиціями — математичною, інженерною та емпіричною. Другий розділ

описує правове й організаційне поле української науки: закони, підзаконні акти, строки, органи, процедури присудження ступеня, фінансування й права на результати. Це довідковий розділ, до якого повертаються за конкретною нормою.

Розділи 3–5 — задум і методологія. Третій розділ проводить читача шляхом від проблемної ситуації до дослідницької програми: тема, актуальність, об'єкт і предмет, мета й завдання, гіпотези, наукова новизна, календарний план і реєстр ризиків. Четвертий розділ подає систему загальнонаукових методів — емпіричних, теоретичних, модельних, експертних — із межами застосовності кожного. П'ятий розділ присвячено методологіям, специфічним саме для інформаційних технологій: design science research, контрольованому експерименту, дослідженню випадку, бенчмаркінгу, коректному оцінюванню моделей машинного навчання та особливим вимогам кібербезпекових досліджень.

Розділи 6–8 — робота з інформацією та даними. Шостий розділ навчає шукати й систематизувати наукову інформацію та виконувати систематичний огляд літератури за протоколом, із звітуванням за PRISMA 2020. Сьомий розділ розглядає наукометрію: як улаштовані показники, що вони насправді вимірюють, як ними маніпулюють і чому міжнародні рамки відповідального оцінювання вимагають не відмовлятися від метрик, а перестати підміняти ними судження. Восьмий розділ — статистичне опрацювання експериментальних даних: від шкал вимірювання й розрахунку обсягу вибірки до розмірів ефекту, поправок на множинність і коректного порівняння класифікаторів.

Розділи 9–10 — умови, за яких результат чогось вартий. Дев'ятий розділ присвячено відтворюваності, керуванню дослідницькими даними, принципам FAIR і практикам відкритої науки: як опублікувати артефакт так, щоб ним можна було скористатися; там-таки розглянуто структуру наукової статті (IMRaD та альтернативні схеми) і типи публікацій — від дослідницької статті до data paper і зареєстрованого звіту. Десятий розділ — академічна доброчесність і дослідницька етика: види порушень і процедури реагування на них, авторство, конфлікт інтересів, етика досліджень за участю людей і кібербезпекових досліджень, а також правила використання генеративного штучного інтелекту.

Як користуватися посібником

Кожен розділ побудовано однаково: вступний абзац пояснює місце розділу в загальній логіці; основний виклад супроводжується рисунками, таблицями й формальними виразами; розділ завершується висновками, питаннями для самоперевірки та *практичними завданнями.* Останні принципово відрізняються від питань: їх виконують не «на папері», а на матеріалі *вла-*

сного дослідження здобувача. Скласти матрицю «завдання — метод — результат», побудувати й оцінити рядок пошуку для своєї теми, обчислити обсяг вибірки для запланованого експерименту, скласти план керування даними, оцінити власний артефакт за FAIR — виконавши ці завдання послідовно, здобувач отримує не конспект, а робочі документи свого дослідження.

Розділи пов'язані перехресними посиланнями й не вимагають суцільного читання. Аспірантові першого року доцільно опрацювати розділи 1–3 і 6 — вони потрібні негайно; розділи 5, 8 і 9 стають актуальними з початком експериментальної роботи; розділ 2 використовують як довідник, а розділ 10 — перед кожним поданням тексту.

Застереження та домовленості

Виклад спирається на чинні станом на 2026 рік нормативні акти України, національні та міжнародні стандарти й рецензовані джерела. Нормативна база змінюється: наведені номери, дати й строки слід звіряти з офіційними текстами на момент використання. Російські та російськомовні джерела в посібнику не використовуються принципово.

Термінологія узгоджена з чинним правописом і галузевою практикою; зокрема вживається форма «безпроводовий», а не калькована «бездротовий». Англомовні терміни, що не мають усталеного українського відповідника, наводяться в дужках при першому вживанні. Посилання на джерела оформлено за ДСТУ 8302:2015; перелік скорочень подано перед основним текстом.

Автори свідомі того, що найкорисніша частина методологічної підготовки здобувається не з посібника, а в роботі з науковим керівником і в спільноті. Це видання не замінює ані того, ані того — воно лише скорочує шлях і зменшує кількість помилок, які інакше довелося б робити самостійно.

РОЗДІЛ 1

Наука, наукове знання та методологія науки

Розділ закладає понятійний каркас посібника. Його ключова теза проста: наука — це не сукупність фактів і не набір технологій, а відтворюваний спосіб здобуття й обґрунтування знання, а методологія — сукупність правил, які визначають, чи є отриманий результат науковим. Здобувач ступеня доктора філософії з інформаційних технологій щодня ухвалює рішення, методологічні за суттю: чи вважати виміряну різницю ефектом, чи достатньо порівняння з однією базовою моделлю, чи є розроблений застосунок науковим результатом. Практичне продовження цих питань — у розділах 3, 4, 5 і 9.

1.1 Наука як система знань, діяльність і соціальний інститут

1.1.1 Три проєкції поняття «наука»

Поняття «наука» має три взаємодоповнювальні проєкції, і змішування їх — джерело більшості непорозумінь у дискусіях про науковість ІТ-досліджень. **Наука як система знань** — упорядкована, внутрішньо узгоджена й обґрунтована сукупність тверджень, організована за рівнями (факти, емпіричні узагальнення, закони, теорії) і придатна до перевірки. **Наука як вид діяльності** — процес отримання нового знання з власними нормами: постановка проблеми, перевірка гіпотез, документування процедур, оприлюднення результату у формі, придатній до критики й відтворення. **Наука як соціальний інститут** — сукупність організацій, ролей, процедур і норм: наукові установи, журнали й конференції, рецензування, наукові ступені, фінансування, етичні комітети; в Україні правові засади цього інституту встановлює Закон «Про наукову і науково-технічну діяльність» [34], а якість підготовки дослідників — Закон «Про вищу освіту» [24] та система забезпечення якості на чолі з НАЗЯВО [16].

Наслідок конкретний: дисертація одночасно є текстом, що додає твердження до системи знань; звітом про діяльність, яку хтось інший має змогти повторити; і документом, який проходить інституційні процедури. Слабка робота задовольняє лише третю проєкцію.

1.1.2 Наукове, буденне й технічне знання

Буденне знання формується в повсякденному досвіді, не систематизоване, спирається на авторитет і особисте спостереження, не має явних меж застосовності. Твердження «Wi-Fi у гуртожитку повільний, бо багато ко-

ристувачів» — буденне: воно не фіксує ані способу вимірювання швидкості, ані альтернативних пояснень (інтерференція, вибір каналу, обмеження провайдера).

Технічне (інженерне) знання — це знання про те, *як зробити*, кодифіковане у стандартах і документації; воно результативне, але не зобов'язане пояснювати *чому*. Рекомендація «використовувати канал 1, 6 або 11 у діапазоні 2,4 ГГц» працює незалежно від того, чи вміє інженер вивести її з моделі спектральної маски.

Наукове знання відрізняється не предметом, а способом обґрунтування: воно фіксує умови отримання, припускає перевірку, вказує межі застосовності й має пояснювальну або предиктивну силу. Те саме твердження про канали стає науковим, щойно сформульоване як модель сукупної інтерференції з передбаченням, яке можна спростувати.

1.1.3 Критерії науковості

Жоден окремий критерій не є достатнім, але їхня сукупність утворює робочий фільтр для кожного власного твердження: *предметність* (ідеться про визначений об'єкт і предмет, а не «взагалі безпеку»); *обґрунтованість* (спирається на дані або виведення, а не на авторитет); *перевірюваність і спростовність*; *відтворюваність* (опис достатній, щоб незалежна група отримала той самий результат, — розділ 9); *системність* (узгодженість із корпусом знань або явна вказівка на конфлікт); *інтерсуб'єктивність*; *визначеність термінів* (величини мають шкали й похибки, — розділ 8); *новизна*.

> **Робоча перевірка.** Візьміть головне твердження свого дослідження й дайте відповідь: *Яких даних вистачить, щоб визнати його хибним? Хто зможе повторити перевірку? Що саме тут нового?* Немає відповіді на перше — твердження ще не наукове; на друге — воно не перевірне; на третє — це огляд, а не дослідження.

1.1.4 Структура наукового знання

Наукове знання неоднорідне: воно має рівні, між якими діють різні правила обґрунтування (рис. 1.1). Помилки початківців — це перескакування рівнів: висновок про закономірність за одним вимірюванням або декларування «нової парадигми» без емпіричного підтвердження.

1.1.5 Етос науки

Інститут науки тримається й на нормах поведінки. Р. Мертон описав їх як чотири імперативи (CUDOS): **комуналізм** (результати належать спільноті), **універсалізм** (твердження оцінюють за змістом, а не за походженням автора), **безкорисливість** і **організований скептицизм** [230]. Це не опис реальної поведінки всіх науковців, а критерій, за яким спільнота відрізняє

Рис. 1.1 — Рівні наукового знання та напрями руху між ними

прийнятне від неприйнятного; кодекс ALLEA [54] і практики COPE [95] є його операційним розгортанням (розділ 10).

1.2 Демаркація науки і псевдонауки

1.2.1 Верифікація і її межі

Логічний позитивізм першої третини XX ст. запропонував критерій **верифікованості**: осмисленими є лише твердження, які можна звести до емпірично перевірних. Він виявився водночас надто суворим — жодна скінченна кількість спостережень не верифікує висловів «усі реалізації протоколу X уразливі» — і надто слабким, бо підтвердження знаходиться майже для будь-якої теорії, якщо шукати саме підтвердження.

1.2.2 Фальсифікованість за Popper

К. Поппер змістив наголос: наукову теорію вирізняє не можливість її підтвердити, а можливість її *спростувати* [270]. Формально, позначивши через E множину логічно можливих результатів спостереження, а через $\models$ — відношення логічного наслідку:

$$\mathrm{Sci}(H) \iff \exists\, e \in E :\ H \models \neg e, \tag{1.1}$$

тобто гіпотеза H науково змістовна тоді й лише тоді, коли забороняє щонайменше один спостережуваний результат e. Чим більше результатів теорія забороняє, тим вона змістовніша й ризикованіша; твердження, сумісне з будь-яким результатом, порожнє незалежно від того, наскільки переконливо воно звучить.

Перевірка відбувається за схемою **modus tollens**: з гіпотези H і допоміжних умов C (модель вимірювання, налаштування середовища, засоби) виводять передбачення P; якщо спостереження суперечить P, спростовано кон'юнкцію:

$$\big[(H \wedge C) \to P\big],\ \neg P \ \vdash\ \neg(H \wedge C). \tag{1.2}$$

Вираз (1.2) містить застереження, відоме як проблема Дюема–Куайна: спростовано не саму гіпотезу, а її кон'юнкцію з допоміжними умовами. Тому негативний результат експерименту в ІТ завжди допускає альтернативне пояснення — хибний код вимірювання, неправильна конфігурація, непридатний набір даних, — і опис умов C у статті не є формальністю.

1.2.3 Чому це критично для ІТ-дослідника

Розглянемо типове твердження: «наш алгоритм кращий». Воно не задовольняє (1.1), бо не визначено ані показника, ані умов, ані порогу, а отже, не забороняє жодного результату вимірювань. Щоб зробити його науковим, твердження операціоналізують як перевірну конструкцію

$$\Theta = \langle A, B, D, m, \delta, \alpha \rangle, \qquad H_0 :\ \mu_m(A, D) - \mu_m(B, D) \le \delta, \tag{1.3}$$

де A — запропонований метод, B — чітко названа базова лінія (baseline), D — зафіксований набір даних або клас навантажень, m — показник із визначеною процедурою вимірювання, μ_m — очікуване значення показника, δ — практично значуща різниця, α — рівень значущості. Тепер твердження забороняє конкретний результат: якщо на D різниця не перевищує δ, його спростовано. Будь-який неназваний компонент перетворює результат на нефальсифіковне гасло. Найпоширеніші претензії рецензентів до ІТ-статей стосуються саме невизначених B, D і m: порівняння з «солом'яним опудалом», підбір набору даних під метод і зміна метрики після перегляду результатів.

> **Нефальсифіковні формулювання та їх виправлення.** «Метод підвищує рівень безпеки системи» $\to$ «частка виявлених атак класу X на наборі Y зростає щонайменше на 5 в. п. за незмінної частки хибних спрацьовувань». «Архітектура є масштабованою» $\to$ «затримка відгуку зростає не більш ніж лінійно до 10 000 одночасних сесій на стенді зі специфікацією Z». «Модель краще враховує контекст» $\to$ «на вибірці D значення F_1 перевищує базову модель B на $\delta \ge 0{,}02$ за $\alpha = 0{,}05$».

1.2.4 Науково-дослідницькі програми за Lakatos

Учені не відмовляються від теорії після першого спростування — і це раціонально. І. Лакатош запропонував за одиницю аналізу брати не окрему теорію, а **науково-дослідницьку програму**: *тверде ядро* (базові припущення, які не переглядають), *захисний пояс* допоміжних гіпотез і *евристики* —

негативну й позитивну [211]. Програма *прогресивна*, якщо її модифікації передбачають нові факти, які підтверджуються, і *регресивна*, якщо вони лише пояснюють відомі невдачі заднім числом.

Для ІТ-дослідника це найпрактичніша з трьох концепцій. Дисертація майже завжди розгортається всередині чинної програми: «виявлення вторгнень на основі машинного навчання», «формальна верифікація протоколів», «конфіденційні обчислення». Тверде ядро (наприклад, припущення, що аномалія в трафіку статистично відрізняється від норми) не перевіряють — його приймають; перевіряють захисний пояс: вибір ознак, архітектуру моделі, спосіб нормалізації. Корисне самозапитання: *мої модифікації передбачають щось нове чи лише пояснюють, чому попередній результат не відтворився?* Друга відповідь — ознака регресивного зсуву.

1.2.5 Парадигми й наукові революції за Kuhn

Т. Кун описав розвиток науки як чергування періодів **нормальної науки** — розв'язування «головоломок» у межах спільної **парадигми** (зразків розв'язання задач, методів, критеріїв оцінки) — і **наукових революцій**, коли накопичення аномалій змінює парадигму [210]. Парадигма визначає не лише відповіді, а й перелік запитань, які взагалі вважають осмисленими.

В інформаційних технологіях зміна парадигм відбувається швидше, ніж у природничих науках: від експертних систем на правилах — до статистичного навчання; від сигнатурного виявлення шкідливого коду — до поведінкового; від периметрової моделі захисту — до zero trust. Практичний наслідок подвійний. «Очевидні» критерії якості насправді є конвенцією спільноти: показник F_1 на певному наборі даних вважають доречним не тому, що це доведено, а тому, що так прийнято — і рецензент оцінюватиме роботу саме за цією конвенцією. Водночас робота, що суперечить панівній парадигмі, потребує значно сильнішої доказової бази.

> **Три концепції разом.** Popper дає критерій *формулювання* гіпотези: вона має забороняти результат. Lakatos — критерій *вибору напряму*: програма має бути прогресивною. Kuhn — критерій *комунікації*: результат оцінюватимуть за конвенціями конкретної спільноти, які треба знати заздалегідь.

1.3 Емпіризм, раціоналізм і логіка наукового виведення

1.3.1 Дві традиції обґрунтування

Емпіризм виводить джерело знання з досвіду: надійним є те, що спостережено й виміряно. **Раціоналізм** наголошує на дедуктивному виведенні з очевидних засновків: надійним є те, що доведено. В інформатиці обидві традиції присутні буквально: доведення коректності алгоритму й оцінка

складності належать раціоналістичній лінії, вимірювання продуктивності реалізації — емпіричній. Зрілий результат поєднує обидві: доведена властивість плюс експериментальна перевірка того, що припущення доведення виконуються на практиці.

1.3.2 Індукція, дедукція, абдукція

Дедукція — виведення, яке зберігає істинність: якщо засновки істинні, висновок істинний обов'язково. **Індукція** — перехід від окремих випадків до загального твердження; вона не зберігає істинності й дає лише правдоподібний висновок. **Абдукція** — виведення до найкращого пояснення: спостерігаючи факт E, обирають гіпотезу, яка пояснює його найкраще [121]. У ймовірнісному формулюванні

$$\hat{H} = \arg\max_{H \in \mathcal{H}} P(H \mid E) = \arg\max_{H \in \mathcal{H}} P(E \mid H)\, P(H), \tag{1.4}$$

де $\mathcal{H}$ — *явно окреслена* множина конкурентних гіпотез. Вираз (1.4) вказує на найчастішу помилку ІТ-досліджень: висновок «найкраще пояснення» коректний лише щодо розглянутих альтернатив, тож якщо в $\mathcal{H}$ внесено єдину гіпотезу — ту, що подобається дослідникові, — абдукція вироджується в самопідтвердження. Тому в розділі «Обговорення» альтернативні пояснення ефекту перелічують явно й показують, чому їх відкинуто.

Типовий ланцюжок в ІТ: точність моделі на новому наборі даних різко впала (факт); можливі пояснення — зсув розподілу, витік даних, помилка попереднього оброблення, інша версія бібліотеки ($\mathcal{H}$, абдукція); якщо причиною є витік, то після коректного розділення вибірок падіння відтвориться й на початковому наборі (P, дедукція); перевірка передбачення (експеримент); відтворення ефекту на кількох наборах даних дає емпіричне узагальнення (індукція).

1.3.3 Гіпотетико-дедуктивна модель

Наведений ланцюжок є окремим випадком **гіпотетико-дедуктивної моделі** — базової схеми емпіричного дослідження (рис. 1.2). Принципово, що рішення в ній асиметричне: розбіжність спростовує (з урахуванням застереження Дюема–Куайна), а збіг лише *підкріплює* гіпотезу, не доводячи її. Формулювання «експеримент довів гіпотезу» в дисертації є методологічною помилкою; коректно — «результати узгоджуються з гіпотезою» або «гіпотезу не спростовано».

Не всі дослідження в ІТ вкладаються в цю схему: побудова артефакту, формальне доведення, систематичний огляд, розвідувальний аналіз даних мають власну логіку (розділи 5, 6). Проте кожне з них у якийсь момент потребує перевірного кроку — і тоді схема стає обов'язковою.

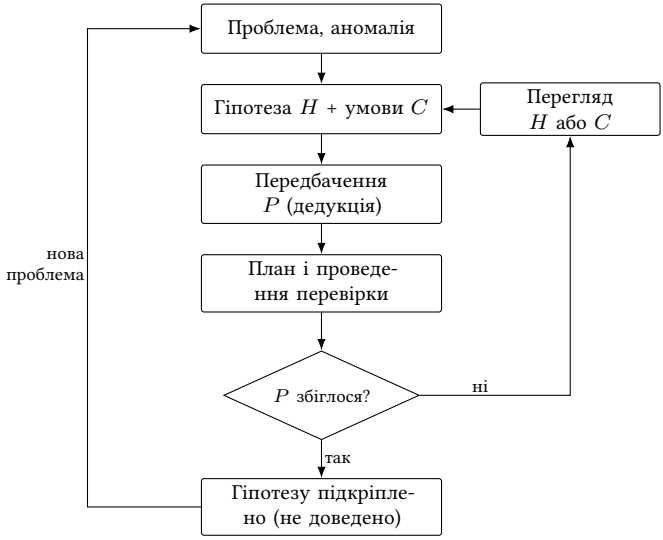


Рис. 1.2 — Цикл гіпотетико-дедуктивного методу

1.4 Класифікація наук і місце інформаційних технологій

1.4.1 Міжнародна класифікація OECD

Міжнародною рамкою для статистики досліджень і розробок є Frascati Manual OECD [249] з класифікацією галузей науки й технологій (Fields of Research and Development, FOS) [250]: шість галузей першого рівня, деталізованих до другого (табл. 1.1). Для галузі знань «Інформаційні технології» принципові дві позиції — `1.2 Computer and information sciences` у складі природничих наук і `2.2 Electrical, electronic, information engineering` у складі технічних.

Розділ між 1.2 і 2.2 не є формальністю: дослідження в межах 1.2 оцінюють переважно за теоретичною строгістю, у межах 2.2 — за характеристиками технічного рішення та відтворюваністю вимірювань. Кібербезпекова робота часто лежить на межі: модель порушника й доведення стійкості — це 1.2, вимірювання реальних мереж і розробка засобу — 2.2. Явне самовизначення допомагає обрати журнал і критерії якості (розділ 7).

1.4.2 Український класифікатор галузей знань

Українська система донедавна користувалася переліком, затвердженим постановою Кабінету Міністрів України від 29.04.2015 № 266, де галузь 12 «Інформаційні технології» охоплювала спеціальності 121–126 [27]. Постановою від 30.08.2024 № 1021 перелік наближено до Міжнародної стандартної

Таблиця 1.1 — Класифікація галузей досліджень OECD FOS і позиція IT (за [250, 249])

Код	Галузь	Релевантні для IT підгалузі
1	Природничі науки	1.1 математика; **1.2 computer and information sciences**; 1.3 фізика
2	Технічні науки та технології	2.1 цивільна інженерія; **2.2 електрична, електронна, інформаційна інженерія**; 2.3 механічна інженерія; 2.11 інші
3	Медичні науки та науки про здоров'я	3.3 науки про здоров'я (медична інформатика)
4	Сільськогосподарські науки	4.1 агротехнології (сенсорні мережі)
5	Суспільні науки	5.2 економіка; 5.8 медіа й комунікації
6	Гуманітарні науки	6.2 мови (обробка природної мови)

класифікації освіти: кількість галузей скорочено, а галузь IT дістала літерний код F [25]. Відповідність основних позицій така:

F1 — Прикладна математика (колишня 113); F2 — Інженерія програмного забезпечення (121); F3 — Комп'ютерні науки (122); F4 — Системний аналіз та наука про дані (124); F5 — Кібербезпека та захист інформації (125); F6 — Інформаційні системи і технології (126); F7 — Комп'ютерна інженерія (123).

Позначення «F/12» у посібнику вживається як спільне для чинного й попереднього кодів, оскільки на перехідний період обидва трапляються в документах. Практичне значення класифікатора прямолінійне: шифр спеціальності визначає склад разової спеціалізованої вченої ради, вимоги до публікацій і формулювання наукової новизни (розділ 2). Аналіз стратегій побудови освітніх програм другого й третього рівнів зі спеціальності 125 показує, що методологічна складова підготовки є найбільш варіативним елементом таких програм [44].

1.5 Типи досліджень і рівні готовності технологій

1.5.1 Фундаментальні, прикладні, експериментальні розробки

Frascati Manual визначає три типи діяльності у сфері досліджень і розробок за критерієм мети, а не за складністю [249]; українське законодавство використовує близьку термінологію [34]. Порівняння — у табл. 1.2.

Межі рухомі. Помилка — не в тому, щоб виконувати розробку, а в тому, щоб видавати її за фундаментальне дослідження. Дисертація в галузі F/12 здебільшого є прикладним дослідженням з елементом експериментальної розробки; тоді наукова новизна полягає в методі або моделі, а прототип є засобом її перевірки, а не самоціллю.

Таблиця 1.2 — Типи діяльності у сфері досліджень і розробок (за [249, 34])

Ознака	Фундаментальні	Прикладні	Експериментальні розробки
Мета	нове знання без конкретного застосування	нове знання для визначеної практичної мети	створення нових або вдосконалення наявних продуктів і процесів
Результат	теорія, модель, теорема, закономірність	метод, алгоритм, обґрунтована технологія	прототип, засіб, дослідний зразок, документація
Горизонт	10 і більше років	3–10 років	1–3 роки
Приклад в ІТ	межі обчислюваності, нижні оцінки складності	метод виявлення аномалій для мереж IoT	реалізація й випробування системи моніторингу
Фінансування	НФДУ, ERC	НФДУ, МОН, Horizon Europe	гранти інноваційного циклу, замовники

1.5.2 Рівні готовності технологій

Шкала **рівнів готовності технологій** (Technology Readiness Levels, TRL) описує зрілість технічного рішення на шляху від спостереження принципу до промислової експлуатації. Її визначено стандартом ISO 16290:2013 [185] і закріплено в Загальних додатках до робочих програм Horizon Europe [142, 280]; формулювання для ІТ — у табл. 1.3.

Для дисертації типовий діапазон — TRL 2–5; заявляти TRL 7–9 без актів упровадження й даних тривалої експлуатації не варто. Шкала корисна як інструмент планування: вона робить очевидним, який саме доказ потрібен для наступного кроку, і не дає підмінити перехід 4→5 демонстрацією на тому самому синтетичному наборі даних. У заявках Horizon Europe цільовий TRL є формальною вимогою конкурсу [280], а невідповідність йому — поширена причина відхилення.

1.6 Рівні методології: метод, методика, методологія

1.6.1 Розмежування понять

Три близькі терміни позначають різні речі. **Метод** — спосіб досягнення мети, впорядкована сукупність пізнавальних дій із визначеними межами застосовності (експеримент, моделювання, формалізація). **Методика** — конкретизація методу для певної задачі: послідовність операцій, інструменти, параметри, форми фіксації результату; методика вимірювання завантаженості каналу — це метод вимірювання плюс конкретний аналізатор, тривалість вибірки, крок дискретизації, спосіб усереднення. **Методологія** — учення про методи: система принципів, яка обґрунтовує *вибір* методів і критерії оцінювання знання.

Таблиця 1.3 — Рівні готовності технологій TRL та їх інтерпретація для IT-досліджень (за [142, 185])

TRL	Визначення	Типовий доказ в IT
1	спостережено базові принципи	огляд, постановка задачі
2	сформульовано концепцію технології	модель, схема методу, оцінка складності
3	експериментальне підтвердження концепції	прототип функції, синтетичні дані
4	перевірка компонента в лабораторії	стенд, бенчмарк на відкритому наборі
5	перевірка компонента в релевантному середовищі	тестова мережа з реальним трафіком
6	демонстрація в релевантному середовищі	пілот у сегменті інфраструктури
7	прототип в операційному середовищі	дослідна експлуатація в діючій системі
8	система завершена й кваліфікована	сертифікація, аудит
9	система перевірена в операційному середовищі	промислова експлуатація

Найчастіша помилка — підрозділ «Методологія», у якому переліковано методи без обґрунтування, чому обрано саме їх. Коректний виклад відповідає на три питання: *яку властивість об'єкта має виявити метод*, *чому саме цей метод здатен її виявити* і *за яких умов його висновки недійсні*.

1.6.2 Чотири рівні методологічного знання

Методологічне знання розглядають за чотирма рівнями (табл. 1.4). Рівні не конкурують: у будь-якому дослідженні присутні всі чотири.

Проблеми виникають, коли рівні розриваються. Робота, сильна на технологічному рівні й порожня на конкретно-науковому, залишає враження технічного звіту; зворотна ситуація — декларації про «системний підхід» без жодної методики — дає текст, який неможливо відтворити. Конкретно-науковий рівень для галузі F/12 розгорнуто в розділі 5, загальнонауковий — у розділі 4.

1.7 Специфіка інформатики як науки

1.7.1 Три традиції

Звіт робочої групи ACM «Computing as a Discipline» зафіксував, що обчислювальна дисципліна історично формувалася трьома паралельними традиціями [111]:

Таблиця 1.4 — Рівні методології та їх вияв у дослідженні з інформаційних технологій

Рівень	Зміст	Вияв у роботі з ІТ
Філософський	загальні принципи пізнання, критерії науковості, демаркація	чому артефакт є науковим результатом; що вважати поясненням
Загальнонауковий	підходи й методи, спільні для багатьох наук: системний підхід, моделювання, експеримент	системний аналіз об'єкта, побудова моделі, планування експерименту
Конкретно-науковий	методи, властиві галузі	Design Science Research, контрольований експеримент у програмній інженерії, кейс-стаді, моделювання порушника
Технологічний	методики, процедури, інструменти, техніка отримання даних	протокол вимірювань, скрипти обробки, налаштування стенда, ведення журналу

Математична (теорія).
Об'єкти визначають аксіоматично, властивості доводять. Результат — теорема, оцінка складності, доведення коректності протоколу; критерій якості — строгість доведення.

Інженерна (проєктування).
Будують артефакт, що задовольняє вимоги за наявних обмежень. Результат — архітектура, реалізація, специфікація; критерій — відповідність вимогам і характеристики рішення.

Емпірична (абстрагування, моделювання).
Систему досліджують як явище: висувають гіпотези про її поведінку, вимірюють, будують моделі. Результат — емпірична закономірність або перевірена модель; критерій — план експерименту й відтворюваність.

А. Ньюелл, А. Перліс і Г. Саймон наполягали, що інформатика є *емпіричною* наукою в буквальному сенсі: «машина — не просто апаратура, а запрограмована, жива машина — є тим організмом, який ми вивчаємо» [244]; згодом А. Ньюелл і Г. Саймон розгорнули цю тезу в Тюрінгівській лекції [245]. П. Деннінг показав, що суперечка «чи є computer science наукою» зводиться до підміни: критики порівнюють дисципліну лише з математичною традицією, ігноруючи емпіричну [110]. Для здобувача тріада є діагностичним інструментом: перш ніж обирати методи, слід визначити, у якій традиції лежить очікуваний результат. Змішування традицій породжує найчастіший дефект ІТ-дисертацій — доведена теоретична властивість підмінює відсутній експеримент або вдалі вимірювання видають за доведення загальної властивості методу.

1.7.2 Артефакт як науковий результат

Особливість галузі F/12 полягає в тому, що результатом часто є не твердження, а **артефакт** — конструкт, модель, метод або реалізація. Таке дослідження є науковим лише тоді, коли дає не сам артефакт, а *обґрунтоване знання про нього*: за яких умов він працює, які властивості забезпечує, чим перевершує альтернативи, де його межі. Парадигму систематизовано в межах Design Science Research [170, 262], її процедурне розгортання — у розділі 5. Методологічний мінімум: артефакт без оцінювання — це розробка, з оцінюванням за визначеними критеріями та базовими лініями — дослідження.

1.7.3 Криза відтворюваності в IT

В. Тихі ще 1998 р. показав, що частка статей в інформатиці з бодай елементарною експериментальною перевіркою істотно нижча, ніж в інших дисциплінах, і систематизував шістнадцять типових виправдань цього [322]. Через два десятиліття К. Колберг і Т. Пребстінг спробували відтворити результати статей із провідних видань із комп'ютерних систем: код виявився доступним і працездатним лише для меншої частини робіт [93]. Аналіз досліджень зі шкідливого програмного забезпечення виявив систематичні проблеми з набором даних і майже повну відсутність опису запобіжних заходів під час експериментів [284]. Відповіддю стали контрольні списки й програми відтворюваності конференцій з машинного навчання [269], а в ACM — система бейджів артефактів [52].

Висновок: в IT низький бар'єр відтворення поєднується з високою чутливістю результату до дрібниць (версія бібліотеки, початкове значення генератора випадкових чисел, розділення вибірок), тож саме тут вимоги до опису умов C у (1.2) мають бути найсуворішими. Практики керування даними й публікації артефактів — у розділі 9.

1.7.4 Кейс: від інженерної розробки до наукової дисципліни

Еволюція досліджень безпеки безпроводових мереж показує, як технічна галузь набуває рис наукової дисципліни. Умовно можна виокремити чотири стадії.

Стадія 1 — інженерна розробка (1997–2000). Стандарт IEEE 802.11 містив механізм WEP, спроєктований інженерно: обрано потоковий шифр, контрольну суму й схему керування ключами, які здавалися достатніми. Обґрунтування було конструктивним («ми забезпечили шифрування»), а не науковим: не визначено ані моделі порушника, ані властивостей, які система має гарантувати. Твердження «WEP забезпечує конфіденційність» не задовольняло критерію (1.1).

Стадія 2 — поява моделі й перших спростувань (2001). Н. Борисов, І. Ґолдберґ і Д. Ваґнер сформулювали явну модель порушника й показали, що

WEP не досягає жодної з чотирьох заявлених цілей: конфіденційності, контролю доступу, цілісності й захисту від відтворення [71]. Того самого року С. Флурер, І. Мантін і А. Шамір описали слабкості розкладу ключів RC4, що уможливлюють пасивне відновлення ключа [151]. Ці роботи ввели формалізовану модель загроз, чітко названі властивості й відтворюваний експеримент.

Стадія 3 — нормальна наука в новій парадигмі (2004–2020). Парадигмою стало проєктування та аналіз протоколів у явній моделі порушника, а дослідження набули характеру «головоломок» за Куном: аналіз чотиристороннього рукостискання WPA2 виявив можливість повторного встановлення ключа (атака KRACK) [331], аналіз рукостискання Dragonfly у WPA3 — побічноканальні витоки й вразливість до зниження версії [332]. Кожна з цих робіт містила всі елементи наукового результату: гіпотезу, передбачення, експериментальну перевірку на конкретних реалізаціях, відкриті інструменти та відповідальне розкриття.

Стадія 4 — самостійна дослідницька програма (з 2020). Сформовано тверде ядро (безпека протоколу доводиться у визначеній моделі), захисний пояс (припущення про можливості порушника, канал, реалізацію) і позитивну евристику: переносити аналіз на керуючі кадри, механізми енергозбереження, фрагментацію. Українська складова програми представлена дослідженнями захисту обміну даними в безпроводових мобільних мережах з автентифікацією й протоколами обміну ключами [12] у фаховому виданні «Кібербезпека: освіта, наука, техніка» [10].

> **Що з кейсу випливає для власної роботи.** Перехід від розробки до дослідження стався не тоді, коли з'явилися складніші алгоритми, а тоді, коли з'явилися *явна модель порушника* й *явно названі властивості*, які можна спростувати. Доки немає моделі та властивостей, є розробка; щойно вони з'явилися — є дослідження.

1.8 Наукова школа й самоорганізація науки

1.8.1 Наукова школа

Наукова школа — неформальне об'єднання дослідників навколо спільної дослідницької програми, лідера й способу роботи. Її ознаки: спільне тверде ядро припущень; відтворюваний спосіб підготовки молодих дослідників; наскрізна тематика публікацій протягом принаймні двох поколінь; власний інструментарій і набори даних; визнання зовнішньої спільноти. Школа не тотожна кафедрі: організаційна одиниця може не мати школи, а школа — охоплювати кілька установ і країн.

Вибір наукової школи за наслідками поступається хіба що вибору теми. Школа дає готовий захисний пояс методик, доступ до даних і стенда, ме-

режу рецензентів, а головне — усталені критерії якості, за якими робота оцінюватиметься. Ризик симетричний: усередині школи легко перейняти регресивний зсув і не помітити, що програма вичерпалася; протиотрута — читання робіт конкурентних груп і участь у конференціях за межами власного кола [43].

1.8.2 Самоорганізація науки

Наука не має центрального органу, який визначав би істинність тверджень; її впорядкованість забезпечують механізми самоорганізації: *рецензування* — попередній фільтр якості, реалізований самими дослідниками; *цитування* — розподілений механізм визнання й водночас джерело метрик, які деформують поведінку, щойно стають метою (розділ 7); *відтворення й спростування* — механізм виправлення помилок, дієвий лише за умови доступності артефактів [52]; *відкрита наука* — інституціалізація доступності даних, коду й публікацій [327, 346] (розділ 9); *кодекси доброчесності* — ALLEA [54], COPE [95] (розділ 10); *конкурсне фінансування* — НФДУ [19], Horizon Europe [280], що спрямовує ресурси через експертне оцінювання, а не адміністративний розподіл.

Усі ці механізми спираються на норми етосу науки: універсалізм робить можливим сліпе рецензування, комуналізм — обмін даними, організований скептицизм — відтворення. Там, де норми порушуються (купівля авторства, хижацькі видання, фальсифікація даних), механізми самоорганізації не просто дають збій, а починають працювати проти науки, надаючи недоброчесним результатам зовнішні ознаки визнання.

1.9 Висновки до розділу

Наука розглядається у трьох проєкціях — як система знань, вид діяльності й соціальний інститут; дисертація має задовольняти всі три. Науковість твердження визначається не предметом і не складністю апарату, а способом обґрунтування.

Ключовим механізмом демаркації є фальсифікованість (1.1): наукове твердження забороняє щонайменше один спостережуваний результат. Для ІТ-дослідника це означає обов'язкову операціоналізацію претензії на перевагу (1.3) через базову лінію, набір даних, показник і практично значущу різницю. Концепція науково-дослідницьких програм Лакатоша дає критерій вибору напряму (прогресивний чи регресивний зсув), а концепція парадигм Куна пояснює, чому критерії оцінювання є конвенціями спільноти й чому їх слід вивчати до початку роботи, а не після рецензування. Логіку дослідження утворює поєднання абдукції, дедукції та індукції в гіпотетико-дедуктивний цикл; його асиметрія — спростування остаточне, підтвердження лише тимчасове — визначає коректні формулювання висновків.

Класифікації OECD FOS і українського переліку галузей знань F/12 задають систему координат для позиціювання роботи, типологія «фундаментальне — прикладне — розробка» й шкала TRL 1–9 — для планування результату та добору доказів. Чотири рівні методології розмежовують філософські засади, загальнонаукові підходи, галузеві методології та конкретні методики; розрив між рівнями є найчастішою причиною того, що технічно бездоганна робота не сприймається як наукова.

Інформатика поєднує математичну, інженерну й емпіричну традиції, і результатом дослідження тут часто є артефакт; науковим його робить оцінювання за визначеними критеріями, а не факт створення. Криза відтворюваності робить точний опис умов експерименту й публікацію артефактів умовою науковості. Кейс еволюції досліджень безпеки безпроводових мереж показує момент переходу від інженерної розробки до наукової дисципліни: поява явної моделі порушника й явно названих властивостей, які можна спростувати.

Питання для самоперевірки

1. Назвіть три проєкції поняття «наука» і поясніть, якій із них відповідає кожен обов'язковий елемент дисертації.
2. Чим наукове знання відрізняється від технічного? Наведіть приклад твердження з власної предметної області в обох формах.
3. Чому критерій верифікованості виявився непридатним для універсальних тверджень? Проілюструйте прикладом з ІТ.
4. Поясніть зміст виразу (1.1). Які результати вимірювань забороняє ваша гіпотеза?
5. У чому полягає проблема Дюема–Куайна і які наслідки вона має для опису експериментального стенда у статті?
6. Чому твердження «наш алгоритм кращий» не є науковим? Які шість компонентів із (1.3) треба визначити, щоб воно стало таким?
7. Що таке тверде ядро й захисний пояс науково-дослідницької програми? За якими ознаками можна розпізнати її регресивний зсув?
8. Наведіть приклад зміни парадигми в інформаційних технологіях за останні 15 років і поясніть, як вона вплинула на критерії оцінювання результатів.
9. Порівняйте індукцію, дедукцію й абдукцію за збереженням істинності та за роллю у вашому дослідженні.
10. Чому висновок «експеримент довів гіпотезу» є методологічною помилкою? Запропонуйте коректне формулювання.

11. Де в класифікації OECD FOS розташована ваша тема і чим експериментальна розробка відрізняється від прикладного дослідження за Frascati Manual?
12. Визначте цільовий рівень TRL вашого результату й назвіть доказ, потрібний для переходу на наступний рівень.
13. Розмежуйте поняття методу, методики й методології на прикладі одного з методів, які ви застосовуєте.
14. Аналітичне: які з трьох традицій інформатики поєднує ваша робота і чи не підмінюється в ній експеримент доведенням (або навпаки)?
15. Дослідницьке: спираючись на кейс безпроводових мереж, визначте, на якій із чотирьох стадій перебуває ваша предметна підгалузь, і обґрунтуйте відповідь ознаками, а не враженням.

Практичні завдання

1. *Фальсифікація власної гіпотези.* Сформулюйте головну гіпотезу дослідження у вигляді $\Theta = \langle A, B, D, m, \delta, \alpha \rangle$ за (1.3). Окремо випишіть перелік результатів вимірювання, які її спростовують. Якщо перелік порожній — переформулюйте гіпотезу.
2. *Аудит нефальсифіковних формулювань.* Візьміть три статті за темою дослідження з фахових видань категорії «Б» або зі Scopus. Випишіть із анотацій по два твердження про перевагу запропонованого рішення й класифікуйте кожне: фальсифіковне / нефальсифіковне. Для нефальсифіковних запропонуйте операціоналізацію.
3. *Карта дослідницької програми.* Побудуйте схему програми, у межах якої працюєте: тверде ядро (3–5 припущень), захисний пояс, позитивна евристика. Визначте, які з очікуваних результатів належать до ядра, а які — до поясу.
4. *Позиціювання роботи.* Складіть таблицю з чотирьох рядків: код OECD FOS; шифр спеціальності за [25] і попередній за [27]; тип дослідження за Frascati Manual; поточний і цільовий TRL. Для кожного рядка наведіть обґрунтування.
5. *Чотири рівні методології.* Для власного дослідження заповніть табл. 1.4 конкретним змістом: філософський рівень (що вважаєте науковим результатом), загальнонауковий, конкретно-науковий і технологічний. Позначте рівень, який наразі найслабший.

РОЗДІЛ 2

Організація та нормативно-правове регулювання наукової діяльності й підготовки докторів філософії в Україні

Розділ 1 розглянув науку як систему знань, діяльність і соціальний інститут. Цей розділ присвячено третій проєкції: правовому й організаційному полю, у якому працює дослідник в Україні. Теза прагматична — знання цього поля економить роки. Здобувач, який знає, що індивідуальний план наукової роботи затверджується протягом двох місяців з дня зарахування, а стаття з чотирма співавторами у фаховому виданні зараховується як пів публікації, планує роботу інакше, ніж той, хто дізнається про це за півроку до завершення навчання. Розділ побудовано як довідник — із номерами актів, строками й назвами документів; методологічне наповнення процедур — у розділі 3, вимоги доброчесності до кваліфікаційної роботи — у розділі 10, зразки бібліографічних описів — у додатку В.

2.1 Законодавчі засади наукової діяльності

2.1.1 Два базові закони

Закон «Про наукову і науково-технічну діяльність» від 26.11.2015 № 848-VIII [34] регулює статус ученого й наукової установи, наукові ступені, державні гарантії та фінансування науки. Закон «Про вищу освіту» від 01.07.2014 № 1556-VII [24] визначає рівні вищої освіти, зокрема третій (освітньо-науковий), і систему забезпечення якості. Закон «Про освіту» від 05.09.2017 № 2145-VIII [37] додає наскрізні норми про академічну доброчесність (розділ 10).

Розмежування не формальне: ступінь доктора філософії присуджується за Законом «Про вищу освіту», а ступінь доктора наук — за Законом «Про наукову і науково-технічну діяльність» (ст. 28) [34, 24]. Звідси — дві різні процедури, два типи рад і два підзаконні акти (підрозділ 2.4).

2.1.2 Суб'єкти, ступені, гарантії

Стаття 4 Закону № 848-VIII зараховує до **суб'єктів наукової і науково-технічної діяльності** наукових і науково-педагогічних працівників, *аспірантів, ад'юнктів і докторантів*, інших учених, наукові установи, університети, академії, інститути та громадські наукові організації [34]. Наслідок прямий: аспірант є повноцінним суб'єктом, а не лише «особою, що навчається», — він може самостійно подавати заявки на гранти й наукові стипен-

дії [29], а також повідомлення про невідповідність складу разової ради чи про порушення процедури захисту [30].

Стаття 28 встановлює вичерпний перелік: наукові ступені — **доктор філософії** і **доктор наук**; учені звання — **старший дослідник**, **доцент**, **професор** [34]. Ступінь кандидата наук більше не присуджується: підготовка, розпочата до 1 вересня 2016 р., завершується за попереднім законодавством [29]. Ступінь доктора філософії присуджує *сам заклад* через разову раду [30], диплом доктора наук видає МОН на підставі рішення атестаційної колегії; ступінь доктора наук здобувається не раніше ніж через п'ять років після ступеня доктора філософії (кандидата наук), крім випадків вагомого особистого внеску, підтвердженого трьома патентами на винахід, трьома публікаціями у виданнях квартилів Q1–Q2 або державними нагородами [6].

Державні гарантії, якими здобувачі користуються рідко лише через необізнаність: наукове відрядження (ст. 33) і стажування, зокрема закордонне (ст. 34); зарахування часу підготовки в аспірантурі до стажу наукової роботи (ст. 35); захищеність бюджетних видатків на науку й норма про її фінансування на рівні не менш як 1,7 % валового внутрішнього продукту (ст. 48); права на об'єкти інтелектуальної власності (ст. 64) [34]. Порядок № 261 прямо забороняє покладати на здобувача обов'язки, не пов'язані з виконанням освітньо-наукової програми: залучення аспіранта до чергувань чи профорієнтаційної роботи є порушенням і типовим предметом опитувань експертних груп під час акредитації [29, 28].

2.1.3 Фінансування й державна атестація

Стаття 48 розрізняє два канали [34]. **Базове фінансування** підтримує основну діяльність державних наукових установ і дослідження закладів вищої освіти; **конкурсне (грантове)** розподіляється через Національний фонд досліджень України та конкурси МОН (підрозділ 2.5). Базове фінансування прив'язане до державної атестації: за ст. 11 вона проводиться для наукових установ державної й комунальної форм власності не рідше ніж раз на п'ять років (для новоутворених — не пізніш як через три роки після утворення) за кадровим забезпеченням, станом матеріально-технічної бази та якістю наукової діяльності на основі експертної оцінки з використанням наукометричних показників [34]. Процедуру деталізує Порядок, затверджений постановою Кабінету Міністрів України від 19.07.2017 № 540, чинний і для закладів вищої освіти в частині їх наукової діяльності [31], а механізм розподілу коштів — постанова від 28.01.2026 № 116: обсяг залежить від групи закладу (коефіцієнт 1 для групи А, 0,7 для групи Б) і середньорічної чисельності наукових та науково-педагогічних працівників [23].

2.2 Інституційна система організації науки

Система має чотири функційні контури: *політика й бюджет* (Верховна Рада, Кабінет Міністрів, МОН), *експертна координація* (Національна рада України з питань розвитку науки і технологій), *забезпечення якості й атестація* (НАЗЯВО, МОН), *конкурсне фінансування* (НФДУ). Виконавці — заклади вищої освіти, наукові установи, НАН України та національні галузеві академії (рис. 2.1).

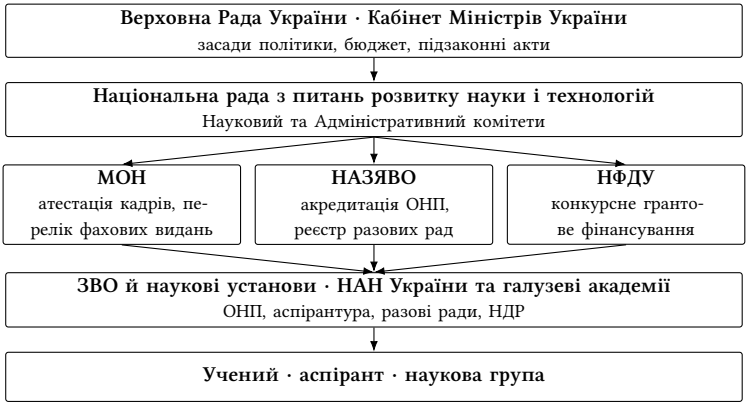


Рис. 2.1 — Інституційна система організації наукової діяльності в Україні

Таблиця 2.1 зводить повноваження до формату «орган — повноваження — правова підстава» і слугує маршрутною картою.

Таблиця 2.1 — Органи системи організації науки: повноваження та правова підстава

Орган	Ключові повноваження	Підстава
Кабінет Міністрів	порядки підготовки й атестації здобувачів, перелік галузей знань	ст. 41 № 848-VIII; [29, 30, 6]
МОН	вимоги до дисертацій і публікацій, перелік фахових видань, докторські ради, перевірка складу разових рад, позбавлення ступенів	ст. 42 № 848-VIII; [6, 32, 36]
Національна рада	засади політики, пріоритети, інтеграція до Європейського дослідницького простору	ст. 20–21 № 848-VIII; [40]
НАЗЯВО	акредитація ОНП, інформаційна система разових рад, скасування їх рішень	ст. 18, 25 № 1556-VII; [30, 28]
НФДУ	індивідуальні, колективні та інституційні гранти; експертиза заявок	ст. 49–55 № 848-VIII; [35, 19]
ЗВО, установи, академії	освітня діяльність на третьому рівні, аспірантура, індивідуальні плани, разові ради, диплом	ст. 17–19 № 848-VIII; [29, 30]

Найчастіша помилка — звернення «не за адресою». МОН не присуджує ступінь доктора філософії й не скасовує рішення разової ради після видачі диплома: перше робить заклад, друге — НАЗЯВО протягом шести місяців з дня наказу про видачу диплома [30, 16]. Натомість саме МОН перевіряє склад ради на відповідність пунктам 14–16 Порядку № 44 протягом місяця з дня оприлюднення інформації про її утворення і може зупинити роботу ради [30]. Перелік фахових видань веде МОН, а не НАЗЯВО [32].

Правило трьох питань. Перед будь-яким зверненням дайте відповідь: *який орган має це повноваження; у якій нормі воно записане; який строк розгляду встановлено.* Без відповіді на друге питання з номером акта й пункту звернення буде переадресоване.

2.3 Третій рівень освіти та підготовка доктора філософії

2.3.1 Рівень і програма

Стаття 5 Закону № 1556-VII встановлює чотири рівні вищої освіти: початковий (короткий цикл), перший (бакалаврський), другий (магістерський) і третій (освітньо-науковий / освітньо-творчий) [24]. Другий рівень — здатність розв'язувати задачі дослідницького та (або) інноваційного характеру; третій — здатність розв'язувати *комплексні проблеми*, що передбачає оволодіння методологією наукової та педагогічної діяльності й проведення власного дослідження з науковою новизною, теоретичним і практичним значенням. Третій рівень відповідає восьмому рівню Національної рамки кваліфікацій; за відсутності стандарту вищої освіти вимоги до здобувача визначають безпосередньо за цим рівнем [24, 30].

Освітньо-наукова програма (ОНП) — єдиний комплекс освітніх компонентів (дисципліни, індивідуальні завдання, практики, контрольні заходи) і наукових (дослідження, публікації, виступи на конференціях), спрямованих на підготовку та публічний захист дисертації; окремі елементи ОНП може забезпечувати інший заклад [29]. Обсяг освітньої складової нормовано жорстко (п. 21 Порядку № 261):

$$30 \le C_{\text{edu}} \le 60, \qquad C_{\text{el}} \;\ge\; 0{,}25\, C_{\text{edu}}, \tag{2.1}$$

де C_{edu} — обсяг освітніх компонентів у кредитах ЄКТС, C_{el} — обсяг дисциплін за вибором здобувача [29]. Для програми з 40 кредитами це щонайменше 10 кредитів вибіркових дисциплін. Наукова складова в кредитах не вимірюється: вона реалізується через індивідуальний план наукової роботи. Нормативний строк підготовки доктора філософії в аспірантурі (ад'юнктурі) незалежно від форми здобуття освіти — *чотири роки*; підготовка поза аспірантурою для осіб, які професійно провадять наукову діяль-

ність за основним місцем роботи, — також чотири роки за кошти закладу; докторантура — два роки [29].

2.3.2 Вступ, науковий керівник, індивідуальні плани

Вступ здійснюється на конкурсній основі за Порядком № 261, Умовами прийому МОН і правилами прийому закладу — за результатами єдиного вступного іспиту, що проводить Український центр оцінювання якості освіти, вступного іспиту зі спеціальності в обсязі програми рівня магістра та інших форм випробувань (іспити, співбесіди, презентації дослідницьких пропозицій). Результати дійсні для вступу до цього закладу протягом одного календарного року [29]. Дві норми критичні: перелік *акредитованих і неакредитованих* ОНП у правилах прийому обов'язковий, а *атестація можлива лише за акредитованою програмою.* Тому перевірка акредитаційного статусу освітньо-наукової програми в Реєстрі НАЗЯВО [16] — перший крок вступника.

Науковий керівник призначається *протягом місяця* з дати наказу про зарахування з числа працівників закладу, які мають науковий ступінь і відповідну наукову кваліфікацію; він контролює виконання індивідуального плану наукової роботи і *відповідає перед вченою радою* за виконання своїх обов'язків. Кількісні межі: доктор наук — не більше п'яти здобувачів одночасно, з них не більше трьох здобувачів ступеня доктора наук; доктор філософії (кандидат наук) — не більше трьох здобувачів ступеня доктора філософії. Пункт 20 наводить вичерпні підстави зміни керівника, серед яких — невиконання ним обов'язків, відмова надати висновок, притягнення до академічної відповідальності та *спільне бажання керівника і здобувача*: зміна керівника за взаємною згодою є передбаченою законом процедурою, а не конфліктом [29].

Порядок № 261 розрізняє два документи, які систематично плутають [29]:

Індивідуальний навчальний план.
Визначає послідовність і форму засвоєння освітніх компонентів; містить дисципліни за вибором здобувача в обсязі не менш як 25 % кредитів освітньої складової (2.1); змінюється за погодженням із керівником.

Індивідуальний план наукової роботи.
Визначає *зміст, строки виконання та обсяг етапів і завдань* наукової роботи; погоджується з керівником, обговорюється кафедрою (відділом, лабораторією) і затверджується вченою радою разом із темою дисертації.

Обидва плани затверджує вчена рада *протягом двох місяців з дня зарахування.* Невиконання індивідуального навчального плану, порушення строків без поважних причин, умов договору або академічної доброчесності

є підставою для відрахування. Здобувач зобов'язаний *двічі на рік* звітувати про виконання плану наукової роботи на засіданні кафедри; протоколи цих звітів стають доказовою базою для висновку наукового керівника [29].

2.3.3 Передзахисний етап

Пункти 25–27 Порядку № 261 описують найщільніший за строками відрізок [29]. *Не пізніше ніж за дев'ять місяців* до завершення нормативного строку навчання здобувач отримує *довідку про виконання освітньо-наукової програми* (результати навчання, дисципліни, оцінки, кредити ЄКТС) і *висновок наукового керівника*, після чого звертається до базового структурного підрозділу із заявою про *висновок про наукову новизну, теоретичне та практичне значення результатів дисертації. Не пізніше ніж через місяць* проводиться публічна презентація результатів на засіданні підрозділу, а висновок надається *протягом двох тижнів* після неї; негативний висновок дозволяє повторне звернення не пізніше ніж за шість місяців до завершення навчання. За позитивного висновку здобувач *протягом двох тижнів* звертається до вченої ради із заявою про утворення разової ради [29, 30].

Числа не орнаментальні: за дев'ять місяців дисертація має бути фактично готова, а публікації — опубліковані, а не прийняті до друку. Звідси робочий дедлайн подання останньої статті — приблизно за 18 місяців до формального завершення навчання. Якщо заклад не отримав сертифіката про акредитацію ОНП або строк його дії закінчився, здобувач *переводиться* до закладу з акредитованою програмою; на першому році навчання переведення заборонено [29].

2.4 Присудження ступенів доктора філософії та доктора наук

2.4.1 Разова спеціалізована вчена рада

Присудження ступеня доктора філософії регулює Порядок, затверджений постановою Кабінету Міністрів України від 12.01.2022 № 44 [30]. **Разова спеціалізована вчена рада** утворюється закладом, у якому здобувач виконав *акредитовану* ОНП, з правом разового захисту однієї дисертації. Вчена рада закладу утворює її *не пізніше ніж протягом двох місяців* з дня отримання заяви, у складі *п'яти осіб*; рішення вводиться в дію наказом керівника протягом п'яти робочих днів (табл. 2.2).

Базова конфігурація — голова, два рецензенти, два опоненти; за неможливості призначити двох рецензентів — голова, один рецензент і три опоненти. **Компетентність** члена ради визначено формально: не менше трьох наукових публікацій за тематикою дослідження здобувача, опублікованих протягом останніх п'яти років і після присудження цій особі ступеня доктора філософії (кандидата наук); до них зараховують одноосібні монографії та статті у виданнях Переліку фахових видань України, Web of Science Core

Таблиця 2.2 — Склад разової спеціалізованої вченої ради (за [30])

Роль	К-сть	Вимоги
Голова	1	доктор наук; основне місце роботи — заклад, що утворив раду; компетентний учений за тематикою
Рецензент	2 (або 1)	науковий ступінь; основне місце роботи — цей самий заклад
Офіційний опонент	2 (або 3)	науковий ступінь; не працює в цьому закладі; опоненти не з одного закладу

Collection і Scopus, причому монографія обсягом не менш як п'ять авторських аркушів або публікація у виданні квартилів Q1–Q3 прирівнюється до двох публікацій [30].

Пункт 16 містить вісім підстав несумісності [30]: особа є науковим керівником здобувача, керівником закладу або співавтором його публікацій; має реальний чи потенційний конфлікт інтересів; притягувалася до академічної відповідальності; працювала на керівних посадах в організаціях, що незаконно провадили діяльність на тимчасово окупованих територіях; не володіє мовою захисту; отримала диплом доктора філософії (кандидата наук) *менше ніж за три роки* до утворення ради. Опоненти не можуть мати спільних публікацій за останні п'ять років із головою, рецензентами та науковим керівником; одна особа протягом календарного року може бути членом не більш як восьми разових рад.

2.4.2 Вимоги до дисертації та публікацій

Дисертація має містити нові науково обґрунтовані результати, які виконують *конкретне наукове завдання*, що має істотне значення для галузі знань; виконується державною або англійською мовою й подається у вигляді спеціально підготовленого рукопису; обсяг основного тексту встановлює ОНП закладу, вимоги до оформлення — МОН [30, 36] (зразки описів — у додатку В). Наукові результати мають бути висвітлені *не менше ніж у трьох наукових публікаціях*:

$$P = \sum_{i=1}^{n} w_i \;\geq\; 3, \qquad w_i \in \{0{,}5;\; 1;\; 2\}, \tag{2.2}$$

де w_i — вага i-ї публікації за табл. 2.3 [30]. Вираз (2.2) пояснює, чому «три статті» і «три публікації» — різні речі: три статті у фахових виданнях, кожна з чотирма співавторами, дають $P = 1{,}5$ і вимоги не задовольняють.

Додаткові умови (п. 9 Порядку № 44) [30]: в одному випуску видання зараховується не більше однієї статті; статті, опубліковані після набрання чинності Порядком, зараховуються *лише за наявності активного DOI*; для співавторських статей у дисертації зазначається особистий внесок кожного автора; публікації, що видаються в Україні, виходять державною мовою,

Таблиця 2.3 — Вага публікацій здобувача ступеня доктора філософії (за [30])

Тип публікації	Вага	Умова
Стаття у виданні квартилів Q1–Q3 (SJR або JCR)	2	квартиль за роком публікації або за останнім рейтингом
Одноосібна монографія	2	рекомендована вченою радою, пройшла рецензування
Стаття у виданні Web of Science Core Collection та/або Scopus	1	крім видань держави-агресора
Стаття у виданні Переліку фахових видань України	1	до двох співавторів разом зі здобувачем
Та сама стаття за трьох і більше співавторів	0,5	не стосується видань Scopus і WoS
Патент на винахід	1	не більше одного; кваліфікаційна експертиза

англійською або іншими офіційними мовами ЄС. Окремо зафіксовано: використання власних раніше опублікованих праць у тексті дисертації без посилання на них *не вважається самоплагіатом*, якщо вони опубліковані для висвітлення основних результатів дисертації і вказані в її анотації (пор. розділ 10).

Перелік наукових фахових видань України формує МОН за Порядком, затвердженим наказом від 15.01.2018 № 32 (у редакції наказу від 19.01.2026 № 56) [32]. Передбачено дві категорії: **«А»** — видання, що індексуються у Web of Science Core Collection та (або) Scopus; **«Б»** — видання, що відповідають вимогам п. 7 Порядку. Заявки на категорію «Б» приймаються з 1 лютого до 30 квітня *один раз на три роки* й оцінюються за шкалою 1–5 балів за кожним критерієм (максимум 30 балів), а рейтинговий список формується в межах кожного кластера споріднених галузей знань строком на три роки. Видання категорії «Б» представлене граничною кількістю спеціальностей одного кластера — тому перед поданням статті перевіряють не лише наявність журналу в Переліку, а й те, чи охоплює він спеціальність здобувача.

2.4.3 Процедура захисту й скасування рішення

Послідовність дій після утворення ради жорстко хронометрована [30]. Протягом *5 робочих днів* з дати наказу заклад оприлюднює електронну копію дисертації у форматі PDF/A з текстовим шаром і кваліфікованим електронним підписом здобувача, висновок про наукову новизну, склад ради й посилання на трансляцію; вносить інформацію до інформаційної системи НАЗЯВО; передає друкований примірник до бібліотеки; подає електронний примірник до УкрІНТЕІ та локального репозитарію [17]. Протягом *15 днів* з дня оприлюднення будь-який суб'єкт наукової діяльності може по-

дати до МОН повідомлення про невідповідність складу ради; МОН перевіряє склад протягом місяця. Протягом *45 календарних днів* кожен рецензент подає рецензію, кожен опонент — відгук із кваліфікованим електронним підписом. Протягом *3 робочих днів* з дня надходження останньої рецензії призначається дата захисту — не раніше ніж через два і не пізніше ніж через чотири тижні. Засідання правоможне лише за повного складу ради; заклад забезпечує трансляцію й повний відеозапис. Наказ про видачу диплома доктора філософії та додатка європейського зразка видається *не раніше ніж через 15 і не пізніше ніж через 30 календарних днів* з дня захисту; рішення ради набирає чинності з дати набрання чинності цим наказом.

Голосування відкрите: рішення про присудження ухвалюється, якщо його підтримали *не менше чотирьох* членів ради; про відмову — якщо відмову підтримали *два або більше* [30]. Член ради може письмово викласти окрему думку, що додається до рішення; невід'ємною частиною рішення є й відеозапис трансляції. Здобувач має право *до початку голосування* зняти дисертацію із захисту — лише один раз і лише якщо рада не виявила порушень доброчесності; у разі виявлення академічного плагіату, фабрикації чи фальсифікації заява не приймається, а рада відмовляє *без права повторного подання.* В інших випадках повторний захист можливий не раніше ніж через рік і лише після повторного висновку про наукову новизну. Матеріали захисту мають бути доступними для вільного перегляду не менш як шість місяців.

Скасування рішення можливе у двох режимах [30]. *До видачі диплома* його скасовує сам заклад у зв'язку з порушенням процедури захисту: повідомлення подається протягом 15 календарних днів з дня захисту, вчена рада утворює комісію з трьох працівників закладу (без членів разової ради), яка розглядає повідомлення протягом двох тижнів і за п'ять робочих днів готує висновок; порушення вимог *до складу* ради підставою для скасування не є, якщо роботу ради не зупинило МОН. *Після видачі диплома* рішення скасовує НАЗЯВО — протягом шести місяців з дня наказу про видачу диплома; повідомлення розглядає Комітет з питань діяльності разових спеціалізованих вчених рад протягом трьох місяців [16].

2.4.4 Доктор наук і позбавлення наукових ступенів

Порядок присудження та позбавлення наукового ступеня доктора наук затверджено постановою Кабінету Міністрів України від 17.11.2021 № 1197 [6]. Відмінності принципові: докторська рада утворюється *МОН* у складі *11–19 осіб* за науковою спеціальністю (не більше трьох); члени ради — доктори наук зі стажем роботи на наукових і науково-педагогічних посадах не менше п'яти років. Докторська дисертація має обсяг основного тексту не менше 10 авторських аркушів (або не менше трьох — у вигляді наукової доповіді

за наявності монографії чи сукупності статей) і реферат обсягом 1–2 авторських аркуші; наявність публікацій у Web of Science Core Collection та (або) Scopus обов'язкова, як і апробація на конференціях; попередня експертиза триває три місяці, висновок про наукову новизну чинний протягом року.

Норма, внесена постановою від 30.07.2025 № 928, заслуговує на окрему увагу: здобувач засвідчує підписом на титульній сторінці дисертації, що наукові положення є його власним напрацюванням, усі ідеї й цитати супроводжуються посиланнями, а *всі частини тексту, під час написання яких використовувалися технології штучного інтелекту, перевірені та відредаговані ним особисто* [6]. Це перша пряма вимога українського підзаконного акта щодо генеративного ШІ в кваліфікаційній роботі (межі допустимого — розділ 10).

Позбавлення наукового ступеня здійснює МОН на підставі рішення атестаційної колегії за зверненнями фізичних або юридичних осіб, закладу чи НАЗЯВО; повторний розгляд можливий лише за рішенням суду, апеляція подається до МОН не пізніше ніж через два місяці [6]. Наслідки виходять за межі здобувача: науковий консультант і опоненти позбавляються права участі в підготовці та (або) атестації наукових кадрів на два роки, а заклад — права утворювати докторську раду за спеціальністю на один рік.

2.4.5 Кейс: маршрут аспіранта IT-спеціальності

Розглянемо чотирирічний маршрут здобувача спеціальності «Кібербезпека та захист інформації» (F5; за попереднім переліком — 125) з темою про виявлення аномалій у трафіку безпроводових сенсорних мереж (рис. 2.2).

Вступ і перші два місяці. Перед поданням документів здобувач перевіряє, чи акредитована ОНП. Після зарахування призначено наукового керівника — доктора філософії, який керує ще двома здобувачами (межа — три). Затверджено тему й обидва плани; у плані наукової роботи зафіксовано: рік 1 — систематичний огляд і побудова стенда; рік 2 — збір датасету й базові експерименти; рік 3 — основна серія експериментів і дві статті; рік 4 — доопрацювання й захист [29].

Публікації. Стратегію побудовано за (2.2): стаття у фаховому виданні категорії «Б» з одним співавтором ($w = 1$) — наприклад, у виданні «Кібербезпека: освіта, наука, техніка» [10], де публікуються й суміжні роботи з освітніх програм і практикумів з кібербезпеки [11]; стаття у виданні, що індексується Scopus поза квартилями Q1–Q3 ($w = 1$); стаття у виданні квартиля Q3 ($w = 2$). Разом $P = 4 \geq 3$. Запас потрібен: якщо журнал випаде з квартиля Q3 у рік публікації, P знизиться до 3, і вимогу все одно буде виконано. *Типова помилка:* план із трьох статей у фаховому виданні з колективом із чотирьох авторів дає $P = 1{,}5$, а часу на нові публікації вже немає [30].

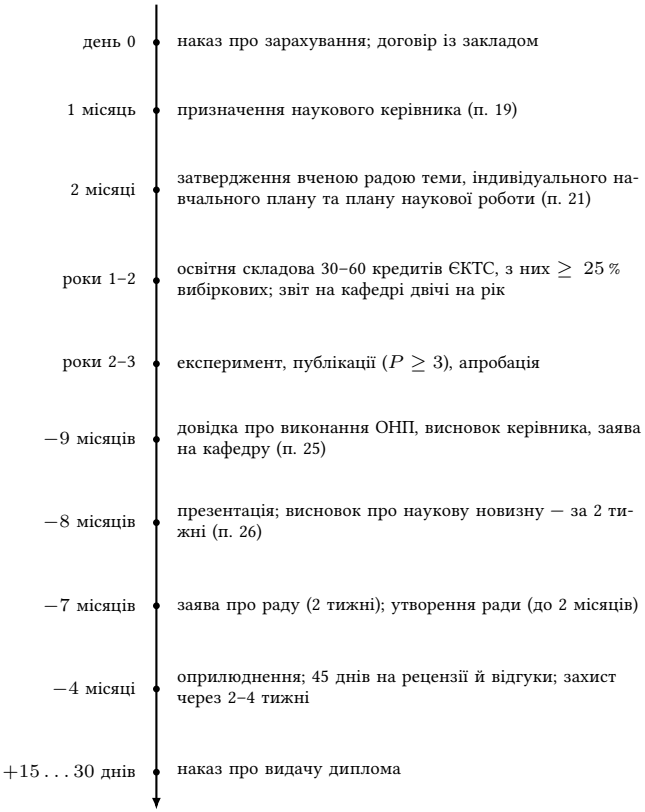


Рис. 2.2 — Життєвий цикл підготовки доктора філософії: контрольні точки й строки (за [29, 30])

Передзахист і захист. За дев'ять місяців до завершення — довідка й висновок керівника, презентація на кафедрі, позитивний висновок про наукову новизну, заява до вченої ради протягом двох тижнів. Раду утворено у складі: голова — доктор наук закладу; два рецензенти — доктори філософії закладу з трьома профільними публікаціями за п'ять років; два опоненти з різних закладів без спільних публікацій із головою, рецензентами й керівником. За 45 днів надійшли рецензії й відгуки; захист призначено через три тижні; рішення підтримали п'ять членів ради; через 20 днів видано наказ про диплом [30].

Що перенести у власний план. (1) Обчисліть P за (2.2) на початку другого року й закладіть запас щонайменше в одну одиницю. (2) Сплануйте подання останньої статті не пізніше ніж за 18 місяців до завершення навчання. (3) Перевірте наперед, чи є в закладі два потенційні рецензенти з трьома профільними публікаціями за останні п'ять років — без них рада не утвориться у базовій конфігурації.

2.5 Наукові проєкти, фінансування та Horizon Europe

2.5.1 НФДУ і конкурси МОН

НФДУ утворено постановою Кабінету Міністрів України від 04.07.2018 № 528 [35]; статус, завдання й органи управління визначено ст. 49–55 Закону № 848-VIII [34]. Фонд надає *індивідуальні*, *колективні* та *інституційні* гранти (ст. 51); напрями підтримки охоплюють виконання досліджень, розвиток матеріально-технічної бази, наукову співпрацю й мобільність, *наукове стажування аспірантів, ад'юнктів і докторантів, зокрема за кордоном*, розвиток дослідницької інфраструктури, трансфер знань, проєкти молодих учених і популяризацію науки. Наукова рада Фонду визначає тематичні напрями конкурсів і затверджує результати; експертиза проєктів незалежна й залучає зовнішніх, зокрема закордонних, експертів [35, 19]. Умови кожного конкурсу — строки, обсяг гранту, тривалість проєкту, вимоги до колективу — оприлюднюються окремим документом [18]. Паралельно діють конкурси МОН на виконання досліджень і розробок та державне замовлення на найважливіші розробки (ст. 57) [34, 14]. Порівняння джерел — у табл. 2.4.

Таблиця 2.4 — Джерела фінансування досліджень, доступні здобувачеві та його науковій групі

Джерело	Тип	Особливості	Підстава
Базове фінансування	інституційне	за результатами державної атестації; коефіцієнт групи А/Б	[34, 23]
НФДУ	гранти трьох видів	конкурс, незалежна експертиза; конкурси молодих учених	[35, 19]
Конкурси МОН	проєкти НДР, державне замовлення	дво- і трирічні проєкти ЗВО	[34, 14]
Horizon Europe	гранти ЄС	Україна — асоційована країна з 09.06.2022	[280, 53]
ERC	індивідуальні гранти	єдиний критерій — наукова досконалість	[145]
MSCA	докторські мережі, стипендії	обов'язкова міжнародна мобільність	[220]
Госпдоговірна тематика	замовлення підприємств	прикладні розробки; питання прав ІВ	[34]

2.5.2 Horizon Europe і участь України

Horizon Europe — рамкова програма ЄС з досліджень та інновацій на 2021–2027 рр., створена Регламентом (ЄС) 2021/695 [280]. *Стовп I «Досконала наука»*: Європейська дослідницька рада (ERC) — фронтирні дослідження окремих учених за єдиним критерієм наукової досконалості [145]; дії Марії Склодовської-Кюрі (MSCA) — підготовка дослідників і мобільність, зокрема докторські мережі [220]; дослідницькі інфраструктури. *Стовп II* — тематичні кластери, серед яких кластер 3 «Громадянська безпека для суспільства» і кластер 4 «Цифрові технології, промисловість і космос», основні для досліджень з ІТ і кібербезпеки. *Стовп III* — Європейська рада з інновацій, інноваційні екосистеми, Європейський інститут інновацій і технологій.

Україна є *асоційованою* країною: Угода про участь у Horizon Europe та програмі Євратому набрала чинності 9 червня 2022 р. і застосовується ретроактивно з 1 січня 2021 р. [53]. Наслідок: українські установи беруть участь у конкурсах на тих самих умовах, що й організації держав-членів, зокрема як координатори консорціумів, а аспіранти можуть бути найняті в проєктах MSCA. Інфраструктура підтримки — мережа **національних контактних пунктів** (NCP) за тематичними напрямами [14] і портал **EURAXESS**, що агрегує вакансії, стипендії та інформацію про умови мобільності [136].

2.5.3 Структура й оцінювання проєктної заявки

Заявка має канонічну тричастинну структуру, і за тими самими вимірами її оцінюють експерти [141]. *Excellence* — чіткість і обґрунтованість цілей, відповідність темі конкурсу, новизна «поза станом справ», придатність методології, зокрема відкрита наука й керування даними (розділ 9). *Impact* — очікувані результати й шлях до впливу, масштаб внеску, заходи з поширення, використання й комунікації результатів, керування правами інтелектуальної власності. *Implementation* — робочий план, пакети робіт, етапи (milestones) і контрольні результати (deliverables), розподіл ресурсів, якість консорціуму, керування ризиками. Оцінювання формалізоване:

$$S = S_{\mathrm{E}} + S_{\mathrm{I}} + S_{\mathrm{M}}, \qquad S_{\bullet} \in [0;\,5], \qquad S_{\bullet} \geq 3, \qquad S \geq 10, \tag{2.3}$$

де S_{E}, S_{I}, S_{M} — бали за критеріями excellence, impact та implementation із кроком 0,5 [141]. Вираз (2.3) містить два пороги: *індивідуальний* — не менш як 3 бали за кожним критерієм, і *сукупний* — не менш як 10 балів із 15. Проєкт з оцінками $5 + 5 + 2{,}5$ набирає 12,5 бала, але не проходить через провал за третім критерієм. Це найчастіша причина відхилення сильних наукових ідей: методологічно бездоганна заявка з нечітким робочим планом відсікається порогом $S_{\mathrm{M}} \geq 3$. Звідси рекомендація — писати заявку у зворотному порядку: спочатку робочий план і пакети робіт, потім розділ про вплив, і лише наприкінці наукову частину (пор. розділ 3).

2.6 Європейський дослідницький простір і кар'єра дослідника

Європейський дослідницький простір (European Research Area) — спільний простір, у якому вільно циркулюють дослідники, знання й технології; інтеграція України до нього є однією з функцій Національної ради з питань розвитку науки і технологій (ст. 20 Закону № 848-VIII) [34, 40].

Базовим документом щодо умов праці дослідників була Рекомендація Європейської комісії 2005 р. про Європейську хартію дослідників і Кодекс працевлаштування наукових працівників. Її оновлено: Рекомендація Ради від 18 грудня 2023 р. про європейську рамку залучення й утримання дослідницьких, інноваційних і підприємницьких талантів запровадила *нову Європейську хартію дослідників* із 20 принципами у чотирьох блоках — етика, доброчесність, гендерна рівність і відкрита наука; оцінювання, добір і кар'єрне просування; умови праці; кар'єри та розвиток талантів [99, 138]. Для здобувача Хартія — перелік того, чого він має право очікувати від установи: прозорий і відкритий добір, визнання з першого дня докторської підготовки *професіоналом*, а не студентом, доступ до наставництва, визнання мобільності частиною кар'єри, оцінювання за широким спектром результатів, а не лише за кількістю публікацій (пор. [88] і розділ 7). Механізм упровадження — стратегія HRS4R і відзнака **HR Excellence in Research** для установ, що демонструють прогрес в імплементації принципів Хартії; процедура має три фази — початкову (заява з gap-аналізом і планом дій), проміжну (оцінювання прогресу приблизно через два роки) і фазу поновлення. Наявність відзнаки є перевагою в конкурсах Horizon Europe й сигналом під час вибору місця підготовки [143].

Європейську рамку докторської освіти сформульовано у Зальцбурзьких принципах (2005) та Рекомендаціях «Зальцбург II» (EUA, 2010) [146]. Три наскрізні тези: докторська освіта посідає окреме місце в Європейському дослідницькому просторі, бо ґрунтується на *практиці дослідження*, що відрізняє її від першого й другого циклів (звідси — обмеження освітньої складової 60 кредитами ЄКТС (2.1)); докторант потребує *незалежності й гнучкості*, бо докторська освіта високоіндивідуальна (український відповідник — індивідуальний план наукової роботи); докторська освіта розвивається *автономними й підзвітними* установами за гнучкого регулювання — саме цю логіку реалізує модель разових рад, де держава задає рамку, а відповідальність за якість несе заклад [29, 30]. Порівняння діагностичне: якщо освітня складова витісняє дослідження, індивідуальний план є формальністю, а керівник виконує роль контролера відвідуваності — програма відхиляється від європейської рамки, і це буде помітно за критерієм 10 [28].

2.7 Інтелектуальна власність у науковому дослідженні

2.7.1 Чотири режими охорони

Стаття 64 Закону № 848-VIII відсилає набуття, охорону й захист прав інтелектуальної власності на науковий результат до спеціального законодавства [34]. Для дослідника галузі F/12 значущі чотири режими. *Авторське право* виникає з моменту створення твору й не потребує реєстрації (презумпція авторства); охоплює статті, дисертацію, документацію, комп'ютерні програми й бази даних [21]. *Патентне право* на винахід і корисну модель виникає *з моменту державної реєстрації* за Законом № 3687-XII [38]. *Комерційна таємниця (ноу-хау)* охороняється як нерозкрита інформація за главою 46 Цивільного кодексу України (ст. 505–508), доки зберігається секретність [47]. *Право особливого роду (sui generis)* стосується неоригінальних об'єктів, згенерованих комп'ютерною програмою без безпосередньої участі фізичної особи (ст. 33 Закону № 2811-IX); особисті немайнові права при цьому не виникають, а строк становить 25 років [21].

2.7.2 Комп'ютерна програма й межі охорони

Стаття 20 Закону № 2811-IX формулює межі точно [21]: охорона поширюється на програми, виражені у *вихідному або об'єктному кодах*, якщо вони оригінальні, і надається *формі вираження*. Не є формами вираження графічний інтерфейс користувача, набір виконуваних функцій і формат файлів даних. Не охороняються *ідеї та принципи*, на яких ґрунтується будь-який елемент програми, зокрема логічні схеми, *алгоритми* й мови програмування.

Наслідок фундаментальний: *алгоритм як такий не охороняється авторським правом*. Захистити можна конкретну реалізацію (код) — авторським правом або технічне рішення, що відповідає умовам патентоздатності, — патентом. Наукова новизна алгоритму й правова охорона — різні системи координат. Строк чинності майнових авторських прав — 70 років з 1 січня року, наступного за роком смерті автора (для співавторства — останнього співавтора); особисті немайнові права охороняються безстроково (ст. 31, 11) [21].

2.7.3 Службовий твір і розподіл прав

Найважливіша для здобувача норма — ст. 14 [21]. **Службовий твір** — твір, створений працівником у зв'язку з виконанням трудових обов'язків. Особисті немайнові права (авторство, ім'я, недоторканність) належать працівникові — авторові; майнові права переходять до роботодавця *з моменту створення твору у повному складі*, якщо інше не передбачено законом, трудовим договором (контрактом) або іншим договором; працівник має право

на винагороду, яка за договором може бути включена до заробітної плати, якщо посадові обов'язки прямо передбачають створення таких творів.

Звідси ключове розмежування. Якщо здобувач навчається в аспірантурі й *не перебуває у трудових відносинах* із закладом, створені ним твори (стаття, дисертація, код) службовими не є, і майнові права належать йому. Якщо ж він водночас працює асистентом, інженером чи виконавцем НДР, результати, створені в межах трудових обов'язків, є службовими, і майнові права за загальним правилом переходять до закладу. Конфігурація часто змішана, тому питання прав урегульовують *до* початку роботи — у договорі з закладом або в договорі про виконання НДР.

> **Чекліст прав на результати.** (1) Чи перебуваю я у трудових відносинах із закладом і чи охоплюють посадові обов'язки створення цього результату? (2) Чи є договір, який змінює загальне правило ст. 14? (3) Хто фінансував роботу — бюджет, грант, замовник — і які умови щодо ІВ містить договір? (4) Чи є серед результатів патентоздатне технічне рішення і чи не оприлюднив я його передчасно? (5) Під якою ліцензією публікується код і датасет? (6) Чи погоджено це з керівником і з умовами конкурсу?

2.7.4 Патенти, ліцензування, відкриті ліцензії

Для дослідника критичні три обставини патентування [38]. По-перше, умовою патентоздатності винаходу є *новизна*, яка втрачається внаслідок оприлюднення суті рішення: публікація статті чи препринта до подання заявки може унеможливити патентування (пільговий строк розкриття інформації автором обмежений, і його умови перевіряють у чинній редакції закону). По-друге, патент на винахід видається після кваліфікаційної експертизи, тоді як корисна модель охороняється без неї, а отже, її охоронна сила слабша. По-третє, у процедурі присудження ступеня доктора філософії зараховується *не більше одного* патенту на винахід, що пройшов кваліфікаційну експертизу; корисні моделі не зараховуються [30].

Розпоряджання майновими правами здійснюється через передання (відчуження) прав або ліцензійний договір (ст. 48–50) [21]. Окремий інструмент — **публічна ліцензія** (ст. 51): дозвіл на використання об'єкта *будь-якою особою* на визначених правовласником умовах, що видається шляхом оприлюднення умов разом із наданням дистанційного доступу до об'єкта. Три норми важливі для відкритої науки: публічна ліцензія може бути безвідкличною і тоді діє протягом усього строку чинності майнових прав; вона *невиключна*, а виключні та одиничні публічні ліцензії нікчемні; ліцензія без обов'язкової винагороди вважається безоплатною. Саме ст. 51 є українською правовою підставою для ліцензій Creative Commons на публікації й дані та відкритих ліцензій (MIT, Apache 2.0, GPL) на код; практика депонува-

ння артефактів — у розділі 9, а технічні засоби захисту прав і їх обмеження розглянуто, зокрема, в українських фахових дослідженнях [46].

Мінімальний порядок дій здобувача: на початку підготовки з'ясувати статус (трудові відносини чи лише навчання) і зафіксувати розподіл прав у договорі; перед першою публікацією перевірити наявність патентоздатного рішення; перед оприлюдненням коду й даних узгодити ліцензію з керівником і умовами гранту; у дисертації зазначити особистий внесок у кожну співавторську публікацію [30] і задекларувати використання технологій штучного інтелекту, якщо воно було [6].

2.8 Висновки до розділу

Правове поле української науки будується на двох законах: № 848-VIII визначає статус ученого, наукові ступені, державні гарантії та фінансування, № 1556-VII — рівні вищої освіти й систему забезпечення якості. Розмежування породжує дві процедури атестації: ступінь доктора філософії присуджує заклад через разову спеціалізовану вчену раду, ступінь доктора наук — МОН на підставі рішення докторської ради. Інституційна система має чотири контури — політика й бюджет, експертна координація, забезпечення якості та конкурсне фінансування; кожен орган має обмежені повноваження з конкретною правовою підставою.

Підготовка доктора філософії триває чотири роки за Порядком № 261 і жорстко регламентована строками: від призначення керівника й затвердження теми до передзахисного етапу за дев'ять місяців до завершення; освітня складова — 30–60 кредитів ЄКТС (2.1), атестація можлива лише за акредитованою програмою. Порядок № 44 формалізує захист через разову раду з п'яти осіб із вісьмома підставами несумісності, а вимога «не менше трьох публікацій» насправді є зваженою сумою (2.2), у якій стаття у фаховому виданні з трьома й більше співавторами важить удвічі менше ($w = 0{,}5$) за статтю з не більш ніж двома співавторами ($w = 1$), а публікація у виданні квартилів Q1–Q3 ($w = 2$) — удвічі більше за останню й учетверо більше за першу.

Фінансування надходить двома каналами: базовим, прив'язаним до результатів державної атестації, і конкурсним — через НФДУ, конкурси МОН та Horizon Europe, до якого Україна асоційована з 9 червня 2022 р.; заявка оцінюється за трьома критеріями (2.3) з подвійним порогом. Європейська рамка — Хартія дослідників, відзнака HR Excellence in Research і Зальцбурзькі принципи — задає орієнтир, за яким докторська освіта є практикою дослідження, а не продовженням магістратури. У сфері інтелектуальної власності ключовими є три розмежування: алгоритм не охороняється авторським правом, а його реалізація — охороняється; майнові права на службовий твір переходять до роботодавця з моменту створення, якщо договором не вста-

новлено інше; публічна ліцензія за ст. 51 Закону № 2811-IX є правовою підставою для відкритих ліцензій на код і дані.

Питання для самоперевірки

1. Які два закони формують основу регулювання науки в Україні і за яким із них присуджується кожен науковий ступінь?
2. Перелічіть суб'єктів наукової і науково-технічної діяльності за ст. 4 Закону № 848-VIII. Які права дає аспірантові цей статус?
3. Чим відрізняється базове фінансування від конкурсного і як результати державної атестації впливають на обсяг коштів закладу?
4. Назвіть повноваження МОН і НАЗЯВО щодо разових спеціалізованих вчених рад. Хто і в який строк може скасувати рішення такої ради?
5. Який нормативний строк підготовки доктора філософії та який обсяг освітньої складової встановлює Порядок № 261? Поясніть вираз (2.1).
6. Чим індивідуальний навчальний план відрізняється від індивідуального плану наукової роботи? У який строк їх затверджують?
7. Скількома здобувачами може одночасно керувати доктор філософії? доктор наук? Які підстави для зміни наукового керівника?
8. Які документи має отримати здобувач не пізніше ніж за дев'ять місяців до завершення навчання і до кого він звертається далі?
9. Опишіть склад разової ради та вимоги до компетентності її членів. Хто не може входити до її складу?
10. Обчисліть P за (2.2) для набору: дві статті у фаховому виданні з чотирма співавторами та одна стаття у виданні квартиля Q2. Чи виконано вимогу Порядку № 44?
11. Скільки голосів потрібно для присудження ступеня і за якої умови ухвалюється рішення про відмову? Коли здобувач не може зняти дисертацію із захисту?
12. Чому спроможність закладу сформувати разову спеціалізовану вчену раду є практичним критерієм під час вибору освітньо-наукової програми?
13. Які дві категорії Переліку наукових фахових видань України передбачено і за яких умов їх присвоюють?
14. Аналітичне: які два пороги містить вираз (2.3) і чому заявка з оцінками $5 + 5 + 2{,}5$ не проходить?
15. Дослідницьке: зіставте українську модель підготовки доктора філософії з трьома тезами Зальцбурзьких рекомендацій і назвіть елемент, який найбільше відхиляється від європейської рамки.

Практичні завдання

1. *Індивідуальний план наукової роботи.* Складіть план власного дослідження на чотири роки за структурою «етап — завдання — вимірний результат — строк — форма звітування». Для кожного року зазначте щонайменше одну контрольну точку перевірки гіпотези й цільове видання із зазначенням категорії Переліку [32]. Позначте на шкалі рис. 2.2 місяць подання останньої статті.
2. *Оцінювання відповідності публікацій вимогам.* Зведіть усі свої публікації (наявні та заплановані) у таблицю «видання — категорія або квартиль — кількість співавторів — наявність DOI — вага w_i» й обчисліть P за (2.2). Якщо $P < 3$ або запас менший за одиницю, запропонуйте план коригування з конкретними виданнями й строками.
3. *Аудит правового статусу дослідження.* Заповніть чекліст прав на результати (підрозділ 2.7). Визначте, які результати є службовими творами, які ліцензії ви застосуєте до коду й даних і чи є серед результатів потенційно патентоздатне технічне рішення.
4. *Перевірка складу разової ради.* Доберіть для своєї теми гіпотетичний склад разової ради з п’яти осіб. Для кожного кандидата перевірте наявність трьох профільних публікацій за останні п’ять років, відсутність спільних публікацій із науковим керівником і відповідність п. 16 Порядку № 44 [30].
5. *Ескіз проєктної заявки.* Підготуйте односторінковий ескіз заявки на конкурс НФДУ [18] або Horizon Europe за структурою excellence / impact / implementation: по три тези на розділ. Оцініть їх за шкалою (2.3) і визначте найнижчий бал.

РОЗДІЛ 3

Наукове дослідження: від проблемної ситуації до дослідницької програми

Розділ описує найвідповідальніший етап роботи — той, на якому дослідження ще не почалося, але вже визначено, чи має воно шанс на успіх. Правильно поставлена проблема — половина дисертації: вона задає об'єкт і предмет, диктує мету й завдання, обмежує набір прийнятних методів і наперед окреслює, що можна буде заявити як наукову новизну. Розділ спирається на понятійний апарат розділу 1 і правове поле розділу 2, а його результат — дослідницька програма — стає входом для розділів 4 і 5. Виклад побудовано інструментально: замість переказу означень подано шаблони формулювань, критерії самоперевірки й пари «погане формулювання → добре формулювання». Наскрізний кейс — розгортання теми «виявлення аномалій у трафіку пристроїв Інтернету речей» — проходить через усі підрозділи.

3.1 Проблемна ситуація, наукова проблема і наукова задача

3.1.1 Розрив між потрібним і наявним

Проблемна ситуація — стан, за якого наявні знання й засоби не дають змоги досягти потрібного результату, але саме утруднення ще не сформульоване мовою науки. Її зручно описати трійкою

$$\mathcal{P} = \langle S_{\text{req}},\, S_{\text{act}},\, K \rangle, \qquad \Delta = S_{\text{req}} \setminus S_{\text{act}}, \tag{3.1}$$

де S_{req} — властивості (вимоги), яких потребує практика або теорія; S_{act} — властивості, що їх фактично забезпечують наявні рішення; K — корпус знань, методів і засобів. Проблемна ситуація існує, якщо розрив Δ непорожній.

Цього, однак, ще не досить. **Наукова проблема** виникає лише тоді, коли розрив не закривається застосуванням уже відомого: позначивши через $\vdash$ відношення виводимості засобами корпусу K,

$$\text{Prob}(\Delta) \iff \Delta \neq \varnothing \;\wedge\; K \nvdash \Delta. \tag{3.2}$$

Вираз (3.2) — робочий фільтр, а не формальність. Якщо потрібну властивість можна отримати, коректно застосувавши наявний метод, ідеться про *інженерну (проєктну) задачу*: її розв'язують, а не досліджують. Р. Вірінга розрізняє саме ці дві сутності — *design problems* і *knowledge questions* — і зазначає, що змішування їх є найчастішою вадою рукописів з програмної

інженерії [344, 343]. Наукова проблема — це «знання про незнання»: вона фіксує не лише те, чого бракує, а й те, чому наявні засоби цього не дають. **Наукова задача** — частина проблеми, для якої відомий клас методів розв'язання і яку можна довести до перевірного результату в межах однієї роботи; спроба «розв'язати проблему цілком» — типова причина розмитої роботи без чіткого результату.

3.1.2 Протиріччя як джерело проблеми

Проблему не вигадують — її виявляють у протиріччі. Практично корисними є п'ять типів:

- *між потребою практики й відсутністю засобу*: вимога існує, а методу, що її задовольняє, немає;
- *між суперечливими емпіричними результатами*: дві групи на подібних даних дістають несумісні висновки — діє неврахований чинник;
- *між моделлю і спостереженням*: теоретична оцінка систематично розходиться з виміряним;
- *між конкурентними вимогами*: точніше виявлення скорочує час автономної роботи пристрою, сильніший захист збільшує затримку;
- *між масштабом явища і масштабом методу*: метод, коректний для десятків вузлів, непридатний для десятків тисяч.

Кожен тип задає власну форму запису проблеми: «потрібно X, наявні засоби дають лише Y, причина — Z»; «за умов A спостерігають R_1, за умов B — R_2, хоча теорія передбачає однакове»; «жодне з відомих рішень не забезпечує X і Y одночасно за обмеження C».

3.1.3 Ознаки псевдопроблеми

Псевдопроблема виглядає як наукова, але не задовольняє (3.2). Розпізнавальні ознаки:

1. *Розрив не визначено кількісно*: «недостатня ефективність» без показника, значення й умов вимірювання.
2. *Розрив закривається наявним засобом*: потрібне вже реалізовано у відкритій бібліотеці чи продукті.
3. *Проблема термінологічна*: суперечність зникає після узгодження означень.
4. *Твердження нефальсифіковне*: жоден результат вимірювання не спростовує очікуваної відповіді (розділ 1).
5. *Проблема зведена до інструмента* («відсутність програмного засобу для …»), *нерозв'язна за принципом* або *не має адресата*.

Тест «і що з того?». Сформулюйте проблему одним реченням і двічі поставте запитання: *і що з того, якщо її не розв'язати?* Якщо відповідь зводиться до «буде трохи гірше», проблема не має ваги, достатньої для дисертації [70].

Розрізняють два способи знайти проблему. **Пошук прогалини** (gap-spotting) надійний, але дає передбачувані, дрібні результати. **Проблематизація** — випробування *припущення*, яке література приймає без перевірки [56]: «трафік стаціонарний у межах години», «розмічений набір даних репрезентує реальний розподіл атак», «часовий порядок вибірки не має значення» [60]. Робота, що спростовує одне таке припущення, перевершує десяток робіт, які заповнюють прогалини.

3.1.4 Наскрізний кейс: крок 1

Ситуація. Оператор розумної будівлі має кілька тисяч гетерогенних пристроїв Інтернету речей. Засоби виявлення вторгнень навчено на еталонних наборах даних; в експлуатації частка хибних спрацьовувань робить систему непридатною — оператор не встигає опрацьовувати сповіщення. **Розрив:** S_{req} — не більше як 2 % хибних тривог за незмінної повноти; S_{act} — опубліковані моделі на власному трафіку дають 8–15 %.

Перевірка за (3.2). Перенавчання на власних даних розриву не закриває: через 4–6 тижнів показники повертаються до початкових; отже, $K \nvdash \Delta$ — маємо наукову проблему, а не задачу налаштування. **Протиріччя** — між якістю моделей на еталонних наборах і їх нестійкістю в експлуатації; джерело — неврахований дрейф профілю трафіку, тобто мовчазне припущення про стаціонарність «нормальної» поведінки.

Логіку розгортання кейсу в дослідницьку програму подано на рис. 3.1: кожен наступний елемент виводиться з попереднього, а зворотний зв'язок означає, що уточнення на пізньому етапі змушує переглянути ранній.

3.2 Вибір і обґрунтування теми дослідження

3.2.1 Джерела тем

Тема не «спадає на думку» — її здобувають систематичним пошуком. Продуктивні джерела, впорядковані за якістю результату:

- *Прогалини, названі в оглядових статтях*: розділ «open issues» — вже відфільтрована спільнотою множина тем (розділ 6).
- *Розділ «future work» чужих статей* — найшвидше джерело конкретних формулювань; перевірте, чи автори самі не виконали цю роботу в наступній публікації.

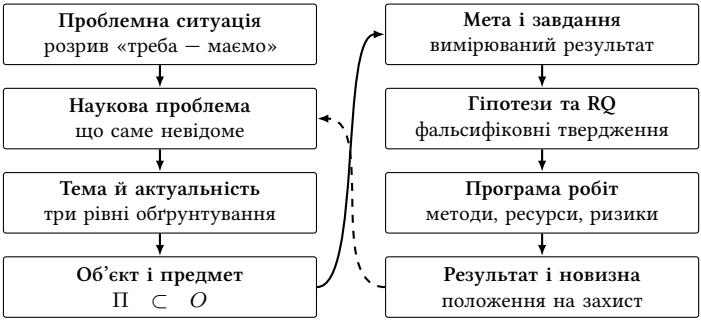


Рис. 3.1 — Логічна схема розгортання дослідження від проблемної ситуації до результату (штрихова лінія — уточнення проблеми за підсумками отриманого результату)

- *Суперечливі результати*: дві роботи з несумісними висновками — готова наукова проблема. Споріднене джерело — *невідтворені результати* (розділ 9).
- *Потреби практики*: регуляторні вимоги, інциденти, звіти команд реагування — потребують перетворення інженерної задачі на наукову.
- *Нові технології й нові дані*: поява класу пристроїв, протоколу чи відкритого набору даних відкриває вікно на 2–4 роки.
- *Перенесення методу* з іншої області — цінне лише тоді, коли нетривіальне й супроводжується аналізом меж застосовності.
- *Тематика кафедри*: теми, забезпечені обладнанням, даними й компетенцією керівника.

3.2.2 Критерії добору й формальна оцінка

Кандидатні теми (5–8) оцінюють за узгодженим набором критеріїв. Якщо $c_i \in [0;1]$ — нормована оцінка i-го критерію, а w_i — його вага, $\sum_i w_i = 1$, то інтегральна оцінка теми

$$F(T) = \sum_{i=1}^{n} w_i\, c_i(T) \;\cdot\; \prod_{j \in \mathcal{B}} \mathbb{1}\big[c_j(T) \geq c_j^{\min}\big], \tag{3.3}$$

де $\mathcal{B}$ — множина *блокувальних* критеріїв, для яких недосягнення порога $c_j^{\min}$ знімає тему з розгляду: відсутність доступу до даних не компенсується жодною актуальністю. Ваги узгоджують із керівником *до* оцінювання, а не після.

Критерій публікаційності недооцінюють, хоча він визначає виконання вимог до публікацій здобувача (розділ 2). Перевірка: знайти три статті за

Таблиця 3.1 — Критерії добору теми дослідження

Критерій	Питання для оцінювання	w_i	Блок.
Актуальність	Чи потрібна відповідь комусь, крім автора?	0,20	—
Новизна	Що залишиться новим після препринтів цього року?	0,25	так
Здійсненність	Чи буде результат за 3–4 роки наявними силами?	0,20	так
Доступ до даних	Чи є дані або стенд — легально й на весь термін?	0,15	так
Ресурсність	Обчислення, обладнання, ліцензії	0,10	—
Публікаційність	Чи є 3–4 видання, для яких тема профільна?	0,10	—

останні два роки, які *цитували б* вашу майбутню роботу; якщо таких немає, тема або надто вузька, або поза полем зору спільноти.

3.2.3 Узгодження теми

Тема існує у трьох контекстах одночасно, і невідповідність будь-якому з них дорого коштує.

Науковий керівник відповідає за здійсненність і методичну якість. Продуктивна форма першої розмови — не «дайте тему», а подання 2–3 опрацьованих кандидатів із заповненою табл. 3.1, переліком по 5–7 джерел і відповіддю на запитання «що саме тут нового».

Науковий напрям кафедри. Тема, що спирається на наявні стенд, дані й компетенції, рухається вдвічі швидше. Перевірте тематику виконуваних НДР і захищених за п'ять років дисертацій: збіг предмета з уже захищеною роботою — ризик для новизни [2].

Спеціальність. У назві має домінувати складова, що відповідає шифру спеціальності, за якою роботу представлятимуть до захисту, а разова спеціалізована вчена рада — бути спроможною її оцінити [30].

Назва теми — стислий запис предмета: 8–14 значущих слів, без абревіатур, незрозумілих поза вузькою спільнотою, без порожніх слів «дослідження», «аналіз», «деякі аспекти».

> **Погано → добре.** «Дослідження питань безпеки IoT» → «Методи виявлення аномалій у мережевому трафіку пристроїв Інтернету речей з обмеженими обчислювальними ресурсами в умовах дрейфу профілю поведінки». *Чому:* перше не називає предмета й не обмежує умов; з другого видно об'єкт, предмет і клас очікуваного результату.

3.3 Обґрунтування актуальності

3.3.1 Три рівні обґрунтування

Актуальність — не риторична прикраса вступу, а доказ того, що розрив Δ з (3.1) існує і що його варто закривати. Повне обґрунтування розгортається на трьох рівнях.

Суспільний рівень відповідає на запитання «чому це важливо поза наукою»: статистика інцидентів, вимоги законодавства, економічні або безпекові наслідки. Обсяг — один абзац, джерела перевірювані.

Галузевий (технологічний) рівень показує, що потреба конкретизується в технічній задачі: який клас систем зачеплено, які обмеження діють, чому наявні промислові рішення недостатні.

Науковий рівень — головний: він доводить, що задачу не розв'язано *у літературі* — називає конкретні роботи, показує, що в них зроблено, і вказує, чого бракує. Це єдиний рівень, де обґрунтування будується цитатами, а не міркуваннями; його слабкість — найчастіша причина зауваження «актуальність не доведено».

3.3.2 Робочі мовні шаблони

Шаблони нижче — каркас, який наповнюють фактами й використовують як перевірку повноти аргументу.

```
(1) За даними [джерело, рік], [явище] охоплює [кількісна
оцінка], що зумовлює [наслідок].
(2) Вимоги [нормативний акт / стандарт] передбачають [вимога],
проте [клас систем] не забезпечує її через [обмеження].
(3) Питанню [тема] присвячено праці [автори]. У роботах [A]
досягнуто [результат], у роботах [B] --- [результат]. Водночас
поза увагою залишилося [конкретна прогалина].
(4) Нерозв'язаною частиною загальної проблеми є [формулювання],
що й визначає актуальність цього дослідження.
```

Речення (3) — найважливіше: воно має називати *конкретні* результати, а не прізвища. Формулювання «дослідженням цього питання займалися …» з переліком із п'ятнадцяти прізвищ без змісту — ознака того, що літературу не опрацьовано.

3.3.3 Типові помилки

1. *Підміна актуальності описом технології* — абзаци про те, що таке Інтернет речей, замість доказу нерозв'язаності задачі. Перевірка: викресліть речення, що пояснюють загальновідоме; якщо лишилося менш ніж третина — актуальності немає.

2. *Підміна актуальності популярністю*: «кількість публікацій зростає» доводить радше протилежне — треба показати, чому ці публікації не закривають розриву.
3. *Обґрунтування чужою актуальністю*: аргументи стосуються ширшої проблеми, а робота — вузької задачі; не названо адресата.
4. *Неперевірювані джерела* (реклама, новинні перекази) або аргумент, що вичерпується часом.

Погано → добре.
Погано: «Інтернет речей стрімко розвивається, кількість пристроїв зростає щороку, тому питання їх безпеки є актуальним».
Добре: «Понад 60 % сповіщень системи виявлення вторгнень у досліджуваній інфраструктурі є хибнопозитивними (частка хибних *серед сповіщень*, тобто $1 - \mathrm{precision}$), що унеможливлює опрацювання черги інцидентів наявним штатом операторів. Відомі методи [умовні джерела] дають 1–3 % хибнопозитивних спрацьовувань (FPR) на еталонних наборах, які не відтворюють дрейфу профілю трафіку; стійкості таких методів до дрейфу в літературі не досліджено».
Увага до показників: 60 % і 1–3 % — *різні* величини. Перша — частка хибних серед сповіщень, і вона різко залежить від базової частки атак; друга — частка хибних спрацьовувань серед негативів. Плутати їх не можна (розділ 8): за частки атак 1 % навіть $\mathrm{FPR} = 1\ \%$ дає більшість хибних сповіщень.
Чому: перше формулювання не містить ані розриву, ані посилань, ані адресата; друге називає показник, значення, умову й місце прогалини.
Увага: позначку «[умовні джерела]» у власному тексті заміняють на номери *тих* праць, де справді наведено зазначені значення. Посилатися «для солідності» на нормативні акти чи випадкові позиції списку літератури не можна: рецензент перевіряє саме такі місця, і невідповідність посилання твердженню знецінює весь аргумент.

3.4 Об’єкт і предмет дослідження

3.4.1 Співвідношення понять

Об’єкт дослідження — процес, явище або система, що породжує проблемну ситуацію й існує незалежно від дослідника. **Предмет дослідження** — та частина, властивість або аспект об’єкта, які вивчають у роботі. Співвідношення однозначне:

$$\Pi \subset O, \qquad \Pi = \{\, x \in O \mid \varphi(x) \,\}, \tag{3.4}$$

де φ — предикат, що виокремлює предмет: аспект (стійкість), умови (дрейф профілю), клас систем (пристрої з обмеженими ресурсами). Предмет завжди *вужчий* за об’єкт і завжди задається обмеженням, яке можна назвати словами; якщо назвати не вдається — предмет не визначено.

Функції понять різні: об'єкт вказує на *місце* роботи серед інших, предмет визначає *зміст* новизни. Звідси правило самоперевірки: назва теми, предмет і формулювання новизни мають бути про одне й те саме, лише з різним ступенем деталізації.

3.4.2 Типові помилки в IT-темах

1. *Предмет дорівнює об'єктові*: об'єкт — «безпека мереж Інтернету речей», предмет — «методи забезпечення безпеки мереж Інтернету речей»; порушено (3.4).
2. *Об'єктом названо артефакт автора* («розроблена система виявлення аномалій»): створена автором система не існує незалежно від дослідження, об'єктом є процес, який вона обслуговує.
3. *Об'єктом названо технологію, а не процес*: «нейронні мережі» замість «процес виявлення відхилень мережевої поведінки».
4. *Предмет сформульовано як перелік робіт* («розроблення, реалізація й тестування модуля») — це план, а не аспект об'єкта.
5. *Предмет не пов'язаний із метою*: мета — про частку хибних тривог, предмет — про архітектуру системи; новизна формулюється не про те, що вимірюють.
6. *Об'єкт надмірно широкий*: під означення «інформаційна безпека» підпадає будь-яка робота галузі.

3.4.3 Наскрізний кейс: крок 2

Тема. «Метод адаптивного виявлення аномалій у мережевому трафіку пристроїв Інтернету речей з обмеженими обчислювальними ресурсами в умовах дрейфу профілю поведінки». **Об'єкт** — процес виявлення відхилень мережевої поведінки пристроїв Інтернету речей у промисловій експлуатації. **Предмет** — моделі та метод виявлення аномалій, стійкі до дрейфу профілю нормального трафіку пристроїв з обмеженими ресурсами. **Перевірка за** (3.4): предикат φ — «стійкість до дрейфу» $\wedge$ «обмежені ресурси»; обидва обмеження звужують об'єкт, і обидва названо в темі — інакше назва обіцяла б ширший результат, ніж дає предмет.

3.5 Мета, завдання й дослідницькі питання

3.5.1 Мета дослідження

Мета — очікуваний науковий результат, сформульований як стан, а не як дія. Канонічна конструкція:

```
Мета --- [підвищення / зменшення / забезпечення] [вимірюваної
властивості] [об'єкта] за рахунок [класу засобу, що є
предметом] за умов [обмеження].
```

Мета одна: дві мети означають дві роботи. Мета не може бути процесом («розробити», «дослідити»): розроблення — це завдання, а метою є те, заради чого розробляють. Перевірка — підставити мету в речення «дослідження буде успішним, якщо наприкінці буде показано, що …»; якщо продовження не вимірюване, мету треба переписати.

> **Погано → добре.** «Дослідити методи виявлення аномалій у трафіку IoT» → «Зменшення частки хибних спрацьовувань виявлення аномалій у трафіку пристроїв Інтернету речей за незмінної повноти в умовах дрейфу профілю за рахунок методу адаптивного оновлення моделі нормальної поведінки». *Чому:* у другому названо показник, обмеження й засіб — видно, за чим вимірюватиметься успіх.

3.5.2 Завдання

Завдання — кроки, сукупне виконання яких досягає мети G. Для завдань $T_1, \dots, T_m$ умови *повноти та релевантності* записуються як

$$\bigcup_{k=1}^{m} R(T_k) \;\supseteq\; G, \qquad \forall k:\; R(T_k) \cap G \neq \varnothing, \tag{3.5}$$

де $R(T_k)$ — результат k-го завдання. Перша умова — завдання *покривають* мету (нічого не бракує); друга — кожне працює на мету (немає завдання, що не дає нічого для мети). Завдання, яке не задовольняє другої умови, — баласт.

Чого (3.5) *не гарантує.* Ці умови не виключають *дублювання*: два завдання з однаковим результатом, кожне з яких саме по собі покриває мету, їх задовольняють. Якщо потрібна саме ненадмірність, додають умову незамінності — внесок кожного завдання має бути таким, що його неможливо отримати без нього:

$$\forall k:\quad \bigcup_{l \neq k} R(T_l) \;\not\supseteq\; G, \tag{3.6}$$

тобто вилучення будь-якого завдання руйнує покриття мети. Разом (3.5) і (3.6) дають мінімальне покриття: нічого не бракує і нічого не можна викинути. Канон для дисертації з IT-спеціальності — 4–6 завдань за структурою роботи: аналіз стану питання й формування вимог (розділ 1); побудова моделі (розділ 2); розроблення методу — ядро новизни (розділи 2–3); експериментальна перевірка й порівняння з базовими рішеннями (розділ 3–4); упровадження та оцінювання практичного значення (розділ 4, додатки).

3.5.3 Перевірка завдань на вимірюваність

Кожне завдання проходить чотири запитання: *Що буде на виході?* (артефакт, таблиця, модель, доведення); *Як буде видно, що виконано?* (критерій приймання); *Яким методом?*; *Де в тексті це буде?*. Відповідь «буде проаналізовано» означає, що завдання сформульоване як дія, а не як результат.

> **Погано → добре.** «Дослідити наявні методи виявлення аномалій» → «Виконати порівняльний аналіз методів за критеріями стійкості до дрейфу, обчислювальної складності та потреби в розмічених даних; сформувати вимоги до методу». «Розробити програмне забезпечення» → «Реалізувати метод у вигляді прототипу й підтвердити відтворюваність публікацією коду та конфігурацій експерименту». *Чому:* перша форма не має критерію приймання — аналіз можна вважати виконаним завжди.

3.5.4 Дослідницькі питання

Дослідницьке питання (research question, RQ) — операційна форма завдання: те, на що робота дає відповідь даними. У методології емпіричних ІТ-досліджень усталено п'ять типів [127, 303, 313]:

- *Описові* (RQ-D): «Які характеристики має …?»; відповідь — виміряний розподіл, таксономія, опис явища.
- *Пояснювальні* (RQ-E): «Який чинник зумовлює …?»; потребують моделі та контролю змішувальних чинників.
- *Проєктні* (RQ-P): «Як побудувати артефакт, що забезпечує …за обмежень …?» — основний тип у Design Science (розділ 5).
- *Порівняльні* (RQ-C): «Чи перевершує метод A метод B за показником m на даних D?»; потребують названої базової лінії та статистичного висновку (розділ 8).
- *Оцінювальні* (RQ-V): «Наскільки придатний метод A в умовах C?»; відповідь — оцінка за критеріями, визначеними до вимірювання.

Тип питання зумовлює метод: описове питання не перевіряють контрольованим експериментом, а порівняльне не розв'язують опитуванням. Неузгодженість типу й методу — часта причина відхилення статей [303].

3.5.5 Шаблон PICOC

Щоб питання не залишалося розмитим, його розкладають на компоненти. До шаблону PICO М. Петтікрю й Г. Робертс додали компонент контексту, утворивши PICOC [267]; у програмній інженерії його рекомендовано настановами із систематичних оглядів [206] (табл. 3.2).

Заповнена таблиця дає формулювання питання: «Чи зменшує [I] порівняно з [$\mathrm{C_{omp}}$] показник [O] на сукупності [P] за умов [$\mathrm{C_{ont}}$]?» — дві літери C

Таблиця 3.2 — Шаблон PICOC для IT-досліджень

Компонент	Зміст	Приклад із наскрізного кейсу
Population	Сукупність, система або клас даних	Трафік пристроїв Інтернету речей класу «розумна будівля»
Intervention	Метод, технологія, втручання	Адаптивне оновлення моделі нормальної поведінки за виявленим дрейфом
Comparison	Базова лінія	Автокодувальник зі статичною моделлю [228]; поріг за ковзним середнім
Outcome	Показник, що його вимірюють	Частка хибних спрацьовувань за фіксованої повноти; затримка виявлення
Context	Умови, обмеження	12 тижнів безперервного трафіку; шлюз з 2 ГБ ОЗП

позначають різні компоненти (Comparison і Context), тому в шаблоні їх розрізняють. Таблиця ж є основою пошукового запиту для огляду літератури (розділ 6) і плану експерименту (розділ 8). Порожня клітинка «Comparison» — сигнал, що робота зводиться до демонстрації власного рішення без порівняння.

Завдання й дослідницькі питання не дублюють одне одного: завдання описує *роботу*, питання — *знання*. Відображення між ними «багато до багатьох» із двома обмеженнями: жодне питання не залишається без завдання, що дає відповідь, і жодне завдання не існує без питання чи артефакту, який воно забезпечує (розд. 3.8).

3.5.6 Наскрізний кейс: крок 3

Завдання. T_1 — аналіз методів і формування вимог; T_2 — побудова моделі дрейфу профілю трафіку та критерію його виявлення; T_3 — розроблення методу адаптивного оновлення моделі; T_4 — експериментальна перевірка на відкритих [243, 228] і власних даних; T_5 — оцінювання обчислювальної вартості на шлюзі та впровадження.

Питання. RQ1 (D): як змінюються статистичні характеристики трафіку пристроїв за 12 тижнів експлуатації? RQ2 (E): які чинники зумовлюють зростання частки хибних тривог із часом? RQ3 (P): як побудувати процедуру оновлення моделі без розмічених даних? RQ4 (C): чи зменшує метод частку хибних тривог порівняно з базовими за фіксованої повноти? RQ5 (V): яка його обчислювальна вартість на шлюзі з 2 ГБ ОЗП?

3.6 Гіпотези дослідження

3.6.1 Види гіпотез

Гіпотеза — обґрунтоване припущення, істинність якого перевіряють даними. За пізнавальною функцією розрізняють *описові* («профіль трафіку пристроїв нестаціонарний у горизонті тижнів»), *пояснювальні* («зростання частки хибних тривог зумовлене дрейфом профілю, а не зміною складу атак») та *прогностичні* («адаптивне оновлення втримає частку хибних тривог нижче за 2 % протягом 12 тижнів»). Дисертація зазвичай містить одну–три гіпотези, кожна — до окремого питання пояснювального або порівняльного типу.

3.6.2 Вимоги до гіпотези

Фальсифікованість. Гіпотеза має забороняти щонайменше один спостережуваний результат (розділ 1). Формально, якщо E — множина можливих результатів вимірювання, а $E_H \subset E$ — підмножина, сумісна з гіпотезою, то

$$\mathrm{Fals}(H) \iff E \setminus E_H \neq \varnothing, \qquad \mathrm{Inf}(H) = 1 - \frac{|E_H|}{|E|}, \tag{3.7}$$

де $\mathrm{Inf}(H)$ — змістовність гіпотези: частка результатів, які вона забороняє. Гіпотеза, сумісна з будь-яким результатом ($E_H = E$), має нульову змістовність і не є науковою.

Область застосування формули. Умова фальсифікованості $E \setminus E_H \neq \varnothing$ загальна, а от відношення $|E_H|/|E|$ як *число* визначене лише для скінченної непорожньої множини E. Для нескінченного чи неперервного простору результатів (наприклад, $\mathrm{FPR} \in [0;1]$) потужності множин потрібної частки не дають: обидві множини незліченні, і відношення кардинальностей нічого не вимірює. У такому разі або дискретизують простір до скінченного переліку можливих результатів вимірювання (що й відбувається насправді: прилад має скінченну роздільність), або задають на E явну міру μ і беруть $1 - \mu(E_H)/\mu(E)$. Отже, (3.7) — робочий інструмент порівняння гіпотез на скінченному переліку результатів, а не універсальне кількісне означення змістовності будь-якої наукової гіпотези.

Операціоналізованість. Кожне поняття гіпотези має процедуру вимірювання: не «якість виявлення», а «частка хибнопозитивних спрацьовувань, обчислена за розміткою L на вікні w».

Нетривіальність. Гіпотеза, істинність якої випливає з означень або теореми, перевірки не потребує; гіпотеза, що суперечить доведеному, потребує виправлення. Суперечність із опублікованими емпіричними результатами називають прямо й пояснюють її причину.

3.6.3 Від дослідницької гіпотези до нуль-гіпотези

Дослідницька гіпотеза H_1 формулюється змістовно; перевіряють же статистично **нуль-гіпотезу** H_0 — припущення про відсутність ефекту. Перехід виконують у чотири кроки: (1) обрати показник m і процедуру його вимірювання; (2) назвати базову лінію B; (3) визначити практично значущу різницю δ *до* експерименту; (4) записати пару гіпотез. Для наскрізного кейсу:

$$H_0:\ \mathrm{FPR}_A - \mathrm{FPR}_B \geq -\delta, \qquad H_1:\ \mathrm{FPR}_A - \mathrm{FPR}_B < -\delta, \tag{3.8}$$

де FPR — частка хибних спрацьовувань за фіксованої повноти, A — запропонований метод, B — базовий, δ — мінімальне зменшення, яке оператор вважає відчутним; його визначають із міркувань предметної області, а не з отриманих даних. Вибір критерію, рівня значущості, обсягу вибірки й поправок на множинність — предмет розділу 8. Гіпотеза, для якої (3.8) не записується, ще не операціоналізована.

Погано → добре. «Запропонований метод покращує виявлення аномалій» → «На наборі [243] та на власному 12-тижневому трафіку частка хибних спрацьовувань методу A за повноти $0{,}95$ менша за показник методу B щонайменше на $\delta = 0{,}02$». «Дрейф впливає на якість моделі» → «Через 6 тижнів експлуатації без оновлення моделі частка хибних тривог зростає щонайменше вдвічі порівняно з першим тижнем». *Чому:* перші формулювання не забороняють жодного результату — $\mathrm{Inf}(H) = 0$ за (3.7).

3.7 Наукова новизна і практичне значення

3.7.1 Три рівні новизни

Українська практика усталила три рівні формулювань, що відповідають різній мірі приросту знання: *уперше*, *удосконалено*, *набуло подальшого розвитку* [13]. Рівень визначає не бажання автора, а співвідношення результату з опублікованим станом питання (табл. 3.3); міжнародна традиція оперує спорідненою класифікацією типів внеску [303].

3.7.2 Як формулювати положення новизни

Кожне положення починають із *рівня* новизни (табл. 3.3) і будують за сталою схемою з чотирьох елементів — об’єкт новизни, аналог, відмінність, ефект:

```
[рівень] [1: об'єкт новизни] , яке (який) , на відміну від [2:
названий аналог] , [3: суттєва відмінність] , що дало змогу [4:
вимірюваний ефект із числом та умовами].
```

Таблиця 3.3 — Рівні наукової новизни

Рівень	Коли застосовний	Приклад формулювання
Уперше	Результату такого класу немає; потрібен доказ відсутності (огляд)	Уперше запропоновано критерій виявлення дрейфу профілю трафіку пристроїв Інтернету речей, що не потребує розмічених даних і спирається на зміну ентропії міжпакетних інтервалів
Удосконалено	Відомий метод змінено так, що поліпшено вимірювану властивість	Удосконалено метод виявлення аномалій на основі автокодувальника [228] уведенням адаптивного порога, що зменшило частку хибних тривог з 8 % до 2 %
Набуло подальшого розвитку	Підхід поширено на новий клас об'єктів або умов	Дістали подальшого розвитку методи оцінювання стійкості моделей до зсуву розподілу в частині обмежень ресурсів крайових шлюзів

Відсутність будь-якого з елементів робить положення неперевірним. Найчастіше бракує *аналога* («на відміну від відомих» без назви) і *числа* («що підвищує ефективність»). Кількість положень — 2–4; кожне підтверджують публікацією здобувача й співвідносять із конкретним завданням.

Порожні формулювання. «Удосконалено метод виявлення аномалій, що дозволяє підвищити ефективність системи» — немає аналога, відмінності й числа. «Уперше розроблено програмний комплекс моніторингу» — комплекс не є науковим результатом; новизною може бути покладений в його основу метод. «Дістало подальшого розвитку поняття кіберстійкості» — не названо, у чому розвиток. *Виправлення:* перенести акцент з дії на відмінність — *що* змінено, *порівняно з чим*, *з яким наслідком.*

3.7.3 Новизна і практичне значення

Новизна — приріст *знання*: твердження, метод або закономірність, яких не було в літературі. *Практичне значення* — приріст *користі*: можливість застосувати результат, підтверджена актом упровадження, опублікованим артефактом, економічним чи безпековим ефектом. Метод може мати високу новизну й нульове практичне значення і навпаки; у дисертації потрібні

обидва твердження, але акт упровадження не доводить новизни, а публікація не доводить практичного значення.

> **Перевірка новизни трьома запитаннями.** *Хто зробив це раніше?* (назвіть роботу або доведіть оглядом, що її немає); *чим моє відрізняється?* (механізм відмінності); *наскільки краще і за яких умов?* (число, показник, набір даних). Немає відповіді бодай на одне — положення ще не готове.

3.8 Планування дослідження

3.8.1 Матриця відповідності

Програму починають із матриці, що зв'язує завдання, дослідницьке питання, метод, очікуваний результат і місце в тексті (табл. 3.4). Вона доводить повноту за (3.5), показує узгодженість типу питання й методу і задає структуру роботи; порожня клітинка — дефект програми.

Таблиця 3.4 — Матриця «завдання — метод — результат — розділ» (наскрізний кейс)

T	RQ	Метод	Результат	Р.
T_1	—	Систематичний огляд [206]	Таксономія методів, вимоги	1
T_2	RQ1, RQ2	Вимірювання, аналіз часових рядів [41]	Модель дрейфу, критерій його виявлення	2
T_3	RQ3	Design science [343]	Метод адаптивного оновлення	2
T_4	RQ4	Контрольований експеримент [350]	Порівняння з базовими, перевірка H_0	3
T_5	RQ5	Вимірювання на стенді, кейс-стаді [285]	Оцінка вартості, акт упровадження	4

3.8.2 Декомпозиція робіт і календарний план

Декомпозиція робіт (work breakdown structure, WBS) розкладає кожне завдання на пакети робіт до рівня, на якому пакет має одного виконавця, обсяг 1–4 тижні й перевірюваний результат [186]. Правило зупинки: дробити далі не потрібно, якщо можна назвати документ або артефакт, поява якого означає завершення пакета. Типова помилка — пакет «робота над розділом 2».

Календарний план розгортає WBS у часі з урахуванням залежностей, контрольних точок атестації та вимог до публікацій (розділ 2); зручне подання — діаграма Ґанта за семестрами (рис. 3.2), де окремо показано резерв часу.

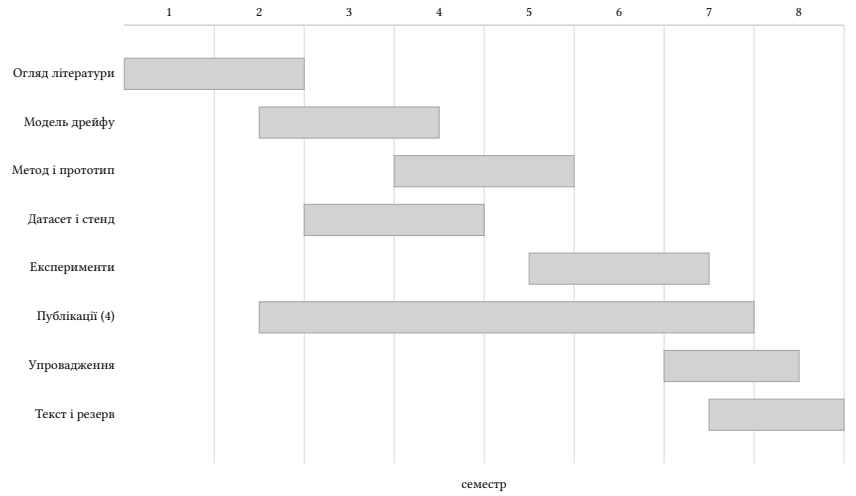


Рис. 3.2 — Календарний план аспіранта у вигляді діаграми Ґанта (за семестрами)

3.8.3 Ризики дослідження

Ризик — вплив невизначеності на цілі [188]. Реєстр складають *до* початку робіт і переглядають щосеместру: для кожного ризику оцінюють імовірність p і вплив I за шкалою 1–5, а пріоритет визначають як $R = p \cdot I$. Рівні задають наперед: $R \le 5$ — низький, $6 \le R \le 11$ — середній, $R \ge 12$ — високий. Реєстр упорядковують за спаданням R (табл. 3.5), і кожен ризик високого рівня обов'язково має захід пом'якшення з виконавцем і терміном. Пом'якшення має бути дією з терміном, а не наміром «бути уважним».

Таблиця 3.5 — Реєстр ризиків дослідження (за спаданням $R = p \cdot I$)

Ризик	p	I	R	Пом'якшення
Затримка публікацій	4	4	16	Подання до 3 видань за різними результатами; препринт
Втрата доступу до даних оператора	3	5	15	Паралельно готувати відкритий набір [243]; угода на весь термін
Тему закрито чужою публікацією	3	4	12	Щомісячний моніторинг препринтів; запас новизни в умовах експлуатації
Результат не відтворюється	2	5	10	Фіксація середовища й seed, публікація коду (розділ 9)
Відмова обладнання стенда	2	3	6	Резервний шлюз; щотижневе резервне копіювання

Три перші позиції мають високий рівень ($R \geq 12$), дві наступні — середній; ризиків низького рівня в цьому реєстрі немає. Якщо реєстр узагалі не містить високих ризиків, це не привід їх вигадувати — це привід перевірити, чи не занижено оцінки.

3.8.4 Ресурси

Плануванню підлягають чотири види ресурсів. *Обчислювальні*: прискорювачі, пам'ять, бюджет обчислень. *Дані*: джерело, обсяг, правові підстави оброблення, зокрема персональних даних [279, 33], зберігання й резервування. *Доступ*: угоди з оператором, дозволи на вимірювання в живій мережі, погодження етичного комітету (розділ 10). *Час*: реальний фонд часу аспіранта, зайнятого викладанням, становить 40–60 % від номінального, і план без цієї поправки зривається на другому році.

3.8.5 Ітеративна організація роботи: CDIO та eduScrum

Календарний план фіксує *що* й *коли*, але не відповідає на питання, *як* організувати щоденну роботу — власну або дослідницької групи. Дві рамки, адаптовані українською в Київському столичному університеті імені Бориса Грінченка, дають готові відповіді.

CDIO (Conceive — Design — Implement — Operate: задумати — спроєктувати — реалізувати — ввести в дію) — міжнародна ініціатива з реформування інженерної освіти, започаткована наприкінці 1990-х років і оформлена як відкрита модель із дванадцяти стандартів [9]. Її ядро — теза, що інженер опановує фах, проходячи повний життєвий цикл продукту, а не вивчаючи дисципліни ізольовано. Для дослідника в інформаційних технологіях цінність CDIO не в освітніх стандартах, а саме в цьому циклі: він майже збігається з інженерним циклом design science research (розділ 5), де артефакт задумують з огляду на проблему середовища, проєктують, реалізують і випробовують у реальних умовах. Відповідність очевидна: *Conceive* — це підрозділи 3.1 і 3.5 (проблема, мета, вимоги); *Design* — модель або метод; *Implement* — програмна чи апаратна реалізація; *Operate* — випробування, вимірювання, впровадження й супровід. Дисертація, що зупиняється на *Design*, залишає нерозкритими саме ті властивості артефакту, які виявляються лише в експлуатації — масштабованість, стійкість до відмов, вартість супроводу.

eduScrum — адаптація гнучкої методології Scrum до навчального й проєктного процесу: студентська (або дослідницька) команда сама планує роботу, розподіляє ролі й відповідає за результат, а викладач чи керівник виступає власником продукту, який формулює вимоги, але не диктує спосіб їх досягнення [5]. Робота розбивається на **спринти** фіксованої тривалості (для аспіранта зручний горизонт — два-чотири тижні), кожен спринт починає-

ться плануванням і завершується оглядом результату та **ретроспективою** — розбором того, що сповільнило роботу. Перелік завдань упорядковується за пріоритетом у **беклозі**, а видимість прогресу забезпечує проста дошка «заплановано — у роботі — виконано».

Для індивідуального дослідження з eduScrum варто запозичити три практики. *Перша* — спринт як одиниця планування між великими віхами: семестрові смуги календарного плану (рис. 3.2) надто тривалі, щоб вчасно помітити відхилення, а спринт дає сигнал уже через місяць. *Друга* — **критерій готовності** (definition of done), сформульований до початку спринту: не «попрацювати над експериментом», а «отримати таблицю результатів для трьох датасетів із довірчими інтервалами». *Третя* — ретроспектива, яку доречно проводити перед кожним звітуванням науковому керівникові: вона перетворює звіт із переліку зробленого на аналіз причин відставання. Обидві рамки не замінюють методології дослідження (розділи 4 і 5) — вони впорядковують *організацію* роботи, без якої коректна методологія лишається нереалізованою.

3.9 Дослідницька пропозиція

Дослідницька пропозиція (research proposal) — документ на 5–10 сторінок, що подає задум як цілісну програму. Його готують для вступу до аспірантури, щорічної атестації, конкурсу НФДУ [19] або проєктної заявки (розділ 2). Структура: назва й ключові слова; проблема й актуальність; стан питання з 15–25 джерелами; мета, завдання, RQ і гіпотези; методологія; очікувані результати й новизна; план робіт із контрольними точками; ризики; ресурси й бюджет; етичні аспекти та керування даними (розділи 9, 10); перелік джерел.

Критерії оцінювання зводяться до трьох груп, найчіткіше сформульованих у Horizon Europe [140]: *excellence* (ясність цілей, вихід за межі сучасного стану, обґрунтованість методології), *impact* (очікуваний ефект і шляхи його досягнення) та *implementation* (якість плану, ресурсів і керування ризиками). Та сама тріада діє й під час затвердження теми на кафедрі: пропозиція із сильними першими двома блоками й без плану робіт оцінюється нижче, ніж помірно амбітна, але забезпечена програма.

3.10 Висновки до розділу

Дослідження починається не з методу й не з інструмента, а з розриву між потрібним і наявним. Наукова проблема виникає лише тоді, коли цей розрив не закривається засобами наявного корпусу знань (3.2); усе інше — інженерна задача. Тему обирають за набором критеріїв із блокувальними умовами (3.3), актуальність доводять на трьох рівнях, і вирішальний з них — науковий, побудований на цитатах. Предмет є власною частиною об’єкта

(3.4), і саме про предмет формулюють новизну. Мета одна й вимірювана, завдання покривають мету без надлишку (3.5), дослідницькі питання розкладаються за шаблоном PICOC, а гіпотези мають забороняти спостережуваний результат (3.7) і допускати перехід до нуль-гіпотези (3.8). Новизну формулюють на одному з трьох рівнів, обов'язково називаючи аналог, відмінність і число. Задум стає програмою лише тоді, коли розгорнутий у матрицю «завдання — метод — результат — розділ», декомпозицію робіт, календарний план, реєстр ризиків і перелік ресурсів, зведені в дослідницьку пропозицію. Вибір методів — предмет розділів 4 і 5.

Питання для самоперевірки

1. Чим проблемна ситуація відрізняється від наукової проблеми? Запишіть власну проблему у вигляді (3.1).
2. Назвіть п'ять типів протиріч і вкажіть, який із них лежить в основі вашої теми.
3. Перелічіть ознаки псевдопроблеми. Які з них загрожують вашому формулюванню?
4. Чим пошук прогалини відрізняється від проблематизації? Яке припущення у вашій області приймають без перевірки?
5. Які критерії добору теми є блокувальними у (3.3) і чому їх не компенсують інші?
6. Назвіть три рівні обґрунтування актуальності. Чому науковий рівень неможливо замінити суспільним?
7. Сформулюйте об'єкт і предмет дослідження та покажіть предикат φ із (3.4).
8. Чому мета не може бути сформульована як дія? Які дві умови має задовольняти система завдань за (3.5)?
9. Перелічіть п'ять типів дослідницьких питань і назвіть метод, придатний для кожного.
10. Що означає $\mathrm{Inf}(H) = 0$ у (3.7)? Як перетворити таку гіпотезу на наукову?
11. Опишіть чотири кроки переходу до нуль-гіпотези (3.8). Звідки береться значення δ?
12. Чим рівень «уперше» відрізняється від «удосконалено»? Який доказ потрібен для першого?
13. Аналітичне: чому акт упровадження не є доказом наукової новизни, а публікація — практичного значення?

Практичні завдання

1. *Паспорт проблеми.* Опишіть проблемну ситуацію власного дослідження за (3.1): перелічіть S_{req} і S_{act} з показниками, доведіть, що $K \nvdash \Delta$, і перевірте формулювання за ознаками псевдопроблеми.
2. *Добір теми.* Сформуйте 5 кандидатних тем із різних джерел, оцініть їх за табл. 3.1 і обчисліть $F(T)$ за (3.3). Обґрунтуйте вибір і поясніть, чому відсіялися блоковані теми.
3. *Понятійний каркас.* Запишіть назву, об'єкт, предмет, мету, 4–6 завдань, 3–5 дослідницьких питань із зазначенням типу та 1–3 гіпотези. Перевірте узгодженість за (3.4), (3.5) і (3.7); для однієї гіпотези запишіть пару H_0/H_1 за (3.8) з обґрунтуванням δ. Окремо сформулюйте 3 положення новизни за шаблоном із чотирьох елементів, розподіливши їх за рівнями табл. 3.3, і запишіть практичне значення.
4. *Програма, ризики, пропозиція.* Побудуйте матрицю «завдання — RQ — метод — результат — розділ» за зразком табл. 3.4, розкладіть завдання на пакети робіт, накресліть календарний план за семестрами, складіть реєстр із 6 ризиків з оцінками p, I та діями з пом'якшення. Зведіть результати завдань 1–3 у пропозицію на 5–7 сторінок за структурою розд. 3.9 і самооцініть її за трьома критеріями [140] за шкалою 0–5 з обґрунтуванням.

РОЗДІЛ 4

Методи наукового дослідження

Розділ відповідає на запитання, яке постає одразу після того, як задум дослідження оформлено в програму (розділ 3): *чим саме здобувати результат*. Ключова теза — метод обирають під завдання, а не навпаки, і кожен метод має межі, за якими він дає не просто неточний, а систематично хибний результат. Тому виклад побудовано не як словник означень, а як інструмент добору: для кожного методу подано робоче означення, приклад застосування в інформаційних технологіях і кібербезпеці та умови, за яких його висновки недійсні. Галузеві методології (Design Science Research, контрольований експеримент у програмній інженерії, кейс-стаді) є предметом розділу 5, статистичне опрацювання даних — розділу 8.

4.1 Класифікація методів наукового дослідження

4.1.1 Три незалежні осі класифікації

Метод визначено в розділі 1 як упорядковану сукупність пізнавальних дій із визначеними межами застосовності. Методів сотні, і жодна лінійна класифікація їх не охоплює; практично корисними є три *незалежні* осі (рис. 4.1).

Вісь 1 — рівень спільності (табл. 4.1): **філософські** методи процедур не дають, але визначають, що вважати поясненням; **загальнонаукові** застосовні в багатьох науках; **конкретно-наукові** властиві галузі; нижче лежить *технологічний* рівень — методики й техніки.

Вісь 2 — етап пізнання. **Емпіричні** методи працюють з об'єктом безпосередньо й дають дані; **теоретичні** — зі знаковими репрезентаціями й дають узагальнення, моделі та виведення; **метатеоретичні** — із самим знанням: методологічний аналіз, систематичний огляд (розділ 6), наукометричне картування (розділ 7).

Вісь 3 — характер даних. **Кількісні** методи оперують числовими величинами й вимагають шкали та процедури вимірювання; **якісні** — текстами, спостереженнями, кодуванням категорій; **змішані** (mixed methods) поєднують обидва типи в одному дизайні [101]. В ІТ якісні методи недооцінюють: причини, з яких адміністратори вимикають механізм захисту, лічильником не виміряти.

Осі незалежні: контрольований експеримент — загальнонауковий, емпіричний, кількісний; напівструктуроване інтерв'ю — загальнонауковий, емпіричний, якісний; аналіз обчислювальної складності — загальнонауко-

вий, теоретичний, неемпіричний. Характеристика власного методу за всіма трьома осями — перша перевірка методологічної свідомості дослідника.

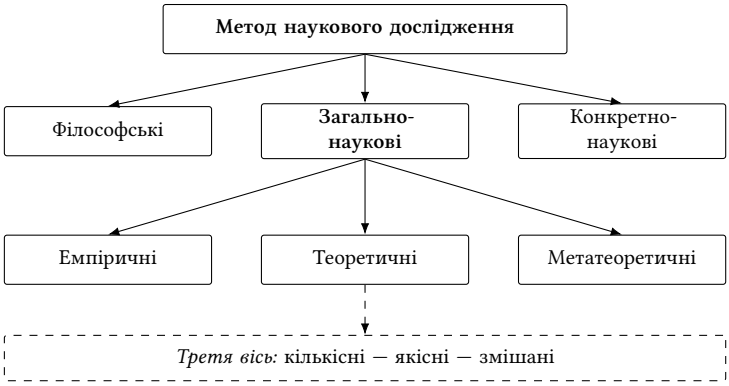


Рис. 4.1 — Класифікація методів наукового дослідження за трьома незалежними осями

Табл. 4.1 розгортає першу вісь із прикладами для галузі F/12. Критерії якості на рівнях різні: внутрішня узгодженість — коректність процедури — відповідність галузевим настановам — відтворюваність.

Таблиця 4.1 — Рівні спільності методів і їх вияв в IT-дослідженні

Рівень	Типові методи	Приклад в IT і кібербезпеці
Філософський	Діалектичний підхід, феноменологія, герменевтика	Що вважати поясненням поведінки системи: механізм чи статистичну залежність
Загальнонауковий	Спостереження, вимірювання, експеримент, моделювання, аналіз і синтез, формалізація	Вимірювання завантаженості радіоканалу; імітаційна модель черги на шлюзі
Конкретнонауковий	Методи галузі та суміжних дисциплін	Фаззинг, статичний аналіз коду, формальна верифікація протоколу, крос-валідація
Технологічний	Методики, процедури, техніки, інструменти	Протокол захоплення трафіку; скрипт обробления; фіксація версій стенда

4.1.2 Критерії добору методу

Метод обирають за чотирма критеріями: *адекватність предметові* (здатність виявити саме ту властивість, про яку сформульовано дослідницьке питання); *відповідність типу питання* (описове не розв'язують контрольованим експериментом, порівняльне — опитуванням, проєктне — статистичним тестом); *здійсненність* (дані, обладнання, експерти, правові підстави); *перевірність результату.*

Перевірка добору методу трьома запитаннями. *Яку властивість об'єкта має виявити метод? Чому саме цей метод здатен її виявити? За яких умов його висновки будуть хибними?* Якщо відповідь на третє запитання відсутня, у розділі «Методи» ви переказуєте підручник, а не обґрунтовуєте вибір.

4.2 Емпіричні методи

4.2.1 Спостереження, опис, вимірювання, порівняння

Спостереження й опис.

Спостереження — цілеспрямоване систематичне сприйняття об'єкта без навмисного втручання; від побутового відрізняється наявністю *протоколу* (що, коли, як часто, яким засобом фіксується), *одиниці спостереження* і способу реєстрації, незалежного від пам'яті спостерігача. *Приклад:* пасивний моніторинг трафіку сегмента протягом 12 тижнів; збір телеметрії з honeypot-вузлів. *Межі:* причинності не встановлює; діє *ефект спостерігача* (сенсор змінює затримки, поінформований адміністратор — поведінку) і *селективність* — відсутність події в журналі частіше означає обмеження засобу, ніж відсутність явища. **Опис** фіксує результат у системі понять і є проміжним продуктом; типова помилка — *підміна пояснення описом* (розд. 4.9).

Вимірювання.

Експериментальне одержання значень, які обґрунтовано можна приписати величині [197]; передбачає величину, шкалу, засіб із відомими метрологічними характеристиками й оцінку невизначеності. Шкали (номінальна, порядкова, інтервальна, відношень) визначають *допустимі операції* (розділ 8). *Приклад:* затримка відгуку (шкала відношень), RSSI (умовна шкала, залежна від реалізації радіомодуля), критичність за CVSS (порядкова за суттю, попри числовий вигляд). *Межі:* арифметика над порядковими шкалами (усереднення балів CVSS по активах дає число без змісту); ігнорування невизначеності засобу (різниця 2 % нічого не означає за розкиду повторень 5 %); **підміна конструкту показником** — «зручність конфігурування» замінюють кількістю рядків файлу налаштувань.

Порівняння.

Встановлення подібності й відмінності за спільною ознакою; потребує зіставності об'єктів, тотожності ознаки й однаковості умов. *Приклад:* дві системи виявлення вторгнень на одному наборі даних, за однакового порога й однакового розподілу вибірок. *Межі:* неназвана або застаріла базова лінія («солом'яне опудало»); набір даних, дібраний після отримання результатів; різна ретельність налаштування (власний метод —

пошуком за сіткою гіперпараметрів, чужий — зі стандартними значеннями).

4.2.2 Експеримент: види, планування, відтворюваність

Експеримент — метод, за якого дослідник *активно й контрольовано* змінює незалежну змінну та фіксує зміну залежної за інших умов, утримуваних сталими. Саме контрольоване втручання дає підстави для причинних висновків; види розрізняють за середовищем і природою об'єкта впливу (табл. 4.2).

Таблиця 4.2 — Види експерименту в IT-дослідженні

Вид	Приклад в IT	Головне обмеження
Лабораторний	Стенд із керованою генерацією трафіку	Низька зовнішня валідність
Натурний	Шлюз на реальному обладнанні в тестовій зоні	Обмежена керованість умов; вартість повторення
Обчислювальний (*in silico*)	Імітаційна модель мережі; прогін на відкритому наборі	Перевіряється модель, а не дійсність
Польовий	Прототип у діючій інфраструктурі	Змішувальні чинники; етичні й правові обмеження
Природний	Зміна без втручання (оновлення політики, інцидент)	Немає рандомізації

Межі. Експеримент не рятує від поганої операціоналізації: якщо показник не відповідає конструктові, контроль умов лише робить хибний висновок точнішим. Польовий експеримент у кібербезпеці майже завжди зачіпає третіх осіб, тому потребує етичного розгляду за принципами Menlo Report [119] (розділ 10).

Мінімальний план фіксує *до* вимірювань: змінні й спосіб їх вимірювання; чинники, утримувані сталими, і чинники, що рандомізуються; кількість повторень і обґрунтування обсягу вибірки; базові лінії; критерій рішення (α, практично значуща різниця δ); процедуру опрацювання даних. Така фіксація усуває зміну показника після перегляду результатів і добір підмножини даних під бажаний висновок (розділ 8). Відтворюваність обчислювального експерименту забезпечується фіксацією версій бібліотек, початкових значень генераторів випадкових чисел, розподілу вибірок, конфігурації й команд запуску (розділ 9); галузеві вимоги до плану та загрози валідності — у розділі 5.

4.3 Теоретичні методи

Теоретичні методи не звертаються до об'єкта безпосередньо: вони працюють із його знаковими репрезентаціями, і їх результат — не дані, а структура (поняття, модель, виведення, класифікація). Помилка тут не виявляється вимірюванням, тому контроль коректності має бути вбудований у сам виклад.

4.3.1 Операції з поняттями й виведення

Аналіз і **синтез** — розчленування об'єкта на складники й зворотне об'єднання з урахуванням зв'язків (стійкість кожної фази протоколу автентифікації окремо → стійкість композиції фаз). *Межа:* аналіз хибний там, де властивість *емерджентна* (розд. 4.5): стійкість композиції криптографічних примітивів не випливає зі стійкості кожного з них — найчастіша вада «аналізу підсистем» у роботах з кібербезпеки.

Індукція — перехід від окремих випадків до загального твердження, що не зберігає істинності; **дедукція** — виведення, яке її зберігає (розділ 1). Вимірювання на 40 моделях пристроїв дає узагальнення «пристрої класу не перевіряють ланцюжок сертифікатів»; з властивостей режиму шифрування виводиться межа кількості блоків за заданої ймовірності колізії. *Межі:* індуктивний висновок дійсний лише в охопленій вибірці — перенесення з 40 пристроїв одного виробника на «усі пристрої Інтернету речей» є **надузагальненням**; дедуктивний дійсний лише за істинності засновків, тож доведення в моделі, припущення якої не виконуються в реалізації, на практиці не гарантує нічого.

Абстрагування — відволікання від несуттєвих для завдання властивостей, **ідеалізація** — побудова об'єкта з властивостями, яких не буває (канал без втрат, порушник з необмеженими ресурсами), **узагальнення** — перехід до родового поняття. Модель порушника Долева–Яо вважає канал повністю підконтрольним, а примітиви — абсолютно стійкими: ця ідеалізація уможливила формальну верифікацію протоколів і водночас приховала клас побічноканальних атак. *Межа:* хибно відкинути саме ту властивість, що породжує явище, тому кожну ідеалізацію супроводжують переліком відкинутого й умовами, за яких воно стає суттєвим.

Аналогія — висновок про властивість об'єкта за подібністю до іншого: не доводить, а підказує гіпотезу. Епідемічна модель поширення мережевого черв'яка продуктивна, бо спільним є механізм контактного передавання; аналогія «імунна система — система виявлення вторгнень» лишається метафорою, бо в біологічному механізмі немає аналога навмисного адаптивного супротивника. *Межа:* **некоректна аналогія** — перенесення за зовнішньою подібністю; якщо спільний механізм назвати не вдається, аналогія годиться лише для ілюстрації.

4.3.2 Побудова та розгортання теорії

Формалізація — подання змісту в штучній мові з точним синтаксисом і семантикою; **аксіоматичний метод** — побудова теорії з невеликої множини недоведуваних тверджень. Приклади: опис протоколу в темпоральній логіці з перевіркою моделі (model checking), семантика Гоара для доведення коректності програм. *Межі:* формалізація не додає знання, а робить явним наявне; типові вади — формалізація заради вигляду строгості (позначення вводять, виведення не роблять) і *втрата предмета*, коли модель точна, але описує не той об'єкт. «Аксіомами» в прикладних роботах часто називають довільні припущення, і метод вироджується в риторику.

Гіпотетико-дедуктивний метод — побудова гіпотез, з яких дедуктивно виводять перевірні наслідки (розділ 1). **Сходження від абстрактного до конкретного** — розгортання теорії від бідної вихідної абстракції («потік подій із розподілом F») через послідовне введення гетерогенності джерел, добової періодичності, дрейфу профілю й ресурсних обмежень до моделі, придатної для інженерного застосування; саме такою має бути логіка теоретичного розділу дисертації. *Межа:* метод дає ілюзію строгості, якщо вихідну абстракцію дібрано довільно.

Історичний метод відтворює розвиток об'єкта в часі з усіма випадковостями, **логічний** — його структуру в «очищеному» вигляді. Розуміння того, *чому* протокол має саме таку структуру (сумісність із попередньою версією, обмеження обладнання 1990-х рр., компроміс у комітеті зі стандартизації), часто пояснює вразливість краще, ніж формальна модель. *Межі:* історичний виклад вироджується в хронологію публікацій без аналітичного стрижня, логічний — подає випадковий шлях розвитку як необхідний.

Мисленнєвий експеримент — міркування, у якому уявну ситуацію доводять до логічного завершення, щоб виявити наслідки припущень: машина Тюрінга, задача візантійських генералів, «чи довіряти довірі» К. Томпсона. Перш ніж будувати стенд, корисно подумки надати порушникові максимальні можливості: якщо властивість руйнується за перших реалістичних припущень, потрібен не експеримент, а переформулювання гіпотези. *Межа:* результат є висновком про логічну сумісність припущень, а не про дійсність.

4.4 Моделювання

4.4.1 Типи моделей і адекватність

Модель — об'єкт, що заміщує оригінал у дослідженні й зберігає суттєві для завдання властивості. Звідси головне: моделі «взагалі» не існує — існує модель *для певного завдання*, і поза ним вона не має ані точності, ані сенсу [74].

Аналітичні.

Система рівнянь, розв'язувана в замкненій формі (черга $M/M/1$ для середньої затримки на шлюзі): загальність висновку ціною сильних припущень, які реальний трафік порушує.

Імітаційні.

Поведінку відтворюють покроково в часі (дискретно-подієва модель мережі, агентна модель поширення шкідливого коду): довільна складність, але результат є вибіркою й потребує статистичного опрацювання й аналізу чутливості.

Натурні (фізичні).

Спрощена копія оригіналу — полігон із реальним обладнанням: реалістичні деталі реалізації ціною масштабу й вартості.

Знаково-математичні.

Формальні описи структури й обмежень: граф атак, автомат протоколу, система лінійних обмежень. Дають змогу доводити властивості, а не лише обчислювати.

Емпіричні (навчені).

Модель побудовано з даних без явного механізму (регресія, дерево рішень, нейронна мережа): працює там, де механізм невідомий, але дійсна лише в області, покритій навчальними даними.

Адекватність — відповідність моделі оригіналові в межах завдання й заданої точності; її не декларують, а перевіряють. Якщо y_i — виміряне значення показника, $\hat{y}_i$ — обчислене за моделлю, а ε — допустима похибка, узгоджена *до* перевірки, то критерій адекватності

$$\max_{i=1..N} \frac{|y_i - \hat{y}_i|}{|y_i|} \le \varepsilon \quad \text{або} \quad \sqrt{\frac{1}{N}\sum_{i=1}^{N}\bigl(y_i - \hat{y}_i\bigr)^2} \le \varepsilon_{\mathrm{RMS}}. \tag{4.1}$$

Перший запис — гарантія «у найгіршому випадку», другий — усереднена оцінка; вибір між ними змістовий: для пропускної здатності достатньо середньої похибки, для гарантованої затримки в системі реального часу потрібна перша форма. Значення ε беруть із вимог предметної області, а не з отриманих нев'язок.

4.4.2 Верифікація, валідація, межі перенесення висновків

Верифікація (verification) відповідає на запитання «чи правильно побудовано модель»: чи відповідає реалізація концептуальній моделі, чи немає помилок кодування, чи коректна чисельна схема; це перевірка *внутрішньої* коректності, зовнішніх даних вона не потребує. **Валідація** (validation) відповідає на запитання «чи побудовано правильну модель»: чи відповідає

поведінка моделі поведінці оригіналу в межах завдання; потребує незалежних експериментальних даних [293, 248]. Повний цикл V&V (рис. 4.2) містить також *валідацію концептуальної моделі* — перевірку припущень і структури до реалізації — та *перевірку достовірності даних*, без якої обидві попередні втрачають сенс.

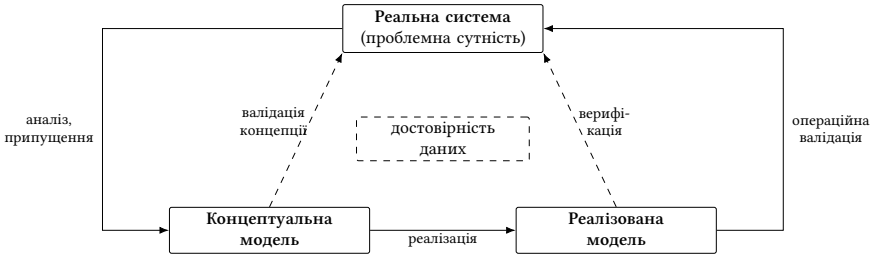


Рис. 4.2 — Цикл верифікації та валідації моделі (перемальовано за [293])

Практичний мінімум для дисертації: описати, *як* модель верифіковано (тести на вироджених режимах, звіряння з аналітичним розв'язком в окремому випадку, перевірка балансів) і *на яких незалежних даних* її валідовано — обов'язково не на тих, за якими її калібрували.

Область дійсності моделі — множина значень вхідних величин, для якої адекватність підтверджено; перенесення висновку за її межі є **помилкою екстраполяції** й дає результат, хибність якого неможливо виявити всередині самої моделі. *Приклад:* модель, валідована на трафіку 100 пристроїв, застосована до сегмента з 10 000 — механізми, нехтовні за малого масштабу (колізії, вичерпання таблиць стану, синхронізація періодичних передавань), стають визначальними; модель виявлення аномалій, навчена на трафіку 2023 р., дає хибні спрацьовування на трафіку 2026 р. через дрейф розподілу — та сама помилка, лише в часі.

Три обов'язкові речення про модель. *Для якого завдання модель побудовано* (поза ним вона не оцінюється); *у якій області вхідних величин її валідовано* (за межами — екстраполяція); *які припущення відкинуто і коли вони стають суттєвими.*

4.5 Системний підхід і системний аналіз

Система — сукупність взаємопов'язаних елементів, що утворює цілість із властивостями, яких немає в жодного елемента окремо [67]; стандарт життєвого циклу визначає її як поєднання взаємодійних елементів, організованих для досягнення визначеної мети [192]. **Системний підхід** — настанова

розглядати об'єкт як систему; **системний аналіз** — сукупність процедур, що реалізують цю настанову для конкретної слабкоструктурованої задачі. Робочий апарат утворюють шість понять.

Межа системи.
Рішення про те, що належить системі, а що — середовищу. Межа не «існує», її *проводить дослідник*: якщо користувач опинився поза межею системи захисту, соціотехнічні атаки стають методологічно невидимими.

Емерджентність.
Властивості системи, не виведені з властивостей елементів: стійкість інфраструктури, живучість, спостережність. Їх не можна дослідити, аналізуючи компоненти поодинці.

Декомпозиція й агрегування.
Розклад цілого на підсистеми за обраною основою (функційною, структурною, життєвого циклу) і зворотне об'єднання. Коректна декомпозиція задовольняє покриття (підсистеми вичерпують ціле) і мінімальність взаємодії.

Ієрархія.
Рівні з власними закономірностями: пакет, сесія, застосунок, організація. Перенесення закономірності між рівнями без обґрунтування — методологічна помилка.

Зворотні зв'язки.
Від'ємний зв'язок стабілізує систему, додатний — підсилює відхилення: посилення правил → зростання хибних тривог → зниження уваги операторів → пропуск справжньої атаки [227].

Мета й критерій.
Без явно названого критерію ефективності «системний аналіз» перетворюється на опис.

Системний аналіз розгортається як упорядкована процедура, кожен етап якої має перевірюваний результат: (1) формулювання проблеми й мети — що не влаштовує, для кого, за яким показником (розділ 3); (2) визначення межі системи й переліку стейкхолдерів; (3) декомпозиція з побудовою структури підсистем і зв'язків (результат — схема, а не перелік); (4) побудова моделі та її V&V (розд. 4.4); (5) формування множини альтернатив; (6) вибір критеріїв і оцінювання альтернатив, зокрема експертне (розд. 4.6) і багатокритеріальне (розд. 4.8); (7) аналіз чутливості й ризиків; (8) рекомендації, впровадження й оцінювання наслідків із поверненням до п. 1 за зміни умов.

П. Чекленд показав, що для задач із людським чинником жорстка послідовність етапів працює погано, і запропонував «м'яку» методологію, у якій

сама постановка задачі є предметом узгодження між учасниками [84]: чим більше в системі організаційних елементів, тим раніше слід узгоджувати критерії.

Межі застосовності. Системний аналіз дає хибний результат, якщо межу системи проведено так, що суттєві чинники опинилися в середовищі; якщо декомпозицію виконано за основою, що розриває механізм явища; якщо критерій ефективності сформульовано після отримання результатів. Окремо слід назвати найпоширеніше зловживання: декларацію «застосовано системний підхід» без жодної з перелічених процедур — це риторична фігура, а не метод.

4.6 Методи експертного оцінювання

Експертні методи застосовують, коли пряме вимірювання неможливе або невиправдано дороге, а рішення все одно треба обґрунтувати: пріоритизація загроз, оцінювання критичності активів, добір критеріїв. Це повноцінні наукові методи за умови, що процедура формалізована, а узгодженість суджень виміряна; без цього «експертне оцінювання» є думкою керівника, оформленою в таблицю.

4.6.1 Анкетування, інтерв'ю, добір експертів

Анкетування — письмове опитування за наперед сформульованим переліком запитань: дає стандартизовані дані ціною втрати контексту. **Інтерв'ю** — усне опитування; напівструктуроване є основним якісним методом в ІТ, коли треба з'ясувати, *чому* практики чинять усупереч рекомендаціям. *Межі:* обидва методи вимірюють *висловлювання про* поведінку, а не поведінку; діють ефект соціальної бажаності, вплив формулювання запитання й зсув добору. Опитування за участю людей потребують інформованої згоди й захисту персональних даних [279, 33] (розділ 10).

Групу експертів (7–15 осіб) формують не за посадами, а за критеріями: предметний досвід, підтверджений публікаціями або практикою; незалежність від результату; неперетинання джерел знання. Якщо x_{ij} — оцінка i-го експерта за j-м об'єктом, а k_i — коефіцієнт компетентності ($\sum_i k_i = 1$), то групова оцінка та її розкид

$$\bar{x}_j = \sum_{i=1}^{m} k_i\, x_{ij}, \qquad s_j^2 = \frac{\sum_{i=1}^{m} k_i \big(x_{ij} - \bar{x}_j\big)^2}{1 - \sum_{i=1}^{m} k_i^2}, \qquad V_j = \frac{s_j}{\bar{x}_j}. \tag{4.2}$$

Обидві величини мають бути узгоджені за схемою зважування: якщо середнє зважене, то й дисперсію обчислюють зваженою, зі знаменником $1 - \sum_i k_i^2$ (аналог поправки Бесселя для ваг, що сумуються до одиниці). Змішувати схеми — брати зважене середнє й незважену суму квадратів —

не можна: така оцінка не є незміщеною ні для однієї з них, і коефіцієнт варіації втрачає інтерпретацію. За рівних компетентностей ($k_i = 1/m$) вираз зводиться до звичайного $s_j^2 = \frac{1}{m-1}\sum_i (x_{ij} - \bar{x}_j)^2$.

Коефіцієнт варіації $V_j \le 0{,}3$ умовно вважають ознакою прийнятної узгодженості. Для матриць попарних порівнянь індивідуальні судження агрегують *геометричним* середнім $a_{jl} = \left(\prod_{i=1}^{m} a_{jl}^{(i)}\right)^{1/m}$: лише воно зберігає обернену симетрію $a_{lj} = 1/a_{jl}$ [288].

4.6.2 Ранжування й узгодженість експертів

Ранжування — упорядкування об'єктів за спаданням важливості; результат — порядкова шкала. Узгодженість m експертів, які ранжують n об'єктів, вимірюють **коефіцієнтом конкордації Кендалла** [202]:

$$W = \frac{12\,S}{m^2\,(n^3 - n)}, \qquad S = \sum_{j=1}^{n}\Big(R_j - \frac{m(n+1)}{2}\Big)^2, \qquad \chi^2 = m\,(n-1)\,W, \tag{4.3}$$

де R_j — сума рангів j-го об'єкта. Значення $W \in [0;1]$: нуль — повна неузгодженість, одиниця — повний збіг ранжувань; значущість перевіряють за χ^2 із $n-1$ ступенями вільності (розділ 8). Орієнтир: за $W < 0{,}3$ групову оцінку використовувати не можна, треба шукати причину розбіжності (різне розуміння критерію, неоднорідна група, некоректне формулювання).

4.6.3 Метод Delphi

Delphi — процедура структурованого заочного узгодження експертних суджень, розроблена в RAND Corporation [106] і систематизована в [215]. Її визначають чотири ознаки: *анонімність* експертів один для одного (усуває тиск авторитету), *ітеративність*, *контрольований зворотний зв'язок* (між турами повідомляють групову статистику й аргументи, а не персональні позиції) і *статистична агрегація* відповідей [167]. Процедура: (1) формулювання питання й добір експертів; (2) тур 1 — індивідуальні оцінки з письмовим обґрунтуванням; (3) надсилання учасникам медіани, міжквартильного розмаху та анонімних аргументів; (4) тур 2 — перегляд оцінок, причому експерт, чия оцінка лишається поза міжквартильним розмахом, обґрунтовує її окремо; (5) повторення до виконання критерію збіжності.

Критерій зупинки визначають *до* початку: найчастіше поєднують *стабільність* (оцінки перестали змінюватися між турами) і *консенсус* (розкид достатньо малий) [105]:

$$\mathrm{IQR}^{*}_{(r)} = \frac{Q_3^{(r)} - Q_1^{(r)}}{\mathrm{Me}^{(r)}} \le \theta \quad \wedge \quad \left|\mathrm{Me}^{(r)} - \mathrm{Me}^{(r-1)}\right| \le \Delta, \tag{4.4}$$

де θ і Δ — наперед узгоджені пороги (типово $\theta = 0{,}25$, Δ — один крок шкали). Разом із порогами в протоколі *обов'язково* фіксують конвенцію обчислення квартилів і те, чи є порівняння строгим: від цього безпосередньо залежить, у якому турі спрацює правило зупинки. У посібнику всюди застосовано тип 7 (лінійна інтерполяція за позицією $h = (n-1)p+1$; стандартна поведінка `quantile()` у R) — див. розділ 8. Альтернативний критерій — частка експертів, чия оцінка потрапляє в задану околицю медіани, не менша за 70–80 %. Огляд 100 Delphi-досліджень показав, що означення консенсусу різняться радикально й часто взагалі не наводяться, тому критерій збіжності обов'язково описують у розділі «Методи» [114].

Межі застосовності. Delphi дає *узгодженість*, а не *істинність*: збіжна група може бути одностайно помилковою, і процедура це лише замаскує; ризик посилюється за однорідної групи. Метод непридатний там, де можливе пряме вимірювання: його результат є гіпотезою або пріоритизацією, а не встановленим фактом.

4.6.4 Метод аналізу ієрархій (AHP)

Метод аналізу ієрархій (Analytic Hierarchy Process, AHP) Т. Сааті розкладає задачу вибору на ієрархію «мета — критерії — альтернативи» й отримує ваги елементів кожного рівня з матриці попарних порівнянь [287, 289]. Ключова ідея: людині значно легше порівняти два об'єкти, ніж приписати абсолютні ваги одразу всім; обмеження на розмір групи (7 ± 2) випливає з обсягу оперативної пам'яті людини [232]. Судження виражають у дев'ятибальній **шкалі Сааті** (табл. 4.3); матриця $A = [a_{jl}]$ обернено симетрична: $a_{jl} = 1/a_{lj}$, $a_{jj} = 1$.

Таблиця 4.3 — Шкала Сааті для попарних порівнянь (за [289, 288])

a_{jl}	**Ступінь переваги**	**Тлумачення**
1	Однакова важливість	Обидва елементи однаково впливають на мету
3	Помірна перевага	Судження дають незначну перевагу j над l
5	Суттєва перевага	Перевага j сильна й підтверджена досвідом
7	Значна перевага	Домінування j підтверджене практикою
9	Абсолютна перевага	Найвищий ступінь переконливості переваги
2, 4, 6, 8	Проміжні значення	Компроміс між сусідніми градаціями
$1/a_{jl}$	Обернені оцінки	Якщо j переважає l на a, то l переважає j на $1/a$

Вектор вагів $\mathbf{w}$ — нормований головний власний вектор матриці; на практиці його точно апроксимують нормованим середнім геометричним рядків:

$$A\mathbf{w} = \lambda_{\max}\mathbf{w}, \qquad w_j = \frac{\left(\prod_{l=1}^{n} a_{jl}\right)^{1/n}}{\sum_{k=1}^{n}\left(\prod_{l=1}^{n} a_{kl}\right)^{1/n}}, \qquad \sum_{j=1}^{n} w_j = 1. \tag{4.5}$$

Судження людини не бувають ідеально транзитивними, тому відхилення $\lambda_{\max}$ від n править за міру неузгодженості: обчислюють **індекс узгодженості** CI і **відношення узгодженості** CR — відношення CI до середнього CI випадково заповнених обернено симетричних матриць того самого порядку (**випадковий індекс** RI, табл. 4.4):

$$\lambda_{\max} = \frac{1}{n}\sum_{j=1}^{n}\frac{(A\mathbf{w})_j}{w_j}, \qquad CI = \frac{\lambda_{\max} - n}{n-1}, \qquad CR = \frac{CI}{RI(n)}. \tag{4.6}$$

Таблиця 4.4 — Випадковий індекс $RI(n)$ (за [289, 55])

n	2	3	4	5	6	7	8	9	10
RI (Saaty)	0	0,58	0,90	1,12	1,24	1,32	1,41	1,45	1,49
RI (Alonso–Lamata)	0	0,52	0,88	1,11	1,25	1,34	1,41	1,45	1,49

Умова прийнятності — $CR \le 0{,}1$. Якщо $CR > 0{,}1$, матрицю *не* «виправляють» підгонкою чисел: повертаються до експерта, показують найсуперечливішу трійку суджень і просять переглянути саме її.

Межі застосовності. AHP критикують за **інверсію рангів** (rank reversal): додавання або вилучення альтернативи може змінити порядок решти [65, 126]; часткові розв'язки — режим ideal mode або перехід до рейтингового оцінювання за наперед визначеними шкалами. Друге обмеження: $CR \le 0{,}1$ свідчить лише про внутрішню несуперечливість суджень, а не про їх правильність — послідовно помилковий експерт дасть $CR = 0$. Третє: кількість порівнянь зростає як $n(n-1)/2$, тож на кожному рівні ієрархії має бути не більш ніж 7–9 елементів.

4.6.5 Ранжувальні процедури групового вибору

Коли експерти дають упорядкування, а не числові оцінки, потрібне правило агрегації. **Метод Борда** приписує альтернативі, що посіла r-те місце з n, $n - r$ балів і підсумовує бали за всіма експертами: враховує всю інформацію про порядок, але чутливий до появи нових альтернатив. **Критерій Кондорсе** визнає переможцем альтернативу, що перемагає кожну іншу в попарному голосуванні: стійкіший, але переможець може не існувати —

парадокс Кондорсе, коли колективна перевага циклічна ($A \succ B \succ C \succ A$) попри транзитивність індивідуальних переваг. Теорема К. Ерроу про неможливість показує, що жодна процедура агрегації впорядкувань не задовольняє одночасно всіх природних вимог [61]. Звідси правило: обране правило агрегації *називають і обґрунтовують*, а стійкість результату перевіряють, застосувавши два різні правила.

4.7 Кейс: пріоритизація загроз методами Delphi й AHP

4.7.1 Постановка й етап Delphi

Дослідницька група університету разом з оператором розумної будівлі визначає черговість упровадження контрзаходів. Прямого вимірювання ризику немає: статистики інцидентів на цій інфраструктурі недостатньо, а оцінювання ймовірностей за методикою [246] все одно спирається на судження. Розглядають чотири класи загроз: T_1 — компрометація облікових даних адміністратора шлюзу; T_2 — експлуатація відомих вразливостей мікропрограми; T_3 — відмова в обслуговуванні каналу керування; T_4 — витік телеметрії незахищеним каналом. Залучено $m = 7$ експертів (двоє інженерів оператора, троє дослідників кафедри, двоє зовнішніх аудиторів), які оцінюють значущість кожної загрози за шкалою 1–9 з письмовим обґрунтуванням. Пороги (4.4) узгоджено наперед: $\theta = 0{,}25$, $\Delta = 1$. Динаміку для загрози T_3 подано в табл. 4.5.

Таблиця 4.5 — Збіжність оцінок загрози T_3 за турами Delphi ($m = 7$, шкала 1–9; квартилі — тип 7)

Тур	Оцінки	Me	Q_1	Q_3	IQR*	Частка ± 1
1	3, 4, 5, 6, 7, 8, 9	6	4,5	7,5	0,50	43 %
2	4, 5, 5, 6, 6, 7, 8	6	5,0	6,5	0,25	71 %
3	5, 5, 6, 6, 6, 6, 7	6	5,5	6,0	0,08	100 %

Уже після другого туру $\mathrm{IQR}^* = 1{,}5/6 = 0{,}25$, тобто *рівно* на порозі $\theta = 0{,}25$, а $|\mathrm{Me}^{(2)} - \mathrm{Me}^{(1)}| = 0 \le 1$. Оскільки протокол задав нестроге порівняння ($\le \theta$), формально обидві умови (4.4) виконано вже тут, і процедуру можна було зупиняти. Група провела ще один, заздалегідь передбачений підтверджувальний тур, після якого $\mathrm{IQR}^* = 0{,}5/6 \approx 0{,}08$ — уже з очевидним запасом.

Чому цей приклад тут. Якби квартилі обчислювали за типом 6 ($h = (n+1)p$, поведінка SPSS і Minitab), для другого туру вийшло б $Q_1 = 5$, $Q_3 = 7$ і $\mathrm{IQR}^* = 2/6 \approx 0{,}33 > 0{,}25$ — зупинка не спрацювала б. Обидві конвенції коректні; недопустима саме непозначена зміна методу, бо правило зупинки Delphi виявляється чутливим до неї. Тому конвенцію квартилів і

знак порівняння оголошують у протоколі до першого туру, а в статті — у розділі «Методи».

Паралельно експерти ранжували загрози. Суми рангів: $R_1 = 9$, $R_2 = 13$, $R_3 = 22$, $R_4 = 26$ (контроль: $\sum R_j = 70 = m\,n(n+1)/2$). Середнє $m(n+1)/2 = 17{,}5$, тому за (4.3)

$$S = 8{,}5^2 + 4{,}5^2 + 4{,}5^2 + 8{,}5^2 = 185, \qquad W = \frac{12 \cdot 185}{7^2\,(4^3 - 4)} = \frac{2220}{2940} = 0{,}755.$$

Статистика $\chi^2 = m(n-1)W = 21 \cdot 0{,}755 = 15{,}86$ за трьох ступенів вільності перевищує критичне значення $7{,}81$ ($\alpha = 0{,}05$), отже, узгодженість групи статистично значуща й групове ранжування $T_1 \succ T_2 \succ T_3 \succ T_4$ можна використовувати далі.

4.7.2 Етап AHP

Delphi дає порядок, але не ваги, потрібні для розподілу бюджету. Тому проводять попарні порівняння за шкалою табл. 4.3; індивідуальні судження агреговано геометричним середнім. Отриману матрицю A та обчислення подано в табл. 4.6.

Таблиця 4.6 — Матриця попарних порівнянь загроз і обчислення вагів за (4.5)

	T_1	T_2	T_3	T_4	r_j	w_j	$(A\mathbf{w})_j / w_j$
T_1	1	3	5	7	3,201	0,564	4,129
T_2	1/3	1	3	5	1,495	0,263	4,100
T_3	1/5	1/3	1	3	0,669	0,118	4,104
T_4	1/7	1/5	1/3	1	0,312	0,055	4,135
Сума					5,678	1,000	—

Тут $r_j = \left(\prod_l a_{jl}\right)^{1/4}$; наприклад, $r_1 = \sqrt[4]{1 \cdot 3 \cdot 5 \cdot 7} = \sqrt[4]{105} = 3{,}201$ і $w_1 = 3{,}201/5{,}678 = 0{,}564$. Добуток $A\mathbf{w}$ для першого рядка дорівнює $1{\cdot}0{,}564 + 3{\cdot}0{,}263 + 5{\cdot}0{,}118 + 7{\cdot}0{,}055 = 2{,}328$, звідки $2{,}328/0{,}564 = 4{,}129$. Усереднення четвірки відношень дає $\lambda_{\max} = (4{,}129 + 4{,}100 + 4{,}104 + 4{,}135)/4 = 4{,}117$, а за (4.6)

$$CI = \frac{4{,}117 - 4}{3} = 0{,}039, \qquad CR = \frac{0{,}039}{0{,}90} = 0{,}043 \le 0{,}1.$$

Матриця узгоджена, тож ваги прийнятні: T_1 — 56 %, T_2 — 26 %, T_3 — 12 %, T_4 — 6 % бюджету контрзаходів. Порядок той самий, що й за Delphi, але ваги показують, що T_1 важливіша за T_2 більш ніж удвічі — висновок, якого ранжування не давало. Подібну схему попарних порівнянь використано в українській методиці розрахунку рівня кіберзахисту об'єктів критичної

інфраструктури [50]; альтернативний підхід — баєсівські мережі з аудитом зрілості — описано в [48].

Що треба показати в дисертації. Склад групи й критерії добору експертів; анкету; критерій збіжності, визначений *до* процедури; таблицю турів; W і його значущість; повну матрицю попарних порівнянь; $\lambda_{\max}$, CI, CR; аналіз чутливості вагів до зміни одного судження на одну градацію. Без останнього пункту результат виглядає точнішим, ніж є насправді.

4.8 Кваліметричні, оптимізаційні та багатокритеріальні методи

Кваліметрія — кількісне оцінювання якості складних об'єктів через ієрархію одиничних, комплексних та інтегральних показників з нормуванням і зважуванням. Показники f_i зводять до безрозмірного вигляду $\tilde f_i \in [0;1]$, після чого застосовують **згортку** — адитивну $S = \sum_i w_i \tilde f_i$, мультиплікативну $S = \prod_i \tilde f_i^{\,w_i}$ або мінімаксну $S = \min_i w_i \tilde f_i$. Вибір згортки змінює результат, і найчастіше його описують неправильно: мовляв, адитивна згортка допускає компенсацію (низька захищеність «окупається» високою швидкодією), а мультиплікативна — ні. Насправді мультиплікативна згортка теж компенсаційна, вона лише змінює *характер* компенсації. Контрприклад: за $w_1 = w_2 = 0{,}5$ набори $(0{,}25;\ 1)$ і $(0{,}5;\ 0{,}5)$ дають той самий результат, бо $\sqrt{0{,}25 \cdot 1} = \sqrt{0{,}5 \cdot 0{,}5} = 0{,}5$: падіння першого показника вчетверо повністю компенсоване зростанням другого. Різниця з адитивною згорткою в тому, що «ціна» компенсації зростає в міру наближення показника до нуля: щоб урятувати $\tilde f_1 = 0{,}01$, потрібне вже недосяжне $\tilde f_2$. Некомпенсаційною є лише мінімаксна згортка (результат визначає найгірший показник), і навіть вона не боронить від низького значення, якщо всі показники низькі.

Висновок практичний: для обов'язкових вимог безпеки згортка будь-якого типу не є гарантією, тому поряд із нею вводять *явні порогові обмеження* $\tilde f_i \ge \tilde f_i^{\min}$ для кожного критичного показника, причому альтернатива, що порушує хоча б одне з них, вибуває з розгляду незалежно від значення S. Вибір згортки й пороги обґрунтовують явно.

Коли критерії суперечливі, коректним первинним результатом є не одна «найкраща» альтернатива, а множина недомінованих (**Парето-оптимальних**) розв'язків [130]: за максимізації всіх критеріїв $f_1, \dots, f_p$

$$X^* = \bigl\{ x \in X \mid \nexists\, y \in X :\ \forall i\ f_i(y) \ge f_i(x) \ \wedge\ \exists j\ f_j(y) > f_j(x) \bigr\}. \tag{4.7}$$

Множина X^* не залежить від вагів і є об'єктивною частиною задачі; усе, що звужує її до однієї точки, спирається на *суб'єктивні* переваги. Тому спершу будують фронт Парето, а вже потім застосовують правило вибору — і показують, як зміна вагів рухає вибрану точку вздовж фронту.

Поширене таке правило — **TOPSIS**: альтернативи впорядковують за близькістю до ідеального розв'язку A^+ (найкращі значення за всіма критеріями) і віддаленістю від антиідеального A^- [177, 64]. Якщо d_i^+ і d_i^- — зважені евклідові відстані i-ї альтернативи до цих точок, то показник близькості

$$C_i = \frac{d_i^-}{d_i^+ + d_i^-} \in [0;1], \qquad \text{краща альтернатива} \;\Leftrightarrow\; \max_i C_i. \tag{4.8}$$

Межі застосовності. Усі ці методи чутливі до нормування й до вагів; TOPSIS, як і AHP, допускає інверсію рангів після вилучення альтернативи. Жоден із них не створює інформації: якщо ваги взято «з міркувань здорового глузду», точність трьох знаків після коми в підсумковому індексі є фікцією. Обов'язковий елемент звітування — аналіз чутливості.

4.9 Типові помилки застосування методів

Підміна методу технікою або інструментом.
«Методи дослідження: Python, Wireshark, MATLAB» — перелічено засоби, а не методи. Інструмент реалізує методику, методика конкретизує метод; рівні не взаємозамінні (розділ 1).

«Метод заради методу».
Застосування складного апарату там, де завдання розв'язується простіше: нейронна мережа для задачі, у якій порогове правило дає ту саму якість; AHP для вибору з двох альтернатив. Ознака — метод названо в меті дослідження, а не в засобах.

Некоректна аналогія.
Перенесення моделі за зовнішньою подібністю без спільного механізму.

Узагальнення поза межами вибірки.
Висновок про клас об'єктів за вибіркою, що його не репрезентує; окремий випадок — екстраполяція моделі за межі області валідації (розд. 4.4).

Підміна пояснення описом.
Докладна характеристика явища подається як відповідь на запитання «чому»: «зростання частки хибних тривог зумовлене зростанням їх частки в нічні години» — це уточнення опису, а не механізм.

Cherry-picking методу.
Добір методу, шкали, показника чи статистичного критерію після перегляду результатів — так, щоб висновок був бажаним. Протидія — фіксація плану аналізу до отримання даних і звітування про всі виконані перевірки (розділи 8, 9).

Підсумкова табл. 4.7 придатна як шаблон підрозділу «Методи дослідження» власної роботи: рядок, у якому не заповнюється стовпець меж, свідчить, що метод дібрано без розуміння його застосовності.

Таблиця 4.7 — Метод — вхід — вихід — межі

Метод	Вхід	Вихід	Межі
Спостереження	Протокол, засіб реєстрації	Ряд даних, опис	Немає причинності; ефект спостерігача
Вимірювання	Величина, шкала, засіб	Значення з невизначеністю	Операції, недопустимі для шкали
Експеримент	План, змінні, базові лінії	Перевірена різниця	Хибна операціоналізація; зовнішня валідність
Моделювання	Припущення, дані для калібрування	Модель і прогноз	Область валідації; екстраполяція
Системний аналіз	Мета, межа системи, критерій	Структура, альтернативи	Межу проведено так, що чинник — у середовищі
Delphi	Анкета, група, критерій збіжності	Узгоджена оцінка	Узгодженість $\neq$ істинність
AHP	Ієрархія, попарні порівняння	Ваги, CR	Інверсія рангів; CR не доводить правильності

4.10 Висновки до розділу

Методи класифікують за трьома незалежними осями — рівнем спільності, етапом пізнання й характером даних. Емпіричні методи дають дані, але кожен має жорсткі межі: спостереження не встановлює причинності, вимірювання втрачає зміст за недопустимих для шкали операцій, порівняння знецінюється неназваною базовою лінією, експеримент не рятує від хибної операціоналізації. У теоретичних методів аналіз безсилий перед емерджентністю, індукція — поза межами вибірки, аналогія — без спільного механізму, формалізація — без предмета.

Моделювання потребує трьох обов'язкових тверджень: для якого завдання модель побудовано, у якій області її валідовано й які припущення відкинуто. Верифікація і валідація — різні перевірки, а перенесення висновку за межі області валідації є помилкою екстраполяції, хибність якої неможливо виявити всередині самої моделі. Системний підхід працює лише як послідовність процедур із явно проведеною межею системи й наперед названим критерієм.

Експертні методи є науковими за умови формалізованої процедури й виміряної узгодженості: коефіцієнт конкордації (4.3) для ранжувань, критерій збіжності (4.4) для Delphi, відношення узгодженості (4.6) для AHP. Кейс пріоритизації загроз дав повний розрахунок: $W = 0{,}755$ за $\chi^2 = 15{,}86$; ваги $(0{,}564;\ 0{,}263;\ 0{,}118;\ 0{,}055)$ за $\lambda_{\max} = 4{,}117$, $CI = 0{,}039$, $CR = 0{,}043$. За багатокритеріального вибору об'єктивною частиною задачі є множина Парето (4.7), а звуження її до одного розв'язку спирається на суб'єктивні

переваги й потребує аналізу чутливості. Усі типові помилки застосування методів усуваються однією вимогою: для кожного методу називати умови, за яких його висновки недійсні.

Питання для самоперевірки

1. Назвіть три осі класифікації методів і схарактеризуйте за ними один метод власного дослідження.
2. Чим метод відрізняється від методики та інструмента? Наведіть приклад усіх трьох рівнів для однієї процедури вашої роботи.
3. Що саме відрізняє експеримент від спостереження і чому лише перший дає підстави для причинних висновків?
4. Порівняйте лабораторний, натурний, обчислювальний, польовий і природний експерименти за внутрішньою та зовнішньою валідністю.
5. Чому усереднення балів CVSS по активах є методологічною помилкою? Який тип шкали тут порушено?
6. Наведіть приклад емерджентної властивості у вашій предметній області й поясніть, чому її не можна дослідити аналізом компонентів.
7. У чому різниця між верифікацією й валідацією моделі? Що таке область дійсності моделі та як виявляється помилка екстраполяції?
8. Які шість понять утворюють робочий апарат системного аналізу? Як вибір межі системи змінює перелік досліджуваних загроз?
9. Чому індивідуальні матриці попарних порівнянь агрегують геометричним, а не арифметичним середнім?
10. Поясніть зміст $\lambda_{\max}$, CI і CR у (4.6). Чому $CR = 0$ не означає, що оцінка правильна?
11. Які дві умови утворюють критерій зупинки Delphi (4.4) і чому їх визначають до початку процедури?
12. Що показує коефіцієнт конкордації Кендалла і як перевіряють його значущість?
13. Аналітичне: чому множина Парето є об'єктивною частиною багатокритеріальної задачі, а згортка — суб'єктивною?
14. Дослідницьке: назвіть метод, який ви застосовуєте, і сформулюйте умови, за яких він дасть систематично хибний результат саме на вашому об'єкті.

Практичні завдання

1. *Паспорт методів дослідження.* Складіть для власної роботи таблицю за зразком табл. 4.7: для кожного з 4–6 методів зазначте вхід, вихід і межі застосовності та схарактеризуйте метод за трьома осями рис. 4.1. Позначте рядки, у яких стовпець меж заповнити не вдалося.
2. *Обчислення вагів за AHP.* Експерти порівняли критерії добору засобу виявлення вторгнень і дали матрицю $a_{12}=3$, $a_{13}=7$, $a_{14}=9$, $a_{23}=3$, $a_{24}=5$, $a_{34}=3$ (решта — обернені). Обчисліть r_j, w_j, $\lambda_{\max}$, CI і CR за (4.5) і (4.6). Чи прийнятна матриця? Змініть a_{14} на 5 і покажіть, як змінилися ваги й CR.
3. *Узгодженість експертів.* Шестеро експертів ($m=6$) проранжували п'ять контрзаходів ($n=5$) від 1 (найважливіший) до 5; зв'язаних рангів немає:

Експерт	C_1	C_2	C_3	C_4	C_5
E_1	1	2	3	4	5
E_2	1	3	2	5	4
E_3	2	1	3	4	5
E_4	1	2	4	3	5
E_5	2	1	3	5	4
E_6	1	3	2	4	5
R_j	8	12	17	25	28

   Перевірте контрольну суму $\sum_j R_j = m\,n(n+1)/2$, обчисліть W за (4.3) і перевірте значущість за $\chi^2 = m(n-1)W$ з $n-1$ ступенями вільності. Поясніть, які дії потрібні за $W<0{,}3$, і як змінилася б процедура за наявності зв'язаних рангів.
4. *Протокол Delphi.* Розробіть протокол для одного питання власного дослідження: анкета, критерії добору 7–9 експертів, правило зворотного зв'язку, критерій збіжності за (4.4) з обґрунтуванням θ і Δ, межа кількості турів, план звітування результату.
5. *Модель і її межі.* Для моделі, яку ви будуєте, запишіть три обов'язкові речення (завдання, область валідації, відкинуті припущення), сформулюйте критерій адекватності за (4.1) з обґрунтуванням ε і складіть перелік процедур верифікації та валідації. Окремо візьміть дві статті за темою дослідження з фахового видання категорії «Б» або зі Scopus і перевірте, чи названо в них межі застосованих методів.

РОЗДІЛ 5

Дослідницькі методології в інформаційних технологіях і кібербезпеці

Розділ 4 дав загальнонауковий інструментарій. Цей розділ переходить на конкретно-науковий рівень: у галузі F/12 результатом дослідження часто є не твердження про світ, а *артефакт* — метод, алгоритм, архітектура, інструмент, набір даних. Для таких результатів вироблено власні строгі методології, і саме за ними рецензенти ухвалюють рішення про публікацію. Для кожної методології подано умови застосовності, артефакти, критерії оцінювання рецензентами й типові підстави для відхилення; наскрізний кейс — дослідження стійкості моделі виявлення вторгнень.

5.1 Науковий результат в інформаційних технологіях

5.1.1 Сім типів результату

Питання «що саме тут є науковим результатом» в IT має більше відповідей, ніж у природничих науках; практично розрізняють сім типів. **Теорія** пояснює клас явищ і дає перевірні передбачення; доказ — виведення й узгодження з даними. **Метод** — процедура з названими межами; доказ — перевага над наявними методами за наперед визначеним критерієм, а не новизна ідеї. **Алгоритм** — метод, поданий формально; доказ — доведення коректності й оцінка складності. **Архітектура** — компоненти, інтерфейси, обмеження; доказ — досяжність атрибутів якості й зафіксовані компроміси. **Інструмент** — реалізація для використання іншими; рівні оцінювання формалізовано бейджами ACM [52]. **Набір даних** є результатом, якщо має схему, ліцензію, ідентифікатор і опис збирання [346]. **Емпірична закономірність** — стійка залежність, виявлена вимірюванням; оцінюють обсяг вибірки, контроль змішувальних чинників і відтворюваність. Найчастіша вада робіт галузі — розбіжність між заявленим і фактичним типом: заявлено «розробити метод», а перевірено одну реалізацію на одному наборі даних.

5.1.2 Класифікація статей за Shaw і Wieringa

М. Шоу показала, що ймовірність прийняття рукопису визначається узгодженістю трійки «тип питання — тип результату — тип валідації» [303]. Валідація буває аналітичною (доведення), оцінювальною, досвідною, демонстраційною та переконувальною; остання — «ми переконані, що підхід корисний» — найслабша й найтиповіша для відхилених робіт.

Р. Вірінга запропонував класифікацію, що стала стандартом де-факто для рецензування [344, 343]. Її основа — розрізнення **дослідницьких питань** (knowledge questions), на які відповідають твердженням про світ, і **проєктних задач** (design problems), які розв'язують побудовою артефакту. Звідси два класи статей: *проблемно-орієнтовані* — дослідження оцінювання наявного засобу в практиці, вивчення проблеми, філософські; і *рішення-орієнтовані* — пропозиція рішення та дослідження перевірки (validation research), тобто вивчення властивостей ще не впровадженого рішення в лабораторних умовах. Найпоширеніша помилка — подати пропозицію рішення як дослідження перевірки (*solution proposal without validation*). За класифікацією ABC [313] жодна методологія не максимізує одночасно узагальненість, реалізм і точність контролю.

5.1.3 Технологічне правило як форма результату

Прикладний результат Вірінга пропонує подавати **технологічним правилом** — твердженням, що зв'язує втручання з ефектом у заданому контексті [343]. Формально це четвірка

$$\mathcal{R} = \langle C,\ A,\ E,\ M \rangle : \quad \text{«у контексті } C \text{ застосування } A \text{ дає ефект } E \text{ через механізм } M\text{»}, \tag{5.1}$$

де C — умови (клас систем, навантаження), A — артефакт або втручання, E — вимірюваний ефект зі шкалою, M — механізм, що пояснює, *чому* ефект виникає. Без C результат заявлено ширше, ніж доведено; без E він неперевірний; без M ви маєте інженерне рішення, а не науковий результат.

5.2 Design Science Research

5.2.1 Парадигма і сім настанов

Design Science Research (DSR, **проєктувальне дослідження**) — парадигма, у якій знання здобувають *побудовою й оцінюванням артефакту*, призначеного розв'язати актуальну проблему практики. С. Марч і Ґ. Сміт протиставили її природничо-науковій парадигмі: природнича наука пояснює те, що є, **наука проєктування** (design science) створює те, чого немає [219].

> **Про український відповідник.** Усталеного перекладу термін ще не має, тому в літературі трапляються щонайменше три варіанти. *«Дослідження через проєктування»* — калька з *research through design*, а це *інша* традиція, що виросла в дизайнерських школах і не вимагає ані строгого оцінювання артефакту, ані внеску до бази знань; вживання цього варіанта для DSR змішує дві різні методології. *«Дизайн-орієнтоване дослідження»* хибує на те, що українське «дизайн» тяжіє

до візуального й промислового проєктування, тоді як артефактом DSR є метод, модель, алгоритм чи архітектура. Тут ужито *«проєктувальне дослідження»*: прикметник «проєктувальний» однозначно вказує на проєктування як діяльність, а не на проєкт як форму фінансування, і утворює послідовну пару з «наукою проєктування». У власному тексті достатньо один раз навести англійську назву з абревіатурою, а далі вживати обраний відповідник.

А. Гевнер зі співавторами формалізували парадигму [170] і дали сім настанов: (1) життєздатний *артефакт* як результат; (2) *релевантність* — артефакт розв'язує важливу нерозв'язану проблему практики; (3) *оцінювання* визначеними методами; (4) *внесок* в артефакт, підвалини проєктування або методологію оцінювання; (5) *строгість*; (6) *проєктування як пошук*; (7) *комунікація* результату. Найчастіше відхиляють через настанови 2 і 3.

5.2.2 Три цикли

DSR-проєкт структурується трьома циклами (рис. 5.1) [169]. **Цикл релевантності** пов'язує середовище застосування з дослідженням: звідти надходять вимоги й критерії приймання, назад повертається артефакт для польового випробування. **Цикл строгості** пов'язує дослідження з базою знань: звідти беруть теорії, методи й метрики, туди повертають нове знання — саме він відрізняє DSR від розробки, бо якщо з бази знань нічого не взято й нічого не додано, це консалтинг. **Цикл проєктування** — внутрішній цикл «побудова → оцінювання → уточнення», що триває до виконання критеріїв.

Рис. 5.1 — Три цикли Design Science Research (перемальовано за [169, 170])

5.2.3 Процес DSRM, артефакти, рівні внеску

Послідовність дій задає **DSRM** — процес із шести кроків [262]: (1) ідентифікація проблеми; (2) визначення цілей розв'язку; (3) проєктування й розроблення артефакту; (4) демонстрація на показовій задачі; (5) оцінювання — вимірювання того, наскільки артефакт досягає цілей кроку 2, з поверненням до кроку 3 за потреби; (6) комунікація. Принципово важливі *чотири точки входу*: з проблеми (problem-centered), з цілей (objective-centered), з наявного прототипу (design-centered) або з клієнтського контексту. Тож послідовність «спершу прототип, потім проблема» легітимна, але вимагає ретроспективно виконати кроки 1–2 із реальним обґрунтуванням. Крок 2 — найслабша ланка дисертацій: ціль «підвищити ефективність виявлення» не є ціллю, бо немає показника, базового рівня й порогу успіху. Робоче формулювання — «досягти повноти не нижче 0,95 за частки хибнопозитивних не більш ніж 0,01 на трафіку класу C»; лише воно робить крок 5 можливим.

Марч і Сміт виокремили чотири типи артефактів (табл. 5.1) [219]. Ортогонально до них розрізняють три рівні абстракції внеску: рівень 1 — інстанціація, рівень 2 — метод, модель чи конструкт для повторного використання, рівень 3 — **теорія проєктування** [161]; дисертація рівня 1 без піднесення результату до рівня 2 рідко проходить захист.

Таблиця 5.1 — Типи артефактів DSR

Артефакт	Що це	Приклад і спосіб оцінювання
Конструкт	Поняття й словник галузі	Таксономія технік порушника; повнота, взаємовиключність, узгодженість кодувальників
Модель	Зв'язки між конструктами	Модель загроз; адекватність, покриття сценаріїв, прогностичність
Метод	Процедура досягнення результату	Метод виявлення аномалій; порівняння з базовими лініями, вартість, межі
Інстанціація	Реалізація в середовищі	Прототип шлюзу; польове випробування, продуктивність

5.2.4 Оцінювання артефакту й два вкладені цикли

Найслабшою частиною DSR-робіт лишається оцінювання. **FEDS** — система координат для його планування [333]. *Вісь призначення:* **формувальне** (formative) оцінювання дає підстави для доопрацювання артефакту — його результат повертається у проєктування; **підсумкове** (summative) оцінювання дає підстави для судження про вже створений артефакт — про його придатність, якість і межі застосовності. *Вісь середовища:* **штучне** (artificial) — лабораторія, імітація, синтетичні дані: висока внутрішня валідність, низька зовнішня; **натуралістичне** (naturalistic) — реальні користувачі, задачі й системи: навпаки. Стратегій руху квадрантами чотири: *швидка й проста*;

людський ризик (рано переходити до натуралістичного оцінювання); *технічний ризик* (тривале штучне, натуралістичне наприкінці) — доречна для кібербезпеки; *чиста технічна* (виключно штучне). Планування зводиться до питань: навіщо, коли, як і що саме оцінювати.

Не плутати з часом проведення. Поділ на **ex ante** (оцінювання до створення працездатного артефакту — за ескізом, специфікацією чи моделлю) і **ex post** (оцінювання артефакту в дії) — це окрема характеристика, а не друга вісь FEDS у цій редакції. Вона з призначенням корелює, але не збігається: підсумкове оцінювання концепції може бути ex ante, а формувальне випробування робочого прототипу — ex post.

Вірінга розділяє дві вкладені структури [343]. **Инженерний цикл** відповідає на проєктну задачу: дослідження проблеми → проєктування розв'язку → перевірка розв'язку → упровадження → оцінювання. **Цикл емпіричного дослідження** відповідає на дослідницьке питання: аналіз контексту → проєктування дослідження → перевірка плану → виконання → аналіз. Другий вкладається в перший на кроці «перевірка розв'язку»; типова вада — підміна цього кроку впровадженням, бо акт упровадження не є емпіричним доказом ефекту.

> **Коли обирати DSR.** Якщо питання — «як побудувати засіб, що в контексті C дає ефект E». *Рецензенти оцінюють:* наявність артефакту, релевантність проблеми, строгість оцінювання, внесок у базу знань. *Підстави для відхилення:* демонстрація замість оцінювання; відсутність базових ліній; артефакт рівня 1 без узагальнення; невизначені критерії успіху.

5.3 Контрольований експеримент в емпіричній інженерії

5.3.1 Змінні, об'єкти, суб'єкти, плани

Контрольований експеримент — дослідження, у якому дослідник систематично змінює одну чи кілька **незалежних змінних** (факторів) і вимірює вплив на **залежні змінні**, утримуючи решту умов сталими або рандомізуючи їх [350]. **Обробка** (treatment) — конкретне значення фактора; **об'єкт** — те, над чим виконують дії (програма, набір даних, конфігурація); **суб'єкт** — той, хто виконує дії (розробник, аналітик SOC); **змішувальна змінна** — чинник, що впливає на залежну змінну й корелює з фактором, роблячи причинний висновок хибним (досвід учасника, складність модуля).

Основні плани. *Один фактор, дві обробки* (міжгрупове порівняння): дві випадкові групи; просто, але потребує більшої вибірки. *Парний план* (within-subjects): кожен суб'єкт виконує обидві обробки, потужність вища; ціна — ефекти навчання й перенесення, які нейтралізують **балансуванням** порядку (половина одержує послідовність AB, половина — BA). *Більш ніж*

дві обробки: дисперсійний аналіз із поправкою на множинність [66]. *Факторний план* (2^k): кілька факторів одночасно, що виявляє *взаємодію*; якщо перевага алгоритму зникає за високого навантаження, звітувати лише «середній ефект» неправильно. *Блоковий план*: заважальний чинник фіксують як **блок** і рандомізують усередині блоків.

Рандомізація — єдиний механізм, що нейтралізує *невідомі* змішувальні змінні; балансування й блокування нейтралізують лише *відомі*. Тому «учасників розподілено за бажанням» знищує підстави для причинного висновку. В обчислювальному експерименті рандомізують порядок прогонів, початкові значення генераторів і розподіл вибірок [306].

5.3.2 Чотири типи загроз валідності

Класифікацію перенесено з методології соціальних наук [299] і адаптовано в [350]. **Внутрішня валідність** — чи справді спостережену зміну спричинила саме обробка (загрози: історія, добір, відсів, інструментарій, ефект навчання). **Зовнішня** — чи переносяться висновки на інші об'єкти, суб'єктів, час і середовища. **Конструктна** — чи вимірює показник той конструкт, про який ідеться в гіпотезі. **Статистична валідність висновку** — чи коректно застосовано апарат і чи достатньо даних. Згадка загрози без зазначення дії, якою її пом'якшено, — ритуал, а не методологія (табл. 5.2).

Таблиця 5.2 — Загрози валідності в ІТ-дослідженні (за [350, 299])

Тип	Приклад загрози в ІТ	Пом'якшення
Внутрішня	Оновлення бібліотеки між прогонами; власний метод налаштовано пошуком за сіткою, чужий — ні	Фіксація версій і образу середовища; однаковий бюджет налаштування; рандомізація прогонів
Зовнішня	Модель перевірено на одному наборі 2017 р.; учасники — студенти замість практиків	Кілька незалежних наборів; перевірка на пізніших даних; опис популяції
Конструктна	«Захищеність» вимірюють кількістю закритих портів	Явне означення конструкту; кілька показників; обґрунтування їх зв'язку з конструктом
Статистична	Порівняння за середнім без розкиду; перекриття згортків; множинні порівняння без поправки	Оцінка потужності; розмір ефекту й довірчі інтервали; поправка на множинність

Коли обирати контрольований експеримент. Питання має форму «чи спричиняє A зміну E за інших рівних умов». *Рецензенти оцінюють:* обґрунтування обсягу вибірки, рандомізацію, повноту опису для відтворення, розділ про загрози валідності. *Підстави для відхилення:* немає рандомізації; причинний висновок із кореляційних даних; замала вибірка; недоступні сирі дані.

5.4 Якісні, польові та змішані методології

5.4.1 Дослідження випадку

Дослідження випадку (case study) — емпіричне дослідження сучасного явища в його природному контексті, коли межа між явищем і контекстом нечітка й дослідник не контролює умов [285]. Це не «приклад застосування»: справжнє кейс-стаді має дослідницькі питання, визначену одиницю аналізу й стратегію збирання даних. За метою розрізняють *розвідувальне* (породити гіпотези), *описове*, *пояснювальне* та *поліпшувальне*.

Протокол дослідження випадку складають *до* збирання даних: мета й питання; одиниця аналізу; критерії добору випадків; джерела даних; план аналізу; етичні домовленості — згода, анонімізація, право організації переглянути текст [285, 286]. Його публікують як додаток: без нього читач не може оцінити, чи не добиралися дані під бажаний висновок. Головний механізм підвищення довіри — **тріангуляція**: одержання одного висновку з незалежних джерел — *за даними* (інтерв'ю, журнали, документи, код), *за дослідником* (незалежне кодування двома аналітиками, розділ 6), *за теорією* і *за методом*.

5.4.2 Опитування, дослідження дією, етнографія

Опитування (survey) — систематичне збирання даних від репрезентативної вибірки популяції для опису її характеристик або перевірки зв'язків [234]; застосовують до питань про поширеність практик: «яка частка українських ЗВО має задокументовану політику реагування на інциденти». Якість визначають означення цільової популяції й основи вибірки, спосіб її формування (випадковий, стратифікований, зручний — останній різко обмежує висновки), пілотування інструмента та звітування частки відповідей із аналізом зсуву неповернення. Найпоширеніша вада — анкета «серед знайомих фахівців» із узагальненням на галузь.

Дослідження дією (action research) — методологія, у якій дослідник разом із практиками свідомо втручається в реальну систему, щоб одночасно розв'язати проблему й здобути знання; цикл має п'ять фаз: діагностування, планування дії, виконання, оцінювання, специфікація знання [316]. Вона доречна, коли явище неможливо вивчати ззовні (побудова SOC у

ЗВО). Ключова вимога — заздалегідь розмежувати дослідницьку мету й мету втручання, інакше робота стає звітом про консультаційний проєкт.

Етнографія — тривале включене спостереження спільноти в її природному середовищі [302]; вона незамінна там, де кількісні дані оманливі (чому інженери обходять політику безпеки), а її внесок — не відсотки, а механізм.

5.4.3 Критерії якості й змішані методології

Якісні дослідження оцінюють за чотирма критеріями: **достовірність** (credibility) — відповідність висновків даним, забезпечена тріангуляцією й перевіркою тлумачень в учасників; **переносність** (transferability) — щільний опис контексту; **надійність** (dependability) — стабільність процедури, забезпечена протоколом та аудиторським слідом; **підтверджуваність** (confirmability) — збереження первинних матеріалів і цитування фрагментів.

Змішані методології (mixed methods) поєднують кількісний і якісний складники [101]: *послідовна пояснювальна* (вимір виявляє ефект, інтерв'ю пояснюють механізм); *послідовна розвідувальна* (якісний етап породжує конструкти, кількісний перевіряє їх); *паралельна конвергентна*. Дизайн виправданий лише тоді, коли складники відповідають на *різні* підпитання одного питання (рис. 5.2, табл. 5.3).

Рис. 5.2 — Вибір методології за типом дослідницького питання

5.5 Бенчмаркінг і відтворювані вимірювання

Бенчмаркінг — порівняльне вимірювання систем або методів на спільному, наперед визначеному наборі задач за спільними метриками. Це основна форма доказу в системних дослідженнях і водночас найлегша для непомітної маніпуляції [68, 174].

Таблиця 5.3 — Методологія — доказ — типова помилка

Методологія	Що є доказом	Типова помилка
DSR / DSRM	Артефакт і результат його оцінювання	Демонстрація замість оцінювання
Експеримент	Різниця між обробками з розміром ефекту	Немає рандомізації; замала вибірка
Бенчмаркінг	Показники на спільних наборах і навантаженнях	Слабка базова лінія; добір наборів
Кейс-стаді	Тріангульовані свідчення за протоколом	Приклад видають за кейс-стаді
Опитування	Оцінка частки з довірчим інтервалом	Зручна вибірка; зсув неповернення
Дослідження дією	Задокументовані цикли й рефлексія	Консультаційний звіт без знання
Етнографія	Щільний опис, польові нотатки	Узагальнення на галузь
Систем. огляд	Протокол, критерії, синтез [206, 257]	Наратив під виглядом огляду

Еталонний набір даних або навантаження має бути репрезентативним для контексту C з (5.1), а не просто доступним; добір набору «на якому метод працює найкраще» після перегляду результатів — різновид cherry-picking. Навантаження описують розподілом, а не середнім: саме хвости визначають поведінку системи під час атаки. *Метрики* фіксують до вимірювань: accuracy на незбалансованих даних вводить в оману, середня затримка приховує хвости (звітують квантилі p95, p99), а прискорення різних наборів агрегують геометричним середнім [68].

Базова лінія (baseline) має бути *сучасною* і *однаково ретельно налаштованою*. Порушення цих вимог дає **«солом'яне опудало»** (straw-man baseline): застарілий метод, чужа реалізація зі стандартними параметрами проти власної, відлагодженої пошуком за сіткою, або базова лінія, оцінена за іншим протоколом розділення вибірок. Протидія — однаковий і явно звітований бюджет налаштування для всіх методів. *Абляційне дослідження* (ablation study) — вилучення складників власного методу з повторним вимірюванням — відповідає на питання «що саме дає ефект»: якщо вилучення ключового компонента не змінює показників, механізм M у (5.1) названо неправильно.

Окрема проблема — **вимірювальний шум**: стан кешів, частота процесора, збирач сміття, фонові процеси, випадкова ініціалізація. Одиничний прогін не є вимірюванням. Кількість повторень n оцінюють через коефіцієнт варіації $\mathrm{CV} = s/\bar{x}$ і бажану відносну півширину довірчого інтервалу δ:

$$n \geq \left(\frac{t_{1-\alpha/2,\,n-1} \cdot \mathrm{CV}}{\delta} \right)^2, \qquad \bar{x} \pm t_{1-\alpha/2,\,n-1} \frac{s}{\sqrt{n}}. \tag{5.2}$$

За $\mathrm{CV} = 0{,}12$, $\delta = 0{,}05$ і $\alpha = 0{,}05$ потрібно щонайменше $n \approx (2{,}1 \cdot 0{,}12/0{,}05)^2 \approx 26$ повторень. Звітувати різницю в 3 %, маючи три прогони з розкидом 12 %, — статистично порожнє твердження [174].

5.6 Специфіка кібербезпекових досліджень

5.6.1 Модель порушника й експериментальні формати

У кібербезпеці твердження про захищеність не має сенсу без явно заданого супротивника. **Модель порушника** — формалізований опис припущень про його можливості: позиція щодо системи (зовнішній, внутрішній, у каналі), ресурси, знання (відкритий код, ключі, навчальна вибірка), дозволені дії і мета. **Модель загроз** — перелік загроз, виведений із моделі порушника й архітектури. Обидві є не вступною формальністю, а *частиною наукового результату*: змінивши модель порушника, можна дістати протилежний висновок про стійкість того самого механізму. Тому її формулюють *до* експериментів і записують переліком припущень, які можна порушити; «розглядається зловмисник високої кваліфікації» моделлю не є. Матриця ATT&CK [315] корисна як спільний словник технік, але формальної моделі не замінює.

Вправи червоної і синьої команд стають науковим експериментом лише тоді, коли мають план: фактори, залежні змінні (час до виявлення, час до стримування, частка виявлених технік), фіксовані сценарії, рандомізований порядок і незалежне оцінювання. Загроза внутрішній валідності — обізнаність синьої команди зі сценарієм; зовнішній — полігон, що не відтворює реальної інфраструктури.

Honeypot-дослідження — розгортання приманок для збирання даних про реальну атакувальну активність [154]. Описують рівень взаємодії, адресний простір, тривалість спостереження, відсіювання фонового сканування та заходи стримування: приманка не має стати плацдармом для атак на третіх осіб.

Вимірювальні дослідження в Інтернеті — активне сканування або пасивне спостереження (телескопи, потоки, DNS). Сканування всього адресного простору стало технічно доступним [124], що одразу поставило етичні питання: сканувати слід з ідентифікованих адрес зі сторінкою з поясненням мети, передбачити відмову (opt-out), обмежувати інтенсивність, не використовувати вразливостей і не публікувати даних, що ідентифікують вразливі системи.

5.6.2 Етика, розкриття вразливостей, dual-use

Menlo Report адаптує Белмонтські принципи до досліджень у сфері ІКТ [119, 120]: повага до осіб; *баланс користі й шкоди* — ідентифікація стейкхолдерів і зважування ризиків, зокрема для третіх осіб, які не є учасниками дослідження; справедливість; *повага до права й суспільного інтересу* (докладніше — у розділі 10).

Скоординоване розкриття вразливостей (coordinated vulnerability disclosure) — процедура, за якою дослідник повідомляє виробника, узгоджує строк усунення й лише потім оприлюднює деталі. ISO/IEC 29147 визначає, як організація приймає й оприлюднює інформацію про вразливості, ISO/IEC 30111 — як вона їх обробляє [190, 191]; практичні настанови подано в [176]. Типовий строк ембарго — 90 днів. Рукопис, у якому описано нову вразливість без згадки про процедуру розкриття, етичний комітет відхиляє незалежно від якості.

Дослідження подвійного призначення (dual-use) — ті, результат яких можна використати і для захисту, і для нападу. Діє принцип пропорційності: зважити приріст можливостей нападника від публікації проти приросту можливостей захисту; публікувати артефакти на рівні, достатньому для відтворення результату, але не готовий до застосування інструмент; розглянути поетапне оприлюднення. Етична частина роботи (ethics statement) має охоплювати: межі дозволених дій; правову підставу роботи із системою (договір, письмовий дозвіл власника, власний стенд); стейкхолдерів і можливі шкоди; заходи стримування; поводження з персональними даними [279, 33]; план розкриття вразливостей.

5.7 Дослідження з машинним навчанням

5.7.1 Розділення вибірок і витік даних

Модель оцінюють за здатністю узагальнювати, тому всі рішення, що спираються на дані, мають бути ізольовані від тестової вибірки: навчальна частина дає параметри, валідаційна — гіперпараметри, пороги й ознаки, тестова використовується *один раз*.

Витік даних (data leakage) — потрапляння до процесу навчання інформації, недоступної в момент застосування [200]: нормалізація або добір ознак на всьому наборі до розділення; ознаки, похідні від мітки; дублікати між навчальною й тестовою частинами; балансування класів до розділення. Багаторазове звертання до тестової вибірки з «підкручуванням» моделі — **data snooping** — формально не є витоком, але дає той самий наслідок: оцінка стає оптимістично зміщеною, а **перенавчання** переходить із моделі на дослідника.

Часова коректність розділення — критична для кібербезпеки вимога: навчальні дані мають передувати тестовим, бо модель не бачить майбутнього. Випадкове перемішування зразків, зібраних за кілька років, дає «неможливу» точність, яка зникає за коректного часового розділення [263]. Неприпустимо й формувати тестову вибірку з нереалістичним співвідношенням класів: базова частка атак у реальному трафіку на порядки нижча, ніж у еталонних наборах (*base rate fallacy* [310]).

5.7.2 Крос-валідація й коректне порівняння моделей

За обмежених даних використовують k**-кратну крос-валідацію**: набір D розбивають на k неперетинних згортків $D_1, \dots, D_k$, навчають модель h_{-i} на $D \setminus D_i$ й оцінюють на D_i:

$$\widehat{\mathrm{err}}_{\mathrm{CV}} = \frac{1}{k}\sum_{i=1}^{k} \mathrm{err}\big(h_{-i}, D_i\big), \qquad \hat{s}^2 = \frac{1}{k-1}\sum_{i=1}^{k}\big(\mathrm{err}_i - \widehat{\mathrm{err}}_{\mathrm{CV}}\big)^2. \tag{5.3}$$

Величина $\hat{s}^2/k$ *не* є незміщеною оцінкою дисперсії $\widehat{\mathrm{err}}_{\mathrm{CV}}$: підвибірки перетинаються, похибки err_i залежні, і наївний t-критерій систематично завищує значущість [115]. За r-кратного повторення k-кратної крос-валідації поправку дає виправлений t-критерій [238]:

$$t = \frac{\bar{d}}{\sqrt{\left(\dfrac{1}{n} + \dfrac{n_2}{n_1}\right)\hat{s}_d^2}}, \qquad n = k\,r, \tag{5.4}$$

де d_j — різниця показників двох моделей на j-му розбитті, $\bar{d}$ і $\hat{s}_d^2$ — їх середнє й дисперсія, n_1 і n_2 — обсяги навчальної й тестової частин. Для *однієї* фіксованої тестової вибірки коректнішим є критерій Мак-Немара за кількостями неузгоджених рішень (n_{01} — зразки, де перша модель помилилася, а друга ні, n_{10} — навпаки):

$$\chi^2_{\mathrm{McN}} = \frac{\big(|n_{01} - n_{10}| - 1\big)^2}{n_{01} + n_{10}}. \tag{5.5}$$

Кілька моделей на кількох наборах даних порівнюють за процедурою Фрідмана з критерієм Немені [109], а не серією парних критеріїв без поправки на множинність [66] (розділ 8). Значуща різниця в частки відсоткового пункта за мільйона зразків може не мати практичного значення, тому звітують розмір ефекту.

5.7.3 Типові вади й контрольні списки

Характерні вади машинного навчання в комп'ютерній безпеці систематизовано в праці Д. Арпа зі співавторами [60]: вибіркове зміщення даних, розмічування за неперевіреним джерелом, витік даних, нереалістична базова частка класів, хибні базові лінії, лабораторні умови, несумісні з експлуатацією, хибне тлумачення показників. Контрольні списки — чекліст NeurIPS [269], бейджі ACM [52], стандарти ACM SIGSOFT [278] — об'єднано в табл. 5.4, яку заповнюють *до* написання статті.

Таблиця 5.4 — Контрольний список ML-експерименту в кібербезпеці

Пункт	Що саме перевіряють
Дані	Походження, версія, ліцензія; процедура розмічування й оцінка її похибки; відомі дефекти набору
Розділення	Три незалежні частини; часова коректність; відсутність дублікатів і групового витоку
Базові лінії	Реалістична базова частка класів; тривіальне порогове правило; сучасні методи з однаковим бюджетом налаштування
Гіперпараметри	Простір пошуку й добір лише на валідаційній частині
Метрики	Precision, recall, PR-AUC, MCC замість accuracy на незбалансованих даних
Невизначеність	Кількість запусків, початкові значення генераторів, довірчі інтервали
Порівняння	Критерій (5.4) або (5.5); поправка на множинність; розмір ефекту; абляція
Артефакти	Код, конфігурації, розділення, дані, DOI репозитарію
Етика	Персональні дані, dual-use, розкриття вразливостей

5.7.4 Генеративний штучний інтелект: предмет і інструмент

Як предмет дослідження генеративні моделі цілком легітимні: оцінювання можливостей і побудова бенчмарків; дослідження безпеки (ін'єкція в підказку, витік навчальних даних, отруєння даних, зловживання агентними можливостями); застосування до задач галузі з порівнянням із простими базовими лініями. Вимоги тут суворіші за звичайні: версія моделі й дата звернення, параметри генерування, повні тексти підказок, кількість повторень (відповіді стохастичні), процедура оцінювання відповідей і показник узгодженості оцінювачів, перевірка на забруднення навчальних даних тестовим набором.

Як інструмент генеративний ШІ допустимий для допоміжних операцій (прототипування коду, мовне редагування), але не є носієм наукової відповідальності. Не делегують: формулювання результату й положень новизни; інтерпретацію даних і висновки; добір та перевірку джерел (фабрикація посилань є систематичним збоєм); етичні рішення. Європейські настанови покладають на дослідника повну відповідальність за такий результат

[137, 281]; авторство, розкриття використання ШІ та межі допустимого в написанні тексту — у розділі 10.

5.8 Кейс: стійкість моделі виявлення вторгнень

Постановка. Група кафедри вивчає, чи зберігає модель виявлення вторгнень якість за умов, для яких її не навчали. Питання розпадається на *проєктне* (як побудувати метод, стійкий до дрейфу трафіку) і *причинне* (чи справді перетворення ознак зменшує падіння повноти на пізніших даних), тому дизайн комбінований: каркас — DSRM [262], перевірка розв'язку — обчислювальний експеримент, фінал — натуралістичне оцінювання за стратегією технічного ризику FEDS [333]. Результат наперед записано за (5.1): у контексті C (корпоративний сегмент із перевагою вебтрафіку, потокові ознаки, зсув розподілу за 12 місяців) перетворення A дає ефект E (падіння повноти на пізнішому періоді менше щонайменше на третину) через механізм M (усунення ознак, залежних від конфігурації полігона).

Дані й базові лінії. Використано два відкриті набори потокових ознак із *різним призначенням*, і саме тому їхні часові межі треба оголосити окремо.

Набір 1 — CIC-IDS2017 [301]: п'ять робочих днів, 3–7 липня 2017 р. Він *не* дає багатомісячного спостереження, тому його використано лише для перевірки методу в межах одного захоплення: навчання й валідація — понеділок–середа (3–5 липня), тест — четвер і п'ятниця (6–7 липня), з перевіркою відсутності спільних сесій між частинами. Ураховано задокументовані дефекти конвеєра розмічування цього набору [135]: без виправлень частина «успіхів» моделі пояснюється артефактами — звідси правило шукати публікації про дефекти еталонного набору до його використання.

Набір 2 — добові трасування MAWILab [153]: відкритий архів із щоденними позначеними трасуваннями магістрального каналу, що триває роками. Саме він забезпечує заявлений 12-місячний дрейф: навчання й валідація — трасування січня–березня 2024 р., три послідовні тестові вікна — квітень–червень 2024 р., жовтень–грудень 2024 р. і квітень–червень 2025 р. Розрив між кінцем навчального періоду (31.03.2024) і початком останнього тестового вікна (01.04.2025) становить рівно 12 місяців, а зростання розриву від вікна до вікна (1, 7 і 12 місяців) дає змогу побачити *динаміку* деградації, а не лише її наявність. Дати розбиття фіксують у протоколі до першого запуску.

Для обох наборів базову частку атак приведено до реалістичної (менш ніж 1 %). Висновок про стійкість до дрейфу роблять лише за набором 2; набір 1 відповідає на вужче питання — чи не втрачає метод якості вже в межах одного тижня. Базових ліній три: порогове правило за обсягом потоку; градієнтний бустинг, налаштований тим самим бюджетом пошуку, що й запропонований метод (підхід, застосовуваний і в українських роботах напряму [1]); автокодувальник.

Вимірювання й порівняння. Метрики — повнота за фіксованої частки хибнопозитивних 10^{-3}, PR-AUC і MCC. Кількість повторень обчислено за (5.2): на п'яти пілотних прогонах виміряно CV $= 0{,}07$, і за $\delta = 0{,}05$, $\alpha = 0{,}05$ формула дає $n \geq (2{,}26 \cdot 0{,}07/0{,}05)^2 \approx 10$, тож кожну конфігурацію запускають 10 разів. Порівняння з бустингом на часовому тесті виконано критерієм Мак-Немара (5.5), а на повторюваній крос-валідації — виправленим t-критерієм (5.4) з поправкою на множинність [66]. Абляція показує, що без нормалізації за підмережею ефект зникає — отже, механізм M названо правильно.

Загрози валідності. Внутрішня — різні версії бібліотек між прогонами (усунено контейнерним образом); *зовнішня* — два набори одного типу мережі; *конструктна* — «стійкість» операціоналізовано як відносне падіння повноти, що не охоплює вартості реагування на хибні тривоги; *статистична* — залежність згортків ураховано (5.4).

Артефакти й етика. Опубліковано код із фіксованими версіями залежностей, скрипти відтворення розділення (зі списками ідентифікаторів потоків), оброблені ознаки, сирі результати всіх прогонів, конфігурації пошуку гіперпараметрів для *всіх* методів і ноутбук, що будує таблиці й рисунки статті; депонування — у репозитарії з постійним ідентифікатором і ліцензією (розділ 9), подання на артефактне оцінювання [52]. Натуралістичний етап виконано на сегменті власного університету за письмовим дозволом; побічно виявлену вразливість конфігурації повідомлено адміністраторові за процедурою CVD [190].

5.9 Висновки до розділу

Науковим результатом в ІТ може бути теорія, метод, алгоритм, архітектура, інструмент, набір даних або емпірична закономірність, і тип результату визначає методологію й критерії рецензування. Узгодженість трійки «питання — результат — валідація» за Шоу та розрізнення дослідницьких питань і проєктних задач за Вірінгою пояснюють більшість відхилень, а запис результату технологічним правилом (5.1) змушує назвати контекст, ефект і механізм.

Design Science Research дає парадигму (сім настанов, три цикли), процес (шість кроків DSRM із чотирма точками входу), типологію артефактів і систему планування оцінювання FEDS; перевірка розв'язку є повноцінним емпіричним дослідженням, а не актом упровадження. Контрольований експеримент дає причинні висновки лише за рандомізації, а чотири типи загроз валідності мають бути названі з конкретними пом'якшеннями. Якісні й польові методології відповідають на питання, недоступні експериментові.

Бенчмаркінг стає доказом лише за репрезентативного навантаження, фіксованих метрик, однаково налаштованих базових ліній, абляції та врахування вимірювального шуму (5.2). Кібербезпекові дослідження вимагають явної моделі порушника, етичної рамки Menlo Report, скоординованого розкриття вразливостей і зваженого рішення щодо подвійного призначення. Дослідження з машинним навчанням потребують ізоляції тестової вибірки, часової коректності розділення, реалістичної базової частки класів і коректних критеріїв порівняння (5.4), (5.5). Генеративний ШІ є законним предметом дослідження й обмежено придатним інструментом, але формулювання результату, інтерпретацію й добір джерел делегувати йому не можна.

Питання для самоперевірки

1. Назвіть сім типів наукового результату в IT і вкажіть, до якого належить результат вашої роботи.
2. Чим інстанціація відрізняється від методу і чому дисертація рівня інстанціації рідко проходить захист?
3. Сформулюйте власний результат за (5.1). Який складник назвати найважче й чому?
4. У чому різниця між дослідницьким питанням і проєктною задачею за Вірінгою? Наведіть приклади з вашої теми.
5. Перелічіть сім настанов Гевнера. Яку з них найчастіше порушують?
6. Опишіть три цикли DSR. Чим DSR відрізняється від розробки?
7. Назвіть шість кроків DSRM і чотири точки входу. Яка з точок входу відповідає вашій роботі?
8. Назвіть дві осі FEDS. Чим формувальне оцінювання відрізняється від підсумкового, а штучне — від натуралістичного? Чому ex ante/ex post не є віссю цієї системи координат? Яку стратегію оберете?
9. Наведіть по одному прикладу загрози кожного з чотирьох типів валідності та спосіб її пом'якшення.
10. Назвіть чотири види тріангуляції й застосуйте два з них.
11. Що таке «солом'яне опудало» і як його створюють ненавмисно?
12. Обчисліть кількість повторень за (5.2) для $\mathrm{CV} = 0{,}2$ і $\delta = 0{,}05$.
13. Перелічіть механізми витоку даних. Чим data snooping відрізняється від витоку?
14. Чому наївний t-критерій за крос-валідацією завищує значущість і коли доречний критерій Мак-Немара?

Практичні завдання

1. *Паспорт методології.* Запишіть головне питання власного дослідження одним реченням, визначте його тип за рис. 5.2, оберіть методологію й заповніть для неї рядок табл. 5.3. Обґрунтуйте, чому дві сусідні методології не підходять.
2. *Технологічне правило й цілі DSRM.* Сформулюйте результат за (5.1), переведіть його в ціль кроку 2 DSRM із показником, базовим рівнем і порогом успіху, укажіть точку входу та стратегію FEDS.
3. *Загрози валідності.* Складіть для власного експерименту таблицю за зразком табл. 5.2: не менш ніж дві загрози кожного з чотирьох типів із заходом пом'якшення; позначте ті, які усунути неможливо, і сформулюйте, як обмежити висновки.
4. *Аудит бенчмарку.* Візьміть дві статті за темою дослідження (фахове видання категорії «Б» або Scopus) і перевірте: чи названо базові лінії та їх версії; чи однаковий бюджет налаштування; скільки повторень і яка міра розкиду; чи є абляція; чи доступний код.
5. *Чекліст ML-експерименту й план артефактів.* Заповніть табл. 5.4 для власної моделі, доведіть часову коректність розділення й обчисліть базову частку позитивного класу; складіть перелік артефактів до депонування та сформулюйте ethics statement.

РОЗДІЛ 6

Пошук, аналіз і систематизація наукової інформації. Систематичний огляд літератури

Розділ 5 показав, як обирають методологію дослідження. Але жодну методологію не запускають «з чистого аркуша»: перед плануванням експерименту чи побудовою артефакту треба встановити, що вже відомо, що суперечливе й що досі не досліджене. Цей розділ трактує огляд літератури не як вступну формальність, а як *самостійне дослідження* з протоколом, відтворюваною процедурою пошуку, вимірюваною якістю відбору й перевірним результатом. Наскрізний кейс — систематичний огляд публікацій про безпеку пристроїв Інтернету речей.

6.1 Наукова інформація: типи документів і джерел

6.1.1 Первинні, вторинні й третинні джерела

Первинне джерело містить оригінальне повідомлення про дослідження, зроблене тими, хто його виконав: стаття з новими даними, доповідь на конференції, дисертація, препринт, патент, набір даних, програмний артефакт. **Вторинне джерело** узагальнює або критично осмислює первинні: оглядова стаття, систематичний огляд, метааналіз, реферативна база. **Третинне джерело** подає усталене знання без претензії на новизну: підручник, енциклопедія, довідник, tertiary study.

Практичне значення шкали просте: *твердження про факт підтверджують первинним джерелом*; посилання на підручник там, де йдеться про результат вимірювання, — типова помилка кваліфікаційних робіт.

6.1.2 Типи наукових документів у галузі F/12

Стаття в журналі. Основна одиниця комунікації: *research article*, *review*, *short paper/letter*, *comment* і *reply*. Для українського здобувача істотні поділ на фахові видання категорій «А» і «Б» та індексація у Scopus чи WoS.

Матеріали конференції. В інформатиці — *повноцінний* науковий канал, а не «тези»: провідні конференції (IEEE S&P, USENIX Security, CCS, NDSS, ICSE) мають частку прийняття 12–20 % і рецензування, суворіше за журнальне.

Препринт. Рукопис у відкритому репозитарії (arXiv, TechRxiv, IACR ePrint) *без рецензування*. Є опублікована версія — цитують її; інакше зазначають версію (`arXiv:2205.01833v2`) і статус.

Монографія. Виклад цілісної концепції; рецензування нерівномірне, тому вагу оцінюють за видавцем.

Дисертація та автореферат. Первинне джерело з докладнішою за статтю методикою (Національний репозитарій академічних текстів, репозитарії ЗВО, НБУВ [15]).

Стандарт і патент. ISO/IEC, IEEE, ETSI, IETF RFC, NIST SP, ДСТУ — не науковий результат, а *нормативна рамка*; патент (Espacenet, Google Patents, база УКРНОІВІ) випереджає публікації на 1–3 роки, але доводить новизну заявки, а не працездатність рішення.

Технічний звіт. Звіт про НДР або агентства (ENISA, NIST, CERT-UA); часто містить дані, яких немає у статтях.

Набір даних і програмний артефакт. Цитовані результати за наявності ідентифікатора, ліцензії й опису (розділ 9); цитують версію, а не репозитарій.

6.1.3 Сіра література та ідентифікатори

Сіра література — документи поза видавничим рецензуванням: технічні звіти, whitepapers виробників, блоги лабораторій, звіти про інциденти, трекери помилок, матеріали галузевих конференцій (Black Hat, DEF CON). В IT ігнорувати її не можна: значна частина знання про реальні атаки з'являється саме там. В. Гароусі зі співавторами запропонували критерії доцільності **багатоголосого огляду** (multivocal literature review) і трирівневу шкалу за контрольованістю джерела [157]: 1-й рівень — книги та звіти урядових агентств; 2-й — корпоративні звіти й вікі; 3-й — блоги й дописи в соцмережах. Кожне джерело оцінюють за авторитетністю автора, методологією, об'єктивністю, датою й підтвердженням з інших джерел.

Відтворюваний пошук спирається на постійні ідентифікатори: **DOI**, **ORCID**, **ROR** (установа), `arXiv ID`, `SWHID` (знімок коду). DOI — запис у реєстрі Crossref чи DataCite з відкритими метаданими; на них побудовані OpenAlex [272] та OUCI [253]. Наслідок: існування джерела перевіряють не пошуком назви в мережі, а розв'язанням самого ідентифікатора. Порядок перевірки: відкрити `https://doi.org/<DOI>` і подивитися, куди він веде; звірити метадані запитом до відповідної реєстраційної агенції — `api.crossref.org/works/<DOI>` для статей і матеріалів конференцій, `api.datacite.org/dois/<DOI>` для наборів даних, програмного коду й препринтів у репозитаріях; нарешті відкрити сторінку статті на сайті видавця. Жоден із цих кроків окремо не є вичерпним: відсутність запису в Crossref не доводить, що публікації не існує (DOI міг бути зареєстрований в іншій агенції, а частина видань узагалі не депонує DOI), тоді як наявність запису не гарантує, що назва, автори чи сторінки відповідають тому, що процитовано (підрозділ 6.8).

6.2 Наукові бази даних і пошукові системи

6.2.1 Критерії придатності бази для систематичного пошуку

М. Ґузенбауер і Н. Геддевей перевірили 28 пошукових систем і показали, що більшість непридатні як *основне* джерело систематичного огляду [166]. Придатна база має: підтримувати булеві оператори й дужки без спрощення запиту; давати пошук за визначеними полями; повертати *повторювані* результати; дозволяти експорт усіх знайдених записів; не персоналізувати ранжування. Google Scholar не відповідає критеріям експорту, повторюваності й довжини запиту, тому слугує *додатковим* джерелом. Покриття баз різниться навіть у межах однієї дисципліни [165], а їх перекриття кількісно оцінено в [222].

Таблиця 6.1 — Наукові бази й пошукові системи: покриття, переваги, обмеження

Ресурс	Покриття	Переваги	Обмеження
Scopus	≈28 тис. активних рецензованих журналів (≈43 тис. джерел усіх типів у Scopus Sources), усі галузі	Строгий синтаксис, `TITLE-ABS-KEY`, експорт	Платний; слабше покриття конференцій ACM
Web of Science Core Collection	≈22 тис. журналів; SCIE, SSCI, CPCI	Найдовші ряди цитувань	Платний; вужчий за Scopus
IEEE Xplore	6 млн записів IEEE, IET	Ключова для IT; повні тексти; IEEE Thesaurus	Лише IEEE/IET; власний синтаксис полів
ACM Digital Library	Корпус ACM; класифікація CCS	Обов'язкова для комп'ютерних наук	Лише ACM; обмежений експорт
ScienceDirect, SpringerLink	Повні тексти Elsevier і Springer	Доступ до текстів	Платформи видавців, а не індекси
Google Scholar	Найширше покриття, сіра література	Повнота; «Cited by»	Непрозоре покриття; неповний експорт
Semantic Scholar	200 млн праць; граф цитувань	Відкритий API; семантичний пошук	Неповні метадані
OpenAlex	250 млн праць [272]	Відкритий граф; безплатний API	Похибки автодедублікації
Dimensions	Публікації, гранти, патенти	Зв'язок із фінансуванням	Аналітика за передплатою
arXiv	2 млн препринтів (`cs.CR`)	Швидкість; повні тексти; версії	Без рецензування
DOAJ	Журнали відкритого доступу [118]	Перевірка легітимності, DOAJ Seal	Не пошукова система для огляду
Crossref	Реєстр DOI, ретракції [102]	Відкритий API; верифікація джерел	Анотації лише для частини записів (фільтр `has-abstract`); повнота залежить від депонованих метаданих

Ресурс	Покриття	Переваги	Обмеження
НБУВ, «Наукова періодика України»	900 тис. статей із 2600 журналів [15]	Головне джерело україномовних публікацій	Обмежені булеві можливості
OUCI	Записи Crossref, пов'язані з Україною [253]	Відкриті цитування; Unpaywall	Залежить від Crossref Cited-by
Репозитарії ЗВО	Дисертації, звіти	Тексти, недоступні деінде	Пошук через BASE, CORE, OpenAIRE

6.2.2 Український сегмент і його особливості

Здобувач в Україні працює у двох інформаційних просторах одночасно: міжнародному (Scopus, Web of Science, IEEE Xplore, ACM DL) і національному (НБУВ [15], OUCI [253], Національний репозитарій академічних текстів, репозитарії ЗВО). Другий важливий із двох причин: значна частина прикладних результатів українських дослідників публікується лише у фахових виданнях категорії «Б» і до Scopus не потрапляє, а огляд стану проблеми в дисертації має показати обізнаність саме з українською школою. М. Назаровець показала, що український журнальний ландшафт погано відображений навіть у спеціалізованих довідниках [241]. Робоче правило: *міжнародні бази задають ядро огляду, українські ресурси — контекстний шар*; обидва документують окремо із зазначенням дати запиту.

> **Що фіксувати для кожної бази:** назву ресурсу; *точний* рядок запиту; поля пошуку; фільтри; роки; дату виконання; кількість записів. Без цього пошук не є відтворюваним [257].

6.3 Стратегія пошуку наукової інформації

6.3.1 Декомпозиція питання на поняття

Пошук починається не з бази, а з питання. Його розкладають на 2–4 **поняття** (concepts), які в запиті сполучають оператором AND, а синоніми кожного поняття — оператором OR; для ІТ зручна схема PICOC (розділ 3). Питання наскрізного кейсу: *які емпірично перевірені методи виявлення вразливостей у прошивках побутових і промислових пристроїв Інтернету речей запропоновано у 2019–2025 рр. і як оцінено їх результативність?* Поняття й синоніми:

- **С1 (об'єкт):** internet of things, IoT, IIoT, smart home, smart device, connected device, embedded device, CPS, cyber-physical system, firmware;
- **С2 (властивість):** security, insecurity, vulnerability, weakness, attack, exploit, threat, compromise;

- **С3 (втручання)**: detection, analysis, fuzzing, static analysis, dynamic analysis, emulation, taint, assessment, penetration testing.

Синоніми збирають із ключових слів 5–10 відомих релевантних статей, контрольованих словників (IEEE Thesaurus, ACM CCS, INSPEC Controlled Terms) і тексту власного питання, враховуючи варіанти написання (`cyber-physical`, `cyberphysical`), абревіатури, однину й множину та національні синоніми для україномовного шару.

6.3.2 Синтаксис запиту: оператори, усічення, поля

Булеві оператори — AND, OR, AND NOT; пріоритет операцій у різних базах різний, тому *завжди* застосовують дужки, а AND NOT уживають обережно: він відкидає й релевантні записи, у яких випадково трапилося виключене слово. *Фразовий пошук*: у Scopus лапки `"internet of things"` дають «вільну фразу» (пунктуація ігнорується, усічення допускається), фігурні дужки `{internet of things}` — точний рядок; в IEEE Xplore і WoS лапки означають точну фразу. *Усічення*: * замінює будь-яку кількість символів (`secur*`: security, secure, securing), `?` — рівно один; надмірне усічення руйнує точність (`net*`: network, netting, nettle), тому усікають не раніше ніж після п'яти-шести символів кореня. *Оператори близькості* (Scopus `W/n` і `PRE/n`, WoS та IEEE Xplore `NEAR/n`, `ONEAR/n`) підвищують точність там, де фраза варіативна: `firmware W/3 analysis`. *Поля пошуку*: у Scopus — `TITLE`, `ABS`, `KEY`, `TITLE-ABS-KEY`, `SRCTITLE`, `DOCTYPE`, `PUBYEAR`, `LANGUAGE`; в IEEE Xplore (Command Search) — `"Document Title"`, `"Abstract"`, `"Author Keywords"`, `"Index Terms"`, `"All Metadata"`. Пошук «усюди» дає надмір шуму, тому базовим полем огляду є `TITLE-ABS-KEY` або його аналог.

Робочий рядок для Scopus (Advanced search) у кейсі:

```
TITLE-ABS-KEY ( ( "internet of things" OR iot OR iiot OR
"smart home" OR "smart device*" OR "connected device*" OR
"embedded device*" OR "cyber-physical system*" ) AND ( secur*
OR vulnerab* OR attack* OR exploit* OR threat* ) AND ( firmware
OR "static analysis" OR "dynamic analysis" OR fuzz* OR emulat*
OR "penetration test*" OR detect* ) ) AND PUBYEAR > 2018 AND
PUBYEAR < 2026 AND ( LIMIT-TO ( DOCTYPE , "ar" ) OR LIMIT-TO (
DOCTYPE , "cp" ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) )
```

Той самий запит для IEEE Xplore (Command Search) виглядає інакше:

```
( "All Metadata":"internet of things" OR "All Metadata":IoT OR
"All Metadata":"smart device*" OR "All Metadata":"embedded
device*" ) AND ( "All Metadata":secur* OR "All
Metadata":vulnerab* OR "All Metadata":attack* ) AND (
"Abstract":firmware OR "Abstract":fuzz* OR "Abstract":"static
analysis" OR "Abstract":detect* )
```

Фільтри за роками, типом і мовою в IEEE Xplore задають не в рядку, а на панелі уточнення; це теж фіксують у протоколі. Типова помилка — механічно перенести рядок Scopus в IEEE Xplore: конструкцій `TITLE-ABS-KEY` і `LIMIT-TO` там немає.

6.3.3 Вимірювання якості пошукового рядка

Пошуковий рядок — вимірюваний артефакт. Нехай Q — множина повернутих записів, а R — **еталонний набір** (quasi-gold standard): перелік свідомо релевантних праць, зібраний вручну суцільним переглядом кількох профільних видань [354]. Тоді

$$S = \frac{|R \cap Q|}{|R|} \cdot 100\,\%, \qquad D = \frac{|R \cap Q|}{|Q|} \cdot 100\,\%, \tag{6.1}$$

де S — **чутливість** (recall) пошуку, оцінена на еталонному наборі, а D — *щільність еталонних записів* у видачі. Х. Чжан зі співавторами рекомендують для огляду в програмній інженерії $S \geq 80\,\%$ [354]; точністю свідомо жертвують, бо відкинути зайве на скринінгу дешевше, ніж не знайти потрібне.

Чого D не вимірює. Величину D часто помилково називають точністю. Це різні речі: R — лише *відомий* перелік релевантних праць із кількох видань і років, тоді як Q охоплює значно ширший пошук, і серед решти записів Q теж є релевантні, просто не внесені до еталона. Тому D дає нижню межу точності, а не її оцінку. **Точність** обчислюють лише за результатами оцінювання релевантності:

$$P = \frac{|\mathrm{Rel} \cap Q|}{|Q|} \cdot 100\,\%, \tag{6.2}$$

де $\mathrm{Rel} \cap Q$ — записи, визнані релевантними після скринінгу *всіх* записів видачі або репрезентативної випадкової вибірки з неї (у другому випадку разом з оцінкою подають довірчий інтервал). Похідний показник — **кількість записів на одне включення** $\mathrm{NNR} = 1/P$, якщо P подано часткою, або $\mathrm{NNR} = 100/P$, якщо P подано у відсотках (number needed to read).

Числовий приклад кейсу. Еталонний набір складено суцільним переглядом трьох джерел (IEEE IoT Journal, ACM TIOT, матеріали ACSAC) за 2021–2023 рр.: $|R| = 30$. Перша версія рядка (без поняття СЗ) повернула $|Q_1| = 9\,410$ записів і 28 еталонних: $S_1 = 93,3\,\%$ за мізерної точності. Друга, з обов'язковим СЗ у полі `ABS`, дала $|Q_2| = 1\,204$ і 21 еталонний запис: $S_2 = 70,0\,\%$ — нижче порога, рядок непридатний. Остаточна (СЗ у `TITLE-ABS-KEY`, розширені синоніми) дала $|Q_3| = 1\,842$ і 26 еталонних записів: $S_3 = 86,7\,\%$; саме її занесено до протоколу. Щільність еталонних

записів $D_3 = 26/1\,842 \approx 1,4\,\%$ означає лише, що на один *знайдений еталонний* запис припадає близько 71 запису видачі, — це не точність рядка. Точність оцінено окремо: з Q_3 узято випадкову вибірку 200 записів, з яких два незалежні скринери визнали релевантними 17, звідки $\widehat{P_3} \approx 8,5\,\%$ (95 % довірчий інтервал Вілсона — приблизно 5,4–13,2 %) і NNR ≈ 12. Саме ці числа, а не D_3, характеризують навантаження на скринінг.

6.3.4 Ітеративність і критерій насичення

Позначимо R_k — сукупну множину релевантних праць після k-ї ітерації (нова база, нова версія рядка, черговий цикл сніжного кому). **Приріст** ітерації

$$\nu_k = \frac{\left|R_k \setminus R_{k-1}\right|}{\left|R_{k-1}\right|}. \tag{6.3}$$

Насичення фіксують, коли $\nu_k \le \varepsilon$ упродовж двох послідовних ітерацій; типове $\varepsilon = 0,05$. У кейсі: після баз $|R_0| = 66$; перша ітерація сніжного кому дала 7 нових праць ($\nu_1 = 0,106$), друга — 2 ($\nu_2 = 0,027$), третя — 0. Насичення — *емпіричний* критерій зупинки, а не доказ повноти; у звіті його подають числами ν_k.

6.3.5 Сніжний ком

Зворотний сніжний ком — перегляд списків літератури включених праць; **прямий** — перегляд праць, які цитують включені (Scopus «Cited by», Google Scholar, Semantic Scholar, OUCI для українського шару). К. Волін описав процедуру як повноцінну альтернативу пошуку в базах [348], а згодом зі співавторами емпірично показав, що найкращий баланс чутливості й точності дає **гібридна стратегія**: пошук у Scopus плюс паралельний або послідовний сніжний ком [349]. Сніжний ком виправляє дві системні вади пошуку за рядком: знаходить праці з нетиповою термінологією і праці в джерелах, не індексованих обраними базами (зокрема українські). Його слабкість — залежність від стартового набору, тому останній формують із різних видань, років і наукових шкіл.

6.4 Види оглядів літератури

М. Ґрант і Е. Бут описали 14 типів оглядів і показали, що вибір типу визначається питанням, а не наявним часом [159]. Для галузі F/12 практично значущі шість (табл. 6.2). **Наративний огляд** — експертний виклад стану проблеми без формалізованої процедури пошуку; доречний у вступі статті, але не є доказом. **Картувальний огляд** (scoping review, systematic mapping study) відповідає на питання «що досліджено і як розподілені дослідження»: будує класифікаційну схему галузі, але не синтезує результа-

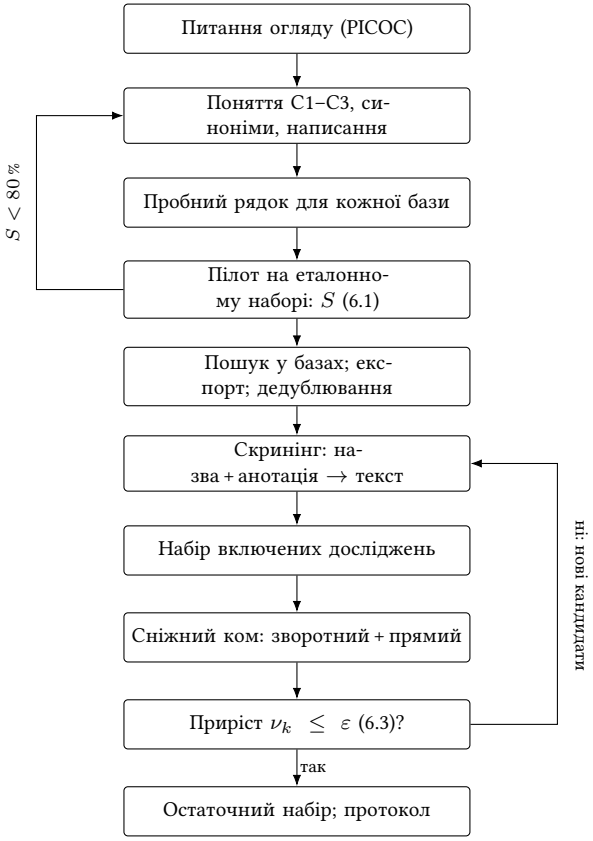


Рис. 6.1 — Ітеративна стратегія пошуку з гібридним сніжним комом (за [354, 348, 349])

тів [266, 236]. **Систематичний огляд** відповідає на конкретне питання про результат за задокументованою відтворюваною процедурою [206]. **Tertiary study** — огляд, первинними дослідженнями якого є самі огляди [205], а **метааналіз** — кількісний синтез ефектів, в ІТ рідкісний через неоднорідність метрик.

Вибір фіксують явно: «виконано картувальний огляд за настановами [266]» — і далі дотримуються саме їх. Найпоширеніша вада дисертаційних оглядів — називати наративний огляд систематичним: щойно з'явилося слово «систематичний», рецензент має право вимагати протокол, рядок пошуку, критерії та діаграму.

Таблиця 6.2 — Види оглядів літератури

Вид	На яке питання відповідає	Обов'язкові вимоги	Типовий масштаб
Наративний	Як склалося розуміння проблеми	Компетентність автора; чесне подання альтернатив	30–80 джерел, 2–6 тижнів
Картувальний	Що і як досліджено; де прогалини	Протокол, схема класифікації, частотний аналіз [266]	100–400 праць, 3–6 місяців
Систематичний	Який результат дає підхід A за умов C	Протокол, критерії, два скринери, якість, PRISMA [206, 257]	30–150 праць, 6–12 місяців
Метааналіз	Яка узагальнена величина ефекту	Однорідність метрик, розміри ефектів, гетерогенність	10–50 досліджень
Tertiary study	Що вже узагальнено оглядами	Первинні одиниці — огляди [205]	20–60 оглядів
Швидкий	Що відомо тут і зараз	Декларування всіх спрощень	20–60 праць, 2–8 тижнів

6.5 Систематичний огляд літератури: протокол і процес

6.5.1 Три фази й протокол

Б. Кітченгем і С. Чартерс поділили процес на три фази [206]: *планування* (потреба, питання, протокол); *виконання* (пошук, відбір, оцінювання якості, вилучення й синтез даних); *звітування*. Принципова вимога — протокол складають *до* початку пошуку.

Протокол огляду містить: обґрунтування потреби; питання огляду; стратегію пошуку (бази, рядки для кожної з них, роки, мови, сніжний ком); критерії включення й виключення; процедуру відбору; критерії оцінювання якості; форму вилучення даних; метод синтезу; графік і розподіл ролей. Найкраща практика — **передреєстрація**: викладення протоколу у відкритому репозитарії (OSF, Zenodo) з отриманням DOI до початку пошуку; вона знімає підозру, що критерії підганялися під бажаний результат (розділ 9).

Питання огляду стосуються *стану знання*. У кейсі їх три. *RQ1:* які класи методів виявлення вразливостей у прошивках пристроїв IoT запропоновано у 2019–2025 рр.? *RQ2:* на яких наборах пристроїв і за якими метриками їх оцінено? *RQ3:* які обмеження декларують автори і які результати відтворено незалежно?

6.5.2 Критерії включення й виключення

Критерії мають бути *перевірними*: два незалежні скринери, застосувавши їх до одного запису, повинні дійти однакового висновку; формулювання «сучасні роботи високої якості» критерієм не є. Кожен критерій кодують, щоб указувати підставу виключення (табл. 6.3).

Таблиця 6.3 — Критерії відбору первинних досліджень (кейс «безпека пристроїв IoT»)

Код	Формулювання	Обґрунтування
I1	Пропонує або оцінює метод виявлення вразливостей у прошивці чи ПЗ пристрою IoT/ІоТ	Відповідає RQ1
I2	Є емпіричне оцінювання на реальних або емульованих пристроях чи образах прошивок	RQ2; відсіює концептуальні праці
I3	Рецензована стаття або повна доповідь (від 6 сторінок)	Мінімум рецензування
I4	Опубліковано у 2019–2025 рр.	Період після ботнетів класу Mirai
I5	Мова — англійська або українська	Робочі мови команди
E1	Тези, постери, редакційні матеріали	Бракує даних
E2	Вторинні дослідження (огляди, мапування)	Використано для сніжного кому
E3	Предмет — мережевий рівень або хмарний бекенд без аналізу самого пристрою	Поза межами RQ1
E4	Немає опису процедури оцінювання (вибірка, метрики)	Неможливо вилучити дані
E5	Дублювальна публікація	Залишають найповнішу версію
E6	Повний текст недоступний після двох спроб	Неможлива оцінка
E7	Видання з ознаками хижацького	Сумнівна достовірність (розділ 7)

6.5.3 Двоетапний скринінг і узгодженість скринерів

Eman 1 — скринінг за назвою й анотацією, за принципом «сумнів — на користь включення», бо помилка виключення тут невиправна. *Eman 2* — за повним текстом, із зазначенням коду критерію виключення. Обидва етапи виконують *два незалежні скринери*; одноосібний скринінг допустимий лише у швидкому огляді.

Узгодженість оцінюють на спільній підвибірці (10–20 % записів). **Відсоток згоди** p_o систематично завищує враження про якість, бо в незбалансованій задачі високу згоду дає навіть суцільне відкидання. Тому використовують **коефіцієнт** κ **Коена** [91], що виправляє згоду на випадкову. Для таблиці спряженості 2×2 з клітинками a (обидва «включити»), b і c (розбіжності), d (обидва «виключити»), $n = a + b + c + d$:

$$\kappa = \frac{p_o - p_e}{1 - p_e}, \quad p_o = \frac{a + d}{n}, \quad p_e = \frac{(a+b)(a+c) + (c+d)(b+d)}{n^2}. \tag{6.4}$$

Числовий приклад кейсу. Пілотну підвибірку 300 записів (9,8 % від 3 065, що надійшли на скринінг) скринували двоє. Результат: $a = 41$, $b = 19$, $c = 14$, $d = 226$. Тоді $p_o = 267/300 = 0,890$; $p_e = (60 \cdot 55 + 240 \cdot 245)/300^2 = (3\,300 + 58\,800)/90\,000 = 0,690$; звідси

$$\kappa = \frac{0,890 - 0,690}{1 - 0,690} = \frac{0,200}{0,310} \approx 0,645.$$

Тобто 89 % формальної згоди відповідають лише $\kappa \approx 0,65$. За шкалою Дж. Лендіса і Ґ. Коха [212]: $\kappa \leq 0$ — згоди немає; 0,01–0,20 — незначна; 0,21–0,40 — посередня; 0,41–0,60 — помірна; 0,61–0,80 — істотна; 0,81–1,00 — майже повна. Шкала умовна, тому в ІТ-оглядах правило зупинки задають трьома смугами: $\kappa \geq 0,60$ — продовжувати основний скринінг; $0,40 \leq \kappa < 0,60$ — провести калібрувальну сесію (розбір розбіжностей, уточнення формулювань критеріїв) і повторити пілот на новій підвибірці, не зупиняючи основного скринінгу, але й не вважаючи його результати остаточними до повторного вимірювання κ; $\kappa < 0,40$ — *зупинити скринінг*, переписати критерії й відкалібрувати скринерів наново.

Процедуру вирішення розбіжностей описують у протоколі *до* їх появи: обговорення двох скринерів із поверненням до тексту критерію; якщо згоди немає — рішення арбітра; у разі систематичної розбіжності — перегляд критерію з повторним скринінгом. У кейсі 33 розбіжності розв'язано так: 27 — обговоренням, 6 — арбітром; за результатами уточнено ЕЗ.

6.5.4 Оцінювання якості, вилучення даних, синтез

Оцінювання якості дає додаткове відсіювання, зважування доказів під час синтезу й опис методологічної культури галузі. Класичний інструмент — анкета Т. Дюби і Т. Дінгсьойра з одинадцяти питань [125] у чотирьох блоках: *звітування* (чітка мета, описаний контекст), *строгість* (обґрунтованість плану, добору даних, збору й аналізу), *достовірність*, *релевантність* для практики; пункти оцінюють у шкалі $\{0;\ 0,5;\ 1\}$. Для кейсу анкету доповнено пунктами про доступність артефактів, порівняння з базовими лініями та координоване розкриття вразливостей. Праці з низьким балом не виключали, а перевіряли чутливість висновків до їх вилучення.

Форма вилучення даних — заздалегідь визначений перелік полів, який пілотують на 3–5 працях: бібліографічні дані й DOI; тип і місце публікації; клас методу за кодованим словником; об'єкт оцінювання; обсяг вибірки; метрики й значення; базові лінії; доступність артефактів; обмеження; бал якості. Її подають у додатку.

Синтез — місце, де огляд перестає бути каталогом. *Наративний* синтез структурує результати за темами; ризик — підміна синтезу послідовністю анотацій («у роботі [12]... у роботі [13]...»). *Табличний* зводить характеристики в одну таблицю з однаковими полями й уже цим виявляє прогалини. *Тематичний* кодує фрагменти, групує коди в теми й формулює твердження вищого рівня; узгодженість кодувальників вимірюють тим самим κ (6.4). *Кількісний* обчислює узагальнений розмір ефекту й гетерогенність;

його сурогат — **підрахунок голосів** («у 7 з 10 робіт метод A кращий») — методологічно хибний, бо ігнорує обсяги вибірок і величини ефектів.

6.5.5 Звітування за PRISMA 2020 і кейс

PRISMA 2020 — чинний стандарт звітування про систематичні огляди [257]: *чекліст із 27 пунктів* у семи розділах (назва; структурована анотація; вступ; методи; результати; обговорення; інша інформація), окремий чекліст для анотації та *потокова діаграма* відбору; пояснення й приклади — у супровідному документі [258]. Ключові пункти методів: критерії прийнятності (п. 5), джерела інформації з датами (п. 6), *повні рядки пошуку для всіх баз* (п. 7), процес відбору й кількість скринерів (п. 8), вилучення даних (п. 9), ризик упередженості (п. 11), методи синтезу (п. 13); «інша інформація» — реєстрація й протокол (п. 24), фінансування (п. 25), конфлікт інтересів (п. 26), доступність даних і коду (п. 27). Для картувальних оглядів застосовують PRISMA-ScR [323]. Найтиповіші порушення: рядок пошуку лише для однієї бази або переказаний словами; не вказано дат пошуку й кількості скринерів; діаграма без чисел; немає переліку виключених за повним текстом праць.

Рис. 6.2 зводить кейс воєдино. Пошук виконано 14.02.2026 у чотирьох базах рядками з підрозділу 6.3; дедублювання — у Zotero за DOI й назвою. Український шар шукали окремо в НБУВ [15] та OUCI [253] запитом «Інтернет речей» AND (безпек* OR вразлив*). Його показано в діаграмі окремою гілкою, бо ці бази не підтримують того самого синтаксису й опрацьовувалися після основного пошуку; далі обидві гілки зводяться в одне дедублювання ($4\,429 + 148 = 4\,577$ записів до нього, 3 065 після). Зі 148 записів українського шару скринінг залишив три: дві первинні праці, які вже були у видачі міжнародних баз (вони зникли на дедублюванні й ураховані серед 66 включених — додавати їх повторно означало б подвійний облік), та огляд І. Опірського зі співавторами [20]. Огляд *не* входить до первинного синтезу — критерій E1 виключає вторинні дослідження; його використано лише як стартову точку зворотного сніжного кому, а знайдені через нього первинні праці враховано серед 9 включених на цьому етапі. Саме тому підсумкові 75 праць не збільшуються на три: множини перетинаються, і звіт має показувати це зіставлення, а не арифметичне додавання гілок.

6.6 Критичне читання наукової статті

6.6.1 Стратегія трьох проходів

С. Кешав описав економну процедуру читання [204]. *Перший прохід* (5–10 хв): назва, анотація, вступ, заголовки, висновки, побіжно — список літератури; мета — п'ять відповідей: до якої категорії належить робота, у якому контексті, чи коректні припущення, який заявлено внесок, чи ясно на-

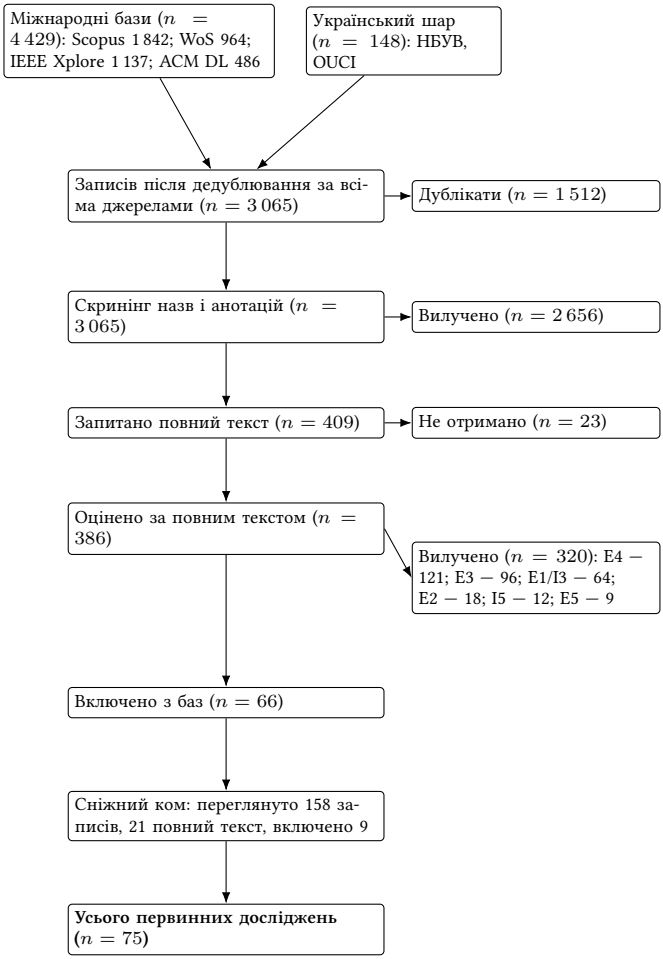


Рис. 6.2 — Потокова діаграма відбору джерел (перемальовано за [257])

писано. Саме цей прохід виконують на скринінгу за назвою й анотацією. *Другий прохід* (близько години): уважне читання без перевірки доведень, з розглядом рисунків, осей і позначень; після нього читач може переказати суть роботи з доказовою базою. *Третій прохід* (кілька годин) — віртуальне відтворення: читач подумки (а в IT — і практично) повторює побудову й перевіряє кожен крок; він потрібен для рецензування й для праць, на яких будується власне дослідження.

6.6.2 Відповідність висновків даним і слабкі докази

Критичне читання — звірка трьох рівнів: що заявлено в анотації й висновках, що показано в результатах, що виміряно в методах. Контрольні питання: чи є заявлений внесок тим, що перевірено; чи охоплює вибірка ту сукупність, на яку поширено висновок; чи налаштована базова лінія так само ретельно; чи повідомлено розкид; чи відповідає метрика заявленій властивості (точність на незбалансованих даних уводить в оману — розділ 8); чи названо обмеження й загрози валідності.

Ознаки **слабкого доказу**: демонстрація замість оцінювання («ми реалізували, і система працює»); порівняння з «солом'яним опудалом»; єдиний набір даних, на якому метод і розроблявся; відсутність повторних запусків для стохастичних методів; висновок про причинність із кореляційних даних; «$p < 0{,}05$» без розміру ефекту.

6.6.3 Конфлікт інтересів і ретракції

Перевіряють розділи *Funding*, *Competing interests*, *Data availability*. Фінансування зацікавленою стороною не знецінює роботи, але змінює вагу висновків у порівняльних оцінюваннях продуктів; відсутність декларації там, де журнал її вимагає, — сигнал недбалості (розділ 10).

Обов'язковий крок — **перевірка на ретракцію**: відкликана стаття цитується роками. Джерела: база Retraction Watch, передана 2023 р. у власність Crossref і відкрита для всіх; позначки *retracted*, *corrected*, *expression of concern* у метаданих і програмному інтерфейсі Crossref [102]; попередження в Zotero та scite. Правило: перед поданням рукопису прогнати весь список літератури через перевірку ретракцій.

6.7 Керування бібліографією та синтез знань

6.7.1 Менеджери посилань

Zotero — відкритий менеджер посилань із розпізнаванням метаданих з PDF і DOI, груповими бібліотеками й нотатками [356]; для LaTeX критично важливе доповнення Better BibTeX, що дає стабільні ключі цитування й автоматично оновлюваний експорт `.bib`. **Mendeley** подібний за функційністю, але є продуктом Elsevier із меншою відкритістю формату. **JabRef** працює

безпосередньо з `.bib` і підходить тим, хто веде бібліографію разом із текстом у системі контролю версій [193].

Робочий процес: імпорт результатів пошуку у форматі RIS або BibTeX → дедублювання (за DOI, потім за назвою й роком) → колекції й теґи за поняттями огляду → нотатки за єдиним шаблоном → експорт `.bib` → підключення в LaTeX. Автоматичні метадані *завжди* перевіряють: бази псують ініціали, діакритику («Höst» → «Host») і сторінки. BibLaTeX, на відміну від BibTeX, підтримує поля `doi`, `urldate`, `langid` і стилі, ближчі до ДСТУ 8302:2015 [7], тому для української дисертації зручніший.

6.7.2 Конспектування, картування, матриця понять

Шаблон нотатки до прочитаної праці: посилання й DOI; питання, на яке вона відповідає; метод і дані; головний результат *числом*; обмеження; що придатне для власного дослідження; цитата з номером сторінки; оцінка якості. Нотатку пишуть *одразу* й своїми словами — це водночас профілактика ненавмисного плагіату (розділ 10).

Дж. Вебстер і Р. Вотсон запропонували перехід від «огляду за авторами» до **матриці понять** [341]: рядки — первинні дослідження, стовпці — поняття, клітинка позначає, чи розглянуто поняття в праці (табл. 6.4). Порожні стовпці й порожні перетини стовпців і є *прогалинами*. Ключове правило: огляд організовують *за поняттями*, а не за працями; фраза «Сміт (2021) показав..., Джонс (2022) показав...» свідчить, що синтезу немає.

Таблиця 6.4 — Матриця понять (фрагмент, кейс «безпека пристроїв IoT»)

Праця	К1	К2	К3	К4	К5	К6
S1	+	+	−	−	+	−
S2	−	+	+	+	−	−
S3	+	−	+	−	+	+
S4	−	−	+	+	−	−
S5	+	+	+	−	−	−
Разом	3	3	4	2	2	1

К1 — емуляція; К2 — фазинг; К3 — статичний аналіз; К4 — реальні пристрої; К5 — відкриті артефакти; К6 — незалежне повторне оцінювання

Матриця показує прогалину кейсу: поєднання емуляції з перевіркою на *реальних* пристроях трапляється рідко, а незалежне повторне оцінювання — лише в одній праці; це й є формулювання проблеми для власного дослідження (розділ 3). Поруч із матрицею будують **таблицю порівняння підходів** (клас методу — припущення — вхідні дані — метрики — обмеження), за потреби — карту термінів у VOSviewer (розділ 7), яка *не заміняє* синтезу.

6.8 ШІ-асистенти огляду літератури

Інструменти на основі великих мовних моделей — Elicit, Consensus, SciSpace, Research Rabbit, Connected Papers, ASReview, чат-моделі з вебпошуком — змінюють трудомісткість окремих кроків огляду, але не заміняють жодного з них [231]. Виправдані застосування: генерація синонімів (з ручною звіркою за тезаурусом); попереднє ранжування записів на скринінгу за ймовірністю релевантності (active learning в ASReview) — *як порядок перегляду*, а не як рішення про виключення; витяг характеристик із повного тексту у форму вилучення даних зі звіркою людиною; побудова карт цитувань для стартового набору сніжного кому.

Головний системний збій — **фабрикація посилань**: модель генерує бібліографічно правдоподібний запис, якого не існує, або приписує реальним авторам вигадану працю, часто із синтаксично коректним, але неіснуючим DOI. Д. Резнік і М. Госсейні доводять, що галюциновані посилання можуть кваліфікуватися як наукове шахрайство, коли посилання функціонують як дані наукової роботи, тобто саме в оглядах [283]. Дотичні ризики: приписування статті висновків, яких у ній немає; «згладжування» суперечностей; упередження навчального корпусу; нездатність відрізнити відкликану публікацію від чинної; невідтворюваність результату.

> **Правило верифікації**. Жодне джерело не потрапляє до списку літератури без незалежного підтвердження: DOI перевірено розв'язанням через `doi.org` або запитом до `api.crossref.org`; збіг авторів, назви, року, видання та сторінок звірено з реєстром; для твердження, приписаного джерелу, знайдено відповідне місце у *повному тексті*; перевірено відсутність позначки про ретракцію [102]. Неперевірене джерело вилучають, а не «залишають про всяк випадок».

Використання генеративного ШІ підлягає *розкриттю* в розділі методів: інструмент і версія, крок, мета, обсяг і спосіб верифікації. Прийнятне формулювання: «Для розширення списку синонімів понять С1–С3 використано <інструмент, версія, дата>; усі терміни звірено з IEEE Thesaurus; відбір і оцінювання досліджень виконано авторами без автоматизації». Європейські «живі» настанови вимагають зберігати відповідальність за результат і не делегувати ШІ оцінних суджень [137]; ШІ не може бути автором (розділ 10).

6.9 Висновки до розділу

Огляд літератури є дослідженням, а не переказом: він має питання, протокол, відтворювану процедуру, вимірювану якість і перевірний результат. Твердження про факт підтверджують первинним джерелом; в ІТ такими є й матеріали конференцій, набори даних і програмні артефакти, а сіра література потребує явних критеріїв допуску.

Жодна база не є самодостатньою: Scopus і Web of Science придатні як основні, IEEE Xplore і ACM DL обов'язкові для галузі, Google Scholar слугує для перевірки повноти, OpenAlex, Semantic Scholar і Crossref дають відкриту інфраструктуру, а НБУВ, OUCI та репозитарії ЗВО формують національний шар, без якого огляд у дисертації неповний. Пошуковий рядок — вимірюваний артефакт: чутливість і точність обчислюють за еталонним набором (6.1) за цільової чутливості не нижче 80 %; пошук зупиняють за критерієм насичення (6.3), а гібридна стратегія «бази + сніжний ком» дає найкращий баланс повноти й трудомісткості.

Назвавши огляд систематичним, автор бере на себе зобов'язання щодо протоколу, критеріїв, двох скринерів, оцінювання якості та звітування за PRISMA 2020 (27 пунктів і потокова діаграма з числами на кожному кроці). Узгодженість скринерів вимірюють коефіцієнтом κ (6.4), а не відсотком згоди. Синтез завершується класифікацією, таблицею порівняння, переліком прогалин і суперечностей — саме вони обґрунтовують власне дослідження.

Критичне читання виконують у три проходи, звіряючи висновки з даними, і обов'язково перевіряють ретракції та декларації конфлікту інтересів. ШІ-асистенти прискорюють допоміжні кроки й систематично фабрикують посилання; чинним лишається просте правило — жодного джерела в списку без перевірки DOI й повного тексту, а факт використання ШІ розкривають у методології огляду.

Питання для самоперевірки

1. Чим первинне джерело відрізняється від вторинного й третинного? Наведіть по два приклади з вашої теми.
2. Чому в комп'ютерних науках матеріали конференцій вважають повноцінним первинним джерелом, а не «тезами»?
3. Що таке сіра література і за якими критеріями вирішують, чи включати її до огляду?
4. Назвіть п'ять вимог до бази даних як основного джерела систематичного огляду. Чому Google Scholar їм не відповідає?
5. Порівняйте покриття та обмеження Scopus, IEEE Xplore і OUCI. Який шар джерел втратить огляд, що спирається лише на Scopus?
6. Розкладіть власне питання на поняття за схемою PICOC і наведіть по три синоніми до кожного.
7. У чому різниця між `"internet of things"` і `{internet of things}` у Scopus? Коли усічення * шкодить? Чому рядок Scopus не можна механічно перенести в IEEE Xplore?
8. Що таке еталонний набір і як за ним обчислюють чутливість і точність рядка (6.1)? Яке порогове значення рекомендовано?

9. Сформулюйте критерій насичення (6.3). Чому насичення не є доказом повноти?
10. Чим зворотний сніжний ком відрізняється від прямого і в чому полягає гібридна стратегія пошуку?
11. Для яких питань доречний картувальний огляд, а для яких — систематичний? Чим tertiary study відрізняється від метааналізу?
12. Обчисліть κ (6.4) для $a = 18$, $b = 12$, $c = 10$, $d = 160$ і витлумачте результат за шкалою Ландіса і Коха. Чи можна продовжувати скринінг?
13. Перелічіть сім груп пунктів чекліста PRISMA 2020 і назвіть п'ять чисел, обов'язкових у потоковій діаграмі.
14. Опишіть три проходи критичного читання за Кешавом і назвіть чотири ознаки слабкого доказу.
15. Чому фабрикацію посилань Ш-моделлю вважають потенційним науковим шахрайством? Опишіть процедуру перевірки джерела.

Практичні завдання

1. *Рядок пошуку й оцінка його чутливості.* Розкладіть тему власного дослідження на 2–4 поняття із синонімами. Складіть еталонний набір із 15–20 свідомо релевантних праць (суцільний перегляд двох профільних видань за три роки). Побудуйте рядок для Scopus і його відповідник для IEEE Xplore, виконайте пошук, обчисліть S і D (6.1), а точність P і NNR (6.2) — за випадковою вибіркою 100–200 записів видачі з довірчим інтервалом. Якщо $S < 80\,\%$, уточніть рядок і повторіть; подайте таблицю «версія рядка — $|Q|$ — S — P» і причини пропусків.
2. *Протокол огляду.* Складіть протокол систематичного або картувального огляду за власною темою: 2–3 питання, бази й рядки, роки й мови, таблиця критеріїв за зразком табл. 6.3 (від 4 критеріїв включення й 5 виключення), процедура скринінгу, форма вилучення даних (від 12 полів), метод синтезу. Викладіть протокол у відкритий репозитарій і отримайте DOI.
3. *Скринінг, κ і потокова діаграма.* Разом із колегою незалежно проскринуйте 100–200 записів, побудуйте таблицю спряженості, обчисліть p_o, p_e і κ (6.4) та витлумачте значення. Виконайте другий етап і цикл сніжного кому з обчисленням ν_1 (6.3); побудуйте діаграму за зразком рис. 6.2 з числами на кожному кроці.
4. *Матриця понять і прогалини.* Для 10–15 включених праць побудуйте матрицю понять за зразком табл. 6.4 (від 6 понять) і таблицю порівня-

ння підходів. Сформулюйте дві прогалини й одну суперечність між дослідженнями, а на їх основі — проєкт мети власної роботи.

5. *Аудит бібліографії.* Імпортуйте результати пошуку в Zotero чи JabRef, дедублюйте, експортуйте `.bib` і підключіть до LaTeX. Перевірте кожен запис: розв'яжіть DOI, звірте авторів, рік, видання, сторінки, позначку про ретракцію. Окремо згенеруйте 10 посилань за темою чат-моделлю й порахуйте, скільки з них не існує або має хибні метадані.

РОЗДІЛ 7

Наукометрія та оцінювання результативності досліджень

Розділ 6 показав, як знаходити й систематизувати наукову інформацію. Але та сама інфраструктура цитувань слугує ще й інструментом *вимірювання* науки: за нею обчислюють показники авторів, журналів, установ і країн, на які спираються конкурси, атестації та рейтинги. Цей розділ пояснює, як улаштовані наукометричні показники, що саме кожен вимірює, де межі їх коректного застосування і якими способами їх спотворюють. Наскрізна теза: *метрика — інструмент, а не мета*; сліпе застосування показників деформує ту наукову поведінку, яку мало б оцінювати. Кейс — наукометричний профіль журналу «Кібербезпека: освіта, наука, техніка».

7.1 Предмет і межі: бібліометрія, наукометрія, інформетрія, альтметрія

7.1.1 Чотири суміжні дисципліни

Бібліометрія — кількісне вивчення документів і їх зв'язків: публікацій, посилань, співавторства, дескрипторів. **Наукометрія** ширша: вимірює науку як діяльність і соціальний інститут, тож поряд із публікаціями враховує кадри, фінансування, патенти, мобільність. **Інформетрія** — статистичні закономірності будь-яких інформаційних об'єктів. **Альтметрія** фіксує сліди наукового повідомлення поза формальним цитуванням: згадки в політичних документах, збереження в менеджерах посилань, покази, завантаження, обговорення у фахових мережах [273].

Це не ієрархія, а різниця в *об'єкті*: бібліометрія вимірює документи, наукометрія — науку, інформетрія — інформацію, альтметрія — увагу. Показник тлумачать лише в межах свого об'єкта: кількість завантажень не свідчить про наукову якість статті, а імпакт-фактор журналу — про якість статті в ньому. Л. Борнманн показав межі альтметрії: слабка кореляція з цитуваннями, залежність від платформ, легкість маніпуляції, нестабільність [72]. Отже, альтметрія — доказ суспільної уваги, а не замінник оцінювання змісту.

7.1.2 Коротка історія

Ю. Ґарфілд 1955 р. запропонував **покажчик цитувань** як інструмент *навігації*: посилання зв'язують працю за змістом краще, ніж рубрикатори, бо зв'язок установлює сам автор [155]. Science Citation Index (1964) мав служи-

ти пошуку; вимірювальні застосування виникли пізніше й усупереч задуму — сам Ґарфілд наполягав, що імпакт-фактор придатний для порівняння *журналів*, а не окремих учених [156].

Д. Прайс у праці «Little Science, Big Science» (1963) заклав емпіричну базу дисципліни: експонентне зростання літератури з періодом подвоєння близько 15 років, принцип кумулятивної переваги, нерівність внеску дослідників [271]. Звідси розуміння, що розподіли в науці принципово *несиметричні*. У 2000-х–2020-х рр. галузь змінили поява Scopus (2004), індекс Гірша (2005) [173] і відкриті метадані (Crossref, OpenCitations [265], OpenAlex [272]), що зробили розрахунки відтворюваними без передплати.

Три питання перед будь-яким показником. *Що* він вимірює — документ, джерело, особу, установу? *На яких даних* (значення для тієї самої особи в Scopus, WoS, Google Scholar і OpenAlex різняться на десятки відсотків)? *За яке вікно*? Показник без цих трьох відповідей не є фактом.

7.2 Емпіричні закономірності й розподіл цитувань

7.2.1 Закони Лотки, Брадфорда і Ципфа

А. Лотка 1926 р. встановив, що кількість авторів, які опублікували рівно n праць, спадає як степенева функція [216]:

$$A(n) = \frac{A(1)}{n^{\alpha}}, \qquad \alpha \approx 2. \tag{7.1}$$

За $\alpha = 2$ приблизно 60 % авторів мають рівно одну працю, 15 % — дві, 7 % — три. *Приклад.* Якщо у вибірці 1200 авторів написали одну статтю, то авторів із двома очікувано $1200/2^2 = 300$, з трьома — $1200/9 \approx 133$, з десятьма — 12. Наслідок: більшість імен у пошуковому наборі випадкові для теми; наукову школу утворює вузький «хвіст» продуктивних авторів.

С. Брадфорд 1934 р. описав розсіювання публікацій за журналами [75]: якщо впорядкувати журнали за спаданням кількості профільних статей і розбити на зони з однаковою кількістю статей, то кількості журналів у зонах співвідносяться як $1 : k : k^2$ (типово $k \approx 3$–5). *Приклад.* Тема дала 300 статей: перша зона — 5 журналів зі 100 статтями, друга — 20 журналів, третя — 80. Отже, суцільний перегляд 5 профільних видань дає третину корпусу, а «довгий хвіст» ловиться лише пошуковим рядком і сніжним комом (розділ 6). Дж. Ципф узагальнив ранговий розподіл: $f(r) \propto r^{-\beta}$ [355]; цю форму мають розподіли цитувань, частот ключових слів, розмірів колаборацій. Усі три закони — прояв однієї сім'ї важкохвостих розподілів, породжених кумулятивною перевагою.

7.2.2 Важкий хвіст і чому середнє вводить в оману

Розподіл цитувань статей у будь-якому журналі різко асиметричний. Ф. Радіккі зі співавторами показали, що після нормалізації на середнє по галузі розподіли різних дисциплін збігаються в одну універсальну криву: важкий хвіст є властивістю *механізму*, а не галузі [277].

Наслідок: *середнє арифметичне для важкохвостого розподілу не є типовим значенням.* У кейсі розділу 902 статті журналу зібрали 1116 цитувань; середнє — 1,24, медіана — нуль, 56 % статей не процитовано жодного разу, а верхні 10 % утримують 54 % цитувань. Твердження «типова стаття отримує близько 1,2 цитування» хибне: типова стаття не отримує жодного. Звідси вимога **розподілу замість середнього**: провідні видавництва ще 2016 р. запропонували публікувати повний розподіл цитувань журналу з медіаною і квартилями [214]. Правило: наводьте n, медіану, міжквартильний розмах і частку нецитованих; середнє — як додаток.

7.2.3 Вікно цитування, старіння, «сплячі красуні»

Вікно цитування — проміжок, за який зараховують посилання; на значення показника воно впливає не менше, ніж якість праць. В інформатиці цитування накопичуються повільніше за перший рік і швидше старіють після п'ятого, ніж у біомедицині: вагома частина комунікації відбувається на конференціях. Тому дворічне вікно імпакт-фактора систематично занижує оцінку ІТ-видань порівняно з чотирирічним вікном CiteScore. **Старіння літератури** вимірюють *періодом напівцитування* (cited half-life) — віком, до якого припадає половина отриманих посилань; для прикладних ІТ-статей це 3–5 років.

Винятком є **«сплячі красуні»** — праці, що роками лишаються непоміченими, а потім різко набирають цитування [330]. Ц. Ке зі співавторами показали, що «сплячість» — не аномалія, а неперервний спектр, значущий в усіх галузях [201]. Звідси: *відсутність цитувань протягом двох-трьох років не є доказом слабкості результату*, а процедури, що вимагають цитувань на свіжі публікації, вимірюють швидкість моди, а не якість.

І. Тахамтан зі співавторами систематизували понад двадцять чинників цитованості; якість — лише один із них, поряд із довжиною статті, кількістю авторів і країн, відкритим доступом, престижем журналу й мовою [319]. Для україномовної публікації мовний бар'єр означає структурно нижчу цитованість за того самого наукового рівня.

7.3 Показники автора

7.3.1 Індекс Гірша: означення, властивості, критика

Первинних величин лише три: N — кількість індексованих публікацій; C — сумарна цитованість; $\bar{c} = C/N$. Решта показників автора — фун-

кції від упорядкованого за спаданням вектора цитувань $c_1 \geq c_2 \geq \cdots \geq c_N$. Наскрізний приклад — умовний профіль $\mathcal{A}$ з вектором $(41, 33, 25, 19, 15, 12, 11, 9, 8, 7, 6, 5, 4, 3, 3, 2, 1, 1, 0, 0)$: $N = 20$, $C = 205$, $\bar{c} = 10{,}25$, медіана 6,5, дві праці нецитовані; три найцитованіші статті дають 48 % доробку.

Дж. Гірш 2005 р. запропонував показник, що поєднує продуктивність і вплив [173]:

$$h = \max\{\, i \in \mathbb{N} \;:\; c_i \geq i \,\}. \tag{7.2}$$

Тобто h — найбільше число, за якого автор має h праць, кожна з яких процитована щонайменше h разів. Для $\mathcal{A}$: $c_8 = 9 \geq 8$, але $c_9 = 8 < 9$, тож $h = 8$. Геометрично h — сторона найбільшого квадрата під кривою ранг–цитування (рис. 7.1).

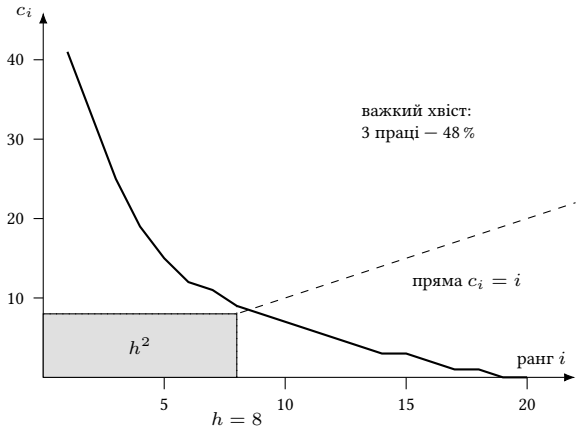


Рис. 7.1 — Крива ранг–цитування профілю $\mathcal{A}$ та положення h-індексу (7.2)

Властивості. h стійкий до одиничних викидів, не спадає з часом (тож зростає з віком кар'єри), $h \leq N$ і $h \leq \sqrt{C}$.

Критика [73, 337]: h *ігнорує хвіст* (для $\mathcal{A}$ це $205 - 8 \cdot 8 = 141$ «невидимих» цитувань, 69 % доробку); *не нормований за галуззю* (у біомедицині типові значення вдвічі–втричі вищі, ніж в інформатиці, а в математиці нижчі); *не нормований за стажем* (порівнювати h аспіранта й професора безглуздо); *залежить від бази* (у Google Scholar вищий, ніж у Scopus, бо індексуються препринти й навчальні матеріали); *нечутливий до внеску* (стаття з 28 співавторами зараховується як одноосібна).

Способи маніпуляції h: нарощування самоцитувань у «пограничних» працях до порога $h + 1$; домовлені взаємні цитування (citation cartel); нарізка результату (salami slicing); включення до списку авторів без внеску

заради збільшення N; вибір видань, що заохочують цитувати «свої» статті — усе це сумнівні дослідницькі практики (розділ 10).

7.3.2 g-індекс та i10

Л. Егґе запропонував g-індекс, чутливий до хвоста [129]:

$$g = \max\Big\{ g \in \mathbb{N} \;:\; \sum_{i=1}^{g} c_i \geq g^2 \Big\}. \tag{7.3}$$

Для $\mathcal{A}$ сума перших 14 значень дорівнює 198 і $198 \geq 14^2 = 196$, а для 15 маємо $201 < 225$; отже, $g = 14$. Різниця $g - h = 6$ показує, наскільки доробок «тримається» на кількох сильних працях. **i10-індекс** (Google Scholar) — кількість праць із щонайменше 10 цитуваннями; для $\mathcal{A}$ це 7. Перевага — прозорість, вада — абсолютний поріг, що не має сенсу поза галуззю.

7.3.3 Самоцитування й гіперавторство

Самоцитування — посилання на власні попередні праці. Помірне самоцитування нормальне: воно вибудовує тяглість дослідницької програми; проблемою стає аномальна *частка*:

$$s = \frac{C_{\text{self}}}{C}, \tag{7.4}$$

де C_{self} — цитування, у яких хоча б один автор цитованої праці є автором цитувальної. Дж. Йоаннідіс зі співавторами побудували базу профілів понад 100 тис. учених із показниками *з урахуванням і без урахування* самоцитувань: для окремих профілів ця частка перевищує 50 % [183]. Орієнтир: $s \leq 0{,}20$ звичайне, $0{,}20 < s \leq 0{,}30$ потребує пояснення, $s > 0{,}30$ — червоний прапорець. *Приклад.* Якщо з 205 цитувань профілю $\mathcal{A}$ 38 є самоцитуваннями, то $s = 0{,}19$, а h падає з 8 до 7; тому коректна довідка подає обидва числа.

Гіперавторство — публікації з десятками авторів; у кейсі розділу серед 870 статей є праці з 21 і 28 співавторами при середньому 2,91. Дж. Йоаннідіс документував тисячі дослідників — авторів понад 72 праць на рік, тобто однієї кожні п’ять днів [184]. За повного (full) підрахунку кожен співавтор отримує одиницю, за дробового — частку $1/k$; Л. Вальтман і Н. ван Ек показали, що вибір методу змінює рейтинги установ і країн сильніше, ніж вибір показника [338]. Правило: у звітності — дробовий підрахунок і таксономія CRediT [58].

Як коректно подати власні показники. «Станом на 21.09.2026 за даними Scopus: 14 публікацій, 96 цитувань (78 без самоцитувань), $h = 6$; за Google Scholar: 31 публікація, 214 цитувань, $h = 9$. Медіана — 3, нецитованих праць — 5; FWCI

— 0,87». Некоректно: «індекс Гірша — 9» без бази, дати й контексту. Вимоги до публікацій здобувача PhD в Україні викладено в розділі 2.

7.4 Показники джерела: журнали й конференції

7.4.1 Імпакт-фактор JCR

Імпакт-фактор (Journal Impact Factor, JIF) журналу за рік Y — середня цитованість у році Y матеріалів двох попередніх років [87]:

$$\mathrm{JIF}_Y = \frac{C_Y(Y-1) + C_Y(Y-2)}{P(Y-1) + P(Y-2)}, \tag{7.5}$$

де $C_Y(y)$ — цитування у році Y матеріалів року y, а $P(y)$ — кількість **цитовних матеріалів** (citable items: статті й огляди) року y. *Приклад.* Журнал випустив 40 статей 2023 р. і 50 статей 2024 р.; у 2025 р. вони зібрали 270 цитувань, тож $\mathrm{JIF}_{2025} = 270/90 = 3{,}0$.

Асиметрія формули (7.5) — головне джерело маніпуляцій: чисельник рахує цитування *усіх* матеріалів, а знаменник — лише цитовні. Звідси прийоми: нарощування частки редакційних матеріалів і листів, які цитуються, але не потрапляють у знаменник; **примусове цитування** (coercive citation) — вимога редактора додати посилання на цей журнал як умову прийняття; **картельне цитування** (citation stacking); збільшення кількості оглядів; маніпуляція датою випуску («online first») задля потрапляння у вікно.

П. Сеґлен ще 1997 р. показав, чому JIF не можна переносити на окрему статтю: розподіл цитувань усередині журналу важкохвостий, цитованість типової статті на порядок нижча за середню, а кореляція між JIF і цитованістю конкретної праці слабка [298]. Це — емпірична основа DORA [291].

7.4.2 CiteScore, SJR і SNIP

CiteScore (Scopus) усуває асиметрію знаменника й розширює вікно до чотирьох років [195]:

$$\mathrm{CiteScore}_Y = \frac{\sum_{y=Y-3}^{Y} C_{[Y-3,\,Y]}(y)}{\sum_{y=Y-3}^{Y} P(y)}, \tag{7.6}$$

де і чисельник, і знаменник охоплюють *однаковий* набір типів документів. *Приклад із кейсу.* Журнал опублікував у 2021–2024 рр. 351 документ, які зібрали 611 цитувань; $611/351 \approx 1{,}74$. Увага: щоб це число було саме $\mathrm{CiteScore}_{2024}$, треба враховувати лише ті цитування, що *надійшли* у 2021–2024 рр.; якщо ж узяти всі накопичені на дату вивантаження, результат буде іншим (див. підрозділ 7.8). Чотирирічне вікно вигідне IT-виданням, але повільніше реагує на падіння якості.

SJR (SCImago Journal Rank) — *рекурсивний* показник: вага цитування залежить від престижу цитувального журналу за логікою, близькою до PageRank; вікно — три роки, враховано тематичну близькість видань, частку самоцитувань обмежено третиною [163, 296]. Рекурсивність ускладнює маніпуляцію: цитування з маловідомого видання майже не додає ваги.

SNIP (Source Normalized Impact per Paper) нормалізує цитованість на «цитувальний потенціал» галузі [233]:

$$\mathrm{SNIP} = \frac{\mathrm{RIP}}{\mathrm{RDCP}}, \tag{7.7}$$

де RIP — середня цитованість статті журналу за трирічне вікно, а RDCP — відношення цитувального потенціалу поля журналу до медіани по базі. Ключовий момент: цитувальний потенціал оцінюють не за списками літератури самого журналу, а за характеристиками *публікацій, які його цитують* — за середньою кількістю їхніх активних посилань, тобто посилань на джерела в межах вікна й бази. У переглянутій 2012 р. методиці CWTS кожне цитування зважують обернено до цитувального потенціалу джерела, звідки воно надійшло [233]; саме тому SNIP чутливий до різної щільності цитування в галузях, а не до редакційної політики щодо обсягу бібліографій. *Приклад.* За $\mathrm{RIP} = 1{,}6$ і $\mathrm{RDCP} = 0{,}8$ маємо $\mathrm{SNIP} = 2{,}0$; значення близько 1,0 — середній для галузі рівень. Саме SNIP робить коректним порівняння журналу з математики й журналу з біоінформатики.

Таблиця 7.1 — Показники джерела: формула, вікно, база, вразливість

Показник	Суть	Вікно	База	Вразливість
JIF	Середнє цитування цитовних матеріалів (7.5)	2 роки	WoS, JCR	Асиметрія знаменника; примусове й картельне цитування; огляди
CiteScore	Симетричне середнє (7.6)	4 роки	Scopus	Нарощування «легкоцитовних» типів документів; самоцитування
SJR	Рекурсивна вага за престижем джерела	3 роки	Scopus	Мережі взаємного цитування «важких» видань
SNIP	Нормалізація на потенціал поля (7.7)	3 роки	Scopus	Залежить від коректності меж поля
h журналу	h за вектором цитувань статей	уся історія	будь-яка	Зростає з віком і обсягом

7.4.3 Квартилі: чому Q1 у Scopus і Q1 у WoS — різні речі

Квартиль — позиція журналу в упорядкованому за показником списку видань *однієї предметної категорії*: Q1 — верхні 25 %, Q2 — 25–50 %, Q3 — 50–75 %, Q4 — нижні 25 %. Джерел розбіжностей між системами чотири.

Різний показник ранжування: у WoS — JIF у категоріях JCR, у Scopus — SJR або перцентиль CiteScore у категоріях ASJC. *Різні класифікатори*: категорії *Computer Science, Information Systems* (JCR) і *Computer Networks and Communications* (ASJC) не збігаються. *Різне покриття*: Scopus індексує помітно більше джерел, ніж WoS Core Collection (≈28 тис. активних рецензованих журналів проти ≈22 тис.), з іншим галузевим і мовним балансом [235]; розрив ще більший, якщо рахувати всі типи джерел. Числа покриття змінюються щороку, тому в роботі їх наводять із датою звірки. *Множинність категорій*: журнал у Scopus може мати Q1, Q2, Q3 і Q3 у чотирьох категоріях, а в довідках указують найкращий. Правило запису: «Q2 (Scopus, CiteScore-перцентиль, категорія *Computer Networks and Communications*, 2025)» — коректно; «журнал Q2» — ні.

7.4.4 Конференції та рейтинг CORE

В інформатиці конференційна публікація часто вагоміша за журнальну: провідні конференції мають частку прийняття 12–20 %, суворіше й швидше рецензування, а результати старіють за два-три роки. Журнальні показники для них незастосовні, тому галузь користується рейтингом **CORE**: A* (флагманські), A (відмінні), B (добрі), C (базовий рівень рецензування); решта — неранговані [98].

Таблиця 7.2 — Інтерпретація квартилів і рівнів CORE

Рівень	Що означає	Чого не означає
Q1	Верхні 25 % категорії за обраним показником	Що стаття в журналі якісна або процитована
Q2–Q3	Середній рівень; для вузьких тем часто оптимальний вибір	Що журнал «слабкий»: у малій категорії Q3 буває єдиним профільним
Q4	Нижні 25 %; потребує перевірки редакційних практик	Автоматичної ознаки хижацтва
CORE A*, A	Флагманська чи відмінна конференція галузі (IEEE S&P, CCS, USENIX Security, NDSS, ICSE)	Що доповідь прирівняно до статті Q1 або що матеріали індексовані у Scopus
CORE B, C	Добрий і базовий рівень; придатні для апробації	Що рецензування формальне

Доповідь на конференції рівня A чи A* є сильним свідченням апробації, але у вітчизняних атестаційних процедурах не заміняє статті (вимоги — у розділі 2). Звідси стратегія: конференція — для швидкої апробації, журнал — для повної версії результату.

7.5 Нормовані показники, ідентифікатори та інфраструктура

7.5.1 FWCI і відсоткові ранги

Середня кількість посилань у математичній статті — близько 20, у статті з молекулярної біології — понад 40. Тому «сире» цитування порівнянне лише в межах однієї галузі, року й типу документа; нормалізація — не вдосконалення, а *умова допустимості* міжгалузевого порівняння [337]. **FWCI** (Field-Weighted Citation Impact) — відношення фактичної цитованості до очікуваної для публікацій того самого року, типу й галузі [134]:

$$\mathrm{FWCI}(S) = \frac{1}{|S|}\sum_{i \in S} \frac{c_i}{e_i}, \qquad e_i = \overline{c}\,(y_i,\, t_i,\, f_i), \tag{7.8}$$

де S — оцінюваний набір праць, c_i — фактичні цитування праці i, а e_i — середня цитованість у світі для року y_i, типу t_i і галузі f_i. Для однієї праці FWCI — це просто відношення c_i/e_i; для сукупності праць методика Elsevier визначає FWCI як *середнє арифметичне індивідуальних відношень* [134], а не як відношення сум. Значення 1,00 — середньосвітовий рівень; 2,11 — на 111 % вище; 0,60 — на 40 % нижче.

Приклад. Здобувач має 5 праць: стаття 2022 р. (12 цитувань, очікувано 5,4), стаття 2023 р. (6 і 4,1), доповідь 2023 р. (2 і 1,8), стаття 2024 р. (3 і 2,6), доповідь 2024 р. (0 і 1,1). Індивідуальні відношення — 2,222; 1,463; 1,111; 1,154; 0; їхня сума 5,951, отже FWCI $\approx 1{,}19$. Нецитована доповідь 2024 р. знижує показник, але не обнуляє: для свіжих доповідей очікування теж низьке.

Типова помилка. Якщо замість (7.8) обчислити відношення сум $\sum c_i / \sum e_i = 23/15{,}0 \approx 1{,}53$, результат відчутно завищений: таке агрегування зважує кожну працю її очікуваною цитованістю, тобто дає більшу вагу старшим статтям у «цитабельних» галузях. Це інший показник, і подавати його як FWCI не можна.

Відсотковий ранг (percentile) показує, яку частку праць свого року, типу й галузі перевершує ця праця; похідний агрегат — **частка публікацій у top-10 %** (очікуване значення — 10 %; 25 % означає у 2,5 раза вищу частку високоцитованих праць, тобто перевищення очікуваного рівня на 15 відсоткових пунктів, або на 150 % у відносному вираженні). Перцентильні показники стійкіші за FWCI, бо не залежать від екстремумів хвоста [338].

Способи маніпуляції нормованими показниками: вибір «зручної» категорії з низьким очікуванням; класифікація статті як огляду чи навпаки; вибір

бази з вигіднішою класифікацією; подання FWCI для 3–5 праць, де одна вдала стаття дає показник 4 і більше. Мінімальна вибірка — 10 праць, краще 20.

7.5.2 Постійні ідентифікатори та відкриті дані про цитування

Розрахунок розпадається, якщо об'єкти не ідентифіковані однозначно. **DOI** — постійний ідентифікатор об'єкта (ISO 26324 [187]) з відкритими метаданими у Crossref або DataCite. **ORCID** — ідентифікатор *дослідника* [256]: відокремлює однофамільців і зберігає атрибуцію при зміні прізвища, мови запису чи установи; для українських авторів критичний, бо одне прізвище трапляється як Shevchenko, Ševčenko і Shevchenko (з кириличною літерою всередині) — три різні особи. **ROR** — ідентифікатор *установи* [282]: без нього «Borys Grinchenko Kyiv University» і «B. Grinchenko KMU» рахуються окремо. **Crossref** — реєстратор DOI та API метаданих, включно з позначками про ретракцію [103]. **OpenAlex** і **OpenCitations** — відкриті графи праць і цитувань [272, 265] для розрахунків без передплати; український сегмент — OUCI [253].

7.5.3 Відтворюваність наукометричного дослідження

М. Віссер зі співавторами порівняли п'ять джерел бібліографічних даних і виявили істотні розбіжності покриття й повноти зв'язків цитування: той самий показник за Scopus, Web of Science, Dimensions, Crossref і Microsoft Academic дає різні значення [335]. Висновок зберігає силу й для новіших відкритих джерел (OpenAlex, що успадкував дані Microsoft Academic після закриття останньої у 2021 р.), але переносити на них саме ці числа не можна — потрібне окреме порівняння. Звідси мінімальні вимоги: зазначити *базу*, *дату вивантаження*, *запит*, *типи документів*, *правила дедублювання й розв'язання омонімії*, *версію інструменту* та опублікувати дані й скрипти (розділ 9). Без цього результат неперевірний.

7.6 Відповідальне оцінювання: критика метрик і реформа

7.6.1 Ефект Гудхарта

Закон Гудхарта у формулюванні М. Стретерн: «щойно показник стає метою, він перестає бути добрим показником». Механізм: раціональний агент оптимізує не приховану величину (наукову якість), а її вимірний замінник. М. Файр і К. Ґестрін на понад 120 млн записів показали, що формальні показники — кількість публікацій на автора, довжина списків літератури, частка самоцитувань — зростали швидше за будь-які ознаки наукового змісту [150]. В Україні ефект документував С. Назаровець: грошове заохочення за публікації в національних журналах без оцінки якості дає зростання

кількості публікацій, а не продуктивності [242]. Отже, *будь-який показник, прив'язаний до винагороди, буде оптимізований; питання лише, якою ціною.*

7.6.2 Leiden Manifesto, DORA, The Metric Tide, CoARA

Лейденський маніфест (2015) — 10 принципів відповідального використання наукометрії [171] (табл. 7.3).

Таблиця 7.3 — Десять принципів Лейденського маніфесту [171]

№	Принцип	Що це означає на практиці
1	Кількісна оцінка підтримує якісну експертну	Метрика — вхідні дані, а не вирок
2	Вимірювати відповідно до місії установи й дослідника	Для кібербезпеки доречні патенти, CVE, стандарти, не лише JIF
3	Захищати локально значущі дослідження	Безпека українських реєстрів вагома, хоч і не дає англомовних цитувань
4	Дані й процедури — відкриті й прості	Запит, база, дата, типи документів — у додатку
5	Оцінювані можуть перевірити дані	Профілі верифікують до, а не після рішення
6	Ураховувати галузеві відмінності цитування	Порівнювати h математика й біолога некоректно
7	Оцінювати особу за судженням про портфель	Читати 3–5 найкращих праць, а не рахувати індекси
8	Уникати хибної конкретності	JIF із трьома знаками після коми — оманлива точність
9	Ураховувати системні ефекти	Показник змінює поведінку тих, кого вимірює
10	Регулярно переглядати показники	Набір метрик фіксують на цикл оцінювання

DORA (2013) містить одну центральну вимогу: *не використовувати показники журналів (передусім імпакт-фактор) як замінник оцінки якості окремих статей* при ухваленні рішень про наймання, просування й фінансування [291]; натомість оцінювати зміст роботи й усі типи результатів — дані, код, артефакти.

The Metric Tide (2015) — незалежний огляд ролі метрик у британській системі оцінювання; його внесок — поняття **відповідальної метрики** з п'ятьма вимірами: *робастність* (найкращі доступні дані), *смиренність* (кількісне не підміняє якісного), *прозорість* (методика відкрита для перевірки), *різноманітність* (показники відображають різні місії й кар'єри), *рефлексивність* (систему переглядають з урахуванням її ж побічних ефектів) [347].

CoARA — коаліція, створена 2022 р. навколо Угоди про реформування оцінювання досліджень [88]. Чотири базові зобов'язання: 1) визнавати рі-

зноманітність внесків і кар'єр; 2) будувати оцінювання передусім на *якісній* експертній оцінці з центральною роллю рецензування, підтриманій відповідальним використанням кількісних показників; 3) відмовитися від недоречного застосування журнальних метрик, зокрема імпакт-фактора й h-індексу; 4) не використовувати рейтинги організацій. Ще шість зобов'язань організаційні: ресурси, перегляд критеріїв, обізнаність, обмін практиками, оцінювання впровадження й звітування.

Протиставлення «метрики або експертиза» хибне. Робоча модель — **informed peer review**: рішення ухвалює людина, а показники слугують контекстом. Правила: показники обчислюють *після* формулювання критерію; жоден не має вирішальної ваги; межі кожного зазначено наперед; у спірних випадках вирішує прочитана робота. Для дисертації в Україні квартиль журналу — формальна умова допуску (розділ 2), а предметом обговорення на захисті є зміст результату.

7.7 Хижацькі видання й конференції

7.7.1 Означення й ознаки

Консенсусне означення 2019 р.: **хижацькі видання** — журнали й видавці, які надають пріоритет власному зиску коштом науковості, подають хибну або невірогідну інформацію, відхиляються від найкращих редакційних практик, бракують прозорості й агресивно чи невибірково залучають авторів [162]. Ключове — *сукупність* ознак: жодна окремо взята риса (плата за публікацію, швидке рецензування, молодий журнал) хижацтва не доводить.

7.7.2 Журнали-клони й фабрики статей

Журнал-клон (hijacked journal) — підроблений сайт, що привласнює назву, ISSN і навіть архів реального видання; А. Абалкіна показала, що клони утворюють мережі з десятків сайтів зі спільною інфраструктурою [51]. Захист: адресу журналу беруть не з листа й не з пошукової видачі, а з реєстру ISSN, DOAJ або сайту наукової бібліотеки. **Фабрики статей** (paper mills) — комерційні структури, що виробляють рукописи й продають авторство [132]; наукометричний слід — аномальні мережі співавторства без спільної афіліації, шаблонні тексти, «замучені фрази» (tortured phrases); докладно — у розділі 10.

7.7.3 Перевірка перед поданням і наслідки помилки

Базовий інструмент — ініціатива **Think. Check. Submit.** із чеклистом перевірки видання за прозорістю політики, ідентифікацією редакції, рецензуванням, платою й індексацією [320]. Доповнюють його DOAJ [118], переліки Scopus Sources і Master Journal List, **Cabells Predatory Reports** [80]. Для кон-

Таблиця 7.4 — Ознаки хижацького видання й спосіб перевірки

Ознака	Прояв	Перевірка
Агресивне залучення	Лист із компліментами, «спецвипуск за вашою темою», дедлайн за 5 днів	Звідки адреса; чи відповідає профілю
Непрозора редакція	Немає афіліацій, ORCID, адрес; науковці в раді без їх згоди	Перевірити 3–4 прізвища у Scopus, написати одному
Фіктивне рецензування	Прийняття за 24–72 год без зауважень	Попросити зразок рецензії
Непрозорий APC	Ціна з'являється після прийняття; приватний рахунок	Пошук APC на сайті *до* подання
Фальшиві показники й неправдива індексація	«Global Impact Factor», «Cosmos IF», заява про Scopus, WoS чи DOAJ без підтвердження	JIF існує лише в JCR; ISSN у Scopus Sources, Master Journal List, DOAJ [118]
Надмірна широта й обсяг	«Journal of Engineering, Medicine and Social Sciences»; сотні статей на випуск	Архів за два роки
Журнал-клон	Копія назви й ISSN легітимного видання на іншому домені [51]	Звірити ISSN у порталі ISSN і домен видавця
Немає архівування й DOI	Немає політики збереження, ліцензії, DOI	Розв'язати кілька DOI статей

ференцій перевіряють історію випусків, реальність програмного комітету, частку прийняття, видавця матеріалів, наявність події в CORE.

Наслідки публікації в хижацькому виданні незворотні: стаття не зараховується до вимог; результат уважають оприлюдненим, тож повторна публікація стає самоплагіатом; видання може зникнути разом зі статтею; згадка в CV створює репутаційний ризик, а кошти не повертають. Правило: *перевірка коштує 20 хвилин, її відсутність — одну публікацію.*

7.8 Наукометрія як предмет дослідження: кейс українського IT-журналу

7.8.1 Що і як досліджує наукометрія

Наукометрія — не лише сервіс оцінювання, а й предмет дослідження. Типові об'єкти аналізу: **тематичні ландшафти** (карти понять за спільним уживанням ключових слів); **мережі співавторства** (вузли — автори, установи чи країни; ребра — спільні праці); **мережі співцитування** (дві праці цитують разом — усталені традиції); **бібліографічне поєднання** (спільні джерела — *актуальні* фронти). Інструменти: VOSviewer [329], `bibliometrix` мовою R [59], CiteSpace [85]; дані — експорт зі Scopus, WoS або API OpenAlex [255].

7.8.2 Профіль журналу «Кібербезпека: освіта, наука, техніка»

Дані отримано запитом до OpenAlex API за ISSN 2663-4023 (вивантаження 21.09.2026; обробка — `bibliometrix`); журнал — електронне фахове видання категорії «Б» у DOAJ, DOI-префікс `10.28925`.

Умови відтворення. Наведені нижче числа отримано запитом

```
GET https://api.openalex.org/works?filter=
primary_location.source.issn:2663-4023
&per-page=200&cursor=*
```

з посторінковим обходом курсором, без фільтра за типом документа; дата знімка — 21.09.2026. Ідентифікацію авторів узято такою, як її подає OpenAlex (поле `authorships.author.id`), без додаткового розв'язання омонімії; дедублювання — за DOI, а за його відсутності — за нормалізованою назвою й роком. Щоб кейс був перевірним, здобувач має разом зі статтею опублікувати архів вихідних записів у форматі JSON, скрипт обробки та контрольну суму вивантаження (розділ 9); без цих матеріалів числа неможливо підтвердити — і саме цього вимагає від інших робіт підрозділ 7.5. Наведені тут значення слід розглядати як ілюстрацію методики: повторне вивантаження на іншу дату дасть інші числа.

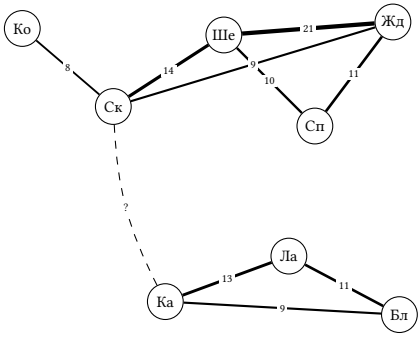
Зріз даних: 902 праці (870 з даними про авторство), 1116 цитувань, середнє 1,24 при медіані 0 і 56 % нецитованих; верхні 10 % праць утримують 54 % цитувань; h-індекс журналу — 10. За 2021–2024 рр. вийшов 351 документ (66, 62, 92 і 131 за роками), які зібрали 611 цитувань, накопичених станом на 21.09.2026, тобто середня накопичена цитованість статей 2021–2024 рр. — 1,74. Унікальних записів авторів — 1278 (без розв'язання омонімії), середня кількість авторів на працю — 2,91 за максимуму 28.

Фрагмент мережі співавторства (рис. 7.2) побудовано за матрицею спільних публікацій: вузол — автор, товщина ребра — кількість спільних праць. Видно дві щільні групи, характерні для українського ландшафту: стабільні внутрішньоуніверситетські колективи й слабкі зв'язки між ними.

7.8.3 Які висновки правомірні, а які ні

Правомірні. Розподіл цитувань важкохвостий, тому середнє 1,24 не описує типову статтю. Видання швидко зростає за обсягом (з 62 праць 2022 р. до 285 у 2025 р.), випереджаючи приріст цитувань, тож середня цитованість свіжих випусків нижча. Співавторська структура кластеризована, а 20 % одноосібних праць свідчать про вагому частку індивідуальних досліджень. Праці з 21 і 28 співавторами — привід перевірити відповідність внеску критеріям авторства (розділ 10).

Неправомірні. Не можна стверджувати, що «журнал має CiteScore 1,74»: це *оцінка за іншими даними й за іншим правилом підрахунку*, бо OpenAlex



І: Ше — Шевченко,
Жд — Жданова, Сп —
Спасітелєва, Ск —
Складанний, Ко — Костюк;
ІІ: Ла — Лахно,
Ка — Касаткін,
Бл — Блозва

Рис. 7.2 — Фрагмент мережі співавторства журналу; цифри — кількість спільних праць; пунктир зі знаком «?» — зв'язок, який OpenAlex подає для однофамільців і який не підтверджено без розв'язання омонімії (OpenAlex, 21.09.2026)

і Scopus різняться покриттям, а CiteScore рахує лише цитування, отримані *всередині* чотирирічного вікна, тоді як 1,74 — це середня цитованість, накопичена до дати знімка [335]. Коректне формулювання — «середня накопичена цитованість статей 2021–2024 рр. станом на 21.09.2026 становить 1,74». Не можна ранжувати авторів за кількістю праць у цьому журналі як за «продуктивністю»: це активність в *одному* виданні. Не можна вважати $h = 10$ свідченням якості: він зростає з віком і обсягом. Не можна тлумачити 56 % нецитованих праць як «половина статей непотрібні»: для україномовного корпусу з молодими випусками це очікуване значення.

7.8.4 Типові помилки наукометричного дослідження

Нечищені дані. У вивантаженні кейсу той самий автор фігурує як `Ivan Opirskyy` (з кириличними `I` та `O` в латинському записі) й окремим записом; прізвище `Zhdanova` закінчується кириличною літерою. Без нормалізації рядків і звіряння з ORCID мережа розпадається на фіктивні вузли. *Омонімія авторів.* `Ігор Козубцов` і `Lesya Kozubtsova` — різні особи, а кирилична й латинська форми одного прізвища — одна; злиття за прізвищем помилкове в обидва боки. *Неповне покриття бази.* Частина статей не має зв'язків цитування, бо цитувальні джерела не депонують списки літератури у Crossref, тож показники занижені. *Змішування типів документів і подвійний облік.* Огляди, редакційні матеріали й доповіді мають різні очікування цитування; дедублювання виконують за DOI, потім за назвою й роком. *Відсутність дати вивантаження.* Показники змінюються щодня; розрахунок без дати неперевірний.

7.9 Висновки до розділу

Наукометрія виросла з інструмента навігації в літературі [155] у систему вимірювання науки, від якої залежать кар'єри й бюджети. Формальний апарат нескладний: показники автора — функції від упорядкованого вектора цитувань, показники джерела — варіації відношення «цитування за вікно до кількості документів»; складним є коректне *тлумачення*.

Розподіли в науці важкохвості, тому середні не описують типовий випадок, а медіана, квартилі й частка нецитованих інформативніші. h-індекс (7.2) стійкий, але ігнорує більшу частину цитувань і не нормований ні за галуззю, ні за стажем, ні за базою. Імпакт-фактор (7.5) має асиметричний знаменник і вимірює журнал, а не статтю [298]; CiteScore (7.6) симетричний, із вигіднішим для IT вікном; SJR рекурсивний, SNIP (7.7) нормований за галуззю. Квартилі Scopus і WoS будуються за різними показниками й класифікаторами, тому не взаємозамінні; в IT до них додається рейтинг конференцій CORE. Міжгалузеві порівняння без нормалізації (FWCI (7.8), перцентилі, top-10 %) некоректні за означенням.

Кожен показник має відомий набір способів маніпуляції й підпорядкований закону Гудхарта [150]. Відповіддю спільноти стали Лейденський маніфест [171], DORA [291], The Metric Tide [347] і Угода CoARA [88]: рішення ухвалює експерт, а показники дають прозорий, перевірний і галузево нормований контекст. Практичний мінімум: мати ORCID і DOI для кожної праці; перевіряти видання за Think. Check. Submit [320]; подавати власні показники з назвою бази, датою й контекстом; не ототожнювати показник журналу з якістю своєї статті. Наукометричне дослідження має бути відтворюваним: запит, дата, база, правила очищення даних і скрипти — обов'язкові елементи звіту (розділ 9).

Питання для самоперевірки

1. Чим наукометрія відрізняється від бібліометрії, інформетрії та альтметрії? Наведіть по одному об'єкту вимірювання для кожної. Яку початкову мету мав Science Citation Index і чому Ґарфілд заперечував застосування імпакт-фактора до окремих учених?
2. Сформулюйте закон Лотки (7.1). Скільки авторів із трьома працями очікувано у вибірці, де 900 авторів мають одну працю? Що описує закон Брадфорда?
3. Чому середня цитованість уводить в оману? Які величини подають замість неї?
4. Що таке вікно цитування й період напівцитування? Чому дворічне вікно систематично невигідне IT-виданням? Що таке «сплячі красуні» і який висновок випливає з їх існування?

5. Дайте означення h-індексу (7.2). Обчисліть h, g (7.3) та i10 для вектора $(20, 14, 9, 7, 6, 5, 3, 2, 1, 0)$. Яка частка цитувань «невидима» для h?
6. Перелічіть п'ять обмежень h-індексу й три способи його штучного нарощування.
7. Запишіть формулу імпакт-фактора (7.5). Чому асиметрія чисельника й знаменника створює простір для маніпуляцій? Назвіть чотири прийоми.
8. Чим CiteScore (7.6) відрізняється від JIF? У чому полягає рекурсивність SJR і що нормалізує SNIP (7.7)? Обчисліть SNIP для RIP $= 2{,}4$, RDCP $= 1{,}2$.
9. Назвіть чотири причини, чому квартиль журналу у Scopus і у Web of Science може не збігатися. Як коректно записати квартиль у CV? Що означають рівні CORE A* і A?
10. Як улаштовано FWCI (7.8)? Обчисліть його за фактичними цитуваннями 8, 4, 0 та очікуваними 3,0, 2,5, 1,5. Чому FWCI для трьох праць неінформативний?
11. Перелічіть принципи 1, 6, 7 і 8 Лейденського маніфесту й поясніть, як кожен порушується фразою «у здобувача $h = 3$, тому робота слабка». Які чотири базові зобов'язання містить Угода CoARA?
12. Що таке ефект Гудхарта і як він виявляється в системах грошового заохочення за публікації?
13. Назвіть шість ознак хижацького видання та спосіб перевірки кожної. Чим журнал-клон відрізняється від хижацького журналу?
14. Які три типи мереж будують у наукометрії і що показує кожна? Чому в кейсі некоректно стверджувати, що журнал «має CiteScore 1,74»?

Практичні завдання

1. *Власний профіль і межі h-індексу.* Побудуйте впорядкований вектор цитувань власних праць (або праць наукового керівника) за Scopus, Web of Science, Google Scholar і OpenAlex на одну дату. Для кожної бази обчисліть N, C, $\bar{c}$, медіану, h (7.2), g (7.3), i10 та частку самоцитувань (7.4); побудуйте таблицю «база $-$ N $-$ C $-$ h» і поясніть розбіжності. Обчисліть частку цитувань, не врахованих h-індексом, і сформулюйте три твердження, яких за цим числом робити *не* можна.
2. *Перевірка видання за Think. Check. Submit.* Для двох журналів і однієї конференції, куди плануєте подати результат, пройдіть чеклист [320]: редакція й ORCID редакторів, рецензування, APC, ліцензія й архівування, індексація (ISSN у Scopus Sources, Master Journal List, DOAJ), до-

мен. Складіть таблицю «ознака — джерело перевірки — результат — висновок».

3. *Показники джерела для власного вибору.* Для трьох кандидатних видань зберіть JIF, CiteScore, SJR, SNIP, квартиль (із зазначенням показника й категорії) та рівень CORE; обчисліть CiteScore (7.6) самостійно за даними OpenAlex, обмеживши цитувальні документи роками вікна, і поясніть розбіжність з офіційним значенням. Окремо наведіть середню накопичену цитованість тих самих статей на дату вивантаження й покажіть, наскільки вона відрізняється від CiteScore. Оформте вибір як обґрунтоване рішення.
4. *Розподіл, мережа й очищення даних.* Вивантажте з OpenAlex усі праці одного журналу за п'ять років. Побудуйте криву ранг–цитування за зразком рис. 7.1, позначте h, обчисліть середнє, медіану, частку нецитованих і частку цитувань у верхніх 10 % праць; напишіть абзац, де ті самі дані подано спершу оманливо (через середнє), а потім коректно. Далі побудуйте мережу співавторства у VOSviewer чи `bibliometrix` [329, 59], виконавши очищення: нормалізація регістру, виявлення змішаних кирилично-латинських записів, злиття варіантів прізвища, звіряння з ORCID; порахуйте записи до і після очищення.
5. *Аудит відповідального оцінювання.* Зіставте чинний документ свого ЗВО про оцінювання наукової діяльності з десятьма принципами Лейденського маніфесту (табл. 7.3) і чотирма зобов'язаннями CoARA [88]: для кожного пункту вкажіть «дотримано / частково / порушено» з цитатою; сформулюйте три пропозиції щодо зміни методики.

РОЗДІЛ 8

Збір, обробка та статистичний аналіз експериментальних даних

Розділ 5 показав, як *спроєктувати* емпіричне дослідження в інформаційних технологіях: обрати план експерименту, розділити вибірки, уникнути витоку даних. Цей розділ починається там, де вимірювання вже виконано: як перетворити сирі числа на обґрунтоване твердження про світ і при цьому не обманути ні рецензента, ні себе. Наскрізна теза: *статистика не робить слабкий експеримент сильним* — вона лише чесно описує невизначеність. Тому виклад побудовано навколо типових помилок здобувача: для кожного блоку наведено характерну ваду й коректну альтернативу. Наскрізний кейс — порівняння двох алгоритмів виявлення вторгнень від розрахунку обсягу вибірки до формулювання висновку.

8.1 Дані, шкали вимірювання та якість вимірювання

8.1.1 Чотири шкали за Стівенсом

Перше рішення аналізу ухвалюють ще до першого обчислення: треба встановити, *що саме* означають числа в таблиці. Класифікація С. Стівенса [312] розрізняє чотири **шкали вимірювання** за тим, які перетворення зберігають зміст даних. **Номінальна**: числа — лише мітки класів (тип атаки, протокол, ідентифікатор вузла); осмислені частоти й мода. **Порядкова**: визначено «більше — менше», але не відстань між градаціями (рівень критичності вразливості, позиція в рейтингу, бал Лайкерта). **Інтервальна**: відстані рівні, нуль умовний (температура за Цельсієм, календарна дата), тож відношення «удвічі більше» позбавлене сенсу. **Шкала відношень**: є абсолютний нуль (час відгуку, пропускна здатність, кількість пакетів, енергоспоживання сенсора) — допустимі всі арифметичні дії. Лічбу подій (кількість хибних тривог за годину) аналізують як шкалу відношень, але моделями для лічильних даних — пуассонівською чи від'ємною біноміальною, а не нормальною апроксимацією.

8.1.2 Типова помилка: усереднення порядкових оцінок

Найпоширеніша вада ІТ-дисертацій — середнє арифметичне за порядковою шкалою. Теза «середня зручність інтерфейсу — 3,7 бала за шкалою Лайкерта» неявно припускає, що відстань між «радше не згоден» і «нейтрально» дорівнює відстані між «нейтрально» і «радше згоден»; це припущення

Таблиця 8.1 — Шкали вимірювання й допустимі статистики

Шкала	Що визначено	Допустимі статистики	Типовий критерій
Номінальна	Рівність	Частоти, мода, χ^2, Cramér's V, ентропія	χ^2, точний критерій Фішера
Порядкова	Порядок	Медіана, квантилі, IQR, рангові кореляції (ρ, τ)	Mann–Whitney, Wilcoxon, Kruskal–Wallis
Інтервальна	Рівність різниць	Середнє, SD, кореляція Пірсона	t-критерій, ANOVA
Відношень	Абсолютний нуль	Усе вище, а також геометричне середнє, коефіцієнт варіації	t-критерій, ANOVA, регресія

зазвичай хибне й ніколи не перевіряється. Ще гірший варіант — усереднення номінальної шкали: «середній тип атаки = 2,4».

> **Типова помилка.** «Середній рівень критичності вразливостей у системі — 2,8 з 4».
> **Коректно.** «Медіана рівня критичності — «високий»; розподіл: низький — 12 %, середній — 31 %, високий — 44 %, критичний — 13 % ($n = 118$)». Для порівняння двох систем — Mann–Whitney або χ^2 за таблицею спряженості, а не t-критерій.

Прийнятний компроміс: *сумарний* бал багатопунктової шкали (наприклад, System Usability Scale) трактують як квазіінтервальний, якщо доведено внутрішню узгодженість інструмента (Cronbach's α) і про це прямо сказано в розділі методів. Окремий пункт шкали залишається порядковим завжди.

8.1.3 Похибки, невизначеність і значущі цифри

Випадкова похибка непередбачувано змінюється від вимірювання до вимірювання; її вплив спадає як $1/\sqrt{n}$ і відбивається в довірчому інтервалі. **Систематична похибка (зміщення)** однаково спотворює всі вимірювання; збільшення вибірки її *не* усуває, а лише надає хибному значенню вужчий інтервал. Ці поняття унормовано словником VIM [197], а правила подання невизначеності — настановою GUM [196].

Характерні джерела зміщення в IT-експериментах: накладні витрати самого засобу вимірювання (профайлер, логер, `tcpdump`) — ефект спостерігача; «непрогріта» система з порожніми кешами й без JIT-компіляції; нестабільне середовище (сусідні віртуальні машини, фонові оновлення, теплове скидання частоти); неоднакові умови для порівнюваних варіантів; недостатня роздільна здатність таймера. Протиотрута — **калібрування** за еталоном, вимірювання «порожньої» операції, відкидання прогрівних ітерацій, ран-

домізація порядку запусків, чергування варіантів (interleaving) і фіксація повної конфігурації середовища.

За GUM результат подають з оцінкою **стандартної невизначеності** u або **розширеної невизначеності** $U = k\,u$ (зазвичай $k = 2$, що приблизно відповідає рівню довіри 95 %). Невизначеність типу A оцінюють статистично ($u_A = s/\sqrt{n}$), типу B — з паспортних даних приладу й роздільної здатності. Для $y = f(x_1, \dots, x_m)$ з незалежними аргументами діє закон поширення невизначеності:

$$u_c(y) = \sqrt{\sum_{i=1}^{m} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i)}, \tag{8.1}$$

звідки практичне правило: для суми й різниці у квадратурі додаються *абсолютні* невизначеності, для добутку й частки — *відносні*.

Правила округлення. Невизначеність округлюють до однієї, рідше двох значущих цифр; значення — до того самого розряду. Запис $94{,}3728 \pm 1{,}2$ Мбіт/с некоректний: правильно $(94{,}4 \pm 1{,}2)$ Мбіт/с, $k = 2$. Так само не варто звітувати точність класифікатора як $0{,}97863$ за тестової вибірки з 200 зразків.

> **Типова помилка.** Перенесення в текст усіх цифр, які видала бібліотека: «$F_1 = 0{,}8734216$», «$p = 0{,}0000000$».
> **Коректно.** «$F_1 = 0{,}873$ (95 % ДІ: 0,851–0,894)», «$p < 0{,}001$». Кількість розрядів визначає точність даних, а не розрядність типу `float64`.

8.2 Генеральна сукупність, вибірка й планування обсягу

8.2.1 Що є генеральною сукупністю в ІТ-дослідженні

Генеральна сукупність — множина всіх об’єктів, на які поширюється висновок; **вибірка** — її підмножина, на якій виконано вимірювання. В ІТ-дослідженнях генеральну сукупність часто залишають невизначеною, і саме тут виникає більшість хибних узагальнень. Перед аналізом треба відповісти письмово: сукупність — це всі можливі *запуски* алгоритму, усі *зразки трафіку*, усі *пристрої* певного класу, усі *розробники* певної кваліфікації чи всі *конфігурації* параметрів? Від відповіді залежить, що вважати незалежним спостереженням.

> **Типова помилка.** Модель навчено й перевірено на одному еталонному наборі, а у висновках стверджують ефективність «для систем виявлення вторгнень загалом».
> **Коректно.** Висновок обмежують сукупністю, з якої фактично взято дані («для

трафіку, подібного до CICIDS2017, за незмінного розподілу класів»), а розширення перевіряють на незалежному наборі іншого походження. Загрози зовнішньої валідності описано в розділі 5.

8.2.2 Види вибірок

Проста випадкова вибірка дає кожному елементові однакову ймовірність відбору, але потребує повного переліку (основи вибірки). **Систематична** бере кожен k-й елемент упорядкованого переліку: дешева, проте небезпечна за прихованої періодичності (кожен 60-й пакет може збігатися з циклом keep-alive). **Стратифікована** ділить сукупність на однорідні страти (типи атак, класи пристроїв) і відбирає всередині кожної — зменшує дисперсію оцінки й гарантує представлення рідкісних класів, тому є базовим вибором для незбалансованих даних кібербезпеки. У **кластерній** одиницею відбору є група (організація, підмережа, добовий зріз трафіку); спостереження всередині кластера корельовані, тож ефективний обсяг вибірки менший за номінальний. **Зручна** (convenience) вибірка — те, що доступно: студенти власної групи, публічний датасет, репозиторії GitHub; припустима лише як явно задекларований обмеження.

Репрезентативність — властивість не обсягу, а *процедури* відбору: тисяча зручно зібраних спостережень зміщена так само, як сто. Типові механізми **зміщення добору** в ІТ — виживання (аналізують лише проєкти, що «дожили» до релізу), самовідбір (на опитування відповідають найактивніші), доступність (відкриті репозиторії відрізняються від промислових) і публікаційне зміщення на рівні датасетів (оприлюднюють набори, на яких метод спрацював).

8.2.3 Потужність і обсяг вибірки

Критерій може помилитися двома способами. **Помилка I роду** (α) — відхилити правильну нуль-гіпотезу (хибне відкриття); **помилка II роду** (β) — не відхилити хибну (пропущений ефект). Величина $1-\beta$ — **статистична потужність**: ймовірність виявити ефект заданого розміру, якщо він справді існує; загальноприйнятий мінімум — $0{,}80$ [92]. Геометрію цих величин показано на рис. 8.1.

Для порівняння двох незалежних середніх за двобічної альтернативи мінімальний обсяг *кожної* групи обчислюють так:

$$n = \frac{2\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2}{d^2}, \qquad d = \frac{|\mu_1 - \mu_2|}{\sigma}, \tag{8.2}$$

де d — **мінімальний істотний ефект** (MDE) у стандартизованих одиницях. За $\alpha = 0{,}05$ ($z = 1{,}96$) і потужності $0{,}80$ ($z = 0{,}842$) чисельник дорівнює $2 \cdot (2{,}8016)^2 = 15{,}70$, тож $n \approx 15{,}70/d^2$: для $d = 0{,}8$ потрібно

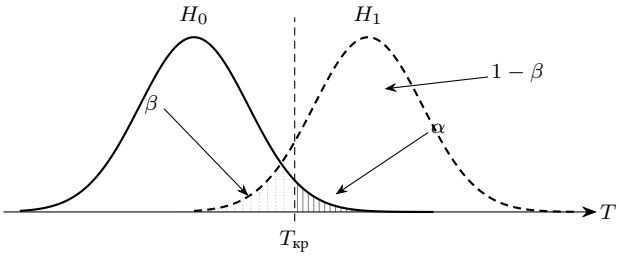


Рис. 8.1 — Розподіли статистики критерію за H_0 (суцільна) і H_1 (штрихова): рівень значущості α, помилка II роду β і потужність $1-\beta$

25 спостережень у групі, для $d=0{,}6$ — 44, для $d=0{,}5$ — 63, для $d=0{,}2$ — 393. Квадратична залежність пояснює, чому «трохи менший» ефект подвоює ціну експерименту.

Розрахунок виконують *до* збору даних, а MDE обґрунтовують практичною значущістю: який приріст F_1 чи зниження затримки має сенс для впровадження. Для непараметричних критеріїв обсяг збільшують на 5–15 % (асимптотична відносна ефективність критерію Манна–Вітні проти t-критерію за нормальності — $3/\pi \approx 0{,}955$).

> **Типова помилка — *post hoc* потужність.** Отримавши $p=0{,}21$, обчислюють потужність *за спостереженим* ефектом і пишуть: «потужність 0,27, тому дослідження недостатньо потужне». Така величина — монотонна функція того самого p і нової інформації не несе.
> **Коректно.** Або планувати потужність заздалегідь, або звітувати довірчий інтервал ефекту: «різниця $-0{,}004$ (95 % ДІ: $-0{,}011$; $0{,}003$) — дані сумісні з відсутністю практично значущої переваги». Хронічно недостатня потужність підвищує частку хибних відкриттів серед опублікованих результатів [78].

8.3 Описова статистика та розвідувальний аналіз даних

8.3.1 Міри положення й розсіяння

Міри положення описують «типове» значення: середнє арифметичне $\bar{x}$, медіана Me (50-й процентиль), мода. **Міри розсіяння** описують мінливість: вибіркова дисперсія $s^2=\frac{1}{n-1}\sum(x_i-\bar{x})^2$, стандартне відхилення s, розмах, міжквартильний розмах $\mathrm{IQR}=Q_3-Q_1$, медіанне абсолютне відхилення $\mathrm{MAD}=\mathrm{Me}(|x_i-\mathrm{Me}(x)|)$.

Середнє й дисперсія мають нульову *точку зламу*: одне довільно велике значення зміщує їх як завгодно далеко. Медіана має точку зламу 50 %, MAD — теж; це **робастні** оцінки. Для зіставності з s використовують нормовану величину $1{,}4826\cdot\mathrm{MAD}$, яка за нормального розподілу є *асимптотично*

узгодженою оцінкою σ: зі зростанням n вона збігається до σ. Це не те саме, що незміщеність: за скінченного обсягу вибірки оцінка має зміщення, помітне для малих n, тож множник 1,4826 забезпечує зіставність шкал, а не відсутність зміщення.

Конвенція квантилів. Значення Q_1, Q_3 і процентилів залежать від алгоритму інтерполяції, і різні пакети дають різні числа. Тут і далі в посібнику використано тип 7 за класифікацією Хайндмана й Фена — лінійну інтерполяцію за позицією $h = (n-1)p+1$; це стандартна поведінка `quantile()` у `R` і `numpy.percentile` з методом `linear`. Конвенцію обов'язково зазначають у методах: від неї залежать і IQR, і межі Тьюкі, і правила зупинки, побудовані на квартилях (наприклад, у методі Дельфі — розділ 4).

Числовий приклад. Затримка виявлення інциденту (мс) за 11 прогонів: 12, 14, 15, 15, 17, 18, 19, 21, 24, 29, 98. Тоді $\bar{x} = 282/11 = 25{,}64$; Me $= 18$; мода = 15; $s = 24{,}49$; $Q_1 = 15{,}0$, $Q_3 = 22{,}5$ (тип 7), IQR $= 7{,}5$; MAD $= 3$, отже $1{,}4826 \cdot 3 = 4{,}45$. Розбіжність $s = 24{,}5$ проти $4{,}45$ — надійний індикатор того, що дані містять викид. Після вилучення значення 98 маємо $\bar{x} = 18{,}4$, Me $= 17{,}5$, $s = 5{,}13$.

Для сильно асиметричних величин (час відгуку, розмір файлу, кількість цитувань) середнє арифметичне вводить в оману. Звітують медіану й квантилі, а для вимог до якості обслуговування — «хвостові» процентилі P_{95}, P_{99}: користувач відчуває саме хвіст розподілу, а не середнє.

8.3.2 Викиди: виявляти, а не видаляти

Класичне правило Тьюкі [325] позначає як викид значення поза межами $[Q_1 - 1{,}5\,\mathrm{IQR};\ Q_3 + 1{,}5\,\mathrm{IQR}]$. Для наведеного прикладу межі становлять $[3{,}75;\ 33{,}75]$, тож значення 98 — викид. Правило «$\pm 3s$» для цього непридатне, бо сам викид роздуває s; робастна версія використовує $|x_i - \mathrm{Me}|/(1{,}4826\,\mathrm{MAD}) > 3$.

Виявлення викиду — початок розслідування, а не підстава для вилучення. Протокол: зафіксувати значення; встановити причину (збій приладу, помилка парсингу логів, одиниці вимірювання, реальна рідкісна подія); ухвалити рішення *за правилом, сформульованим заздалегідь*; задокументувати його й навести аналіз чутливості — результати з викидом і без нього.

> **Типова помилка.** «Три аномальні прогони вилучено як нерепрезентативні» — без пояснення й без правила.
> **Коректно.** «Правило вилучення визначено до аналізу: прогони, у яких профайлер зафіксував конкуренцію за CPU понад 5 %. Вилучено 3 з 44 прогонів; результати без вилучення наведено в додатку (висновок не змінюється)». У кібербезпеці «викид» часто і є шуканою атакою — вилучення знищує сигнал.

8.3.3 Розвідувальний аналіз і чесна графіка

Розвідувальний аналіз даних (EDA) за Дж. Тьюкі [325] передує будь-якій перевірці гіпотез: треба побачити форму розподілу, масштаб, пропуски, дублікати, залежності. Квартет Ф. Енскомба [57] — чотири набори з однаковими середніми, дисперсіями, кореляцією ($r = 0{,}816$) і однаковою лінією регресії, але цілковито різними діаграмами розсіяння — канонічний доказ того, що зведені показники без графіка оманливі.

Робочий мінімум візуалізації: **гістограма** (форма розподілу, мультимодальність), **коробкова діаграма** (медіана, IQR, викиди), **діаграма розсіяння** (зв'язок двох величин), **емпірична функція розподілу** (ECDF — показує всі дані без довільного вибору ширини стовпця), **скрипкова діаграма** (violin — щільність плюс квартилі). Для малих вибірок ($n < 30$) замість коробкової діаграми показують усі точки (strip, beeswarm): boxplot за $n = 6$ приховує більше, ніж показує.

Правила чесної графіки [324]: вісь значень стовпчикової діаграми починається з нуля (обрізана вісь перетворює різницю в 1 % на видиме «подвоєння»); для лінійних графіків і діаграм розсіяння нульова вісь не обов'язкова, але розрив осі позначають явно; кожна вісь має підпис і одиницю вимірювання; поряд із точковою оцінкою показують невизначеність, а в підписі вказують, що саме вона означає (SD, SE чи ДІ); логарифмічна шкала припустима для величин, що змінюються на порядки, але про неї треба сказати; неприпустимі тривимірні ефекти, зрізані пропорції площ і зображення, що не відтворюються в чорно-білому друці.

> **Типова помилка.** Стовпчикова діаграма точності двох моделей з віссю від 0,90 до 0,95: різниця 0,008 виглядає як дворазова перевага.
> **Коректно.** Вісь від нуля або — краще — графік *різниці* з довірчим інтервалом, де видно і величину ефекту, і його невизначеність [104].

Окремий обов'язковий крок EDA — облік **пропущених значень**: скільки їх, чи випадковий механізм пропуску, чим замінено. Мовчазне відкидання рядків із пропусками (listwise deletion) саме є процедурою добору й може створити зміщення.

8.4 Перевірка статистичних гіпотез і p-значення

8.4.1 Логіка перевірки

Нуль-гіпотеза H_0 — твердження про відсутність ефекту («середні значення F_1 двох алгоритмів рівні»). **Альтернативна гіпотеза** H_1 — її заперечення, двобічне ($\mu_1 \neq \mu_2$) або однобічне ($\mu_2 > \mu_1$). Однобічну альтернативу обирають *лише* тоді, коли протилежний напрям ефекту не має практичного

сенсу, і фіксують це до збору даних; перехід на неї після перегляду даних подвоює фактичний рівень значущості.

Процедура: обчислюють статистику критерію T, визначають її розподіл за умови істинності H_0 і знаходять імовірність отримати значення, щонайменше настільки ж екстремальне, як спостережене. Для двох незалежних середніх за рівних дисперсій:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}, \qquad s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}, \tag{8.3}$$

де s_p — об'єднане стандартне відхилення, число ступенів свободи $\nu = n_1 + n_2 - 2$. За нерівних дисперсій застосовують варіант Велча [342] з наближеним ν за формулою Велча–Сатертвейта; він є *кращим типовим вибором*, бо майже не втрачає потужності за рівних дисперсій і зберігає рівень значущості за нерівних.

8.4.2 Що таке p-значення і чим воно не є

p-значення — це ймовірність отримати таке саме або екстремальніше значення статистики критерію за умови, що нуль-гіпотеза *та всі інші припущення моделі* (незалежність, розподіл, коректність добору, відсутність вибіркового звітування) істинні [160]. Наголос на «всіх інших припущеннях» принциповий: мале p сигналізує про несумісність даних із *усією* моделлю, а не обов'язково з H_0.

Заява Американської статистичної асоціації 2016 р. [339] формулює шість принципів.

1. p-значення можуть указувати на несумісність даних із заданою статистичною моделлю.
2. p-значення *не* вимірюють імовірність того, що гіпотеза істинна, ані ймовірність того, що дані отримано випадково.
3. Наукові висновки й рішення не повинні ґрунтуватися лише на тому, чи перетнуло p-значення певний поріг.
4. Належний висновок потребує повного звітування й прозорості аналізу.
5. p-значення не вимірює ані розміру ефекту, ані важливості результату.
6. Саме по собі p-значення не є добрим мірилом сили доказів щодо моделі чи гіпотези.

Продовженням стала редакційна стаття 2019 р. з гаслом «відмовитися від поняття *статистично значуще*»: рекомендовано звітувати точні p-значення як неперервну величину поруч з оцінкою ефекту й інтервалом, а не як дихотомію [340].

Типові хибні тлумачення: «$p = 0{,}03$ означає, що ймовірність помилки 3 %» (ні: це не ймовірність хибності висновку); «отже, H_1 істинна з імовірністю 97 %» (ні: p обчислено *за умови* H_0); «$p > 0{,}05$ доводить відсутність відмінності» (ні: відсутність доказу не є доказом відсутності — потрібні критерії еквівалентності, наприклад TOST, з наперед заданою межею байдужості); «$p = 0{,}049$ і $p = 0{,}051$ — принципово різні результати» (ні: різниця між «значущим» і «незначущим» сама по собі не є значущою).

8.4.3 Спотворення: p-hacking, HARKing, «сад стежок»

P-hacking — добір аналітичних рішень до отримання $p < 0{,}05$: дозбирування даних із проміжними перевірками, вибіркове вилучення спостережень, перебір метрик і трансформацій, включення й виключення коваріат. Моделювання Дж. Сіммонса зі співавторами показало, що лише чотири звичні «ступені свободи дослідника» підвищують частку хибних відкриттів із 5 % до понад 60 % [304].

HARKing (hypothesizing after the results are known) — подання гіпотези, сформульованої після перегляду даних, як апріорної [203]: результат розвідувального аналізу видають за підтверджувальний, і рівень значущості втрачає сенс.

Сад стежок, що розгалужуються — тонша проблема [158]: навіть без свідомого перебору дослідник ухвалює десятки залежних від даних рішень (як згрупувати, що вважати викидом, яку метрику взяти). Кожна стежка виглядає єдино можливою, але сукупність потенційних стежок робить фактичний рівень помилки I роду значно вищим за номінальний. **Публікаційне зміщення** завершує картину: «позитивні» результати приймають охочіше, тому опублікована література систематично переоцінює ефекти — середній приріст нового методу над базовим за літературою завищений.

> **Протиотрути.** Передреєстрація плану аналізу (розділ 9); чітке розмежування підтверджувальної й розвідувальної частин у тексті («аналіз є розвідувальним, гіпотези підлягають перевірці на незалежних даних»); звітування *усіх* виконаних порівнянь, а не найкращого; публікація коду й даних; поправка на множинність (п. 8.6).

8.5 Розмір ефекту, довірчі інтервали й вибір критерію

8.5.1 Розмір ефекту

p-значення залежить і від сили ефекту, і від обсягу вибірки: за мільйона спостережень статистично значущою стає різниця, позбавлена практичного сенсу. **Розмір ефекту** — стандартизована або натуральна міра сили явища, не залежна від n. Для різниці двох середніх це d Коена:

$$d = \frac{\bar{x}_2 - \bar{x}_1}{s_p}, \qquad g = d\left(1 - \frac{3}{4(n_1+n_2)-9}\right), \tag{8.4}$$

де g — виправлена на зміщення оцінка Хеджеса [168], істотна за $n < 20$. Для зв'язку двох кількісних величин ефектом є r Пірсона, для часток дисперсії в ANOVA — η^2 і менш зміщене ω^2, для таблиць спряженості — V Крамера, для порівняння шансів — **відношення шансів** OR, для рангових порівнянь — ймовірність переваги $A = U/(n_1 n_2)$ (common-language effect size).

Таблиця 8.2 — Орієнтири інтерпретації розміру ефекту

Показник	Що вимірює	Малий	Серед.	Великий
d Коена, g Хеджеса	Різниця середніх у SD	0,20	0,50	0,80
r Пірсона	Лінійний зв'язок	0,10	0,30	0,50
η^2, ω^2	Частка поясненої дисперсії	0,01	0,06	0,14
V Крамера ($df = 1$)	Зв'язок у таблиці спряженості	0,10	0,30	0,50
OR	Відношення шансів	1,5	2,5	4,0
A (ймовірність переваги)	Частка пар, де $x_2 > x_1$	0,56	0,64	0,71

Межі за [92] — орієнтири, а не закон. У зрілих галузях типовий ефект менший, і «малий» $d = 0{,}2$ може мати велике практичне значення (зниження частки хибних тривог на 2 в. п. для оператора SOC), натомість $d = 1{,}2$ у порівнянні з навмисно ослабленим базовим методом не означає нічого. Тому поряд зі стандартизованим ефектом завжди наводять *натуральну* різницю в одиницях предметної галузі: мілісекунди, відсоткові пункти, кількість інцидентів за добу.

8.5.2 Довірчі інтервали

Довірчий інтервал рівня $1-\alpha$ — інтервальна оцінка, що за нескінченного повторення дослідження накриє істинне значення параметра в $(1-\alpha)$ частці випадків: для різниці середніх — $(\bar{x}_2 - \bar{x}_1) \pm t_{1-\alpha/2,\nu}\, s_p \sqrt{1/n_1 + 1/n_2}$. Хибним є тлумачення «істинне значення лежить у цьому інтервалі з імовірністю 95 %»: параметр не випадковий, випадковий інтервал [160].

ДІ інформативніший за p, бо показує напрям, величину й точність оцінки [104]. Вузький інтервал навколо нуля — доказ відсутності практично значущого ефекту; широкий інтервал, що перетинає нуль, — доказ невизначеності, а не відсутності. Обидва випадки дають $p > 0{,}05$, але протилежні висновки.

8.5.3 Вибір критерію й перевірка передумов

Параметричні критерії (t, ANOVA, лінійна регресія) припускають: незалежність спостережень; приблизну нормальність (для t — нормальність *розподілу вибіркових середніх*, тож за $n \gtrsim 30$ центральна гранична теорема робить критерій стійким); однорідність дисперсій. Перевіряють їх так: нормальність — Q–Q-графіком і критерієм Шапіро–Вілка [300]; однорідність — критерієм Брауна–Форсайта (робастною версією критерію Лівіня) [77]; незалежність — аналізом плану експерименту, а не статистикою.

> **Типова помилка.** Відхилення нормальності за Шапіро–Вілком на вибірці $n = 2000$ ($p = 0{,}004$) і перехід на непараметричний критерій. За великих n критерій виявляє мізерні, практично нерелевантні відхилення.
> **Коректно.** Оцінювати *ступінь* відхилення (Q–Q-графік, асиметрія, ексцес), а не лише p. Найсерйозніша передумова — це незалежність: її порушення (кілька вимірювань з одного пристрою, повторні запуски на тому самому зразку) не лікується вибором непараметричного критерію, а потребує змішаної моделі або агрегування.

Якщо передумови порушено, порядок дій такий: критерій Велча замість класичного t за нерівних дисперсій; перетворення шкали (логарифм для часу відгуку, $\log(1 + x)$ для лічби) з поверненням до натуральних одиниць у висновках; непараметричний ранговий критерій; бутстреп або перестановковий критерій, що не потребує припущень про форму розподілу [128]; узагальнена лінійна модель з адекватним розподілом. Уточнення: критерії Манна–Вітні [218], Вілкоксона для парних вимірів [345] і Крускала–Волліса [209] перевіряють не рівність медіан, а стохастичну однорідність розподілів, тож висновок «медіани рівні» з них не випливає, якщо форми розподілів різні.

8.6 Множинні порівняння та контроль хибних відкриттів

8.6.1 Сімейна помилка

Якщо за рівня $\alpha = 0{,}05$ виконати m незалежних перевірок, то ймовірність хоча б одного хибного відкриття (**сімейна помилка**, FWER) дорівнює $1 - (1 - \alpha)^m$: для $m = 5$ це 0,23, для $m = 10$ — 0,40, для $m = 20$ — 0,64. Порівнюючи новий метод із базовим на десяти датасетах за трьома метриками, дослідник виконує 30 перевірок і майже гарантовано знайде «значущу» перевагу навіть за повної відсутності ефекту.

Множинність породжують не лише явні критерії: кожна додаткова метрика, підгрупа, поріг класифікації чи варіант препроцесингу — теж перевірка; саме тому «сад стежок» [158] вважають прихованою множинністю.

Таблиця 8.3 — Вибір критерію за типом даних і планом

План	Кількісні, передумови виконано	Кількісні/порядкові, передумови порушено	Категорійні
Дві незалежні групи	t Велча, d Коена	Mann–Whitney U, ефект A	χ^2, точний Фішера, OR
Дві залежні (парні)	Парний t, d_z	Wilcoxon signed-rank	McNemar
$k > 2$ незалежних груп	Однофакторна ANOVA, η^2, Tukey HSD	Kruskal–Wallis, Dunn із поправкою	χ^2 на таблиці $r \times c$
$k > 2$ залежних вимірів	ANOVA з повторними вимірами	Friedman, Nemenyi	Cochran Q
Зв'язок двох величин	Кореляція Пірсона, лінійна регресія	Spearman ρ, Kendall τ	χ^2, V Крамера

8.6.2 Контроль FWER: Бонферроні й Голм

Поправка Бонферроні порівнює кожне p_i з α/m (еквівалентно $p_i^{\text{adj}} = \min(1,\, m\,p_i)$): проста й завжди коректна, але консервативна — за великих m потужність падає катастрофічно. **Процедура Голма** [175] послідовно відхильна й рівномірно потужніша за того самого контролю FWER: p-значення впорядковують за зростанням $p_{(1)} \le \dots \le p_{(m)}$, порівнюють $p_{(i)}$ з $\alpha/(m-i+1)$ і зупиняються на першому невідхиленні.

8.6.3 Контроль FDR: Бенджаміні–Хохберг

Жорсткий контроль FWER доречний, коли ціна *будь-якого* хибного відкриття висока. У розвідувальних задачах (скринінг сотень ознак, порівняння багатьох конфігурацій) раціональніше контролювати **частку хибних відкриттів** (FDR) — очікувану частку помилкових серед *відхилених* гіпотез. Процедура Бенджаміні–Хохберга [66]: знайти найбільше k, для якого

$$p_{(k)} \le \frac{k}{m}\, q, \tag{8.5}$$

і відхилити всі гіпотези з рангами $1, \dots, k$; q — прийнятний рівень FDR (часто 0,05 або 0,10). Еквівалентні скориговані значення: $p_{(i)}^{\text{adj}} = \min_{j \ge i}\{\min(1,\, m\,p_{(j)}/j)\}$.

Числовий приклад. Це той самий експеримент, що й у кейсі підрозділу 8.9: новий детектор порівняно з базовим на шести незалежно зібраних наборах D1–D6, на кожному — за виправленим критерієм (8.8). Сирі p-значення: 0,0023 (D1); 0,0071 (D4); 0,0193 (D2); 0,0402 (D6); 0,0610 (D3); 0,2500 (D5) — $m = 6$, $\alpha = q = 0{,}05$.

Без поправки «значущими» виглядають чотири результати з шести. Бонферроні й Голм відхиляють дві гіпотези (Голм зупиняється на $i = 3$,

Таблиця 8.4 — Поправки на множинність: шість порівнянь

i	Набір	$p_{(i)}$	α/m	$\alpha/(m-i+1)$	iq/m	Рішення
1	D1	0,0023	0,0083	0,0083	0,0083	Bonf.+Holm+BH
2	D4	0,0071	0,0083	0,0100	0,0167	Bonf.+Holm+BH
3	D2	0,0193	0,0083	0,0125	0,0250	лише BH
4	D6	0,0402	0,0083	0,0167	0,0333	—
5	D3	0,0610	0,0083	0,0250	0,0417	—
6	D5	0,2500	0,0083	0,0500	0,0500	—

бо $0{,}0193 > 0{,}0125$), а Бенджаміні–Хохберг знаходить найбільше k з $p_{(k)} \le k \cdot 0{,}05/6$: це $k = 3$ ($0{,}0193 \le 0{,}0250$), тож відхиляє три. Скориговані значення за BH: 0,0138; 0,0213; 0,0386; 0,0603; 0,0732; 0,2500; за Голмом — 0,0138; 0,0355; 0,0772; 0,1206; 0,1220; 0,2500.

Типова помилка. Виконати 30 парних порівнянь, звітувати три з $p < 0{,}05$ і не згадати про решту.
Коректно. Задекларувати сімейство гіпотез *до* аналізу, навести всі m значень (зокрема в додатку), застосувати поправку, відповідну до мети (FWER для підтверджувальних висновків, FDR для розвідувальних), і вказати її в тексті: «застосовано процедуру Бенджаміні–Хохберга, $q = 0{,}05$, $m = 30$».

8.7 Кореляція, регресія та причинність; байєсівський погляд

8.7.1 Коефіцієнти зв'язку

Кореляція Пірсона r вимірює *лінійний* зв'язок кількісних величин; вона чутлива до викидів і не бачить нелінійних залежностей (квартет Енскомба [57]). ρ **Спірмена** — та сама формула на рангах: придатна для порядкових шкал і монотонних нелінійних зв'язків. τ **Кендалла** спирається на частку узгоджених пар, стійкіша за малих вибірок і має пряме ймовірнісне тлумачення, тому для порядкових даних і малих n обирають саме її.

8.7.2 Регресія

Лінійна модель $y = \beta_0 + \sum_j \beta_j x_j + \varepsilon$ оцінює умовне середнє відгуку. **Коефіцієнт детермінації** R^2 — частка поясненої дисперсії; він автоматично зростає з кожним доданим предиктором, тому звітують скоригований R^2_{adj}, а моделі порівнюють за AIC/BIC або похибкою на відкладених даних. Діагностика обов'язкова: графік залишків проти передбачених значень (нелінійність, гетероскедастичність), Q–Q-графік залишків, відстань Кука (впливові точки), автокореляція залишків для часових рядів.

Мультиколінеарність — сильна лінійна залежність між предикторами — не псує прогноз, але робить коефіцієнти нестабільними й непридатними

для інтерпретації. Індикатор — фактор інфляції дисперсії $\mathrm{VIF}_j = 1/(1-R_j^2)$, де R_j^2 — коефіцієнт детермінації регресії x_j на решту предикторів; значення понад 5–10 потребують дій (вилучення чи об'єднання ознак, регуляризація).

Логістична регресія моделює ймовірність бінарної події через логіт $\ln \frac{P}{1-P} = \beta_0 + \sum_j \beta_j x_j$, де e^{β_j} — відношення шансів за зміни x_j на одиницю. Це базова модель для задач «атака/норма», і її коефіцієнти інтерпретують у шансах, а не у ймовірностях.

8.7.3 Кореляція не є причинністю

Три типові пастки. **Конфаундер** — спільна причина обох величин: кореляція між кількістю правил IDS і числом виявлених інцидентів може породжуватися розміром організації. **Зворотна причинність**: не складність коду спричиняє дефекти, а дефектні модулі частіше переписують. **Парадокс Сімпсона** [305] — напрям зв'язку в об'єднаних даних протилежний напряму в кожній підгрупі: метод A виграє і на легких, і на важких датасетах, але програє в об'єднаній таблиці через різну частку задач у групах. Тому агрегувати дані можна лише після перевірки однорідності підгруп.

Причинні твердження потребують або рандомізованого втручання, або явної причинної моделі з перевіркою припущень — графів залежностей, обчислення ефекту втручання, аналізу чутливості до неврахованих конфаундерів [260]. У тексті дисертації обсерваційний результат формулюють мовою асоціації: «пов'язано з», «асоційовано», а не «зменшує», «спричиняє».

8.7.4 Байєсівський погляд

Байєсівський підхід оцінює не ймовірність даних за гіпотези, а **апостеріорний розподіл** параметра: $P(\theta \mid D) \propto P(D \mid \theta)\, P(\theta)$. Замість довірчого інтервалу будують **вірогідний інтервал** (credible interval), який допускає саме те пряме тлумачення, якого хибно очікують від довірчого: «з імовірністю 95 % параметр лежить у цих межах за заданих даних і апріорного розподілу».

Фактор Байєса $BF_{10} = P(D \mid H_1)/P(D \mid H_0)$ показує, у скільки разів дані змінюють відношення правдоподібностей на користь H_1 [199]. На відміну від p-значення, він симетричний: $BF_{10} = 0{,}1$ — це доказ *на користь* H_0, а не просто «незначущий результат» [336]. Орієнтири: 1–3 — ледь помітний доказ, 3–10 — помірний, 10–30 — сильний, понад 100 — вирішальний. Ціна — залежність від апріорного розподілу, тому звітують аналіз чутливості до його вибору.

8.8 Статистичне оцінювання моделей машинного навчання

8.8.1 Матриця плутанини й похідні метрики

Основа оцінювання бінарного класифікатора — **матриця плутанини** з чотирма лічильниками: TP (виявлені атаки), FP (хибні тривоги), FN (пропущені атаки), TN (правильно пропущений нормальний трафік). Усі *порогові* метрики табл. 8.5 — функції цих чотирьох чисел, тому матрицю наводять повністю: за нею читач відтворить будь-яку з них, а за самим F_1 — жодної.

Межі цього твердження слід розуміти точно. Матриця плутанини відповідає *одному* порогу рішення й не містить ранжування прогнозів, тому з неї неможливо відновити показники, що підсумовують поведінку класифікатора за всіма порогами, — ROC-AUC і average precision (AP). Щоб читач міг перерахувати й ці величини, потрібні або істинні мітки разом із числовими оцінками моделі, або таблиця (TP, FP, FN, TN) для всіх порогів, потрібних для побудови кривої. Практичне правило: разом зі статтею публікують файл «мітка — оцінка» (розділ 9).

Таблиця 8.5 — Метрики бінарної класифікації

Метрика	Формула	Коли доречна
Accuracy	$\dfrac{TP+TN}{TP+TN+FP+FN}$	Лише за збалансованих класів
Precision	$\dfrac{TP}{TP+FP}$	Ціна хибної тривоги висока
Recall (TPR)	$\dfrac{TP}{TP+FN}$	Ціна пропуску атаки висока
Specificity (TNR)	$\dfrac{TN}{TN+FP}$	Навантаження на оператора
Balanced acc.	$\dfrac{TPR+TNR}{2}$	Незбалансовані класи
F_1	$\dfrac{2\,TP}{2\,TP+FP+FN}$	Компроміс, TN ігнорується
MCC	див. (8.7)	Збалансована оцінка всіх чотирьох клітинок

$$F_1 = \frac{2\,P \cdot R}{P+R}, \qquad F_\beta = \frac{(1+\beta^2)\,P \cdot R}{\beta^2 P + R}, \tag{8.6}$$

$$\mathrm{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \in [-1;1]. \tag{8.7}$$

Коефіцієнт (8.7) запропонував Б. Меттьюз [223]; він є коефіцієнтом кореляції Пірсона між істинними й передбаченими мітками, тому враховує всі чотири клітинки матриці й не завищує оцінку на незбалансованих даних.

Міра F_β з $\beta = 2$ вдвічі важливішою вважає повноту — типовий вибір для систем виявлення вторгнень, де пропуск дорожчий за хибну тривогу.

8.8.2 Чому accuracy оманлива

Числовий приклад. Тестовий набір: 100 000 потоків, з них 1 000 атак (1 %). Модель дала $TP = 820$, $FN = 180$, $FP = 1\,960$, $TN = 97\,040$, тобто accuracy $= 97\,860/100\,000 = 0{,}9786$, — але тривіальний класифікатор «усе нормальне» дає 0,99. Інформативні метрики: $P = 820/2780 = 0{,}295$; $R = 0{,}820$; $F_1 = 0{,}434$; $TNR = 97\,040/99\,000 = 0{,}980$; збалансована точність $(0{,}820 + 0{,}980)/2 = 0{,}900$; MCC $= 0{,}484$. Розрив між 0,979 і 0,484 і є ціною незбалансованості [86]. Додається й **хибність базової частки**: за 1 % атак навіть 2 % хибних тривог дають майже дві хибні на кожну правильну (розділ 5).

8.8.3 Порогові криві: ROC і PR

ROC-крива показує залежність TPR від FPR за всіх порогів класифікації; AUC дорівнює ймовірності того, що випадковий позитивний зразок отримає вищу оцінку, ніж випадковий негативний [148]. За сильної незбалансованості ROC оптимістична: велика зміна кількості хибних тривог майже не зсуває FPR, бо знаменник $TN + FP$ величезний. PR-крива інформативніша, а її базовий рівень дорівнює частці позитивів [290].

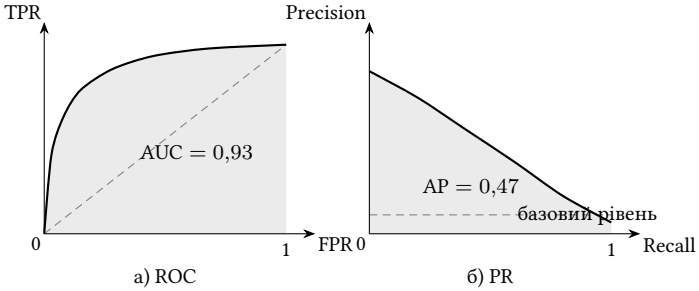


Рис. 8.2 — ROC- і PR-криві того самого класифікатора на незбалансованих даних: площа під кривою (AUC, AP) і базові рівні (за [148, 290])

8.8.4 Невизначеність метрик і порівняння класифікаторів

Метрика, обчислена на скінченній тестовій вибірці, — оцінка з похибкою, тому її подають з інтервалом. Універсальний інструмент — **бутстреп** [128]: з тестової вибірки $B \geq 1000$ разів беруть повторну вибірку того самого обсягу з поверненням, щоразу перераховують метрику й беруть перцентилі 2,5 % і 97,5 %. Для часток працює й аналітичний інтервал Вілсона: для $R = 820/1000$ він дає $[0{,}795;\ 0{,}843]$, для $P = 820/2780$ — $[0{,}278;\ 0{,}312]$.

Два класифікатори на *одній* тестовій вибірці порівнюють критерієм Мак-Немара [226] за кількостями неузгоджених рішень; на кількох розбиттях —

парним критерієм на згортках із поправкою на перетин навчальних вибірок; багато моделей на багатьох датасетах — процедурою Фрідмана з апостеріорним критерієм (формули наведено в розділі 5). Різницю площ під *корельованими* ROC-кривими (та сама вибірка) оцінюють непараметричним критерієм ДеЛонга [108], а не порівнянням двох незалежних інтервалів.

> **Типова помилка.** «Запропонована модель краща: AUC 0,934 проти 0,929».
> **Коректно.** «AUC 0,934 (95 % ДІ: 0,921–0,946) проти 0,929 (0,916–0,941); різниця 0,005 (95 % ДІ: $-0{,}004$–$0{,}014$), критерій ДеЛонга $p = 0{,}27$. Перевага не підтверджена». Три значущі цифри в метриці за тестової вибірки в кілька тисяч зразків — це шум, поданий як результат.

8.9 Кейс: порівняння двох алгоритмів виявлення вторгнень

Задача. Аспірант пропонує детектор B і порівнює його з базовим A за зваженим F_1 на трафіку, подібному до CICIDS2017.

Крок 0. Що саме узагальнюють. Це питання вирішують до планування, бо від нього залежить одиниця спостереження. Розрізняють два різні твердження.

(а) Про цей набір даних. «Чи кращий B за A на трафіку цього набору?» Джерело випадковості — розбиття на навчання й тест і зерна генератора; сукупність, на яку поширюється висновок, — інші розбиття *того самого* набору. Саме на це питання відповідають прогони.

(б) Про новий трафік. «Чи кращий B за A на трафіку, який модель ще не бачила?» Джерело випадковості — вибір мережі, періоду й умов захоплення; одиниця спостереження — *незалежне захоплення трафіку*, а не прогін. Кількість прогонів на одному наборі не наближає до цього висновку: скільки б їх не було, дані ті самі. Твердження (б) обґрунтовують кроком 5 — повторенням експерименту на кількох незалежно зібраних наборах.

Крок 1. Планування до експерименту. Гіпотези для (а): H_0: $\mu_B = \mu_A$; H_1: $\mu_B \neq \mu_A$ (двобічна). Одиниця спостереження — прогін на випадковому стратифікованому розбитті «навчання/тест» (70/30) з фіксованим зерном; страти — типи атак. Розбиття *спільні* для A і B: на кожному j-му розбитті обидва алгоритми навчають і тестують на тих самих даних, тому аналізують парні різниці $d_j = F_{1,B}^{(j)} - F_{1,A}^{(j)}$. Практично значущим визнано приріст F_1 на 0,03 за очікуваного $\sigma \approx 0{,}05$, тобто $d = 0{,}6$. На етапі планування розкид парних різниць s_d ще невідомий (пілотних прогонів не було), тому обсяг обчислено *консервативно* — за (8.2) для двох незалежних середніх, з $\alpha = 0{,}05$ і потужністю 0,80: $n = 15{,}70/0{,}36 = 43{,}6$, тож 44 пари спостережень. Це свідомий запас: для парного плану потрібне $n \geq (z_{1-\alpha/2} + z_{1-\beta})^2 / d_z^2$ без множника 2, де $d_z = \bar{d}/s_d$, і за фактично отриманого $s_d = 0{,}0147$ вистачило б близько десяти пар. Планувати за меншим

числом до вимірювання s_d не можна — оцінка d_z наперед невідома. Порядок прогонів рандомізовано, перші три «прогрівні» ітерації відкинуто за правилом, зафіксованим у протоколі.

Крок 2. Описова статистика. Отримано: A — $\bar{x} = 0{,}871$, $s = 0{,}049$; B — $\bar{x} = 0{,}903$, $s = 0{,}046$ — на тих самих 44 розбиттях, тобто це 44 *пари* спостережень, а не дві незалежні групи по 44. Коробкові діаграми показують симетричні розподіли без викидів за правилом Тьюкі.

Крок 3. Передумови й межі незалежності. Q–Q-графіки парних різниць лінійні, Шапіро–Вілк для d_j: $p = 0{,}37$. Проте головна передумова — незалежність — виконується лише частково, і це треба сказати прямо: 44 розбиття взято з *одного* набору даних, тож вони перекриваються й повторно використовують ті самі спостереження. Різні зерна генератора створюють *випадковість процедури*, а не нові вибірки із сукупності трафіку. Наслідок кількісний: оцінка дисперсії середньої різниці занижена, бо не враховує кореляції між розбиттями, а звичайний t-критерій (як незалежний двовибірковий, так і простий парний) завищує значущість. Перехід на парний критерій сам собою цієї вади не усуває.

Тому для (а) застосовують **виправлений критерій для повторного перевибирання** Надо й Бенжіо [238], який додає до дисперсії доданок за повторне використання даних:

$$t = \frac{\bar{d}}{\sqrt{\left(\dfrac{1}{n} + \dfrac{n_{\text{test}}}{n_{\text{train}}}\right) s_d^2}}, \qquad \nu = n - 1, \tag{8.8}$$

де n — кількість розбиттів, s_d^2 — вибіркова дисперсія парних різниць, а $n_{\text{test}}/n_{\text{train}}$ — частка перекриття, зумовлена планом розбиття. Для *однієї* фіксованої тестової вибірки натомість беруть критерій Мак-Немара за неузгодженими рішеннями.

Крок 4. Критерій, ефект, інтервал. Описово: $\bar{d} = 0{,}0320$, $s_d = 0{,}0147$ — парні різниці набагато стабільніші за самі показники ($s_p = 0{,}0475$), бо складність розбиття діє на обидва алгоритми однаково. Наївний парний t-критерій дав би $t = 14{,}4$ і $p < 10^{-15}$ — саме таке число й з'являється в роботах, де повторне використання даних не враховано. За (8.8) з $n = 44$ і співвідношенням 30/70: стандартна похибка $0{,}0099$, $t = 3{,}24$, $\nu = 43$, $p = 0{,}0023$; 95 % ДІ різниці — $[0{,}0121;\ 0{,}0519]$. Стандартизований ефект за (8.4): $d = 0{,}0320/0{,}0475 = 0{,}67$ (межі, перелічені з інтервалу різниці, — $[0{,}25;\ 1{,}09]$); поправка Хеджеса дає $g = 0{,}67$ — за такого обсягу вона несуттєва. Ефект середній; нижня межа ДІ різниці (0,012) менша за заявлений поріг практичної значущості 0,03, тож дані сумісні і з приростом, меншим

за практично важливий. Важливо: усе це — висновок (а), про цей набір даних.

Крок 5. Множинність і перехід до висновку (б). Щоб говорити про новий трафік, експеримент повторено на шести незалежно зібраних наборах D1–D6 (різні мережі, періоди й умови захоплення); на кожному — та сама процедура з (8.8). Сімейство гіпотез задекларовано до аналізу; застосовано процедуру Бенджаміні–Хохберга, $q = 0{,}05$, $m = 6$ (табл. 8.4; $p = 0{,}0023$ — це набір D1, розглянутий у кроках 1–4): значущою перевага залишилася на трьох наборах із шести, за Голмом — на двох. У статті наведено всі шість сирих і скоригованих значень. Саме ця неоднорідність, а не одне мале p, і є підставою для обережного формулювання про межі узагальнення.

> **Неправильне формулювання.** «Запропонований алгоритм статистично значуще кращий ($p < 0{,}05$) і може бути рекомендований для впровадження в системах виявлення вторгнень».
>
> **Правильне формулювання.** «На 44 спільних стратифікованих розбиттях набору D1 середня парна перевага алгоритму B за зваженим F_1 становила 0,032 (95 % ДІ: 0,012–0,052; виправлений на повторне використання даних критерій (8.8): $t(43) = 3{,}24$; $p = 0{,}002$; d Коена $= 0{,}67$, межі 0,25–1,09). Прогони не є незалежними вибірками із сукупності трафіку, тому цей результат стосується саме набору D1. На шести незалежно зібраних наборах перевага відтворилася на трьох після контролю FDR ($q = 0{,}05$). Оскільки нижня межа інтервалу нижча за заявлений поріг практичної значущості (0,03), висновок обмежено твердженням про наявність ефекту без підтвердження його практичної достатності; узагальнення поширюється лише на трафік, подібний до досліджених наборів».

8.10 Інструменти та звітність статистичних результатів

8.10.1 Програмне середовище

Мова `R` [276] створена для статистики: багатий набір критеріїв «з коробки», пакети для потужності (`pwr`), розміру ефекту (`effectsize`), множинності (`p.adjust`), відтворюваних звітів (`knitr`, `rmarkdown`, `renv`). Екосистема `Python` зручніша, коли аналіз поєднано з машинним навчанням: `pandas` для підготовки даних [225], `SciPy` для критеріїв [334], `statsmodels` для регресій, ANOVA й поправок на множинність [297], `scikit-learn` для метрик і крос-валідації [261]. `Jupyter` поєднує код, результат і пояснення в одному документі [207], але зошит стає відтворюваним лише за фіксації версій, зерен генератора випадкових чисел і виконання «згори вниз» перед здачею (розділ 9).

```
from scipy import stats
t, p = stats.ttest_ind(b, a, equal_var=False)
```

```
from statsmodels.stats.multitest import multipletests
rej, p_adj, *_ = multipletests(pvals, alpha=0.05,
method='fdr_bh')
```

Правило гігієни: аналіз має бути скриптом, а не послідовністю дій у графічному інтерфейсі. Скрипт відтворюваний, придатний до рецензування й дозволяє перерахувати все після виправлення однієї помилки в даних.

8.10.2 Що саме звітувати

Настанови SAMPL [213] формулюють мінімум, який має містити статистичний опис. Контрольний список для статті чи розділу дисертації:

1. одиниця спостереження, спосіб добору, обсяг вибірки та його обґрунтування (розрахунок потужності з наведеними α, $1-\beta$, MDE);
2. кількість і причини вилучених спостережень; опрацювання пропусків і викидів за наперед заданим правилом;
3. описова статистика: для симетричних даних — $\bar{x} \pm s$, для асиметричних — медіана й квартилі; завжди вказувати, що саме стоїть після знака $\pm$;
4. назва критерію, його однобічність/двобічність, перевірені передумови та результат перевірки;
5. точне p-значення (не «$p < 0{,}05$»; «$p < 0{,}001$» — допустимий виняток), розмір ефекту й довірчий інтервал;
6. сімейство гіпотез і застосована поправка на множинність;
7. версії програмного забезпечення й пакетів, зерна генератора, посилання на код і дані (DOI репозитарію);
8. розмежування підтверджувального й розвідувального аналізу;
9. обмеження: на яку сукупність поширюється висновок.

8.11 Висновки до розділу

Статистичний аналіз починається не з обчислень, а з рішення про те, що означають числа: шкала вимірювання визначає допустимі операції, і усереднення порядкових оцінок залишається найпоширенішою вадою IT-дисертацій. Обсяг вибірки й потужність планують *до* експерименту, виходячи з мінімального практично значущого ефекту; розрахунок потужності після отримання p-значення інформації не додає. Систематичну похибку не лікує збільшення вибірки — лише калібрування, рандомізація порядку й контроль середовища; результат подають із невизначеністю та правильною кількістю значущих цифр.

Описова статистика й розвідувальний аналіз передують будь-якій перевірці гіпотез: робастні оцінки та графіки виявляють те, чого не показують зведені показники, а викид — привід для розслідування, а не для мовча-

зного вилучення. p-значення вимірює несумісність даних із *усією* статистичною моделлю, а не ймовірність гіпотези; шість принципів ASA [339] і рекомендація відмовитися від порогової дихотомії [340] мають стати робочою нормою. Звітувати треба розмір ефекту й довірчий інтервал поруч із точним p; практична значущість визначається предметною галуззю, а не порогом 0,05.

Вибір критерію диктують тип даних, план і передумови; порушення незалежності — найсерйозніше й не лікується переходом на непараметричний критерій. Множинні порівняння потребують задекларованого сімейства гіпотез і поправки: FWER (Бонферроні, Голм) для підтверджувальних висновків, FDR (Бенджаміні–Хохберг) — для розвідувальних. Кореляція не є причинністю: конфаундери, зворотна причинність і парадокс Сімпсона руйнують наївні інтерпретації. В оцінюванні моделей машинного навчання accuracy за незбалансованих класів оманлива; матрицю плутанини наводять повністю, метрики супроводжують бутстреп-інтервалами, а порівняння моделей виконують призначеними для цього критеріями. Зрештою жодна процедура не компенсує поганого плану: статистика лише чесно описує невизначеність того, що вже виміряно.

Питання для самоперевірки

1. Назвіть чотири шкали вимірювання за Стівенсом і по дві допустимі статистики для кожної. Чому середнє арифметичне за порядковою шкалою некоректне?
2. Чим репрезентативність вибірки відрізняється від її обсягу? Наведіть три механізми зміщення добору в IT-дослідженні.
3. Що таке помилки I та II роду, рівень значущості й потужність? Чому мінімальним рівнем потужності вважають 0,80?
4. Обсяг вибірки зростає у скільки разів, якщо мінімальний істотний ефект зменшити з $d = 0{,}8$ до $d = 0{,}4$? Обґрунтуйте за (8.2).
5. Чим систематична похибка відрізняється від випадкової і чому збільшення вибірки усуває лише одну з них?
6. Сформулюйте точне означення p-значення. Наведіть три його поширені хибні тлумачення й поясніть, у чому помилка кожного.
7. Перелічіть шість принципів заяви ASA 2016 р. Який із них найчастіше порушують в IT-статтях?
8. Чим p-hacking відрізняється від HARKing і від «саду стежок, що розгалужуються»? Які практики запобігають кожному з них?
9. Чому розмір ефекту й довірчий інтервал інформативніші за p-значення? Наведіть приклад, коли статистично значущий результат практично не значущий, і протилежний приклад.

10. Які передумови має t-критерій і як перевіряють кожну? Що робити, якщо порушено нормальність, однорідність дисперсій, незалежність?
11. У чому різниця між контролем FWER і контролем FDR? За яких дослідницьких завдань доречний кожен і чому Голм завжди кращий за Бонферроні?
12. Поясніть парадокс Сімпсона на прикладі порівняння двох алгоритмів на легких і важких датасетах. Як його виявити?
13. Чим вірогідний інтервал відрізняється від довірчого і що показує фактор Байєса, чого не показує p-значення?
14. Чому accuracy оманлива за частки атак 1 %? Які метрики наводять натомість і чому PR-крива інформативніша за ROC?
15. Складіть перелік із дев'яти позицій, які має містити статистичний опис у розділі результатів, і поясніть призначення кожної.

Практичні завдання

1. *Паспорт змінних власного дослідження.* Складіть таблицю всіх змінних вашого експерименту: назва, означення, шкала, одиниця, діапазон, спосіб вимірювання, джерело систематичної похибки, допустимі статистики. Позначте змінні, для яких у вашому чернетковому аналізі обчислено недопустиму статистику, і замініть її коректною.
2. *Розрахунок обсягу вибірки.* Обґрунтуйте мінімальний практично значущий ефект для вашої задачі (у натуральних одиницях), оцініть σ за пілотними даними чи літературою й обчисліть n за (8.2) для потужності 0,80 і 0,90. Побудуйте таблицю «$d - n$» для $d \in \{0{,}2; 0{,}3; 0{,}5; 0{,}8\}$ і поясніть, який обсяг реально здійсненний за ваших ресурсів.
3. *Повний цикл порівняння.* На власних (або відкритих) даних порівняйте два методи: EDA й перевірка передумов, вибір критерію з обґрунтуванням, p-значення, розмір ефекту з 95 % ДІ (для інтервалу скористайтеся бутстрепом на $B = 2000$ повторень). Оформте результат одним абзацом за зразком кейсу п. 8.9.
4. *Поправка на множинність.* Для серії з $m \geq 8$ порівнянь (різні датасети, метрики або конфігурації) обчисліть скориговані значення за Бонферроні, Голмом і Бенджаміні–Хохбергом; побудуйте таблицю за зразком табл. 8.4 і вкажіть, скільки гіпотез відхиляє кожна процедура. Поясніть, яку з них ви обрали б для підтверджувального висновку й чому.
5. *Метрики на незбалансованих даних.* Для власного класифікатора наведіть повну матрицю плутанини й обчисліть accuracy, precision, recall,

specificity, F_1, F_2, збалансовану точність і MCC (8.6), (8.7). Порівняйте з тривіальним класифікатором «усе нормальне». Побудуйте ROC- і PR-криві й поясніть розбіжність у враженні, яке вони створюють.

6. *Аудит чужої статистики.* Візьміть статтю з фахового видання (наприклад, [10]) і перевірте її за контрольним списком п. 8.10: чи обґрунтовано обсяг вибірки, чи перевірено передумови, чи наведено розмір ефекту й ДІ, чи враховано множинність, чи відповідає формулювання висновку отриманим числам. Оформте висновок як рецензію на одну сторінку з конкретними пропозиціями.

РОЗДІЛ 9

Відтворюваність досліджень, керування даними та відкрита наука

Розділи 5 і 8 відповідали на питання, як спроєктувати емпіричне дослідження в інформаційних технологіях і як коректно опрацювати дані. Цей розділ ставить наступне питання: *чи зможе хтось інший — або ви самі через два роки — одержати той самий результат.* Наскрізна теза: результат, який неможливо відтворити, науковою цінністю не є. Виклад операційний: після розділу здобувач має вміти скласти план керування даними, опублікувати відтворюваний артефакт із постійним ідентифікатором і оцінити його за принципами FAIR, не порушивши вимог щодо персональних даних. Етичні аспекти оприлюднення чутливої інформації розглянуто в розділі 10.

9.1 Криза відтворюваності та її прояви в інформатиці

9.1.1 Масштаб явища та причини

Термін **криза відтворюваності** (reproducibility crisis) закріпився після серії перевірок, що показали: значна частка опублікованих результатів не підтверджується при незалежному відтворенні. Опитування *Nature* 2016 р. серед понад 1500 дослідників дало показові цифри: 70 % респондентів не змогли відтворити експеримент колеги, понад 50 % — власний, а 52 % вважали ситуацію «значною кризою» [63]. Проєкт Open Science Collaboration відтворив 100 досліджень із психології: статистично значущий ефект одержано лише в 36 % реплікацій проти 97 % в оригінальних публікаціях [252]. Пояснення дав Дж. Іоаннідіс: за низької апріорної ймовірності гіпотези, малої потужності та гнучкості аналізу частка хибнопозитивних серед опублікованих «відкриттів» перевищує 50 % [182].

Причини поділяють на чотири групи. *Методологічні*: недостатня потужність, множинні порівняння без поправок, HARKing і p-hacking (розділ 8). *Комунікаційні*: обмеження обсягу не дають навести версії бібліотек, параметри апаратури, передобробку. *Інфраструктурні*: код і дані не збережено, посилання «протухають», середовище виконання зникає разом із версією операційної системи. *Інституційні*: publication bias, відсутність винагороди за реплікацію. В обчислювальних науках домінує третя група: принципової перешкоди для відтворення немає — є лише незбережений стан системи [264].

9.1.2 Стан справ в інформатиці та кібербезпеці

Найвідоміше вимірювання виконали К. Коллберг і Т. Пробстінг: вони проаналізували 601 статтю з провідних конференцій і журналів ACM у галузі комп'ютерних систем і спробували зібрати описані системи з вихідних текстів. Код удалося знайти й зібрати приблизно для третини робіт; ще для частини він існував, але не збирався без втручання авторів, а для більшості був недоступний узагалі; звідси пропозиція запровадити в кожній статті явну «специфікацію поширення» [93]. Пізніший аналіз у *Communications of the ACM* поширив діагноз на емпіричну інформатику загалом: частка реплікаційних досліджень мізерна, а звітність про негативні результати практично відсутня [90]. О. Гундерсен і С. К'єнсмо оцінили документованість 400 робіт AAAI та IJCAI за понад двадцятьма показниками: жодна стаття не описувала експеримент повністю [164].

Кібербезпекові дослідження мають додаткові перешкоди: частина датасетів містить персональні дані або відомості про вразливі системи; «жива» цільова система (вебсервіс, пристрій IoT, безпроводова точка доступу) змінюється між експериментом і рецензуванням. Це не звільняє автора від відтворюваності, а змінює її форму: замість «даних як є» публікують похідні ознаки, синтетичні генератори й опис процедури збирання (п. 9.10). Робоче правило: готуйте дослідження так, щоб відтворення було можливим *для вас самих* через 24 місяці на новому комп'ютері й без доступу до попередньої лабораторії.

9.2 Термінологія відтворюваності й система бейджів ACM

9.2.1 Три поняття і дві несумісні традиції

Терміни *repeatability*, *reproducibility* і *replicability* в різних спільнотах означають різне, тому в дисертації їх уживають із явним означенням. Спільне лише перше: **повторюваність** — та сама група, те саме обладнання й програмне забезпечення дають той самий результат.

За ACM (Artifact Review and Badging, версія 1.1, 2020) **відтворення** (reproducibility) — одержання основних результатів статті *іншою* групою з використанням наданих авторами артефактів, а **реплікація** (replicability) — тих самих результатів іншою групою *без* авторських артефактів [52]. У версії 1.0 значення цих термінів були протилежними, тому в статтях 2016–2019 рр. трапляється зворотне слововживання й посилання на редакцію обов'язкове. За NASEM (2019) **reproducibility** — суто обчислювальне поняття: ті самі дані й код дають той самий числовий результат; **replicability** — узгоджений висновок на *нових* даних [240]. Обидві системи узгоджуються, якщо розрізняти три осі: хто виконує, з якими даними і з чиїм кодом; четвертий рівень — *узагальнення* на іншу популяцію й інші умови.

Практичний наслідок: «результат відтворено» і «результат підтверджено» — не тотожні твердження. Перше означає, що з ваших даних і коду одержано ті самі числа, і не спростовує систематичної похибки збирання; друге — що незалежне дослідження на інших даних дійшло того самого висновку. Видавати перше за друге неприпустимо.

9.2.2 Бейджі ACM і артефактні комітети

Політика ACM передбачає три незалежні *групи* бейджів, які проставляють за результатами окремого рецензування артефактів [52]. Його веде **артефактний комітет** (Artifact Evaluation Committee) уже після ухвалення статті — такі комітети працюють на USENIX Security, ACM CCS, NDSS, PETS, ICSE, FSE.

Таблиця 9.1 — Групи бейджів ACM Artifact Review and Badging v1.1

Група	Бейдж	Умова присвоєння
Artifacts Evaluated	Functional	Артефакти задокументовані, узгоджені, повні, виконуються й відтворюють заявлені результати
	Reusable	Понад Functional: документація й структура дозволяють повторне використання й модифікацію
Artifacts Available	Available	Артефакти в публічному архівному репозитарії з DOI; *перевірки працездатності немає*
Results Validated	Reproduced	Результати одержано іншою групою частково з використанням артефактів авторів
	Replicated	Результати одержано іншою групою без авторських артефактів

Типова помилка здобувача — ототожнювати бейдж Available із відтворюваністю: він засвідчує лише факт архівного депонування, тож його можна одержати й на архів із непрацездатним кодом. Мета підготовки артефакту — не «викласти на GitHub», а пройти сценарій: незнайома людина за наданою інструкцією на чистій машині одержує конкретні таблицю й рисунок статті за кілька годин машинного часу.

9.3 Життєвий цикл дослідницьких даних і план керування ними

9.3.1 Шість стадій і точки ухвалення рішень

Дослідницькі дані — зафіксовані матеріали, потрібні для перевірки результатів: виміри, журнали, конфігурації, анкети, а в IT також код і параметри моделей. Керування ними описують **життєвим циклом даних** — послідовністю стадій, на кожній з яких ухвалюють рішення, що визначає можливості наступних (рис. 9.1).

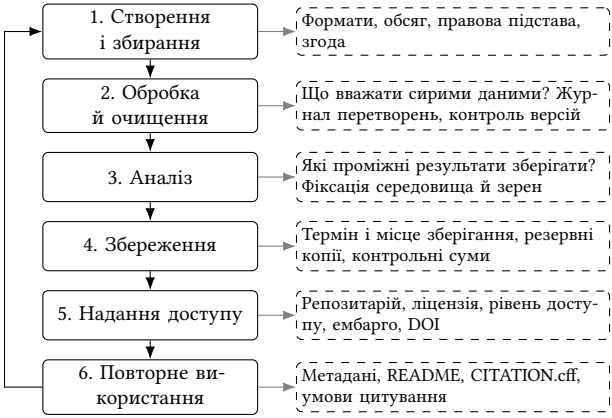


Рис. 9.1 — Життєвий цикл дослідницьких даних і точки ухвалення рішень

Ключова властивість циклу — *незворотність*: помилка на ранній стадії не виправляється на пізній. Якщо під час збирання не зафіксовано версію мікропрограми точки доступу, обробка цього не відновить; якщо згоди учасників опитування не передбачають оприлюднення, дані не можна відкрити згодом. Тому рішення про публікацію ухвалюють *до* експерименту.

9.3.2 План керування даними як документ

План керування даними (Data Management Plan, DMP) — документ, що фіксує відповіді на питання всіх шести стадій. Він обов'язковий за ст. 17 Грантової угоди Horizon Europe: перший варіант подають протягом шести місяців від початку проєкту, оновлення — у середині та наприкінці [280, 139]. НФДУ та програми МОН також вимагають опису роботи з даними в заявці [19], а національний план щодо відкритої науки передбачає поступове поширення вимоги DMP на державне фінансування [26]. Найпоширеніший шаблон — Science Europe Practical Guide, що структурує DMP навколо шести питань і додає рубрику оцінювання якості [295]; інструменти — DMPonline [116] та Argos [254].

Типові помилки: «дані зберігатимуться на комп'ютері виконавця»; «дані буде надано за запитом» (не є відкритим доступом); «персональних даних немає», хоча збираються MAC-адреси; відсутність ліцензії; план без жодної конкретної назви репозитарію чи стандарту метаданих.

Таблиця 9.2 — Контрольний список DMP (за структурою Science Europe)

Блок	Що має бути зафіксовано
1. Опис даних і збирання	Типи (виміри, журнали, код, анкети), формати, обсяг у ГБ, джерела повторного використання
2. Документація і якість	Стандарт метаданих (DataCite, Dublin Core, DCAT), словники, структура README, правила іменування файлів
3. Зберігання й резервування	Сховище в активній фазі, резервні копії (3–2–1), відповідальний, контрольні суми, вартість
4. Правові й етичні питання	Персональні дані, правова підстава за GDPR, згода, анонімізація, права інтелектуальної власності
5. Обмін і збереження	Репозитарій, момент публікації та ембарго, ліцензія, постійний ідентифікатор, термін зберігання
6. Ресурси й відповідальність	Хто відповідає за кожну стадію, кошти (APC, сховище, курація), фінансування після проєкту

9.4 Принципи FAIR і їх практичне втілення

9.4.1 Що саме стверджують принципи

Принципи FAIR сформульовано у 2016 р. консорціумом під керівництвом М. Вілкінсона як відповідь на питання: що має виконуватися, щоб дані були придатними для повторного використання не лише людиною, а й *машиною* [346]. Абревіатура позначає **Findable** (знаходжуваність), **Accessible** (доступність), **Interoperable** (сумісність) і **Reusable** (придатність до повторного використання); разом вони розгортаються у 15 підпринципів.

Два зауваження знімають більшість непорозумінь. *FAIR не дорівнює open*: принцип A вимагає, щоб *умови* доступу були чітко описані й технічно реалізовані, а не щоб доступ був вільним — датасет із персональними даними, метадані якого відкриті, а дані надаються за контрольованою процедурою, цілком відповідає FAIR. I FAIR — властивість *інфраструктури й метаданих*, а не якості дослідження. Лаконічні формулювання тлумачать за GO FAIR [194] і схемою DataCite [107].

Таблиця 9.3 — П'ятнадцять підпринципів FAIR і їх втілення для IT-датасету

Код	Зміст підпринципу	Практичне втілення
F1	Глобально унікальний і постійний ідентифікатор	DOI від Zenodo на кожну версію датасету; ORCID авторів; ROR установи
F2	Дані описано багатими метаданими	Заповнені поля DataCite: автори, рік, тип ресурсу, опис, ключові слова, місце й період вимірювань
F3	Метадані явно містять ідентифікатор даних	У полі `identifier` той самий DOI, що й у README та CITATION.cff
F4	Реєстрація в пошуковому ресурсі	Zenodo передає метадані до OpenAIRE, DataCite Commons, Google Dataset Search

A1	Видобуття за ідентифікатором стандартним протоколом	Роздільник `doi.org`; вивантаження одним запитом
A1.1	Протокол відкритий і універсально реалізовний	HTTPS і OAI-PMH замість власницьких клієнтів чи форм із CAPTCHA
A1.2	За потреби — автентифікація й авторизація	Контрольований доступ: заявка, рішення кураторської ради, журнал
A2	Метадані доступні, навіть коли дані вилучено	Запис-надгробок (tombstone) із причиною вилучення; DOI не перевикористовують
I1	Формальна, поширена мова подання знань	CSV/Parquet замість власницьких дампів; JSON-LD для метаданих; PCAPng для трафіку
I2	Словники, які самі відповідають FAIR	Атаки — за MITRE ATT&CK, вразливості — за CVE/CWE, одиниці — за UCUM
I3	Кваліфіковані посилання на інші (мета)дані	`relatedIdentifier`: `IsSupplementTo` для статті, `IsSupplementedBy` для коду
R1	Багатий і точний опис	Словник змінних з одиницями, діапазонами й кодами пропусків; опис апаратури
R1.1	Чітка й доступна ліцензія	LICENSE і поле `rights`: CC BY 4.0 чи CC0 для даних, MIT чи Apache 2.0 для коду
R1.2	Детальне походження (provenance)	Протокол вимірювань, версії мікропрограм, журнал перетворень сирих даних
R1.3	Відповідність стандартам спільноти	Звичний формат (PCAPng, IPFIX, CSV зі схемою Frictionless Data), структура каталогів за практикою артефактних комітетів

9.4.2 Кількісна оцінка відповідності

Нехай $w_i > 0$ — вага i-го підпринципу (за замовчуванням $w_i = 1$), а $s_i \in \{0; 0{,}5; 1\}$ — ступінь його виконання. Тоді **показник FAIR-відповідності**

$$\mathrm{FAIR} = \frac{\sum_{i=1}^{15} w_i s_i}{\sum_{i=1}^{15} w_i} \cdot 100\,\%. \tag{9.1}$$

Окремо корисний **показник повноти метаданих**: якщо схема (наприклад, DataCite 4.5) містить m рекомендованих полів, а заповнено m_{f} із них непорожнім значенням, то $C_{\text{м}} = m_{\mathrm{f}}/m$. Орієнтир: депозит «за замовчуванням» (назва, автор, файл) дає $C_{\text{м}} \approx 0{,}3$ і FAIR близько 40 %; свідома підготовка за таблицею 9.3 підіймає показники понад 0,8 і 85 %. Автоматизовані оцінювачі (F-UJI, FAIR Evaluator) обчислюють подібні індекси за метаданими, а не за вмістом файлів.

9.5 Репозитарії, ідентифікатори та цитування даних і програм

9.5.1 Вибір репозитарію

Репозитарій даних відрізняється від файлового сховища трьома властивостями: присвоює постійний ідентифікатор, зобов'язується зберігати запис

визначений час і публікує структуровані метадані для збирання агрегаторами. GitHub, Google Drive чи сайт кафедри цим властивостям не відповідають і *не* є репозитаріями в розумінні вимог Horizon Europe або бейджа Artifacts Available. Вибір ведуть за ієрархією: галузевий → інституційний → універсальний; для ІТ галузевих репозитаріїв мало, тому типовим вибором є Zenodo [353, 149, 123].

Таблиця 9.4 — Порівняння репозитаріїв для ІТ-досліджень

Репозитарій	Оператор	DOI / версії	Особливості й типове застосування
Zenodo	CERN, ЄС	Так / так	Безоплатний, ліміт 50 ГБ на запис, інтеграція з GitHub і OpenAIRE; типовий вибір для даних і релізів коду
Figshare	Digital Science	Так / так	Великі обсяги, інституційні інстанції; часто як репозитарій матеріалів видавця
Dryad	Некомерційна	Так / так	Кураторська перевірка, прив'язка до статті, вимога CC0; платне депонування
OSF	Center for Open Science	Так / частково	Проєкт цілком: передреєстрація, матеріали, дані, код
Software Heritage	Inria, ЮНЕСКО	SWHID / так	Архівує *вихідний код* із повною історією Git
Інституційні репозитарії	ЗВО України	Часто ні	Обов'язкові за політиками ЗВО; для текстів, обмежено — для даних
НРАТ	УкрІНТЕІ, МОН	Ні	Державний репозитарій дисертацій і звітів НДР

Національний репозитарій академічних текстів акумулює повні тексти дисертацій, авторефератів і звітів про НДР [17] і розв'язує інше завдання, ніж Zenodo. Software Heritage зберігає *історію* розробки й видає ідентифікатор SWHID, криптографічно прив'язаний до вмісту [112].

9.5.2 Версії та зв'язка GitHub ↔ Zenodo

Zenodo присвоює кожному депозиту два DOI: *версійний* (для конкретної версії) і *концептуальний* (вказує на найновішу). У статті цитують версійний — саме та версія давала опубліковані числа. Послідовність для коду: (1) увімкнути в Zenodo перемикач для потрібного репозитарію GitHub; (2) довести репозитарій до стану релізу (README, LICENSE, CITATION.cff, зафіксовані залежності); (3) створити тег і реліз — Zenodo сформує архів, запис і DOI; (4) додати DOI-значок у README та поле `doi` у CITATION.cff.

9.5.3 Цитування даних і програмного забезпечення

Датасет і програма — повноцінні об'єкти цитування, а не рядок у подяках. Software Citation Principles формулюють шість вимог: важливість, кредит і атрибуція, постійність, доступність, конкретність (версія) та унікальна ідентифікація [307]. Щоб це працювало механічно, у корені репозитарію роз-

міщують `CITATION.cff` — машиночитний YAML із полями `authors`, `title`, `version`, `doi`, `date-released`, `license` [122]; GitHub показує за ним кнопку «Cite this repository», а Zenodo автозаповнює метадані. У списку джерел датасет за ДСТУ 8302:2015 подають за схемою електронного ресурсу [7]: автори, назва, тип «*набір даних*», версія, репозитарій, рік, DOI, який вписують після реального депонування, а не «резервують» наперед.

9.6 Відтворюване обчислювальне середовище

9.6.1 Чотири рівні фіксації

Обчислювальний результат відтворюється, якщо зафіксовано чотири шари: *дані* (незмінні сирі дані з контрольними сумами), *код* (версія в системі контролю версій), *залежності* (точні версії бібліотек та інтерпретатора) і *середовище* (операційна система, драйвери, апаратура). Десять простих правил відтворюваного обчислювального дослідження [292] зводяться до принципу: *кожен результат має бути похідним від скрипту, а не від дії людини.*

Контроль версій: Git фіксує історію коду, але не даних — великі бінарні файли тримають у Git LFS, DVC або зовнішньому репозитарії з DOI; обов'язковий тег релізу, що відповідає поданій статті. **Фіксація залежностей**: розрізняють *декларацію* (`requirements.txt`, `environment.yml` із діапазонами версій) і *замок* (`poetry.lock`, `conda-lock.yml`, `renv.lock`) із точними версіями й хешами транзитивних залежностей; потрібні обидва, а запис `numpy` без версії робить артефакт непрацездатним уже через рік. **Контейнери**: Docker пакує застосунок разом із бібліотеками операційної системи, знімаючи проблему «у мене працює» [69]; базовий образ фіксують за дайджестом, а не міткою `latest`, а `Dockerfile` зберігають у репозитарії. Binder запускає репозитарій у хмарі одним посиланням [274], але не замінює архівного депонування.

9.6.2 Записники, робочі процеси й недетермінізм

Масштабне дослідження 1,4 млн Jupyter-записників із GitHub показало: лише близько чверті вдалося виконати без помилок і лише близько 4 % відтворили початковий результат; основна причина — *нелінійний порядок виконання комірок* [268]. Звідси правила: перед публікацією виконувати *Restart & Run All* і зберігати лише лінійно відтворений записник; не тримати стан у пам'яті між сесіями; винести повторюваний код у модулі. **Системи робочих процесів** усувають цю проблему структурно: Snakemake описує аналіз як правила «вхідні файли → вихідні файли → команда», сам визначає порядок, паралелізує, перезапускає лише застарілі кроки й ізолює кожне правило в контейнері [237]; Nextflow розв'язує ту саму задачу для хмарних середовищ [113]. Для дисертаційного обсягу достатньо `Makefile` або скрипту `run_all.sh`, що відтворює рисунки й таблиці з сирих даних.

Повна повторюваність числа до останнього знака часто недосяжна. Джерела недетермінізму: генератори псевдовипадкових чисел (зерно — *seed* — фіксують окремо для кожної бібліотеки); паралельне зведення чисел із рухомою комою; недетерміновані ядра GPU (cuDNN добирає алгоритм згортки за часом виконання); хеш-сіль інтерпретатора; порядок читання файлів у каталозі. Стратегія: фіксувати все, що фіксується, а решту *вимірювати* — подавати результат як середнє й довірчий інтервал за $n \geq 5$ прогонів із різними зернами (розділ 8), а не як одне «найкраще» число.

9.6.3 Структура репозитарію артефакту

Артефактні комітети очікують передбачуваної структури (рис. 9.2). README має відповідати на п'ять питань: що це; вимоги до апаратури й часу; як установити; як відтворити конкретні рисунки й таблиці статті; як цитувати.

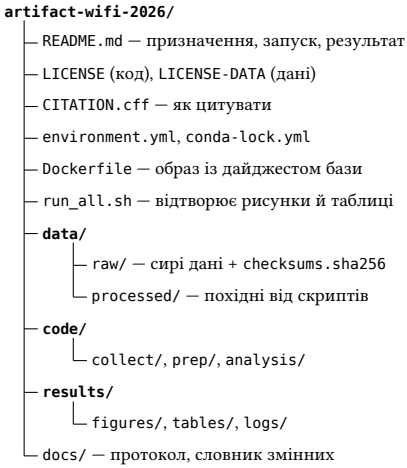


Рис. 9.2 — Структура репозитарію відтворюваного артефакту

9.7 Передреєстрація та зареєстровані звіти

Передреєстрація — депонування плану дослідження в незалежному реєстрі з міткою часу *до* збирання чи аналізу даних. Фіксують: гіпотези, змінні, обсяг вибірки та правило її зупинки, критерії виключення спостережень, план статистичного аналізу з конкретними критеріями й поправками, а також розмежування підтверджувальних і розвідувальних аналізів [247]. Від протоколу дослідження вона відрізняється тим, що є *зовнішньою* (її не мо-

жна змінити заднім числом) і *публічною*. Розбіжності з планом не забороняються — їх описують абзацом «Deviations from preregistration».

Зареєстрований звіт (Registered Report) — сильніша форма: до журналу подають Stage 1 (вступ, методи й план аналізу без результатів). За позитивного рішення надають *принципове схвалення*: статтю буде опубліковано незалежно від того, підтвердиться гіпотеза чи ні [82]. Це усуває головний механізм спотворення літератури — publication bias, коли негативні результати осідають у шухляді (*file-drawer problem*), через що метааналізи системно переоцінюють ефекти. Платформи: OSF Registries (безоплатно, ембарго до 4 років) [81] та AsPredicted; формат Registered Reports приймають понад 300 журналів.

Передреєстрація природна для підтверджувальних експериментів (порівняння алгоритмів, експеримент із користувачами, польове вимірювання з наперед визначеною гіпотезою) і не має сенсу для розвідувальних або конструктивних робіт — побудови артефакту за DSR (розділ 5). У кібербезпеці оприлюднення плану може попередити оператора досліджуваної системи, тож компроміс — передреєстрація з ембарго до публікації.

9.8 Структура наукової статті та типи публікацій

Передреєстрація фіксує *план*, артефакт — *матеріал*, а стаття лишається єдиним документом, який читає більшість. Тому структура статті — не редакційна умовність, а остання ланка відтворюваності: саме вона визначає, чи зможе читач відокремити те, що виміряно, від того, що автор із цього виснував.

9.8.1 IMRaD як контракт відтворюваності

IMRaD (Introduction, Methods, Results and Discussion) — канонічна схема емпіричної статті. У природничих і медичних журналах вона витіснила вільний виклад протягом 1940–1980-х рр. і до кінця століття стала практично єдиною [309]. Кожен розділ відповідає на одне питання:

Introduction — *навіщо*. Проблема, прогалина в знанні, дослідницьке питання або гіпотеза. Закінчується явним формулюванням внеску.

Methods — *як*. Що саме зроблено: дані, процедура, обладнання, параметри, критерії, план аналізу. Це *єдиний* розділ, від повноти якого залежить відтворюваність; писати його треба так, щоб незалежна група повторила роботу без листування з авторами.

Results — *що одержано*. Виміряне, без тлумачення: числа, таблиці, рисунки, інтервали, розміри ефектів.

and Discussion — *що це означає*. Тлумачення, зіставлення з літературою, обмеження й загрози валідності, висновок у межах доведеного.

Розмежування Results і Discussion — головна цінність схеми: вимірювання зберігає чинність, навіть якщо тлумачення згодом спростують. Злиття цих розділів (*Results and Discussion* однією частиною) допустиме в коротких повідомленнях, але позбавляє читача змоги перевірити, де закінчуються дані й починається думка автора.

Геометрично IMRaD описують як **пісковий годинник**: широкий вступ (галузь → проблема), вузькі методи й результати (конкретний експеримент), широке обговорення (результат → галузь). Типові пропорції для журнальної статті — 15 / 25 / 25 / 35 % обсягу; розділ методів, коротший за десяту частину статті, майже завжди означає непридатність роботи до відтворення.

Сучасний IMRaD обріс обов'язковими службовими блоками, і саме вони цікавлять цей розділ найбільше: *Data availability* і *Code availability* з постійними ідентифікаторами (п. 9.5), *Funding*, *Conflicts of interest*, опис внеску за CRediT і заява про використання генеративних моделей (розділ 10). Стаття без цих блоків формально відповідає IMRaD, але не дає змоги ані знайти артефакти, ані оцінити упередженість.

9.8.2 Інші структурні схеми

IMRaD не універсальний: він передбачає *емпіричне* дослідження з вимірюванням. Для інших типів робіт застосовують інші схеми (табл. 9.5).

Перестановка методів. Частина журналів родини *Nature* виносить методи в кінець статті, залишаючи в основному тексті послідовність «вступ — результати — обговорення». Логіка відтворюваності при цьому не змінюється: методи стають довшими, а не коротшими.

CARS (Create a Research Space) Дж. Свейлза описує не статтю загалом, а *вступ* — через три кроки: закріпити територію (чому галузь важлива), показати нішу (чого бракує), зайняти нішу (що робить ця праця) [317]. Це найкорисніший інструмент для здобувача, бо саме вступ переписують найчастіше.

CCC (Context, Content, Conclusion) Б. Менша і К. Кордінга — рекурсивна схема: та сама трійка організує і статтю, і кожен розділ, і кожен абзац; абзац має одну тему, перше речення якої повідомляє його зміст [229].

«Проблема — розв'язок — оцінювання» — фактичний стандарт конференційних статей в інформатиці: вступ, огляд суміжних робіт, постановка й модель загроз, запропонований метод, реалізація, оцінювання, обговорення обмежень, висновки. Для проєктувального дослідження (DSR) цю схему уточнює публікаційний шаблон Ґреґор і Гевнера [161].

Український контекст. Багато вітчизняних фахових видань досі відтворюють у вимогах до авторів перелік із постанови президії ВАК України від 15.01.2003 № 7-05/1 [39]: постановка проблеми та її зв'язок із важливими

завданнями; аналіз останніх досліджень і публікацій із виділенням невирішених раніше частин; формулювання цілей статті; виклад основного матеріалу з обґрунтуванням результатів; висновки й перспективи подальших розвідок. Саму ВАК ліквідовано 2011 р., а її функції перебрали МОН і згодом НАЗЯВО, тож юридична вага документа сумнівна — але вимога живе в редакційній практиці. Змістовно цей перелік є перейменованим IMRaD без окремого розділу методів, і це його головна вада: «виклад основного матеріалу» дозволяє не відокремити процедуру від результатів. Робоча порада: дотримуватися рубрикації, якої вимагає журнал, але всередині «викладу основного матеріалу» зробити явні підрозділи «Методика» і «Результати».

Таблиця 9.5 — Структурні схеми наукового тексту

Схема	Що задає	Коли доречна
IMRaD	Вступ — методи — результати — обговорення	Емпіричне дослідження з вимірюванням
IMRaD із методами в кінці	Те саме, методи винесено після висновків	Журнали з жорстким лімітом основного тексту
CARS [317]	Три кроки вступу: територія, ніша, заняття ніші	Написання й переписування вступу будь-якої статті
CCC [229]	Контекст — зміст — висновок, рекурсивно до абзацу	Планування тексту й правка «важких» розділів
Проблема — розв'язок — оцінювання	Модель загроз, метод, реалізація, оцінювання, обмеження	Конференційні статті в IT та кібербезпеці; проєктувальне дослідження
PRISMA [257]	27 пунктів звітування плюс потокова діаграма	Систематичний огляд і метааналіз (розділ 6)
Registered Report	Stage 1 (план) → Stage 2 (результати)	Підтверджувальні дослідження з ризиком publication bias
Перелік за постановою № 7-05/1 [39]	П'ять обов'язкових елементів без окремих методів	Українські фахові видання, що досі його вимагають

9.8.3 Типи публікацій і що в кожному є результатом

Тип публікації визначає, що саме рецензент вважатиме результатом і яких артефактів вимагатиме (табл. 9.6). Плутанина типів — поширена причина відхилення: позиційну статтю подають як дослідницьку, а опис інструмента — як експериментальну роботу.

Окремо варто розрізняти *тип* і *стадію*: препринт — не тип публікації, а стадія оприлюднення, на якій може перебувати будь-який з наведених типів (розділ 6). Так само редакційна стаття, передмова та анотація випуску науковим результатом не є й до переліку публікацій здобувача не зараховуються (розділ 2).

Таблиця 9.6 — Типи наукових публікацій

Тип	Що є результатом	Що вимагає відтворюваність
Оригінальна дослідницька (research article)	Нове емпіричне або теоретичне твердження	Методи, дані, код, розділення вибірок
Оглядова (review)	Узагальнення стану питання; види й вимоги — у розділі 6	Протокол, рядок пошуку, критерії, потокова діаграма
Систематизація знань (SoK)	Впорядкування розрізнених результатів галузі в узгоджену рамку	Критерії включення, таксономія, відтворюваний перелік праць
Коротке повідомлення (letter, short paper)	Один чіткий результат або спостереження	Ті самі вимоги в стислій формі; методи — у додатку
Стаття про набір даних (data paper)	*Сам набір* та його опис, а не висновки з нього [83]	Депонований набір із DOI, схема, ліцензія, опис збирання, оцінка FAIR
Стаття про програмний засіб (software paper)	Працездатний інструмент і його архітектура [308]	Репозитарій, тести, версія з DOI, інструкція встановлення
Реплікаційне дослідження	Підтвердження або спростування чужого результату	Опис розбіжностей з оригіналом; обидва набори артефактів
Зареєстрований звіт	План (Stage 1) і виконання за ним (Stage 2)	Передреєстрація, опис відхилень від плану
Позиційна (position, vision)	Аргументована теза або програма досліджень	Артефактів не вимагає; не можна подавати як емпіричну
Звіт про досвід (experience report)	Задокументований досвід упровадження	Контекст, обмеження узагальнення, опис середовища
Коментар і відповідь	Уточнення або заперечення до опублікованої праці	Точне посилання на оскаржуване твердження

Типова помилка. Здобувач публікує опис розробленої системи в журналі як «оригінальну дослідницьку статтю», наводить архітектуру й скриншоти, але не має ані базової лінії, ані вимірювання. Рецензент кваліфікує це як *solution proposal without validation* (розділ 5).
Коректно. Обрати тип відповідно до наявного результату: якщо доведено працездатність інструмента — software paper із репозитарієм, тестами й DOI версії; якщо є вимірювання з базовою лінією — дослідницька стаття з розділом методів; якщо готовий лише задум — позиційна стаття або доповідь на воркшопі. Тип оголошують у супровідному листі, і саме за ним редакція добирає критерії.

9.9 Відкрита наука: політики, моделі доступу та ліцензії

9.9.1 Рамкові документи

Рекомендація ЮНЕСКО з відкритої науки, ухвалена 23.11.2021 р. на 41-й сесії Генеральної конференції за підтримки 193 держав-членів, є першим гло-

бальним нормативним документом у цій сфері [327]. Вона визначає відкриту науку через чотири стовпи: *відкриті наукові знання* (публікації, дані, код, освітні ресурси); *відкрита інфраструктура*; *відкрита взаємодія суспільних акторів* (громадянська наука, співпраця з бізнесом і громадами); *відкритий діалог з іншими системами знань*.

Plan S — ініціатива фондів cOAlition S: публікації за профінансованими дослідженнями мають бути в негайному відкритому доступі без ембарго й під відкритою ліцензією [89]. Horizon Europe закріплює це юридично, додаючи «відкриті дані за замовчуванням» [280]. **EOSC** (European Open Science Cloud) — федерована інфраструктура, що об'єднує репозитарії й сервіси ЄС у єдину точку доступу [144]. Український рамковий документ — національний план щодо відкритої науки [26].

9.9.2 Моделі відкритого доступу та ліцензії

Gold — журнал повністю відкритого доступу, витрати часто покриває APC (article processing charge). **Green** — самоархівування прийнятої версії в репозитарії, можливо з ембарго. **Diamond** — відкритий доступ без плати і для автора, і для читача; за цією моделлю працює більшість українських фахових видань, зокрема «Кібербезпека: освіта, наука, техніка» [10]. **Hybrid** — передплатний журнал, у якому окрему статтю відкривають за плату; Plan S визнає цю модель лише в межах трансформаційних угод з обмеженим строком.

Критика APC: типова плата в 2–4 тис. євро (для престижних видань — понад 9 тис.) перетворює бар'єр для читача на бар'єр для автора, дискримінує дослідників із країн з обмеженим фінансуванням і живить хижацькі видання (розділ 7). Альтернативи — diamond-видання, реєстр DOAJ [118], green-шлях через інституційний репозитарій, трансформаційні угоди.

Практичне правило: *код* — MIT або Apache 2.0; *дані* — CC BY 4.0 (або CC0, якщо репозитарій вимагає); *метадані* — CC0; *текст* — CC BY 4.0. Ліцензію зазначають і в метаданих, і у файлі LICENSE: без цього твір захищений авторським правом у повному обсязі й повторне використання незаконне [21, 100, 251].

9.10 Персональні й чутливі дані у відкритій науці

9.10.1 Анонімізація, псевдонімізація і k-анонімність

Формула *as open as possible, as closed as necessary* означає: закритість потребує обґрунтування, відкритість — ні. GDPR не забороняє наукових досліджень: ст. 89 дозволяє відступи від низки прав суб'єкта даних за умови належних запобіжників — мінімізації даних і псевдонімізації [279]; в Україні діє Закон «Про захист персональних даних» [33].

Таблиця 9.7 — Відкриті ліцензії: що дозволяють і чого вимагають

Ліцензія	Дозволяє	Вимагає / обмежує
CC BY 4.0	Будь-яке використання, зміни, комерційне застосування	Зазначення авторства й посилання на ліцензію; стандарт для наукових публікацій і даних
CC BY-SA 4.0	Те саме	Похідні твори — лише під тією самою ліцензією (копілефт); ускладнює комбінування наборів
CC0 1.0	Усе, без умов	Відмова від прав у максимальному обсязі; рекомендована для *метаданих*, вимагається Dryad
CC BY-NC	Некомерційне використання	Межа «комерційного» юридично нечітка, тому ліцензія *не* є відкритою за Plan S і блокує повторне використання
MIT	Використання, зміну, поширення, зокрема у власницькому ПЗ	Збереження тексту ліцензії й повідомлення про авторські права
Apache 2.0	Те саме	Додатково: явне надання патентних прав, фіксація змін; сумісна з GPLv3
GPLv3	Використання й зміну	Поширення похідних робіт лише під GPL (сильний копілефт)
ODbL 1.0	Використання й зміну бази даних	Атрибуція, копілефт для похідних баз; діє там, де є право *sui generis* на базу даних

Псевдонімізація замінює ідентифікатори на коди, зберігаючи ключ відповідності: дані *залишаються* персональними. **Анонімізація** робить повторну ідентифікацію неможливою для будь-якої сторони; лише тоді дані виходять з-під дії регламенту. Видалення прямих ідентифікаторів анонімізацією не є: ризик оцінюють за трьома критеріями — чи можна виокремити особу (singling out), пов'язати записи (linkability), вивести атрибут (inference) [62].

Базова формалізація — k-**анонімність** [318]. Нехай T — опублікована таблиця, а QI — набір квазіідентифікаторів (атрибутів, які окремо не ідентифікують, а в комбінації — так: модель пристрою, виробник, район, часове вікно). Таблиця T задовольняє k-анонімність, якщо

$$\forall t \in T: \quad \left|\{t' \in T \mid t'[QI] = t[QI]\}\right| \;\geq\; k, \tag{9.2}$$

тобто кожен запис невідрізнений щонайменше від $k-1$ інших за квазіідентифікаторами. Досягають цього узагальненням (замість точної координати — район, замість мітки часу — година) та приховуванням поодиноких записів. k-анонімність не захищає від *атаки однорідності*, коли всі k записів мають однакове чутливе значення; звідси посилення — ℓ-різноманіття [217].

9.10.2 Оцінка ризику реідентифікації

Якщо клас еквівалентності j містить f_j записів, ймовірність правильної реідентифікації навмання взятого запису з нього дорівнює $1/f_j$, а середній і максимальний ризики по набору з N записів і J класів становлять

$$R_{\text{сер}} = \frac{1}{N}\sum_{j=1}^{J} f_j \cdot \frac{1}{f_j} = \frac{J}{N}, \qquad R_{\max} = \max_j \frac{1}{f_j} = \frac{1}{k}. \tag{9.3}$$

Звідси орієнтир: за $k = 5$ максимальний ризик становить 0,2, за $k = 11$ — менше 0,1; типовими порогами прийнятності вважають $R_{\max} \le 0{,}1$ для публічного оприлюднення і $R_{\max} \le 0{,}33$ для контрольованого доступу [131]. Ці пороги — правило рішення: набір, що їх не задовольняє, не оприлюднюють відкрито, а або додатково знеособлюють, або переводять у контрольований доступ.

Припущення моделі. Формула (9.3) чинна лише за досить жорстких умов: перелік квазіідентифікаторів вичерпний; зловмисник не має інших відомостей про конкретну особу; записи незалежні, а кожна особа представлена одним записом; набір не поєднується з іншими наборами тих самих осіб. У журналах вимірювань жодна з цих умов не виконується автоматично: одна особа дає тисячі записів, а зовнішні джерела (геолокаційні бази BSSID, розклади занять) додають ознак, яких немає в самому наборі. Тому зростання k зменшує ризик *у межах моделі*, але саме собою не є ані гарантією анонімності, ані доказом того, що дані перестали бути персональними в розумінні GDPR: правову оцінку роблять окремо, за критерієм розумно доступних засобів ідентифікації. Межі «деідентифікації» ілюструє робота А. Нараяна й В. Шматікова: за кількома оцінками фільмів із зовнішнього джерела вдалося ідентифікувати користувачів у формально анонімізованому наборі Netflix [239]. Розріджені дані високої розмірності — журнали трафіку, телеметрія пристроїв, сліди безпроводових з'єднань — реідентифікуються надзвичайно легко.

9.10.3 Режими доступу та дані про вразливості

Замість дихотомії «відкрито/закрито» використовують градацію: (1) повністю відкриті дані; (2) відкриті похідні (агрегати, синтетичні набори, моделі без сирих даних); (3) **контрольований доступ** — метадані відкриті, дані надаються за заявкою після рішення кураторської ради й угоди про використання; (4) закриті дані з відкритим лише описом і кодом. *Метадані публікують завжди.*

Окремий клас — дані про вразливості, конфігурації діючих систем, зразки шкідливого коду: оприлюднення точних координат уразливих точок доступу є інструкцією для атаки. Стандартна практика — відповідальне роз-

криття: повідомлення постачальника, погоджений строк, публікація лише після усунення, а в датасеті — узагальнені або хешовані ідентифікатори. Етичні межі таких публікацій і подвійне використання розглянуто в розділі 10.

9.11 Кейс: публікація датасету вимірювань безпроводових мереж

Вихідна ситуація. Аспірант протягом семестру вимірював завантаженість каналів діапазонів 2,4 і 5 ГГц у трьох навчальних корпусах: сканування ефіру з періодом 60 с, фіксація SSID, BSSID, номера каналу, ширини смуги, RSSI, типу захисту, мітки часу й позиції приймача — 4,1 млн записів (близько 12 ГБ).

Крок 1. Правова й етична перевірка. BSSID може бути прив'язаний до домогосподарства через геолокаційні бази, а SSID часто містить прізвище, тож дані вважають персональними. Рішення: SSID вилучити; BSSID замінити на HMAC із секретним ключем, що знищується після публікації; координати узагальнити до рівня корпусу; мітки часу — до 15-хвилинних інтервалів. Далі перевірено (9.2) за квазіідентифікаторами {корпус, канал, виробник, часовий інтервал}: після вилучення 340 поодиноких записів мінімальний клас еквівалентності становив 7 записів, тобто $k = 7$ і $R_{\max} = 1/7 \approx 0{,}14$. Це *більше* за власний поріг $R_{\max} \leq 0{,}1$, прийнятий для відкритого оприлюднення (підрозділ 9.10), тому відкрито публікувати набір у такому вигляді не можна. Розглянуто дві допустимі відповіді: (а) додаткове узагальнення — часові інтервали з 15 хв до 1 год, виробник — до класу обладнання; (б) контрольований доступ через репозитарій із запитом і угодою про використання, для якого поріг становить $R_{\max} \leq 0{,}33$. Обрано варіант (а), бо втрата роздільності за часом не заважає заявленому призначенню набору — аналізу добової завантаженості каналів. Після узагальнення мінімальний клас еквівалентності — 13 записів, $k = 13$, $R_{\max} \approx 0{,}077 \leq 0{,}1$; додатково вилучено ще 512 записів, що не потрапили в жоден достатньо великий клас. Це рішення, разом із припущеннями моделі ризику, зафіксовано в README: k-анонімність оцінює ризик у межах свого набору квазіідентифікаторів і не скасовує потреби в правовій оцінці.

Крок 2. Структура й документація. Каталог побудовано за рис. 9.2. README описує обладнання (адаптер, версію мікропрограми, антену), методику сканування, формат кожного файлу, відомі обмеження (пропуски під час повітряних тривог) і порядок відтворення рисунків; словник змінних подає для кожного поля тип, одиницю (RSSI — дБм), діапазон і код пропуску. Додано `LICENSE-DATA` (CC BY 4.0) і `LICENSE` (MIT).

Крок 3. Файл CITATION.cff. Мінімальний коректний вміст:

```
cff-version: 1.2.0
title: Wi-Fi channel occupancy dataset
```

```
type: dataset
authors:
- family-names: Sokolov
given-names: Volodymyr
orcid: https://orcid.org/XXXX-XXXX-XXXX-XXXX
version: 1.0.0
date-released: 2026-03-14
license: CC-BY-4.0
doi: 10.5281/zenodo.XXXXXXX
```

Крок 4. Депонування. Створено запис у Zenodo: тип ресурсу — *Dataset*; заповнено опис, ключові слова (Wi-Fi, spectrum occupancy, wireless measurements), період і місце збирання, приналежність авторів (ROR), джерело фінансування; у полі `relatedIdentifier` — `IsSupplementTo` для статті та `IsSupplementedBy` для коду. Одержаний версійний DOI внесено до CITATION.cff і до тексту статті.

Крок 5. Перевірка за FAIR. За таблицею 9.3: F1–F4, A1, A1.1, A2, I1, I3, R1, R1.1, R1.2 — виконано; I2 і R1.3 — частково (локальний словник виробників замість зовнішньої онтології; немає усталеного галузевого формату); A1.2 не застосовується, бо дані відкриті. Отже, з 14 застосовних підпринципів 12 виконано повністю й два наполовину, і за (9.1) при одиничних вагах $\text{FAIR} = (12 + 2 \cdot 0{,}5)/14 \approx 93\ \%$.

Типові помилки, яких уникнуто: публікація «сирих» PCAP із корисним навантаженням; архів `final_v3_new.zip` без структури; ліцензія «для наукових цілей»; посилання на GitHub замість DOI; брак контрольних сум.

9.12 Відтворюваність і збереження даних в умовах війни

Війна перетворила абстрактну вимогу резервування на питання збереження наукового доробку. Ризики: знищення обладнання й серверних приміщень, відключення живлення та зв'язку, переміщення закладу, кібератаки на інформаційні системи ЗВО, втрата доступу до архівів на окупованих територіях. Звідси чотири висновки. *Правило 3–2–1*: три копії даних на двох різних носіях, одна — поза межами регіону, що практично означає закордонний репозитарій або хмару. Депонування в міжнародному репозитарії з DOI (Zenodo, OSF) є не лише виконанням вимог відкритої науки, а й формою страхування. Вихідний код доцільно дзеркалити в Software Heritage [112], який архівує історію розробки незалежно від долі хостингу. Досвід переміщених ЗВО показав, що міграція репозитарію можлива лише тоді, коли збережено не тільки файли, а й *метадані та ідентифікатори*. Ініціатива SUCHO, у межах якої волонтери 2022 р. заархівували понад 5 тис. українських вебсайтів [294], дала урок: розподілене зберігання копій у різних юрисдикціях ефективніше за будь-яке одиничне «надійне» сховище.

9.13 Висновки до розділу

Відтворюваність — не додаткова чеснота, а умова, за якої результат взагалі належить науці. Емпіричні перевірки останнього десятиліття показали системність проблеми: у психології відтворилася близько третини результатів, у комп'ютерних системах код виявився доступним і працездатним приблизно для третини робіт, а в дослідженнях зі штучного інтелекту жодна перевірена стаття не описувала експеримент повністю. Причини в інформатиці переважно інфраструктурні, а отже, усувні технічними засобами.

Розділ дає операційний ланцюг. Термінологія (ACM проти NASEM) дозволяє точно сказати, *що* перевірено; бейджі ACM задають зовнішній критерій якості артефакту. DMP переводить стадії життєвого циклу в перелік рішень, які ухвалюють до початку збирання, а п'ятнадцять підпринципів FAIR — у конкретні дії з метаданими, ідентифікаторами, словниками й ліцензіями. Репозитарій із DOI, версіонування й CITATION.cff роблять артефакт видимим і придатним для цитування; фіксація коду, залежностей і середовища — повторюваним. Передреєстрація захищає літературу від publication bias. Структура статті замикає цей ланцюг: IMRaD відокремлює виміряне від витлумаченого, а розділ методів разом із блоками про доступність даних і коду є тим місцем, де відтворюваність або забезпечується, або втрачається остаточно; тип публікації визначає, що саме вважатимуть результатом і яких артефактів вимагатимуть. Режими доступу й оцінка ризику реідентифікації показують, як бути відкритим настільки, наскільки можливо, і закритим настільки, наскільки необхідно.

Практичний підсумок: відтворюваний артефакт не «дороблюють» після ухвалення статті — його вирощують разом із дослідженням. Закласти структуру репозитарію й скласти DMP варто на початку роботи над дисертацією: це коштує кількох днів і економить місяці.

Питання для самоперевірки

1. Наведіть кількісні свідчення кризи відтворюваності: опитування *Nature* 2016 р., проєкт Open Science Collaboration, дослідження Коллберга й Пробстінга.
2. Чим причини невідтворюваності в інформатиці відрізняються від причин у природничих науках?
3. Розмежуйте repeatability, reproducibility і replicability за ACM v1.1 та NASEM 2019. Чому визначення не збігаються?
4. Назвіть три групи бейджів ACM. Чому Artifacts Available не свідчить про відтворюваність?

5. У чому різниця між «результати відтворено» і «результати підтверджено»?
6. Перелічіть стадії життєвого циклу даних і рішення, яке ухвалюють на кожній.
7. Які блоки має містити DMP? Хто і коли вимагає його в Horizon Europe?
8. Розкрийте зміст підпринципів F1–F4 на прикладі датасету мережевих вимірювань.
9. Чому FAIR не дорівнює open? Наведіть приклад FAIR-датасету з контрольованим доступом.
10. Порівняйте Zenodo, Dryad і Software Heritage. Чим версійний DOI відрізняється від концептуального?
11. Які пастки має аналіз у Jupyter-записниках і як їх усувають? Назвіть п'ять джерел недетермінізму та способи їх контролю.
12. Що фіксує передреєстрація, чим вона відрізняється від протоколу і коли недоречна?
13. Розшифруйте IMRaD і назвіть питання, на яке відповідає кожен розділ. Чому саме розділ методів визначає відтворюваність і чому злиття Results та Discussion небажане?
14. Порівняйте IMRaD із переліком елементів статті за постановою № 7-05/1. Якої частини бракує в українській схемі та як це компенсувати, не порушуючи вимог редакції?
15. Що описують схеми CARS і CCC і на якому етапі роботи над текстом кожна корисна?
16. Чим data paper відрізняється від дослідницької статті за тим, що вважається результатом і яких артефактів вимагає рецензент? Те саме для software paper і SoK.
17. Чому препринт не є типом публікації? Які з перелічених типів не зараховують до публікацій здобувача?
18. Порівняйте моделі gold, green, diamond і hybrid. У чому критика APC і чому CC BY-NC проблемна для науки?
19. Сформулюйте означення k-анонімності, поясніть її обмеження та спосіб обчислення $R_{\max}$.

Практичні завдання

1. Складіть план керування даними для власного дослідження за шістьма блоками таблиці 9.2 (2–3 сторінки), скориставшись шаблоном DMPonline або Argos. Для кожного блоку вкажіть конкретні назви: формати, стандарт метаданих, репозитарій, ліцензію, відповідального, обсяг сховища.

2. Оцініть власний (або опублікований) датасет за п'ятнадцятьма підпринципами FAIR: заповніть таблицю «підпринцип — ступінь виконання — що зробити», обчисліть FAIR за (9.1) і $C_{\text{м}}$, запропонуйте три поліпшення.
3. Сформуйте репозитарій артефакту за структурою рис. 9.2: README з п'яти блоків, LICENSE, CITATION.cff, файл фіксації залежностей, скрипт `run_all.sh`. Перевірте його «сліпим тестом»: попросіть колегу відтворити один рисунок вашої статті на чистій машині та зафіксуйте місця, де він зупинився.
4. Доведіть таблицю з квазіідентифікаторами (журнал підключень до безпроводової мережі) до k-анонімності за $k = 5$. Обчисліть $R_{\max}$ до і після перетворення та частку вилучених записів; опишіть втрату інформативності.
5. Депонуйте у пісочниці `sandbox.zenodo.org` невеликий навчальний датасет: заповніть рекомендовані поля метаданих, зазначте зв'язані ідентифікатори, оберіть ліцензію й одержіть DOI. Складіть його бібліографічний опис за ДСТУ 8302:2015.
6. Підготуйте передреєстрацію одного з експериментів дисертації на OSF: гіпотези, змінні, обсяг вибірки й правило зупинки, план аналізу з поправками, розмежування підтверджувального та розвідувального аналізів.
7. *Аудит структури.* Візьміть власну (або чужу) статтю з фахового видання і розкладіть її за IMRaD: позначте, які абзаци належать до методів, а які — до результатів і обговорення. Порахуйте частку обсягу кожного розділу й зіставте з орієнтиром 15 / 25 / 25 / 35 %. Окремо перевірте наявність блоків *Data availability*, *Funding*, *Conflicts of interest* і опису внеску за CRediT; складіть перелік того, чого бракує для відтворення.
8. *Вибір типу публікації.* Для трьох власних результатів (наявних або запланованих) визначте за таблицею 9.6 тип публікації, обґрунтуйте вибір і випишіть, яких артефактів вимагатиме рецензент саме для цього типу. Для одного з них складіть план статті: рубрикацію за вимогами обраного журналу та внутрішнє розмежування методики й результатів.

РОЗДІЛ 10

Академічна доброчесність та етика наукових досліджень

Розділ 9 показав, що результат, який неможливо відтворити, науковою цінністю не є. Цей розділ додає жорсткішу умову: результат, здобутий або оприлюднений недоброчесно, не є науковим взагалі. Виклад свідомо не моралізаторський, а *процедурний*: означення, ознаки порушення й алгоритм дій. Після розділу читач має вміти кваліфікувати ситуацію (порушення, сумнівна практика чи чесна помилка), знайти застосовну норму — від Закону України «Про академічну доброчесність» до Кодексу ALLEA й настанов COPE — і виконати потрібну процедуру: заяву про внесок, декларацію конфлікту інтересів, запит до етичного комітету, повідомлення про вразливість чи звернення про виправлення публікації.

10.1 Академічна доброчесність: поняття та нормативна рамка

10.1.1 Поняття й законодавче означення

Наука побудована на довірі: жоден дослідник не в змозі перевірити всі результати, на які спирається. Р. Мертон описав це як нормативну структуру науки: універсалізм, комунальність, безкорисливість, організований скептицизм [230], а **академічна доброчесність** є практичним втіленням цих норм. Одна сфабрикована стаття, на яку спираються десятки наступних, «отруює» цілий напрям.

До 2026 р. єдиною нормою прямої дії була ст. 42 Закону України «Про освіту» № 2145-VIII — рамкове положення з переліком порушень, але без процедур [37]. Її замінив спеціальний Закон України «Про академічну доброчесність» № 4742-IX від 18.12.2025, що набрав чинності 1 лютого 2026 р. і *введений у дію 31 липня 2026 р.* [22]. Це перший в історії України окремий закон, присвячений доброчесності; він охоплює всі рівні освіти й науки — від школи до докторантури, а також конкурси, експертизу, рецензування та присудження ступенів і звань. Тим самим законом внесено зміни до законів «Про освіту», «Про вищу освіту» та інших, які набрали чинності тією самою датою.

За ст. 1 Закону **академічна доброчесність** — це сукупність цінностей, принципів і заснованих на них правил, якими мають керуватися суб'єкти академічної діяльності під час провадження такої діяльності [22]. Порівняно з попереднім означенням змінено дві речі: суб'єктом є не лише учасник освітнього процесу, а будь-який суб'єкт академічної діяльності, і до етичних принципів додано *цінності* як окремий рівень. Стаття 18 перелі-

чує тринадцять видів порушень (табл. 10.1) — замість восьми в колишній редакції ст. 42; серед нових — недоброчесне використання результатів штучного інтелекту, академічний саботаж, схиляння до порушення та інституційні порушення самого закладу.

Таблиця 10.1 — Порушення академічної доброчесності за ст. 18 Закону України «Про академічну доброчесність»

Порушення	**Робоча ознака (що доводять)**
Відчуження авторства	Передання власного твору іншій особі для оприлюднення під її іменем
Академічний плагіат	Оприлюднення результату іншої особи як власного, без посилання на джерело
Приписування авторства	Зазначення серед авторів особи, яка не брала участі у створенні твору
Самоплагіат	Оприлюднення власних раніше оприлюднених результатів як нових, без зазначення джерела
Фабрикація	Вигадування даних або фактів
Фальсифікація	Свідома зміна чи модифікація наявних даних
Недоброчесне використання ШІ	Оприлюднення згенерованого результату як власного; приховування суттєвого використання
Недоброчесне оцінювання	Свідоме порушення вимог оцінювання, зокрема упереджене
Несамостійне виконання завдання	Виконання з недозволеними джерелами або сторонньою допомогою
Недозволена допомога	Надання допомоги, не передбаченої умовами завдання
Академічний саботаж	Дії, що перешкоджають іншій особі реалізувати її академічні права
Схиляння до порушення	Прохання, умовляння, підкуп або погроза з метою вчинення порушення
Інституційні порушення	Недостовірна інформація закладу; бездіяльність щодо розбудови системи доброчесності

Заходи реагування й санкції Закон поділяє за суб'єктом. Для здобувачів вони прогресивні: спершу виховні (тематичні заняття, усне чи письмове попередження, повторне виконання завдання), далі — академічні санкції (повторне проходження освітнього компонента, позбавлення стипендії, виключення зі складу органів, відрахування). Для педагогічних, науково-педагогічних і наукових працівників — виключення з органів управління, позбавлення права керівництва аспірантами, недопуск до конкурсів, позбавлення звань і ступенів, звільнення. Повідомлення про порушення, вчинене керівником закладу вищої освіти чи наукової установи, розглядає НАЗЯВО [16]; рішення про початок розгляду ухвалюють упродовж 10 робочих днів, а сам розгляд триває не довше 6 місяців.

Закон не містить відсоткових порогів: «плагіат до 20 %» — вигадка, якої в законодавстві немає (підрозділ 10.4). Водночас його перелік не вичерпує етичних проблем: приховане виключення спостережень, дослідження на живих системах без згоди власника чи недоброчесні практики рецензування оцінюють за міжнародними кодексами.

> **Що перевірити перед застосуванням.** Закон уведено в дію 31.07.2026, тож підзаконні акти й внутрішні положення закладів приводили у відповідність до нього протягом 2026 р. Перед тим, як посилатися на конкретну процедуру, строк чи санкцію, звіряйте їх із чинною редакцією Закону [22] та з положенням свого закладу: внутрішні акти, ухвалені за ст. 42 Закону «Про освіту», могли ще не бути оновлені.

10.1.2 Інституційний рівень: політика закладу та комісії з етики

Методичну рамку університетської системи доброчесності задають лист МОН України № 1/9-650 від 23.10.2018 із розгорнутим глосарієм [49] і рекомендації НАЗЯВО [42, 4]. Працездатна система має шість елементів: кодекс; орган розгляду (комісія з доброчесності або з етики — з постійним складом, правилами самовідводу й письмовим умотивованим рішенням); процедуру розгляду заяв із правом бути вислуханим; технічну перевірку текстів; навчання; прозорість. НАЗЯВО оцінює дієвість такої системи під час акредитації [16]. Здобувачеві варто в перші тижні знайти чотири документи: положення про доброчесність закладу, контакти комісії з етики, порядок перевірки текстів і вимоги обраних журналів щодо авторства, конфлікту інтересів та генеративних моделей.

10.2 Спектр порушень: від сумнівних практик до фабрикації

10.2.1 Груба недоброчесність: FFP

Три діяння беззастережно кваліфікують як **грубу недоброчесність** (research misconduct) і позначають абревіатурою **FFP**; кодекс ALLEA формулює їх так [54]. **Фабрикація** — вигадування результатів і подання їх як справжніх (таблиця затримок, значення якої згенеровано скриптом, а не виміряно). **Фальсифікація** — маніпулювання матеріалами чи процесами, зміна або приховування даних, унаслідок чого результат спотворюється (видалення «незручних» прогонів бенчмарку, після якого перевага власного алгоритму стає значущою). **Плагіат** — використання чужої роботи та ідей без належного посилання (підрозділ 10.3). Спільна ознака FFP — *умисел або груба необережність*: помилку виправляють, а недоброчесність розслідують. Емпірично FFP трапляється рідко: за метааналізом Д. Фанеллі, близько 2 % дослідників зізнаються у фабрикації чи фальсифікації даних і до 14 % повідомляють про такі дії колег [147].

10.2.2 Сумнівні дослідницькі практики (QRP)

Поширенішою є «сіра зона» — **сумнівні дослідницькі практики** (questionable research practices, QRP): дії, що не є прямим обманом, але систематично зміщують науковий запис на користь «красивого» результату. В опитуванні біомедичних дослідників США у приховуванні даних, що суперечать власним висновкам, зізналися 6 %, а в ігноруванні спостережень «за відчуттям» — 15 % [221, 198].

1) **Вибіркове звітування** — публікують лише ті метрики й набори даних, де метод виграє (у статті три датасети, у репозитарії — одинадцять).
2) **HARKing** — подання гіпотези, сформульованої після одержання результатів, як апріорної [203].
3) **Приховане виключення даних** — відкидання прогонів чи респондентів без заздалегідь заявленого критерію.
4) **«Нарізка» публікацій** (salami slicing) — ділення дослідження на мінімальні публікаційні одиниці; тест: чи має частина власне питання й висновок.
5) **Подвійна публікація** — той самий зміст у два видання, зокрема іншою мовою, без повідомлення редакторів.
6) **Почесне авторство** (підрозділ 10.6).
7) **Маніпулювання цитуванням** — примусове самоцитування й картелі цитувань (розділ 7).

QRP шкодять більше за рідкісний обман через добуток частоти на шкоду: якщо частка сфабрикованих робіт $p_{\mathrm{F}} \approx 0{,}02$, а частка робіт із принаймні однією QRP p_{Q} сягає 0,3–0,5, то навіть за меншої шкоди епізоду ($h_{\mathrm{Q}} \ll h_{\mathrm{F}}$) сукупне спотворення літератури $H = p_{\mathrm{F}} h_{\mathrm{F}} + p_{\mathrm{Q}} h_{\mathrm{Q}}$ визначається другим доданком. До того ж QRP майже не виявляються технічно, тому засіб протидії — передреєстрація й публікація протоколу та артефакту (розділ 9).

Спектр не є юридичною класифікацією: перехід між смугами визначають умисел, масштаб і вплив на висновки, і саме тому рішення ухвалює комісія, а не скрипт.

10.3 Плагіат і культура цитування

10.3.1 Види плагіату

Академічний плагіат — оприлюднення наукових результатів, отриманих іншими особами, як результатів власного дослідження та/або відтворення опублікованих текстів інших авторів без зазначення авторства [37]. Розрізняють шість форм: **дослівний** (копіювання без лапок і посилання); **мозаїчний** (patchwriting — абзац з уривків кількох джерел зі зміною окремих слів); **перефразування без посилання** (аргумент відтворено своїми слова-

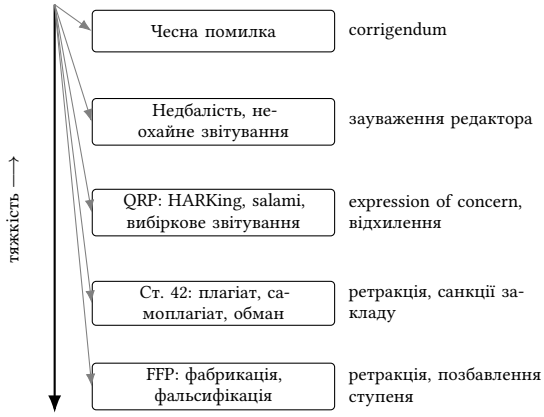


Рис. 10.1 — Спектр відхилень від доброчесності та типові наслідки

ми, джерело не назване — технічні системи цього не бачать); **плагіат ідей** (привласнення задуму, гіпотези, архітектури рішення; типове джерело — рукопис, отриманий на рецензування); **плагіат перекладу**; **самоплагіат** (подання власних раніше оприлюднених результатів як нових). Межа самоплагіату така: повторне використання *опису методу* з посиланням на власну роботу припустиме, повторна подача *результатів* як нових — порушення. Дисертація законно містить результати опублікованих статей автора, але вимагає вказівки, у яких саме публікаціях їх оприлюднено.

10.3.2 Три правила коректного відтворення чужого

Правило 1: цитата. Дослівне відтворення беруть у лапки й супроводжують посиланням із точною локалізацією. Цитують тоді, коли важливе саме формулювання — означення, нормативна вимога, спірна теза, яку далі критикують. Закон «Про авторське право і суміжні права» дозволяє вільне цитування в наукових цілях в обсязі, виправданому метою [21]; цитата понад 3–4 речення майже завжди свідчить, що автор не зрозумів джерела.

Правило 2: перефразування. Переказ своїми словами теж потребує посилання. Достатнім є перефразування, що змінює *синтаксичну структуру й рівень узагальнення*, а не окремі слова.

> **Джерело.** «The authors evaluated three anomaly detection models on the CICIDS2017 dataset and found that the random forest classifier achieved the highest F1-score, although at the cost of a threefold increase in inference latency.»
>
> **Неправильно (мозаїчний плагіат).** Автори оцінили три моделі виявлення аномалій на наборі CICIDS2017 і виявили, що класифікатор «випадковий ліс» до-

сяг найвищого F1-показника, хоча ціною трикратного зростання затримки висновування.

Правильно. У праці [умовне джерело] на наборі CICIDS2017 показано типовий для ансамблевих методів компроміс: приріст F1 супроводжувався потрійним зростанням затримки висновування, що обмежує застосовність таких моделей у режимі реального часу.

Правильний варіант має посилання, власну інтерпретацію замість перекладу й збережену перевірність. Позначку «умовне джерело» тут ужито навмисно: цитований фрагмент вигадано для ілюстрації, тож підставляти замість неї номер зі списку літератури цього посібника не можна. У власному тексті на цьому місці має стояти посилання на працю, у якій справді наведено ці результати, — інакше приклад правильного перефразування сам стає хибною атрибуцією.

Правило 3: вторинне цитування. Посилатися можна лише на те, що прочитано; якщо джерело недоступне, застосовують форму «цитовано за»: `[12, цит. за 34]`. Прихований ланцюжок посилань — коли автор переписує їх із чужого списку, не відкривши першоджерела, — породжує неправильні роки, приписані не тим авторам результати, а з появою генеративних моделей — і неіснуючі публікації (підрозділ 10.10).

Типові помилки здобувача: посилання лише наприкінці абзацу (не видно, де закінчився переказ джерела); копіювання означень із ДСТУ, ISO чи RFC без лапок; запозичення структури чужого огляду й відтворення схеми без підпису «за [номер]»; використання коду зі Stack Overflow без зазначення джерела й сумісної ліцензії [21].

10.4 Технічна перевірка текстів і межі коефіцієнта подібності

10.4.1 Як працюють системи зіставлення текстів

Основний сервіс, яким користуються українські заклади освіти, — StrikePlagiarism, інтегрований із Національним репозитарієм академічних текстів [314]; у міжнародній практиці поширений Turnitin. Раніше широко застосовували також Unicheck [328], однак Turnitin, якому сервіс належав, припинив його роботу 1 січня 2025 р. [326], тому в чинних положеннях про перевірку академічних текстів його вже не згадують — це типовий приклад того, чому перелік сервісів у внутрішніх документах ЗВО потребує регулярного перегляду. Конкретний склад сервісів змінюється, і твердження про поширеність варто робити лише з посиланням на актуальні дані. Усі вони реалізують один каркас: *нормалізація* → *шинглювання* (послідовності з $n = 5 \dots 9$ слів і їх геш-значення) → *пошук* збігів гешів в індексі → *склеювання* суміжних збігів і обчислення показника → *звіт*. Ключова властивість: система виявляє *текстові збіги*, а не плагіат, бо плагіат —

юридично-етична кваліфікація, яку дає людина. Міжнародне тестування 15 сервісів показало, що жоден не є надійним детектором: чутливість істотно падає на перефразованих і перекладених текстах, а звіти різних систем розходяться [152].

10.4.2 Коефіцієнт подібності: означення й межі

Нехай документ складається з N слів, а M — кількість слів у фрагментах, для яких знайдено збіг. **Коефіцієнт подібності** (similarity index)

$$S = \frac{M}{N} \cdot 100\,\%, \qquad S_{\text{екс}} = \frac{M - M_{\text{цит}} - M_{\text{бібл}} - M_{\text{норм}}}{N} \cdot 100\,\%, \tag{10.1}$$

де $M_{\text{цит}}$ — обсяг коректно оформлених цитат, $M_{\text{бібл}}$ — бібліографії, $M_{\text{норм}}$ — нормативних і термінологічних формулювань. Величина $S_{\text{екс}}$ — *експертно скоригований* показник; саме він придатний для обговорення.

Дві застереги до підрахунку. По-перше, віднімання виконують у *чисельнику*: зі знайденого обсягу збігів M вилучають ті збіги, що належать до перелічених категорій; знаменник N — повний обсяг документа — лишається незмінним. По-друге, категорії треба зробити такими, що не перетинаються: фрагмент, який є водночас цитатою й нормативним формулюванням, зараховують лише один раз, інакше можливе подвійне віднімання й навіть від'ємне значення $S_{\text{екс}}$. На практиці кожен збіг експерт відносить рівно до однієї категорії, а протокол перевірки містить перелік фрагментів із цим позначенням.

Три властивості формули (10.1) пояснюють, чому «відсоток» не є вироком: показник не інваріантний щодо індексу (одна робота дасть різні S у двох системах); нечутливий до тяжкості ($S = 25\,\%$ з назв законів і $S = 8\,\%$ зі скопійованого розділу «Методи» — принципово різні ситуації); має ненульову нижню межу, бо для роботи з кібербезпеки типові 5–15 % збігів неминучі.

Джерела збігів поділяють на три групи. *Хибнопозитивні*, які виключають із чисельника: нормативні формулювання (реквізити закону, пункт RFC), стала галузева термінологія («безпроводові сенсорні мережі»), бібліографія й текст у лапках із посиланням. *Не порушення* за наявності посилання — фрагменти власних публікацій, що лежать в основі дисертації. *Порушення* — скопійовані змістові фрагменти (опис методу, результати, висновки) та перефразовані без посилання аргументи, яких система зазвичай не бачить узагалі.

10.4.3 Обхідні маніпуляції та роль експерта

Спроби «обійти» систему доводять умисел і обтяжують відповідальність. *Гомогліфи* виявляються аналізом розподілу кодових точок, *невидимий*

текст і мікропробіли — порівнянням витягнутого тексту з видимим, а *автоматична синонімізація* залишає слід — **«замучені фрази»** (tortured phrases): «counterfeit consciousness» замість «artificial intelligence»; Г. Кабанак зі співавторами показали, що це надійний маркер машинно оброблених рукописів [79].

Процедура має два кроки: машинний звіт, а далі експертний розгляд, під час якого фахівець відкриває кожне джерело зі списку збігів із часткою понад поріг (1–2 % або 150 слів поспіль), перевіряє наявність посилання й лапок і формулює висновок: «загальний коефіцієнт подібності 22 %; після виключення бібліографії (7 %), нормативних формулювань (9 %) і цитат (3 %) скоригований показник становить 3 %; змістових запозичень не виявлено». Такий висновок перевірний, а фраза «оригінальність 78 %, робота допущена» — ні.

10.5 Кодекс ALLEA, практики COPE та виправлення наукового запису

10.5.1 Європейський кодекс поведінки ALLEA

Європейський кодекс поведінки для доброчесності досліджень, ухвалений федерацією академій ALLEA, є еталонним документом для проєктів Horizon Europe: грантова угода зобов'язує бенефіціарів дотримуватися його положень [54, 280]. Редакція 2023 р. врахувала відкриту науку, роботу з даними й автоматизовані інструменти. Кодекс побудовано на чотирьох принципах (табл. 10.2).

Таблиця 10.2 — Чотири принципи ALLEA і їх вияв в ІТ-дослідженні

Принцип	Зміст	Практичний вияв
Надійність (reliability)	Якість дизайну, методів, аналізу й використання ресурсів	Обґрунтований обсяг вибірки, фіксовані зерна, версії бібліотек, чесний baseline
Чесність (honesty)	Прозорість, повнота й неупередженість на всіх етапах	Звітування всіх прогонів і метрик, розкриття фінансування та обмежень
Повага (respect)	Повага до колег, учасників, суспільства, екосистем	Згода учасників тестування, повага до власників систем, коректне рецензування
Підзвітність (accountability)	Відповідальність від задуму до публікації та наставництва	Збереження первинних даних, готовність надати артефакт, відповідь редакторові

Від принципів кодекс переходить до *належних практик*, згрупованих за вісьмома етапами (дослідницьке середовище; навчання й наставництво; процедури; гарантії якості; дані; співпраця; публікація; рецензування), які

зручно використовувати як контрольний список власного проєкту; окремий розділ вимагає від установ мати письмову процедуру розгляду порушень і захищати повідомлювача.

10.5.2 COPE: практики й потокові схеми

Якщо ALLEA задає принципи для *дослідника й установи*, то Committee on Publication Ethics (COPE) регулює поведінку *редакцій*: його базовий документ визначає практики кожного видання — політику авторства, скарг, конфлікту інтересів, захисту даних, етичного нагляду, рецензування, реагування на порушення, виправлень і ретракцій [95]. Прикладна цінність COPE — у **потокових схемах** (flowcharts): алгоритмах для типових ситуацій (підозра на плагіат, зміна авторства, дублювання, сфабриковані дані) [96] (рис. 10.2).

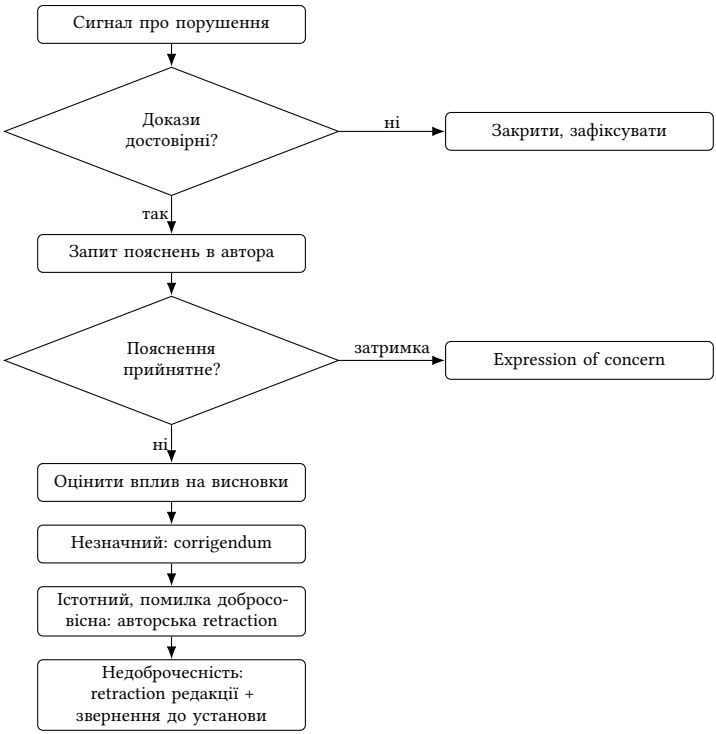


Рис. 10.2 — Спрощене дерево рішень редакції за підозри на порушення (за [95, 96, 97])

10.5.3 Три інструменти виправлення наукового запису

Corrigendum (або erratum) — виправлення помилки, що не змінює висновків: corrigendum виправляє помилку авторів, erratum — видавця. **Expression of concern** — публічне попередження редакції про обґрунтовані сумніви, поки триває розслідування. **Ретракція** — вилучення статті з наукового запису; мета — виправити літературу, а не покарати автора [97]. Підстави: ненадійність результатів, плагіат, дублювання, неетичне дослідження, неможливість підтвердити авторство, маніпуляції з рецензуванням. Стаття не зникає: її позначають знаком «RETRACTED», а повідомлення має вказувати причину. Ретракція не є «клеймом»: автор, який сам виявив помилку й попросив ретрагувати статтю, діє зразково.

10.5.4 Кейс: масові ретракції та фабрики статей

Ресурс Retraction Watch документує ретракції з 2010 р.; у 2023 р. його базу передано Crossref і відкрито для вільного використання [102], тож статус статті, критичної для власної аргументації, слід перевіряти перед цитуванням. **Фабрики статей** (paper mills) — комерційні структури, які виробляють рукописи на замовлення й продають авторство; їхній наукометричний слід описано в розділі 7. Купівля авторства охоплена одразу двома складами ст. 18 Закону «Про академічну доброчесність» — *відчуження авторства* з боку виконавця й *приписування авторства* з боку замовника [22]; це підстава для ретракції та позбавлення ступеня.

Що сталося. Упродовж 2022–2023 рр. видавництво Hindawi (придбане Wiley у 2021 р.) виявило масштабну маніпуляцію рецензуванням у спеціальних випусках своїх журналів, значна частина яких стосувалася обчислювальних наук і машинного навчання: запрошені редактори діяли в змові, рецензенти були підставними, а рукописи походили з фабрик статей. Ретраговано понад 8000 статей — найбільша хвиля ретракцій одного видавця [172].

Процедура. Виявлення спиралося на індикатори: аномальні мережі співавторства, неприродні ланцюжки цитувань, «замучені фрази» [79], збіги в метаданих подання, неправдоподібно короткі цикли рецензування; далі — ручна перевірка кожного випадку, звернення до авторів і їхніх установ, ретракційні повідомлення за настановами COPE [97]. Епізод не унікальний: раніше IEEE ретрагував понад 7300 конференційних публікацій [224].

Наслідки й висновки. Видавець припинив використання бренда, журнали втратили індексацію, а ретракції лишилися в публічному записі назавжди; здобувачі, які «закривали» такими статтями вимоги до публікацій, утратили підстави для захисту. Автор мав перевірити видання за чеклистом Think–Check–Submit і за базою ретракцій; видавець — не делегувати добір рецензентів запрошеним редакторам; заклад — не зводити вимоги до

публікацій до підрахунку індексованих статей, бо саме цей стимул створює попит на фабрики (ефект Гудхарта, розділ 7).

10.6 Авторство й внесок

10.6.1 Критерії ICMJE і таксономія CRediT

Операційний тест авторства запропонував Міжнародний комітет редакторів медичних журналів (ICMJE); його застосовують і в IT. Автором є особа, яка задовольняє *усі чотири* умови одночасно [181]: (1) суттєвий внесок у задум, дизайн, отримання, аналіз чи інтерпретацію даних; (2) написання тексту або критичне доопрацювання його інтелектуального змісту; (3) схвалення версії, що публікується; (4) згода нести відповідальність за всі аспекти роботи. Звідси фінансування, адміністративна підтримка й збір даних без участі в інтерпретації авторства *не* дають — таких осіб згадують у подяках; водночас особа, яка відповідає критерію 1, має одержати змогу виконати критерії 2–4.

Бінарний тест не показує, хто що робив; це завдання виконує **CRediT** (Contributor Roles Taxonomy) — стандарт ANSI/NISO Z39.104-2022 із чотирнадцятьма ролями [58, 76] (табл. 10.3).

Таблиця 10.3 — Матриця CRediT для статті з чотирма співавторами

Роль	А	Б	В	Г
Conceptualization	П	Р	—	—
Methodology	П	Д	—	—
Software	—	П	Д	—
Validation	Д	П	Р	—
Formal analysis	Р	—	П	—
Investigation	Д	Р	П	—
Resources	—	—	—	П
Data curation	—	Д	П	—
Writing — original draft	П	Д	Д	—
Writing — review & editing	Р	Р	Р	Д
Visualization	—	—	П	—
Supervision	П	—	—	Д
Project administration	—	—	—	П
Funding acquisition	—	—	—	П

П — провідна роль, Р — рівна участь, Д — допоміжна

Матриця відразу показує проблему з учасником Г. Провідні ролі він має лише ресурсні й адміністративні (Resources, Project administration, Funding acquisition), а до Writing — review & editing і Supervision долучений допо-

міжно. Висновок роблять не за переліком ролей, а за самими критеріями ICMJE, кожен з яких обов'язковий. Критерій 1 не виконано: Г не зробив істотного внеску ні в задум і план дослідження, ні в збирання, аналіз чи інтерпретацію даних — надання ресурсів, здобуття фінансування й загальне керівництво ICMJE прямо називає підставами, недостатніми для авторства. Критерій 2 теж не виконано: допоміжна участь у редагуванні не є критичним переглядом рукопису щодо важливого інтелектуального змісту. Оскільки критерії застосовують разом, Г не є автором, і його внесок належить до подяк — із зазначенням, у чому саме він полягав.

Внесок зручно й формалізувати: нехай w_r — вага ролі r ($\sum_r w_r = 1$), а $c_{ir} \in \{0;\ 0{,}25;\ 0{,}5;\ 1\}$ — ступінь участі особи i у ролі r; тоді нормований внесок

$$C_i = \frac{\sum_r w_r\, c_{ir}}{\sum_j \sum_r w_r\, c_{jr}}, \qquad \sum_i C_i = 1. \tag{10.2}$$

Вираз (10.2) не замінює критеріїв ICMJE — він робить обговорення предметним: команда наперед (на старті роботи) домовляється про ваги w_r і про поріг $C_{\min}$, нижче якого особа потрапляє до подяк. Порядок застосування однозначний: спершу перевіряють чотири критерії ICMJE, і лише для тих, хто їх задовольняє, число C_i допомагає узгодити порядок авторів та опис внеску. Зворотний порядок — коли високе C_i «дає право» на авторство всупереч критеріям — і є механізмом подарункового авторства.

10.6.2 Патології авторства, порядок авторів і суперечки

Почесне (подарункове) авторство — включення керівника підрозділу чи гранту без кваліфікованого внеску; найпоширеніша патологія української практики, формально — *приписування авторства* за ст. 18 Закону «Про академічну доброчесність» [22]. **Примарне авторство** (ghost authorship) — виключення особи з кваліфікованим внеском: аспіранта, який виконав експеримент, інженера, який написав систему. **Торгівля авторством** — купівля місця у статті, пряма підстава для ретракції.

Порядок авторів не регламентований: у комп'ютерних науках панує внесковий (перший — той, хто зробив найбільше, останній — науковий керівник), у математиці — алфавітний, а позначка «ці автори зробили рівний внесок» легітимно розв'язує суперечку за перше місце. Якщо конфлікт не вдається залагодити: звернення до співавторів із матрицею CRediT → звернення до редактора (COPE має окрему схему зміни авторства [96]) → комісія з етики. У парі «науковий керівник — здобувач» діє те саме правило: керівник, який сформулював задум і критично доопрацював текст, є співавтором, а той, хто лише адміністративно супроводжував роботу, — ні; зведення здобувача до подяк є примарним авторством [54].

Алгоритм: керівник вимагає вписати співавтора, який не працював над статтею. 1. З'ясувати, який саме внесок мається на увазі (він може бути невидимим: ідея, доступ до даних, доопрацювання). 2. Запропонувати заповнити матрицю CRediT на всіх учасників — формальний інструмент знімає особистий характер розмови. 3. Якщо внеску немає, запропонувати реальну участь у наступному дослідженні, а співавторство в ньому визначати за фактичним внеском і тими самими критеріями ICMJE. Обіцянка «впишемо в наступну статтю» без умови про реальний внесок — це лише перенесення подарункового авторства на іншу роботу. Подяка так само має відповідати справжній допомозі (доступ до обладнання, консультація, надання даних), а не бути автоматичною компенсацією за відмову в авторстві. 4. Пояснити ризик: виявлення почесного авторства веде до ретракції й б'є по керівникові насамперед. 5. Якщо тиск триває — письмове звернення (зі збереженням листування) до гаранта програми або комісії з етики.

10.7 Конфлікт інтересів

Конфлікт інтересів — ситуація, у якій сторонній інтерес може неправомірно вплинути (або виглядати таким, що впливає) на професійне судження; сам по собі він не є порушенням — порушенням є його *приховування*. Типів чотири: *фінансовий* (фінансування дослідження компанією, чий продукт оцінюють; володіння часткою чи патентом; оплачене консультування); *особистий* (родинні, дружні або конфліктні стосунки з авторами, рецензентами, редакторами); *інституційний* (оцінювання продукту власного підрозділу); *інтелектуальний* (авторство конкурентного методу).

У статті декларація конфлікту інтересів є окремим розділом; за відсутності конфлікту наводять явну формулу («Автори заявляють про відсутність конфлікту інтересів»), бо мовчання не еквівалентне запереченню, а джерело фінансування зазначають завжди. У рецензуванні правило жорсткіше: рецензент зобов'язаний повідомити редактора й узяти самовідвід, а рукопис є конфіденційним документом, який не можна ні поширювати, ні використовувати у власній роботі, ні завантажувати до зовнішніх сервісів (підрозділ 10.10).

Алгоритм: рецензент вимагає додати посилання на власні статті. 1. Оцінити зауваження по суті: якщо роботи релевантні — додати й процитувати змістовно. 2. Інакше відповісти предметно: «Ми розглянули джерела [X, Y]; вони стосуються іншого класу задач і не впливають на аргументацію розділу 3, тому не додані». 3. Паралельно надіслати редакторові конфіденційне повідомлення: «Рецензент 2 пропонує додати 11 посилань, з яких 9 — на власні публікації». 4. Не погоджуватися «щоб не сперечатися»: примусове цитування прямо згадане серед порушень у настановах COPE [95].

10.8 Етика досліджень за участю людей і персональні дані

10.8.1 Гельсінська декларація, згода, етичні комітети

Сучасна етика досліджень за участю людей виросла з **Гельсінської декларації** Всесвітньої медичної асоціації (чинна редакція — 2024 р.) [351]. Її ядро універсальне й прямо застосовне до ІТ: добробут учасника переважає над інтересами науки; участь добровільна й ґрунтується на інформованій згоді; ризик мінімізують і співвідносять з очікуваною користю; протокол підлягає незалежному етичному розгляду *до* початку дослідження; вразливі групи потребують додаткового захисту; результати оприлюднюють, зокрема негативні. Незалежний розгляд виконує **етичний комітет** — IRB або REC; для Horizon Europe **ethics appraisal** є обов'язковою складовою оцінювання заявки [280, 54], а в українських ЗВО цю функцію виконує комісія з етики.

Згода є інформованою, якщо учасник *до* участі одержав зрозумілою мовою: мету й характер дослідження; що з ним робитимуть і скільки це триватиме; які дані збирають (журнали дій, записи екрана, біометричні сигнали); хто матиме доступ і як довго їх зберігатимуть; які ризики можливі; що участь можна припинити будь-коли; контакти дослідника й органу розгляду скарг. **Вразливі групи** — неповнолітні, особи з обмеженою дієздатністю, особи в залежних стосунках від дослідника (власні студенти, підлеглі), військовослужбовці, внутрішньо переміщені особи. Залежність — найменш помітна проблема в ІТ: опитування власної групи, де відмова здається ризикованою, порушує добровільність.

10.8.2 Чотири типові формати ІТ-досліджень

Опитування й інтерв'ю: не питати ПІБ і місце роботи, якщо вони не потрібні для аналізу; уникати «необов'язкових» демографічних питань, сукупність яких робить респондента унікальним (розділ 9). *Юзабіліті-тестування*: запис екрана та натискань клавіш може захопити сторонні дані (сповіщення, паролі), тож потрібні окремий тестовий обліковий запис і право учасника зупинити запис. *Дослідження соціальних мереж*: публічна доступність даних не означає згоди — експеримент з емоційним «зараженням», у якому без відома користувачів змінювали стрічку новин, спричинив тривалу дискусію саме тому, що відповідність умовам користування не замінює інформованої згоди [208]; агреговане спостереження за публічним контентом зазвичай прийнятне, а втручання в досвід користувача, цитування ідентифіковних дописів і збирання профілів потребують розгляду комітетом. *Краудсорсинг*: виконавці на платформах — учасники, а не «ресурс», тож їм належить справедлива оплата й право відмовитися без штрафу.

10.8.3 Персональні дані: підстава й мінімізація

Етична прийнятність не замінює правової підстави. За GDPR обробка персональних даних потребує однієї з підстав ст. 6 (для досліджень зазвичай згода або суспільний інтерес), для спеціальних категорій — додаткової умови ст. 9; ст. 89 дозволяє відступи для наукових цілей за умови належних гарантій, насамперед мінімізації даних і псевдонімізації [279]. Аналогічні вимоги встановлює Закон «Про захист персональних даних» [33]. Практичний мінімум: записати мету й підставу обробки до збирання; не збирати зайвого; псевдонімізувати на етапі збирання; зберігати ключ відповідності окремо; передбачити режим доступу до датасету ще в плані керування даними. Анонімізацію, k-анонімність і оцінювання ризику реідентифікації розглянуто в розділі 9.

10.9 Етика кібербезпекових досліджень

10.9.1 Menlo Report і оцінювання ризику

Класична модель «дослідник — учасник» погано описує кібербезпекові дослідження: вплив здійснюється на системи, а люди зачіпаються опосередковано. Для цього випадку розроблено **Menlo Report** [119, 120] з чотирма принципами: *повага до осіб* — інформована згода там, де можлива, інакше ідентифікація зачеплених сторін і мінімізація втручання; *благодіяння* — зважування користі й шкоди; *справедливість* — неприпустимо покладати ризик на тих, хто не отримує користі (наприклад, на користувачів заражених пристроїв); *повага до права й публічного інтересу*.

Найскладніший — другий принцип. Якщо B — очікувана суспільна користь, а по кожному сценарію шкоди k відомі ймовірність p_k і збиток h_k, то дослідження прийнятне за умови

$$B \;>\; \sum_{k} p_k h_k \quad \text{за додаткової вимоги} \quad \max_k h_k < h_{\text{доп}}, \tag{10.3}$$

де $h_{\text{доп}}$ — поріг неприйнятної шкоди. Друга умова принципова: жодна очікувана користь не виправдовує катастрофічного наслідку (зупинка медичного обладнання, втрата даних критичної інфраструктури). Оцінки p_k і h_k записують у розділі про етичні міркування, який рекомендують включати до статей із мережевих вимірювань [259].

10.9.2 Типові формати та їх обмеження

Дослідження на живих системах: втручання в чужу систему потребує письмового дозволу власника з окресленим обсягом (scope), вікном часу й контактом для негайної зупинки; без дозволу воно переходить у площину кримінального права. *Сканування Інтернету*: практика «доброго сусі-

да» авторів ZMap — сканувати з виділених адрес зі зворотним DNS, що вказує на дослідження; тримати вебсторінку з поясненням мети; вести й безумовно поважати список виключень (opt-out); обмежувати частоту [124]. *Honeypot і перехоплення трафіку*: не перенаправляти атаки на третіх осіб; не зберігати зайвого корисного навантаження; псевдонімізувати адреси; перехоплення в мережі закладу можливе лише з правовою підставою й повідомленням користувачів. *Витоки й викрадені дані*: дампи баз здобуто протиправно, тож потрібні висновок етичного комітету, робота з агрегатами й відмова від реідентифікації — оцінювати стійкість паролів за витоком припустимо, перевіряти, чи є в ньому конкретна людина, — ні. *Подвійне призначення*: результат, придатний і для захисту, і для нападу, вимагає аналізу зловживання — хто виграє від публікації, хто від затримки.

10.9.3 Координоване розкриття вразливостей

Повне негайне розкриття наражає користувачів на ризик, а безстрокове мовчання залишає вразливість без виправлення. Компромісом є **координоване розкриття** (coordinated vulnerability disclosure, CVD): ISO/IEC 29147:2018 описує *зовнішній* інтерфейс — як постачальник приймає повідомлення й публікує інформацію про виправлення [190], ISO/IEC 30111:2019 — *внутрішній* процес опрацювання [191]. Послідовність: (1) зафіксувати відтворюваний доказ (PoC, версії, умови) і час виявлення; (2) знайти канал постачальника (`security.txt`, bug bounty, CERT-UA як координатор); (3) надіслати повідомлення захищеним каналом із наміром координованого розкриття й строком (типово 90 днів); (4) не публікувати й не тестувати поза узгодженим обсягом; (5) узгодити дату розкриття, номер CVE й текст повідомлення; (6) опублікувати після виправлення або спливу строку без «зброєзованого» експлойта; (7) описати в статті дату повідомлення, реакцію постачальника й узгоджений строк. Готовий експлойт публікують лише тоді, коли він додає наукової цінності порівняно з описом механізму.

Показовий приклад — випадок 2021 р., коли дослідники Університету Міннесоти надсилали супровідникам ядра Linux «лицемірні» патчі, які приховано створювали дефект, не повідомляючи про експеримент; прийняту на IEEE Symposium on Security and Privacy статтю автори відкликали, визнавши, що проводили дослідження над спільнотою без згоди й розгляду [352, 179]. За Menlo тут порушено *повагу до осіб*, *благодіяння* і *справедливість*: шкоду покладено на супровідників, а користь діставали дослідники. Урок: коли об'єктом дослідження є спільнота, процес чи чужа інфраструктура, — це дослідження за участю людей.

10.10 Генеративний штучний інтелект і доброчесність

10.10.1 Чому ШІ не може бути автором

Консенсус видавничої спільноти спирається на четвертий критерій ICMJE: автор має нести *відповідальність* за роботу, а модель її нести не може. Позицію першим сформулював *Science*: інструмент не є автором, бо науковий запис є записом людського пошуку [321]; COPE оформив її як офіційну заяву [94], а великі видавництва — як політики (табл. 10.4).

Таблиця 10.4 — Політики щодо генеративного ШІ (редакції, чинні на 22.09.2026)

Суб'єкт	Ключові вимоги	Де розкривати
COPE [94]	ШІ не автор; автори відповідають за весь зміст, зокрема за коректність джерел	Методи або подяки
Elsevier [133]	Допоміжне використання під наглядом людини (поліпшення мови й читності, а також окремі допоміжні завдання) з розкриттям; згенеровані зображення — лише у визначених випадках, зокрема як пояснювальні, за явного дозволу та позначення	Заява перед списком джерел
Springer Nature [311]	Документування використання LLM; заборона згенерованих зображень	Розділ «Методи»
IEEE [180]	Розкриття ШІ-генерованого тексту, рисунків і коду із зазначенням системи	Acknowledgments
ЄК [137]	Чотири принципи; відповідальність автора; обережність із даними	Методи та звітність

Виняток щодо розкриття треба розуміти вузько. Він охоплює базове мовне редагування — перевірку правопису, пунктуації та граматики; таке використання заяви не потребує. Натомість суттєва перебудова тексту за допомогою ШІ — переписування абзаців, зміна структури викладу, генерування формулювань — розкриття потребує, і політика Elsevier прямо цього вимагає [133]. Межа проходить не між «мовою» і «змістом» узагалі, а за обсягом втручання: чим більше з написаного належить моделі, тим обов'язковіша заява. ШІ як *джерело змісту* — потребує розкриття або заборонене; ШІ замість автора — неприпустимо завжди.

Застереження щодо актуальності. Політики видавництв переглядають щороку, а іноді й частіше. Табл. 10.4 подає редакції, чинні на 22.09.2026; перед поданням рукопису слід відкрити чинну сторінку політики конкретного журналу й звірити вимоги — вимоги журналу можуть бути суворішими за загальні вимоги видавництва.

10.10.2 Живі настанови ЄС, фабрикація джерел і рецензування

«Living guidelines on the responsible use of generative AI in research» Європейської комісії переносять на ШІ чотири принципи ALLEA [137, 54]: *надійність* — розуміти обмеження моделі й перевіряти вихід; *чесність* — прозоро

розкривати використання; *повага* — не завантажувати чужі конфіденційні дані та рукописи; *підзвітність* — відповідальність за результат лишається на дослідникові. Доповнює їх AI Act, який виводить системи, розроблені виключно для досліджень, зі своєї сфери дії [281].

Найчастіший збій — **фабрикація джерел**: модель генерує правдоподібні, але неіснуючі посилання, DOI й цитати. Коли посилання функціонують як дані (в огляді літератури, у бібліометричному аналізі), їх фабрикація кваліфікується як недоброчесність автора, а не «помилка інструмента» [283]; звідси правило: кожне посилання відкрити й звірити (розділ 6), так само як згенеровані числа, формули й код. Окремо настанови застерігають від ШІ в *рецензуванні*: рукопис є непублічним документом, і завантаження його до зовнішнього сервісу є розголошенням, а делегування експертного судження моделі позбавляє процедуру сенсу.

10.10.3 Що дозволено здобувачеві

В Україні орієнтиром є спільні рекомендації МОН і Мінцифри щодо відповідального використання ШІ у ЗВО [3], а на рівні закладу — політика курсу. *Прийнятно*: мовне редагування та переклад власного тексту; варіанти формулювань; пояснення незрозумілої статті; чернетковий код із подальшою перевіркою. *Потребує розкриття*: породження фрагментів тексту, що потрапляють у рукопис; генерування рисунків; істотна частина коду аналізу. *Неприйнятно*: генерування результатів, даних або посилань; видавання згенерованого огляду за власний; завантаження чужого рукопису чи персональних даних у зовнішній сервіс; зазначення моделі співавтором.

10.11 Процедури, повідомлення про порушення й культура доброчесності

Знайшов помилку у власній опублікованій статті. 1. Оцінити вплив: чи змінюються висновки? 2. Відтворити аналіз із виправленням і задокументувати різницю. 3. Повідомити співавторів. 4. Написати редакторові: що помилково, як виявлено, як впливає на висновки, який інструмент пропонуєте (corrigendum за незмінних висновків, ретракція — якщо висновки не витримують). 5. Оновити репозитарій артефакту, додавши виправлену версію з новим DOI (розділ 9). 6. Не чекати, доки помилку знайде хтось інший.

Виявив ознаки плагіату або фабрикації в чужій статті. 1. Не публікувати звинувачень у соціальних мережах. 2. Зібрати перевірні докази — паралельні фрагменти з точними сторінками, посилання на першоджерело. 3. Звернутися до редакції (за настановами COPE розслідування веде редактор [95, 96]) або до установи автора; повідомлення має бути конкретним і безоцінним. 4. Інший канал — коментар на платформі післяпублікаційного рецензування, з фактами. 5. Зберегти листування й дати та бути готовим до тривалого процесу.

Типовий порядок розгляду справи в закладі має шість кроків: подання письмової заяви з доказами; попередня перевірка на прийнятність; розслідування з правом особи надати пояснення й ознайомитися з матеріалами; письмове вмотивоване рішення з посиланням на пункт положення; санкція в межах, передбачених Законом «Про академічну доброчесність» [22]; оскарження. Кодекс ALLEA вимагає від процедури неупередженості (самовідвід осіб із конфліктом інтересів), конфіденційності до ухвалення рішення, пропорційності санкції та права на захист [54].

Повідомлення про порушення (whistleblowing) — інформування відповідального органу про доброчесні підозри. Директива ЄС 2019/1937 встановлює мінімальні стандарти захисту повідомлювачів: внутрішні й зовнішні канали, конфіденційність особи, заборона переслідувальних заходів (звільнення, зниження оцінок, відмова в продовженні аспірантури, виключення зі співавторства) і перенесення тягаря доведення на роботодавця [117, 54]. Для здобувача це означає: повідомляти через офіційний канал і письмово; зберігати докази поза інфраструктурою закладу; пам'ятати, що захист поширюється на того, хто мав обґрунтовані підстави вважати інформацію правдивою, тоді як свідомо неправдиве повідомлення захисту не має.

Поведінка здобувача формується не кодексами, а прикладом керівника, тому кодекс ALLEA відводить наставництву окремий блок практик. Ознаки здорового середовища: склад авторів обговорюють на старті; негативний результат вважається результатом; помилка, знайдена самостійно, не карається; дані й код зберігають так, що їх можна перевірити через рік. Ознаки небезпечного середовища зворотні: тиск «потрібна стаття до атестації», пропозиції «швидких» журналів, прохання не звітувати невдалі прогони.

10.12 Висновки до розділу

Академічна доброчесність є не формальним етапом перевірки тексту, а умовою існування наукового знання: наука працює на довірі, і кожне порушення підриває механізм, що не має альтернативи.

Нормативна рамка для українського дослідника чотирирівнева: Закон України «Про академічну доброчесність» № 4742-IX, уведений у дію 31 липня 2026 р.; політика власного закладу; Європейський кодекс ALLEA (надійність, чесність, повага, підзвітність); настанови COPE. Основну шкоду науковому запису завдають не рідкісні фабрикація й фальсифікація, а поширені сумнівні практики — вибіркове звітування, HARKing, «нарізка», почесне авторство; протидіють їм не детектори, а передреєстрація, повне звітування й публікація артефактів. Коефіцієнт подібності є вимірюванням збігів, а не кваліфікацією плагіату.

Авторство визначають чотири критерії ICMJE, а внесок описують таксономією CRediT; склад авторів фіксують письмово до подання рукопису, а

конфлікт інтересів не є порушенням — порушенням є його приховування. Дослідження за участю людей потребують інформованої згоди й етичного розгляду, і в IT ця вимога поширюється на опитування, юзабіліті-тести, дослідження соціальних мереж і краудсорсинг; персональні дані додатково вимагають правової підстави за GDPR і Законом України «Про захист персональних даних». Кібербезпекові дослідження підпорядковані принципам Menlo Report, а виявлені вразливості розкривають координовано за ISO/IEC 29147 і 30111. Генеративний ШІ не може бути автором, його використання підлягає розкриттю, а відповідальність за результат залишається на дослідникові.

Питання для самоперевірки

1. Наведіть означення академічної доброчесності за ст. 1 Закону України «Про академічну доброчесність» і перелічіть тринадцять видів порушень за ст. 18. Чим цей перелік відрізняється від переліку колишньої редакції ст. 42 Закону «Про освіту»?
2. Чим заходи реагування та санкції для здобувача освіти відрізняються від санкцій для наукового працівника? У яких випадках повідомлення розглядає НАЗЯВО і в які строки?
3. Що таке FFP і за якою ознакою груба недоброчесність відмежовується від добросовісної помилки?
4. Чому сумнівні дослідницькі практики завдають науковому запису більшої сукупної шкоди, ніж фабрикація?
5. Наведіть три приклади QRP, специфічних для досліджень з машинного навчання, і вкажіть, як їм запобігти.
6. Порівняйте мозаїчний плагіат, перефразування без посилання й самоплагіат за ознаками та наслідками.
7. Чому коефіцієнти подібності різних систем не можна порівнювати? Що таке скоригований показник і які збіги з нього виключають?
8. Назвіть чотири принципи Кодексу ALLEA та наведіть приклади їх втілення в IT-дослідженні.
9. У яких випадках застосовують corrigendum, expression of concern і ретракцію? Хто ухвалює рішення?
10. Сформулюйте чотири критерії авторства ICMJE. Чи є автором особа, яка надала обладнання й фінансування?
11. Чим таксономія CRediT доповнює критерії ICMJE? Як розподіляють авторство в парі «науковий керівник — здобувач»?
12. Які чотири типи конфлікту інтересів розрізняють і де їх декларують?

13. Які вимоги Menlo Report обмежують сканування адресного простору Інтернету та приманки? Що таке координоване розкриття вразливості й які стандарти його описують?
14. Чому генеративна модель не може бути співавтором і чому її не застосовують у рецензуванні?

Практичні завдання

1. Складіть перелік нормативних документів власного закладу з питань академічної доброчесності й опишіть процедуру розгляду заяви: хто подає, у який строк, як оскаржується рішення.
2. Візьміть англомовну статтю за темою власного дослідження. Випишіть абзац і підготуйте три версії його використання: цитату в лапках із посиланням, коректне перефразування з посиланням і приклад некоректного мозаїчного перефразування; поясніть різницю.
3. Отримайте звіт системи перевірки на збіги для розділу власної роботи, згрупуйте всі збіги за типом, обчисліть скоригований показник за (10.1) і сформулюйте письмовий висновок експерта.
4. Для власної статті заповніть матрицю CRediT, обчисліть нормовані внески за (10.2) та перевірте кожного учасника за критеріями ICMJE; сформулюйте declaration of interest і заяву про фінансування.
5. Розберіть чотири ситуації й для кожної складіть план дій із зазначенням застосовної норми (конкретна стаття Закону «Про академічну доброчесність», ALLEA, COPE): (а) ви виявили в опублікованій статті дослівні фрагменти свого препринта; (б) керівник вимагає вписати співавтором особу, яка не працювала над статтею; (в) рецензент вимагає додати дев'ять посилань на власні роботи; (г) ви знайшли помилку в обчисленнях власної опублікованої статті, через яку зникає значущість основного ефекту.
6. Для власного дослідження, що передбачає збирання даних про людей або втручання в чужі системи, складіть розділ «Етичні міркування»: зачеплені сторони, правова підстава обробки даних, оцінка p_k і h_k за (10.3), заходи мінімізації, потреба у висновку комісії з етики.

Список використаних джерел

Дата звернення до всіх електронних ресурсів цього переліку — 22.09.2026, якщо в записі не зазначено іншої.

[1] Бойко А. Метод виявлення атак на корпоративні вебдодатки на основі градієнтного бустингу. *Кібербезпека: освіта, наука, техніка.* 2025. Т. 3, № 31. С. 710–716. DOI: 10.28925/2663-4023.2025.31.1062.

[2] Вибір та обґрунтування теми наукового дослідження: конспект лекції з дисципліни «Основи наукових досліджень». Харків: Харківський національний автомобільно-дорожній університет, 2021. URL: https://fmab.khadi.kharkov.ua/fileadmin/F-FUB/3_OND_L1.pdf.

[3] Використання штучного інтелекту у вищій освіті: рекомендації для викладачів, здобувачів, адміністрацій закладів вищої освіти та дослідників / Міністерство цифрової трансформації України, Міністерство освіти і науки України. Київ, 2025. URL: https://cms.thedigital.gov.ua/storage/uploads/files/page/community/docs/Vykorystannya_AI_u_vyshchyy_osviti.pdf.

[4] Глосарій з академічної доброчесності / Національне агентство із забезпечення якості вищої освіти. Київ, 2020. URL: https://naqa.gov.ua/wp-content/uploads/2020/01/Глосарій.pdf.

[5] Делій А., Золінген Р. ван, Вейнандс В. Керівництво по eduScrum. Правила гри. Версія 1.2 / ред. Д. Сазерленд; перекл. з англ. В. Ю. Соколова; ред. перекл. В. Л. Бурячок. Київ: Київський університет імені Бориса Грінченка, 2019. 36 с. URL: https://elibrary.kubg.edu.ua/id/eprint/29432/.

[6] Деякі питання присудження (позбавлення) наукових ступенів: постанова Кабінету Міністрів України від 17.11.2021 № 1197. URL: https://zakon.rada.gov.ua/laws/show/1197-2021-%D0%BF.

[7] ДСТУ 8302:2015. Інформація та документація. Бібліографічне посилання. Загальні положення та правила складання. Київ: ДП «УкрНДНЦ», 2016. 16 с.

[8] ДСТУ 3008:2015. Інформація та документація. Звіти у сфері науки і техніки. Структура та правила оформлювання. Київ: ДП «УкрНДНЦ», 2016. 26 с.

[9] Ініціатива CDIO. Версія 2.1 / перекл. з англ. В. Ю. Соколова; ред. В. Л. Бурячок. Київ: Київський університет імені Бориса Грінченка, 2019. 34 с. DOI: 10.5281/zenodo.2638101. URL: https://elibrary.kubg.edu.ua/id/eprint/27053/.

[10] Кібербезпека: освіта, наука, техніка: науковий журнал Київського столичного університету імені Бориса Грінченка. URL: https://csecurity.kubg.edu.ua/.

[11] Кіцель Н., Борисенко О. Розробка лабораторного практикуму з аналізу та виявлення програм-вимагачів для освітніх програм з кібербезпеки. *Кібербезпека: освіта, наука, техніка.* 2026. Т. 1, № 33. С. 504–511. DOI: 10.28925/2663-4023.2026.33.1145.

[12] Костюк Ю., Бебешко Б., Крючкова Л., Литвинов В., Оксанич І., Складанний П., Хорольська К. Захист інформації та безпека обміну даними в безпроводових мобільних мережах з автентифікацією і протоколами обміну ключами. *Кібербезпека: освіта, наука, техніка*. 2024. Т. 1, № 25. С. 229–252. DOI: 10.28925/2663-4023.2024.25.229252.

[13] Методичні рекомендації до виконання та оформлення дисертаційних робіт. Краматорськ: Інститут проблем штучного інтелекту МОН України і НАН України, 2023. URL: https://icsanaas.com.ua/wp-content/uploads/2023/06/.

[14] Міністерство освіти і науки України: офіційний вебсайт. URL: https://mon.gov.ua/.

[15] Національна бібліотека України імені В. І. Вернадського. Наукова періодика України. URL: http://www.irbis-nbuv.gov.ua/.

[16] Національне агентство із забезпечення якості вищої освіти: офіційний вебсайт. URL: https://naqa.gov.ua/.

[17] Національний репозитарій академічних текстів / Український інститут науково-технічної експертизи та інформації. URL: https://nrat.ukrintei.ua/.

[18] Національний фонд досліджень України. Оголошені конкурси: умови проведення, тематичні напрями, порядок експертизи. URL: https://nrfu.org.ua/en/contests/current-calls/.

[19] Національний фонд досліджень України: офіційний вебсайт. URL: https://nrfu.org.ua/.

[20] Опірський І., Головчак Р., Мойсійчук І., Балянда Т., Гаранюк С. Проблеми та загрози безпеці IoT пристроїв. *Кібербезпека: освіта, наука, техніка*. 2021. Т. 3, № 11. С. 31–42. DOI: 10.28925/2663-4023.2021.11.3142.

[21] Про авторське право і суміжні права: Закон України від 01.12.2022 № 2811-IX. URL: https://zakon.rada.gov.ua/laws/show/2811-20.

[22] Про академічну доброчесність: Закон України від 18.12.2025 № 4742-IX. *Відомості Верховної Ради України*. 2026. URL: https://zakon.rada.gov.ua/laws/show/4742-IX. Дата звернення: 22.09.2026.

[23] Про визначення механізму фінансового забезпечення діяльності державних наукових установ, наукових досліджень та науково-технічних (експериментальних) розробок державних закладів вищої освіти з урахуванням результатів їх державної атестації: постанова Кабінету Міністрів України від 28.01.2026 № 116 (редакція від 16.07.2026). URL: https://zakon.rada.gov.ua/laws/show/116-2026-%D0%BF.

[24] Про вищу освіту: Закон України від 01.07.2014 № 1556-VII. URL: https://zakon.rada.gov.ua/laws/show/1556-18.

[25] Про внесення змін до переліку галузей знань і спеціальностей, за якими здійснюється підготовка здобувачів вищої та фахової передвищої освіти: постанова Кабінету Міністрів України від 30.08.2024 № 1021. URL: https://zakon.rada.gov.ua/laws/show/1021-2024-%D0%BF.

[26] Про затвердження національного плану щодо відкритої науки: розпорядження Кабінету Міністрів України від 08.10.2022 № 892-р. URL: https://zakon.rada.gov.ua/laws/show/892-2022-%D1%80.

[27] Про затвердження переліку галузей знань і спеціальностей, за якими здійснюється підготовка здобувачів вищої освіти: постанова Кабінету Міністрів України від 29.04.2015 № 266. URL: https://zakon.rada.gov.ua/laws/show/266-2015-%D0%BF.

[28] Про затвердження Положення про акредитацію освітніх програм, за якими здійснюється підготовка здобувачів вищої освіти: наказ Міністерства освіти і науки України від 15.05.2024 № 686, зареєстровано в Міністерстві юстиції України 10.07.2024 за № 1013/42358. Чинний з 01.08.2024. URL: https://zakon.rada.gov.ua/laws/show/z1013-24.

[29] Про затвердження Порядку підготовки здобувачів вищої освіти ступеня доктора філософії та доктора наук у закладах вищої освіти (наукових установах): постанова Кабінету Міністрів України від 23.03.2016 № 261. URL: https://zakon.rada.gov.ua/laws/show/261-2016-%D0%BF.

[30] Про затвердження Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії: постанова Кабінету Міністрів України від 12.01.2022 № 44. URL: https://zakon.rada.gov.ua/laws/show/44-2022-%D0%BF.

[31] Про затвердження Порядку проведення державної атестації наукових установ та закладів вищої освіти в частині провадження такими закладами наукової (науково-технічної) діяльності: постанова Кабінету Міністрів України від 19.07.2017 № 540 (редакція від 21.03.2026). URL: https://zakon.rada.gov.ua/laws/show/540-2017-%D0%BF.

[32] Про затвердження Порядку формування Переліку наукових фахових видань України: наказ Міністерства освіти і науки України від 15.01.2018 № 32 (у редакції наказу Міністерства освіти і науки України від 19.01.2026 № 56). URL: https://zakon.rada.gov.ua/laws/show/z0148-18.

[33] Про захист персональних даних: Закон України від 01.06.2010 № 2297-VI. URL: https://zakon.rada.gov.ua/laws/show/2297-17.

[34] Про наукову і науково-технічну діяльність: Закон України від 26.11.2015 № 848-VIII. URL: https://zakon.rada.gov.ua/laws/show/848-19.

[35] Про Національний фонд досліджень України: постанова Кабінету Міністрів України від 04.07.2018 № 528 (редакція від 26.11.2024). URL: https://zakon.rada.gov.ua/laws/show/528-2018-%D0%BF.

[36] Про опублікування результатів дисертацій на здобуття наукових ступенів доктора і кандидата наук: наказ Міністерства освіти і науки України від 23.09.2019 № 1220 (редакція від 09.04.2024). URL: https://zakon.rada.gov.ua/laws/show/z1086-19.

[37] Про освіту: Закон України від 05.09.2017 № 2145-VIII. URL: https://zakon.rada.gov.ua/laws/show/2145-19.

[38] Про охорону прав на винаходи і корисні моделі: Закон України від 15.12.1993 № 3687-XII (редакція від 26.02.2026). URL: https://zakon.rada.gov.ua/laws/show/3687-12.

[39] Про підвищення вимог до фахових видань, внесених до переліків ВАК України: постанова президії Вищої атестаційної комісії України від 15.01.2003 № 7-05/1. URL: https://zakon.rada.gov.ua/rada/show/v05_1330-03.

[40] Про утворення Національної ради України з питань розвитку науки і технологій: постанова Кабінету Міністрів України від 05.04.2017 № 226 (редакція від 15.12.2020). URL: https://zakon.rada.gov.ua/laws/show/226-2017-%D0%BF.

[41] Радівілова Т. А., Кириченко Л. О., Тавалбех М., Ільков А. В. Виявлення аномалій в телекомунікаційному трафіку статистичними методами. *Кібербезпека: освіта, наука, техніка.* 2021. Т. 3, № 11. С. 183–194. DOI: 10.28925/2663-4023.2021.11.183194.

[42] Рекомендації для закладів вищої освіти щодо розробки та впровадження університетської системи забезпечення академічної доброчесності / Національне агентство із забезпечення якості вищої освіти. Київ, 2019. 24 с. URL: https://naqa.gov.ua/wp-content/uploads/2019/10/Рекомендації-ЗВО-система-забезпечення-академічної-доброчесності.pdf.

[43] Соколов В. Ю. Підходи до формування наукового мислення у здобувачів вищої освіти з кібербезпеки. *Кібербезпека: освіта, наука, техніка.* 2022. Т. 2, № 18. С. 124–137. DOI: 10.28925/2663-4023.2022.18.124137.

[44] Соколов В. Ю., Складанний П. М. Порівняльний аналіз стратегій побудови другого та третього рівня освітніх програм зі спеціальності 125 «Кібербезпека». *Кібербезпека: освіта, наука, техніка.* 2023. Т. 4, № 20. С. 183–204. DOI: 10.28925/2663-4023.2023.20.183204.

[45] Український правопис / схвалено Кабінетом Міністрів України (постанова від 22.05.2019 № 437). Київ: Наукова думка, 2019. 392 с. URL: https://mon.gov.ua/ua/osvita/zagalna-serednya-osvita/navchalni-programi/ukrayinskij-pravopis-2019.

[46] Фещак І., Онофрійчук С., Гарасимчук О. Технічні засоби захисту авторських прав, як інструменти боротьби з піратством електронних книг. *Кібербезпека: освіта, наука, техніка.* 2025. Т. 4, № 28. С. 435–451. DOI: 10.28925/2663-4023.2025.28.854.

[47] Цивільний кодекс України від 16.01.2003 № 435-IV. Глава 46. Право інтелектуальної власності на комерційну таємницю (ст. 505–508). URL: https://zakon.rada.gov.ua/laws/show/435-15.

[48] Шевченко С., Жданова Ю., Стороженко В., Рашевська В., Горбач В. Інтегроване оцінювання ризиків інформаційної безпеки на основі баєсівських мереж та аудиту зрілості. *Кібербезпека: освіта, наука, техніка.* 2026. Т. 4, № 32. С. 892–907. DOI: 10.28925/2663-4023.2026.32.1203.

[49] Щодо рекомендацій з академічної доброчесності для закладів вищої освіти: лист Міністерства освіти і науки України від 23.10.2018 № 1/9-650. URL: https://zakon.rada.gov.ua/rada/show/v-650729-18.

[50] Юдіна Д., Юдін Д. Методика розрахунку рівня кіберзахисту об'єктів критичної інформаційної інфраструктури. *Кібербезпека: освіта, наука, техніка*. 2025. Т. 3, № 31. С. 833–859. DOI: 10.28925/2663-4023.2025.31.1083.

[51] Abalkina A. Detecting a network of hijacked journals by its archive. *Scientometrics*. 2021. Vol. 126, no. 8. P. 7123–7148. DOI: 10.1007/s11192-021-04056-0.

[52] ACM. Artifact Review and Badging — Version 1.1. New York: Association for Computing Machinery, 2020. URL: https://www.acm.org/publications/policies/artifact-review-and-badging-current.

[53] Agreement between the European Union, of the one part, and Ukraine, of the other part, on the participation of Ukraine in the Union programme Horizon Europe — the Framework Programme for Research and Innovation and in the Research and Training Programme of the European Atomic Energy Community. *Official Journal of the European Union*. 2022. L 95. P. 1–11. Набрала чинності 09.06.2022, застосовується з 01.01.2021. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:22022A0323(01).

[54] ALLEA — All European Academies. The European Code of Conduct for Research Integrity. Revised Edition 2023. Berlin: ALLEA, 2023. 24 p. DOI: 10.26356/ECOC.

[55] Alonso J. A., Lamata M. T. Consistency in the analytic hierarchy process: a new approach. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. 2006. Vol. 14, no. 4. P. 445–459. DOI: 10.1142/S0218488506004114.

[56] Alvesson M., Sandberg J. Generating research questions through problematization. *Academy of Management Review*. 2011. Vol. 36, no. 2. P. 247–271. DOI: 10.5465/amr.2009.0188.

[57] Anscombe F. J. Graphs in statistical analysis. *The American Statistician*. 1973. Vol. 27, no. 1. P. 17–21. DOI: 10.1080/00031305.1973.10478966.

[58] ANSI/NISO Z39.104-2022. CRediT, Contributor Roles Taxonomy. Baltimore: NISO, 2022. DOI: 10.3789/ansi.niso.z39.104-2022.

[59] Aria M., Cuccurullo C. bibliometrix: an R-tool for comprehensive science mapping analysis. *Journal of Informetrics*. 2017. Vol. 11, no. 4. P. 959–975. DOI: 10.1016/j.joi.2017.08.007.

[60] Arp D., Quiring E., Pendlebury F., Warnecke A., Pierazzi F., Wressnegger C., Cavallaro L., Rieck K. Dos and don'ts of machine learning in computer security. *Proceedings of the 31st USENIX Security Symposium*. Boston: USENIX Association, 2022. P. 3971–3988. URL: https://www.usenix.org/conference/usenixsecurity22/presentation/arp.

[61] Arrow K. J. Social Choice and Individual Values. 2nd ed. New Haven: Yale University Press, 1963. 124 p. ISBN 978-0-300-01364-0.

[62] Article 29 Data Protection Working Party. Opinion 05/2014 on Anonymisation Techniques. WP 216. Brussels: European Commission, 2014. 37 p. URL: https:

//ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

[63] Baker M. 1,500 scientists lift the lid on reproducibility. *Nature*. 2016. Vol. 533, no. 7604. P. 452–454. DOI: 10.1038/533452a.

[64] Behzadian M., Khanmohammadi Otaghsara S., Yazdani M., Ignatius J. A state-of-the-art survey of TOPSIS applications. *Expert Systems with Applications*. 2012. Vol. 39, no. 17. P. 13051–13069. DOI: 10.1016/j.eswa.2012.05.056.

[65] Belton V., Gear T. On a short-coming of Saaty's method of analytic hierarchies. *Omega*. 1983. Vol. 11, no. 3. P. 228–230. DOI: 10.1016/0305-0483(83)90047-6.

[66] Benjamini Y., Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*. 1995. Vol. 57, no. 1. P. 289–300. DOI: 10.1111/j.2517-6161.1995.tb02031.x.

[67] Bertalanffy L. von. General System Theory: Foundations, Development, Applications. Rev. ed. New York: George Braziller, 1968. 295 p. ISBN 978-0-8076-0453-3.

[68] Blackburn S. M., Diwan A., Hauswirth M., Sweeney P. F., Amaral J. N., Brecht T. et al. The truth, the whole truth, and nothing but the truth: a pragmatic guide to assessing empirical evaluations. *ACM Transactions on Programming Languages and Systems*. 2016. Vol. 38, no. 4. Article 15. P. 1–20. DOI: 10.1145/2983574.

[69] Boettiger C. An introduction to Docker for reproducible research. *ACM SIGOPS Operating Systems Review*. 2015. Vol. 49, no. 1. P. 71–79. DOI: 10.1145/2723872.2723882.

[70] Booth W. C., Colomb G. G., Williams J. M., Bizup J., FitzGerald W. T. The Craft of Research. 4th ed. Chicago: University of Chicago Press, 2016. 336 p.

[71] Borisov N., Goldberg I., Wagner D. Intercepting mobile communications: the insecurity of 802.11. *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom '01)*. New York: ACM, 2001. P. 180–189. DOI: 10.1145/381677.381695.

[72] Bornmann L. Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*. 2014. Vol. 8, no. 4. P. 895–903. DOI: 10.1016/j.joi.2014.09.005.

[73] Bornmann L., Daniel H.-D. What do we know about the h index? *Journal of the American Society for Information Science and Technology*. 2007. Vol. 58, no. 9. P. 1381–1385. DOI: 10.1002/asi.20609.

[74] Box G. E. P. Science and statistics. *Journal of the American Statistical Association*. 1976. Vol. 71, no. 356. P. 791–799. DOI: 10.1080/01621459.1976.10480949.

[75] Bradford S. C. Sources of information on specific subjects. *Engineering*. 1934. Vol. 137. P. 85–86.

[76] Brand A., Allen L., Altman M., Hlava M., Scott J. Beyond authorship: attribution, contribution, collaboration, and credit. *Learned Publishing*. 2015. Vol. 28, no. 2. P. 151–155. DOI: 10.1087/20150211.

[77] Brown M. B., Forsythe A. B. Robust tests for the equality of variances. *Journal of the American Statistical Association*. 1974. Vol. 69, no. 346. P. 364–367. DOI: 10.1080/01621459.1974.10482955.

[78] Button K. S., Ioannidis J. P. A., Mokrysz C., Nosek B. A., Flint J., Robinson E. S. J., Munafò M. R. Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*. 2013. Vol. 14, no. 5. P. 365–376. DOI: 10.1038/nrn3475.

[79] Cabanac G., Labbé C., Magazinov A. Tortured phrases: a dubious writing style emerging in science. Evidence of critical issues affecting established journals. arXiv:2107.06751. 2021. DOI: 10.48550/arXiv.2107.06751.

[80] Cabells Predatory Reports: criteria for inclusion. Cabells. URL: https://cabells.com/solutions/predatory-reports/.

[81] Center for Open Science. Open Science Framework (OSF) and Preregistration. URL: https://www.cos.io/initiatives/prereg.

[82] Chambers C. D., Tzavella L. The past, present and future of Registered Reports. *Nature Human Behaviour*. 2022. Vol. 6, no. 1. P. 29–42. DOI: 10.1038/s41562-021-01193-7.

[83] Chavan V., Penev L. The data paper: a mechanism to incentivize data publishing in biodiversity science. *BMC Bioinformatics*. 2011. Vol. 12, suppl. 15. S2. DOI: 10.1186/1471-2105-12-S15-S2.

[84] Checkland P. Systems Thinking, Systems Practice. Chichester: John Wiley & Sons, 1981. 330 p. ISBN 978-0-471-27911-0.

[85] Chen C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*. 2006. Vol. 57, no. 3. P. 359–377. DOI: 10.1002/asi.20317.

[86] Chicco D., Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*. 2020. Vol. 21. Article 6. DOI: 10.1186/s12864-019-6413-7.

[87] Clarivate. Journal Citation Reports: Journal Impact Factor calculation and JCR methodology. URL: https://clarivate.com/academia-government/scientific-and-academic-research/research-funding-analytics/journal-citation-reports/.

[88] Coalition for Advancing Research Assessment. Agreement on Reforming Research Assessment. Brussels: European Commission, 2022. 24 p. URL: https://coara.eu/agreement/the-agreement-full-text/.

[89] cOAlition S. Plan S: Principles and Implementation. 2021. URL: https://www.coalition-s.org/addendum-to-the-coalition-s-guidance-on-the-implementation-of-plan-s/.

[90] Cockburn A., Dragicevic P., Besançon L., Gutwin C. Threats of a replication crisis in empirical computer science. *Communications of the ACM*. 2020. Vol. 63, no. 8. P. 70–79. DOI: 10.1145/3360311.

[91] Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*. 1960. Vol. 20, no. 1. P. 37–46. DOI: 10.1177/001316446002000104.

[92] Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale: Lawrence Erlbaum Associates, 1988. 567 p.

[93] Collberg C., Proebsting T. A. Repeatability in computer systems research. *Communications of the ACM*. 2016. Vol. 59, no. 3. P. 62–69. DOI: 10.1145/2812803.

[94] Committee on Publication Ethics. Authorship and AI tools: COPE position statement. 13 February 2023. URL: https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools.

[95] Committee on Publication Ethics. Core Practices. URL: https://publicationethics.org/core-practices.

[96] Committee on Publication Ethics. Flowcharts for handling cases of suspected misconduct. URL: https://publicationethics.org/guidance/flowcharts.

[97] Committee on Publication Ethics. Retraction Guidelines. Version 2. November 2019. DOI: 10.24318/cope.2019.1.4.

[98] CORE Rankings Portal: ranked list of computer science conferences (A*, A, B, C). Computing Research and Education Association of Australasia. URL: https://portal.core.edu.au/conf-ranks/.

[99] Council Recommendation of 18 December 2023 on a European framework to attract and retain research, innovation and entrepreneurial talents in Europe. *Official Journal of the European Union*. Series C. C/2023/1640. 29.12.2023. Annex II: European Charter for Researchers. URL: https://eur-lex.europa.eu/eli/C/2023/1640/oj.

[100] Creative Commons. About CC Licenses. URL: https://creativecommons.org/share-your-work/cclicenses/.

[101] Creswell J. W., Creswell J. D. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. 5th ed. Thousand Oaks: SAGE Publications, 2018. 304 p.

[102] Crossref. Retraction Watch: retraction data in the Crossref REST API. URL: https://www.crossref.org/documentation/retrieve-metadata/retraction-watch/.

[103] Crossref: metadata infrastructure and the REST API. URL: https://www.crossref.org/documentation/retrieve-metadata/rest-api/.

[104] Cumming G. The new statistics: why and how. *Psychological Science*. 2014. Vol. 25, no. 1. P. 7–29. DOI: 10.1177/0956797613504966.

[105] Dajani J. S., Sincoff M. Z., Talley W. K. Stability and agreement criteria for the termination of Delphi studies. *Technological Forecasting and Social Change*. 1979. Vol. 13, no. 1. P. 83–90. DOI: 10.1016/0040-1625(79)90007-6.

[106] Dalkey N., Helmer O. An experimental application of the Delphi method to the use of experts. *Management Science*. 1963. Vol. 9, no. 3. P. 458–467. DOI: 10.1287/mnsc.9.3.458.

[107] DataCite Metadata Working Group. DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs. Version 4.5. Hannover: DataCite, 2024. DOI: 10.14454/G8E5-6293.

[108] DeLong E. R., DeLong D. M., Clarke-Pearson D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988. Vol. 44, no. 3. P. 837–845. DOI: 10.2307/2531595.

[109] Demšar J. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*. 2006. Vol. 7. P. 1–30. URL: https://www.jmlr.org/papers/v7/demsar06a.html.

[110] Denning P. J. Is computer science science? *Communications of the ACM*. 2005. Vol. 48, no. 4. P. 27–31. DOI: 10.1145/1053291.1053309.

[111] Denning P. J., Comer D. E., Gries D., Mulder M. C., Tucker A., Turner A. J., Young P. R. Computing as a discipline. *Communications of the ACM*. 1989. Vol. 32, no. 1. P. 9–23. DOI: 10.1145/63238.63239.

[112] Di Cosmo R., Zacchiroli S. Software Heritage: why and how to preserve software source code. *Proceedings of the 14th International Conference on Digital Preservation (iPRES 2017)*. Kyoto, 2017. 10 p. URL: https://hal.science/hal-01590958.

[113] Di Tommaso P., Chatzou M., Floden E. W., Barja P. P., Palumbo E., Notredame C. Nextflow enables reproducible computational workflows. *Nature Biotechnology*. 2017. Vol. 35, no. 4. P. 316–319. DOI: 10.1038/nbt.3820.

[114] Diamond I. R., Grant R. C., Feldman B. M., Pencharz P. B., Ling S. C., Moore A. M., Wales P. W. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology*. 2014. Vol. 67, no. 4. P. 401–409. DOI: 10.1016/j.jclinepi.2013.12.002.

[115] Dietterich T. G. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*. 1998. Vol. 10, no. 7. P. 1895–1923. DOI: 10.1162/089976698300017197.

[116] Digital Curation Centre. DMPonline: tool for creating data management plans. URL: https://dmponline.dcc.ac.uk/.

[117] Directive (EU) 2019/1937 of the European Parliament and of the Council of 23 October 2019 on the protection of persons who report breaches of Union law. *Official Journal of the European Union*. 2019. L 305. P. 17–56. URL: https://eur-lex.europa.eu/eli/dir/2019/1937/oj.

[118] Directory of Open Access Journals (DOAJ): criteria for inclusion and the DOAJ Seal. URL: https://doaj.org/apply/guide/.

[119] Dittrich D., Kenneally E. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research. Washington, DC: U.S. Department of Homeland Security, 2012. 17 p. URL: https://www.dhs.gov/sites/default/files/publications/CSD-MenloPrinciplesCORE-20120803_1.pdf.

[120] Dittrich D., Kenneally E., Bailey M. Applying ethical principles to information and communication technology research: a companion to the Menlo Report.

Washington, DC: U.S. Department of Homeland Security, 2013. 27 p. DOI: 10.2139/ssrn.2342036.

[121] Douven I. Abduction. *The Stanford Encyclopedia of Philosophy* / ed. by E. N. Zalta, U. Nodelman. Stanford: Metaphysics Research Lab, Stanford University, 2025. URL: https://plato.stanford.edu/entries/abduction/.

[122] Druskat S., Spaaks J. H., Chue Hong N. et al. Citation File Format. Version 1.2.0. Zenodo, 2021. DOI: 10.5281/zenodo.1003149.

[123] Dryad: open data publishing platform and curation community. URL: https://datadryad.org/.

[124] Durumeric Z., Wustrow E., Halderman J. A. ZMap: fast Internet-wide scanning and its security applications. *Proceedings of the 22nd USENIX Security Symposium*. Washington, DC: USENIX Association, 2013. P. 605–620. URL: https://www.usenix.org/conference/usenixsecurity13/technical-sessions/paper/durumeric.

[125] Dybå T., Dingsøyr T. Empirical studies of agile software development: a systematic review. *Information and Software Technology*. 2008. Vol. 50, no. 9–10. P. 833–859. DOI: 10.1016/j.infsof.2008.01.006.

[126] Dyer J. S. Remarks on the analytic hierarchy process. *Management Science*. 1990. Vol. 36, no. 3. P. 249–258. DOI: 10.1287/mnsc.36.3.249.

[127] Easterbrook S., Singer J., Storey M.-A., Damian D. Selecting empirical methods for software engineering research. *Guide to Advanced Empirical Software Engineering* / ed. by F. Shull, J. Singer, D. I. K. Sjøberg. London: Springer, 2008. P. 285–311. DOI: 10.1007/978-1-84800-044-5_11.

[128] Efron B., Tibshirani R. J. An Introduction to the Bootstrap. New York: Chapman & Hall, 1993. 436 p. ISBN 978-0-412-04231-7.

[129] Egghe L. Theory and practise of the g-index. *Scientometrics*. 2006. Vol. 69, no. 1. P. 131–152. DOI: 10.1007/s11192-006-0144-7.

[130] Ehrgott M. Multicriteria Optimization. 2nd ed. Berlin; Heidelberg: Springer, 2005. 323 p. DOI: 10.1007/3-540-27659-9.

[131] El Emam K., Jonker E., Arbuckle L., Malin B. A systematic review of re-identification attacks on health data. *PLoS ONE*. 2011. Vol. 6, no. 12. e28071. DOI: 10.1371/journal.pone.0028071.

[132] Else H., Van Noorden R. The fight against fake-paper factories that churn out sham science. *Nature*. 2021. Vol. 591, no. 7851. P. 516–519. DOI: 10.1038/d41586-021-00733-5.

[133] Elsevier. Generative AI policies for journals. URL: https://www.elsevier.com/about/policies-and-standards/generative-ai-policies-for-journals.

[134] Elsevier. What is Field-Weighted Citation Impact (FWCI)? Scopus Support Center. URL: https://service.elsevier.com/app/answers/detail/a_id/14894/supporthub/scopus/.

[135] Engelen G., Rimmer V., Joosen W. Troubleshooting an intrusion detection dataset: the CICIDS2017 case study. *2021 IEEE Security and Privacy Workshops (SPW)*. San Francisco: IEEE, 2021. P. 7–12. DOI: 10.1109/SPW53761.2021.00009.

[136] EURAXESS — Researchers in Motion: European Commission portal for researcher mobility, vacancies and funding. URL: https://euraxess.ec.europa.eu/.

[137] European Commission, Directorate-General for Research and Innovation. Living Guidelines on the Responsible Use of Generative AI in Research. Brussels: European Commission, 2024. URL: https://research-and-innovation.ec.europa.eu/document/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en.

[138] European Commission. European Charter for Researchers: EURAXESS. URL: https://euraxess.ec.europa.eu/hrexcellenceaward/european-charter-researchers.

[139] European Commission. Horizon Europe Data Management Plan Template. Brussels, 2021. URL: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/report/data-management-plan_he_en.docx.

[140] European Commission. Horizon Europe Programme Guide. Version 5.0. Brussels: European Commission, 2025. URL: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf.

[141] European Commission. Horizon Europe Proposal Evaluation: Standard Briefing for Experts. Brussels: European Commission, 2024. URL: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/experts/standard-briefing-slides-for-experts_he_en.pdf.

[142] European Commission. Horizon Europe Work Programme 2023–2025. 13. General Annexes. Annex B: Technology readiness levels (TRL). Brussels: European Commission, 2024. URL: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/wp-call/2023-2024/wp-13-general-annexes-history-of-the-updates_horizon-2023-2024_en.pdf.

[143] European Commission. HR Excellence in Research Award (HRS4R): EURAXESS. URL: https://euraxess.ec.europa.eu/jobs/hrs4r.

[144] European Open Science Cloud (EOSC): federated infrastructure for research data and services. URL: https://open-science-cloud.ec.europa.eu/.

[145] European Research Council: official website. URL: https://erc.europa.eu/.

[146] European University Association. Salzburg II Recommendations: European Universities' Achievements since 2005 in Implementing the Salzburg Principles. Brussels: EUA, 2010. 8 p. URL: https://www.eua.eu/publications/positions/salzburg-ii-recommendations.html.

[147] Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS ONE*. 2009. Vol. 4, no. 5. e5738. DOI: 10.1371/journal.pone.0005738.

[148] Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters*. 2006. Vol. 27, no. 8. P. 861–874. DOI: 10.1016/j.patrec.2005.10.010.

[149] Figshare: repository for research outputs. URL: https://figshare.com/.

[150] Fire M., Guestrin C. Over-optimization of academic publishing metrics: observing Goodhart's Law in action. *GigaScience*. 2019. Vol. 8, no. 6. giz053. DOI: 10.1093/gigascience/giz053.

[151] Fluhrer S., Mantin I., Shamir A. Weaknesses in the key scheduling algorithm of RC4. *Selected Areas in Cryptography (SAC 2001)*. Lecture Notes in Computer Science. Vol. 2259. Berlin; Heidelberg: Springer, 2001. P. 1–24. DOI: 10.1007/3-540-45537-X_1.

[152] Foltýnek T., Dlabolová D., Anohina-Naumeca A. et al. Testing of support tools for plagiarism detection. *International Journal of Educational Technology in Higher Education*. 2020. Vol. 17. 46. DOI: 10.1186/s41239-020-00192-4.

[153] Fontugne R., Borgnat P., Abry P., Fukuda K. MAWILab: combining diverse anomaly detectors for automated anomaly labeling and performance benchmarking. *Proceedings of the 6th International Conference on Emerging Networking Experiments and Technologies (CoNEXT 2010)*. New York: ACM, 2010. Article 8. P. 1–12. DOI: 10.1145/1921168.1921179.

[154] Franco J., Aris A., Canberk B., Uluagac A. S. A survey of honeypots and honeynets for Internet of Things, industrial Internet of Things, and cyber-physical systems. *IEEE Communications Surveys & Tutorials*. 2021. Vol. 23, no. 4. P. 2351–2383. DOI: 10.1109/COMST.2021.3106669.

[155] Garfield E. Citation indexes for science: a new dimension in documentation through association of ideas. *Science*. 1955. Vol. 122, no. 3159. P. 108–111. DOI: 10.1126/science.122.3159.108.

[156] Garfield E. The history and meaning of the journal impact factor. *JAMA*. 2006. Vol. 295, no. 1. P. 90–93. DOI: 10.1001/jama.295.1.90.

[157] Garousi V., Felderer M., Mäntylä M. V. Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and Software Technology*. 2019. Vol. 106. P. 101–121. DOI: 10.1016/j.infsof.2018.09.006.

[158] Gelman A., Loken E. The statistical crisis in science. *American Scientist*. 2014. Vol. 102, no. 6. P. 460–465. DOI: 10.1511/2014.111.460.

[159] Grant M. J., Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*. 2009. Vol. 26, no. 2. P. 91–108. DOI: 10.1111/j.1471-1842.2009.00848.x.

[160] Greenland S., Senn S. J., Rothman K. J., Carlin J. B., Poole C., Goodman S. N., Altman D. G. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology*. 2016. Vol. 31, no. 4. P. 337–350. DOI: 10.1007/s10654-016-0149-3.

[161] Gregor S., Hevner A. R. Positioning and presenting design science research for maximum impact. *MIS Quarterly*. 2013. Vol. 37, no. 2. P. 337–355. DOI: 10.25300/MISQ/2013/37.2.01.

[162] Grudniewicz A., Moher D., Cobey K. D. et al. Predatory journals: no definition, no defence. *Nature*. 2019. Vol. 576, no. 7786. P. 210–212. DOI: 10.1038/d41586-019-03759-y.

[163] Guerrero-Bote V. P., Moya-Anegón F. A further step forward in measuring journals' scientific prestige: the SJR2 indicator. *Journal of Informetrics*. 2012. Vol. 6, no. 4. P. 674–688. DOI: 10.1016/j.joi.2012.07.001.

[164] Gundersen O. E., Kjensmo S. State of the art: reproducibility in artificial intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2018. Vol. 32, no. 1. P. 1644–1651. DOI: 10.1609/aaai.v32i1.11503.

[165] Gusenbauer M. Search where you will find most: comparing the disciplinary coverage of 56 bibliographic databases. *Scientometrics*. 2022. Vol. 127, no. 5. P. 2683–2745. DOI: 10.1007/s11192-022-04289-7.

[166] Gusenbauer M., Haddaway N. R. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*. 2020. Vol. 11, no. 2. P. 181–217. DOI: 10.1002/jrsm.1378.

[167] Hasson F., Keeney S., McKenna H. Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*. 2000. Vol. 32, no. 4. P. 1008–1015. DOI: 10.1046/j.1365-2648.2000.t01-1-01567.x.

[168] Hedges L. V. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*. 1981. Vol. 6, no. 2. P. 107–128. DOI: 10.3102/10769986006002107.

[169] Hevner A. R. A three cycle view of design science research. *Scandinavian Journal of Information Systems*. 2007. Vol. 19, no. 2. Article 4. P. 87–92. URL: https://aisel.aisnet.org/sjis/vol19/iss2/4/.

[170] Hevner A. R., March S. T., Park J., Ram S. Design science in information systems research. *MIS Quarterly*. 2004. Vol. 28, no. 1. P. 75–105. DOI: 10.2307/25148625.

[171] Hicks D., Wouters P., Waltman L., de Rijcke S., Ràfols I. Bibliometrics: The Leiden Manifesto for research metrics. *Nature*. 2015. Vol. 520, no. 7548. P. 429–431. DOI: 10.1038/520429a.

[172] Hindawi reveals process for retracting more than 8,000 paper mill articles. *Retraction Watch*. 19 December 2023. URL: https://retractionwatch.com/2023/12/19/hindawi-reveals-process-for-retracting-more-than-8000-paper-mill-articles/.

[173] Hirsch J. E. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*. 2005. Vol. 102, no. 46. P. 16569–16572. DOI: 10.1073/pnas.0507655102.

[174] Hoefler T., Belli R. Scientific benchmarking of parallel computing systems: twelve ways to tell the masses when reporting performance results. *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC 2015)*. New York: ACM, 2015. Article 73. P. 1–12. DOI: 10.1145/2807591.2807644.

[175] Holm S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*. 1979. Vol. 6, no. 2. P. 65–70. URL: https://www.jstor.org/stable/4615733.

[176] Householder A. D., Wassermann G., Manion A., King C. The CERT Guide to Coordinated Vulnerability Disclosure: Special Report CMU/SEI-2017-SR-022. Pittsburgh: Software Engineering Institute, Carnegie Mellon University, 2017. 121 p. URL: https://insights.sei.cmu.edu/library/the-cert-guide-to-coordinated-vulnerability-disclosure/.

[177] Hwang C.-L., Yoon K. Multiple Attribute Decision Making: Methods and Applications. Berlin; Heidelberg: Springer, 1981. 259 p. DOI: 10.1007/978-3-642-48318-9.

[178] IEEE Reference Guide. Piscataway: IEEE, 2024. URL: https://journals.ieeeauthorcenter.ieee.org/wp-content/uploads/sites/7/IEEE-Reference-Guide.pdf.

[179] IEEE Symposium on Security and Privacy 2021 Program Committee. Statement regarding the «Hypocrite Commits» paper. 2021. URL: https://www.ieee-security.org/TC/SP2021/downloads/2021_PC_Statement.pdf.

[180] IEEE. Submission and peer review policies: guidelines for artificial intelligence-generated text. URL: https://journals.ieeeauthorcenter.ieee.org/become-an-ieee-journal-author/publishing-ethics/guidelines-and-policies/submission-and-peer-review-policies/.

[181] International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. URL: https://www.icmje.org/recommendations/.

[182] Ioannidis J. P. A. Why most published research findings are false. *PLoS Medicine*. 2005. Vol. 2, no. 8. e124. DOI: 10.1371/journal.pmed.0020124.

[183] Ioannidis J. P. A., Baas J., Klavans R., Boyack K. W. A standardized citation metrics author database annotated for scientific field. *PLOS Biology*. 2019. Vol. 17, no. 8. e3000384. DOI: 10.1371/journal.pbio.3000384.

[184] Ioannidis J. P. A., Klavans R., Boyack K. W. Thousands of scientists publish a paper every five days. *Nature*. 2018. Vol. 561, no. 7722. P. 167–169. DOI: 10.1038/d41586-018-06185-8.

[185] ISO 16290:2013. Space systems — Definition of the Technology Readiness Levels (TRLs) and their criteria of assessment. Geneva: ISO, 2013. 12 p. URL: https://www.iso.org/standard/56064.html.

[186] ISO 21502:2020. Project, programme and portfolio management — Guidance on project management. Geneva: ISO, 2020. 60 p. URL: https://www.iso.org/standard/74947.html.

[187] ISO 26324:2025. Information and documentation — Digital object identifier system. Geneva: ISO, 2025. URL: https://www.iso.org/standard/85599.html.

[188] ISO 31000:2018. Risk management — Guidelines. Geneva: ISO, 2018. 16 p. URL: https://www.iso.org/standard/65694.html.

[189] ISO 690:2021. Information and documentation — Guidelines for information sources. Geneva: ISO, 2021. URL: https://www.iso.org/standard/72642.html.

[190] ISO/IEC 29147:2018. Information technology — Security techniques — Vulnerability disclosure. Geneva: ISO, 2018. 28 p. URL: https://www.iso.org/standard/72311.html.

[191] ISO/IEC 30111:2019. Information technology — Security techniques — Vulnerability handling processes. Geneva: ISO, 2019. 14 p. URL: https://www.iso.org/standard/69725.html.

[192] ISO/IEC/IEEE 15288:2023. Systems and software engineering — System life cycle processes. Geneva: ISO, 2023. 128 p. URL: https://www.iso.org/standard/81702.html.

[193] JabRef: free reference manager for BibTeX and BibLaTeX. URL: https://www.jabref.org/.

[194] Jacobsen A., de Miranda Azevedo R., Juty N. et al. FAIR principles: interpretations and implementation considerations. *Data Intelligence*. 2020. Vol. 2, no. 1–2. P. 10–29. DOI: 10.1162/dint_r_00024.

[195] James C., Colledge L., Meester W., Azoulay N., Plume A. CiteScore metrics: creating journal metrics from the Scopus citation index. *Learned Publishing*. 2019. Vol. 32, no. 4. P. 367–374. DOI: 10.1002/leap.1246.

[196] JCGM 100:2008. Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM 1995 with minor corrections). Sèvres: Joint Committee for Guides in Metrology, 2008. 120 p. URL: https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf.

[197] JCGM 200:2012. International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM). 3rd ed. Sèvres: Joint Committee for Guides in Metrology, 2012. 92 p. URL: https://www.bipm.org/documents/20126/2071204/JCGM_200_2012.pdf.

[198] John L. K., Loewenstein G., Prelec D. Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*. 2012. Vol. 23, no. 5. P. 524–532. DOI: 10.1177/0956797611430953.

[199] Kass R. E., Raftery A. E. Bayes factors. *Journal of the American Statistical Association*. 1995. Vol. 90, no. 430. P. 773–795. DOI: 10.1080/01621459.1995.10476572.

[200] Kaufman S., Rosset S., Perlich C., Stitelman O. Leakage in data mining: formulation, detection, and avoidance. *ACM Transactions on Knowledge Discovery from Data*. 2012. Vol. 6, no. 4. Article 15. P. 1–21. DOI: 10.1145/2382577.2382579.

[201] Ke Q., Ferrara E., Radicchi F., Flammini A. Defining and identifying sleeping beauties in science. *Proceedings of the National Academy of Sciences*. 2015. Vol. 112, no. 24. P. 7426–7431. DOI: 10.1073/pnas.1424329112.

[202] Kendall M. G., Babington Smith B. The problem of m rankings. *The Annals of Mathematical Statistics*. 1939. Vol. 10, no. 3. P. 275–287. DOI: 10.1214/aoms/1177732186.

[203] Kerr N. L. HARKing: hypothesizing after the results are known. *Personality and Social Psychology Review*. 1998. Vol. 2, no. 3. P. 196–217. DOI: 10.1207/s15327957pspr0203_4.

[204] Keshav S. How to read a paper. *ACM SIGCOMM Computer Communication Review*. 2007. Vol. 37, no. 3. P. 83–84. DOI: 10.1145/1273445.1273458.

[205] Kitchenham B., Brereton P. A systematic review of systematic review process research in software engineering. *Information and Software Technology*. 2013. Vol. 55, no. 12. P. 2049–2075. DOI: 10.1016/j.infsof.2013.07.010.

[206] Kitchenham B., Charters S. Guidelines for performing systematic literature reviews in software engineering: EBSE Technical Report EBSE-2007-01. Keele University; University of Durham, 2007. 65 p.

[207] Kluyver T., Ragan-Kelley B., Pérez F., Granger B., Bussonnier M., Frederic J. et al. Jupyter Notebooks — a publishing format for reproducible computational workflows. *Positioning and Power in Academic Publishing: Players, Agents and Agendas (ELPUB 2016)*. Amsterdam: IOS Press, 2016. P. 87–90. DOI: 10.3233/978-1-61499-649-1-87.

[208] Kramer A. D. I., Guillory J. E., Hancock J. T. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*. 2014. Vol. 111, no. 24. P. 8788–8790. DOI: 10.1073/pnas.1320040111.

[209] Kruskal W. H., Wallis W. A. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*. 1952. Vol. 47, no. 260. P. 583–621. DOI: 10.1080/01621459.1952.10483441.

[210] Kuhn T. S. The Structure of Scientific Revolutions. 4th ed. Chicago: University of Chicago Press, 2012. 264 p.

[211] Lakatos I. The Methodology of Scientific Research Programmes: Philosophical Papers. Vol. 1 / ed. by J. Worrall, G. Currie. Cambridge: Cambridge University Press, 1978. 250 p.

[212] Landis J. R., Koch G. G. The measurement of observer agreement for categorical data. *Biometrics*. 1977. Vol. 33, no. 1. P. 159–174. DOI: 10.2307/2529310.

[213] Lang T. A., Altman D. G. Basic statistical reporting for articles published in biomedical journals: the «Statistical Analyses and Methods in the Published Literature» or the SAMPL Guidelines. *International Journal of Nursing Studies*. 2015. Vol. 52, no. 1. P. 5–9. DOI: 10.1016/j.ijnurstu.2014.09.006.

[214] Larivière V., Kiermer V., MacCallum C. J. et al. A simple proposal for the publication of journal citation distributions. *bioRxiv*. 2016. 062109. DOI: 10.1101/062109.

[215] Linstone H. A., Turoff M. (eds.) The Delphi Method: Techniques and Applications. Reading: Addison-Wesley, 1975. 620 p. ISBN 978-0-201-04294-8.

[216] Lotka A. J. The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*. 1926. Vol. 16, no. 12. P. 317–323. URL: https://www.jstor.org/stable/24529203.

[217] Machanavajjhala A., Kifer D., Gehrke J., Venkitasubramaniam M. ℓ-diversity: privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data*. 2007. Vol. 1, no. 1. 3. DOI: 10.1145/1217299.1217302.

[218] Mann H. B., Whitney D. R. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*. 1947. Vol. 18, no. 1. P. 50–60. DOI: 10.1214/aoms/1177730491.

[219] March S. T., Smith G. F. Design and natural science research on information technology. *Decision Support Systems*. 1995. Vol. 15, no. 4. P. 251–266. DOI: 10.1016/0167-9236(94)00041-2.

[220] Marie Skłodowska-Curie Actions: official website of the European Commission. URL: https://marie-sklodowska-curie-actions.ec.europa.eu/.

[221] Martinson B. C., Anderson M. S., de Vries R. Scientists behaving badly. *Nature*. 2005. Vol. 435, no. 7043. P. 737–738. DOI: 10.1038/435737a.

[222] Martín-Martín A., Thelwall M., Orduna-Malea E., Delgado López-Cózar E. Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. *Scientometrics*. 2021. Vol. 126, no. 1. P. 871–906. DOI: 10.1007/s11192-020-03690-4.

[223] Matthews B. W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) — Protein Structure*. 1975. Vol. 405, no. 2. P. 442–451. DOI: 10.1016/0005-2795(75)90109-9.

[224] McCook A. One publisher, more than 7000 retractions. *Science*. 2018. Vol. 362, no. 6413. P. 393. DOI: 10.1126/science.362.6413.393.

[225] McKinney W. Data structures for statistical computing in Python. *Proceedings of the 9th Python in Science Conference (SciPy 2010)*. Austin, 2010. P. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

[226] McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*. 1947. Vol. 12, no. 2. P. 153–157. DOI: 10.1007/BF02295996.

[227] Meadows D. H. Thinking in Systems: A Primer / ed. by D. Wright. White River Junction: Chelsea Green Publishing, 2008. 240 p. ISBN 978-1-60358-055-7.

[228] Meidan Y., Bohadana M., Mathov Y., Mirsky Y., Shabtai A., Breitenbacher D., Elovici Y. N-BaIoT — network-based detection of IoT botnet attacks using deep autoencoders. *IEEE Pervasive Computing*. 2018. Vol. 17, no. 3. P. 12–22. DOI: 10.1109/MPRV.2018.03367731.

[229] Mensh B., Kording K. Ten simple rules for structuring papers. *PLOS Computational Biology*. 2017. Vol. 13, no. 9. e1005619. DOI: 10.1371/journal.pcbi.1005619.

[230] Merton R. K. The Sociology of Science: Theoretical and Empirical Investigations / ed. by N. W. Storer. Chicago: University of Chicago Press, 1973. 605 p.

[231] Mikriukov A., Senokosov A., Succi G., Tormasov A., Plaksin Y., Trofimova E., Sitnikov V. AI tools for automating systematic literature reviews. *Proceedings of the 2025 International Conference on Software Engineering and Computer Applications*. New York: ACM, 2025. P. 25–30. DOI: 10.1145/3747912.3747962.

[232] Miller G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*. 1956. Vol. 63, no. 2. P. 81–97. DOI: 10.1037/h0043158.

[233] Moed H. F. Measuring contextual citation impact of scientific journals. *Journal of Informetrics*. 2010. Vol. 4, no. 3. P. 265–277. DOI: 10.1016/j.joi.2010.01.002.

[234] Molléri J. S., Petersen K., Mendes E. Survey guidelines in software engineering: an annotated review. *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM 2016)*. New York: ACM, 2016. Article 58. P. 1–6. DOI: 10.1145/2961111.2962619.

[235] Mongeon P., Paul-Hus A. The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*. 2016. Vol. 106, no. 1. P. 213–228. DOI: 10.1007/s11192-015-1765-5.

[236] Munn Z., Peters M. D. J., Stern C., Tufanaru C., McArthur A., Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*. 2018. Vol. 18. Article 143. DOI: 10.1186/s12874-018-0611-x.

[237] Mölder F., Jablonski K. P., Letcher B. et al. Sustainable data analysis with Snakemake. *F1000Research*. 2021. Vol. 10. 33. DOI: 10.12688/f1000research.29032.2.

[238] Nadeau C., Bengio Y. Inference for the generalization error. *Machine Learning*. 2003. Vol. 52, no. 3. P. 239–281. DOI: 10.1023/A:1024068626366.

[239] Narayanan A., Shmatikov V. Robust de-anonymization of large sparse datasets. *Proceedings of the 2008 IEEE Symposium on Security and Privacy*. Oakland: IEEE, 2008. P. 111–125. DOI: 10.1109/SP.2008.33.

[240] National Academies of Sciences, Engineering, and Medicine. Reproducibility and Replicability in Science. Washington, DC: The National Academies Press, 2019. 256 p. DOI: 10.17226/25303.

[241] Nazarovets M. Unlocking the hidden realms: analysing the Ukrainian journal landscape with Ulrichsweb. *Learned Publishing*. 2024. Vol. 37, no. 3. e1605. DOI: 10.1002/leap.1605.

[242] Nazarovets S. Controversial practice of rewarding for publications in national journals. *Scientometrics*. 2020. Vol. 124, no. 1. P. 813–818. DOI: 10.1007/s11192-020-03485-7.

[243] Neto E. C. P., Dadkhah S., Ferreira R., Zohourian A., Lu R., Ghorbani A. A. CICIoT2023: a real-time dataset and benchmark for large-scale attacks in IoT environment. *Sensors*. 2023. Vol. 23, no. 13. 5941. DOI: 10.3390/s23135941.

[244] Newell A., Perlis A. J., Simon H. A. Computer science. *Science*. 1967. Vol. 157, no. 3795. P. 1373–1374. DOI: 10.1126/science.157.3795.1373.c.

[245] Newell A., Simon H. A. Computer science as empirical inquiry: symbols and search. *Communications of the ACM*. 1976. Vol. 19, no. 3. P. 113–126. DOI: 10.1145/360018.360022.

[246] NIST Special Publication 800-30 Revision 1. Guide for Conducting Risk Assessments. Gaithersburg: National Institute of Standards and Technology, 2012. 95 p. DOI: 10.6028/NIST.SP.800-30r1.

[247] Nosek B. A., Ebersole C. R., DeHaven A. C., Mellor D. T. The preregistration revolution. *Proceedings of the National Academy of Sciences*. 2018. Vol. 115, no. 11. P. 2600–2606. DOI: 10.1073/pnas.1708274114.

[248] Oberkampf W. L., Roy C. J. Verification and Validation in Scientific Computing. Cambridge: Cambridge University Press, 2010. 767 p. DOI: 10.1017/CBO9780511760396.

[249] OECD. Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development. Paris: OECD Publishing, 2015. 402 p. DOI: 10.1787/9789264239012-en.

[250] OECD. Revised Field of Science and Technology (FOS) Classification in the Frascati Manual: working paper DSTI/EAS/STP/NESTI(2006)19/FINAL. Paris: OECD, 2007. 12 p. URL: https://one.oecd.org/document/DSTI/EAS/STP/NESTI(2006)19/FINAL/en/pdf.

[251] Open Data Commons. Open Database License (ODbL) v1.0. URL: https://opendatacommons.org/licenses/odbl/1-0/.

[252] Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*. 2015. Vol. 349, no. 6251. aac4716. DOI: 10.1126/science.aac4716.

[253] Open Ukrainian Citation Index (OUCI): пошукова система та база цитувань. Державна науково-технічна бібліотека України. URL: https://ouci.dntb.gov.ua/.

[254] OpenAIRE. Argos: open service for data management plans. URL: https://argos.openaire.eu/.

[255] OpenAlex API: works, sources and authors endpoints. OurResearch. URL: https://docs.openalex.org/.

[256] ORCID: connecting research and researchers. ORCID, Inc. URL: https://orcid.org/.

[257] Page M. J., McKenzie J. E., Bossuyt P. M. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021. Vol. 372. n71. DOI: 10.1136/bmj.n71.

[258] Page M. J., Moher D., Bossuyt P. M. et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*. 2021. Vol. 372. n160. DOI: 10.1136/bmj.n160.

[259] Partridge C., Allman M. Ethical considerations in network measurement papers. *Communications of the ACM*. 2016. Vol. 59, no. 10. P. 58–64. DOI: 10.1145/2896816.

[260] Pearl J. Causality: Models, Reasoning, and Inference. 2nd ed. Cambridge: Cambridge University Press, 2009. 484 p. DOI: 10.1017/CBO9780511803161.

[261] Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O. et al. Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*. 2011. Vol. 12. P. 2825–2830. URL: https://www.jmlr.org/papers/v12/pedregosa11a.html.

[262] Peffers K., Tuunanen T., Rothenberger M. A., Chatterjee S. A design science research methodology for information systems research. *Journal of Management Information Systems*. 2007. Vol. 24, no. 3. P. 45–77. DOI: 10.2753/MIS0742-1222240302.

[263] Pendlebury F., Pierazzi F., Jordaney R., Kinder J., Cavallaro L. TESSERACT: eliminating experimental bias in malware classification across space and time. *Proceedings of the 28th USENIX Security Symposium*. Santa Clara: USENIX

Association, 2019. P. 729–746. URL: https://www.usenix.org/conference/usenix security19/presentation/pendlebury.

[264] Peng R. D. Reproducible research in computational science. *Science*. 2011. Vol. 334, no. 6060. P. 1226–1227. DOI: 10.1126/science.1213847.

[265] Peroni S., Shotton D. OpenCitations, an infrastructure organization for open scholarship. *Quantitative Science Studies*. 2020. Vol. 1, no. 1. P. 428–444. DOI: 10.1162/qss_a_00023.

[266] Petersen K., Vakkalanka S., Kuzniarz L. Guidelines for conducting systematic mapping studies in software engineering: an update. *Information and Software Technology*. 2015. Vol. 64. P. 1–18. DOI: 10.1016/j.infsof.2015.03.007.

[267] Petticrew M., Roberts H. Systematic Reviews in the Social Sciences: A Practical Guide. Oxford: Blackwell Publishing, 2006. 352 p. ISBN 978-1-4051-2110-1.

[268] Pimentel J. F., Murta L., Braganholo V., Freire J. A large-scale study about quality and reproducibility of Jupyter notebooks. *Proceedings of the 16th International Conference on Mining Software Repositories (MSR 2019)*. Montreal: IEEE, 2019. P. 507–517. DOI: 10.1109/MSR.2019.00077.

[269] Pineau J., Vincent-Lamarre P., Sinha K., Larivière V., Beygelzimer A., d'Alché-Buc F., Fox E., Larochelle H. Improving reproducibility in machine learning research (a report from the NeurIPS 2019 reproducibility program). *Journal of Machine Learning Research*. 2021. Vol. 22, no. 164. P. 1–20. URL: https://www.jmlr.org/papers/v22/20 -303.html.

[270] Popper K. The Logic of Scientific Discovery. London; New York: Routledge, 2002. 545 p.

[271] Price D. J. de Solla. Little Science, Big Science. New York: Columbia University Press, 1963. 119 p.

[272] Priem J., Piwowar H., Orr R. OpenAlex: a fully-open index of scholarly works, authors, venues, institutions, and concepts. 2022. arXiv:2205.01833. URL: https: //arxiv.org/abs/2205.01833.

[273] Priem J., Taraborelli D., Groth P., Neylon C. Altmetrics: a manifesto. 2010. URL: http: //altmetrics.org/manifesto/.

[274] Project Jupyter, Bussonnier M., Forde J. et al. Binder 2.0 — reproducible, interactive, sharable environments for science at scale. *Proceedings of the 17th Python in Science Conference*. 2018. P. 113–120. DOI: 10.25080/Majora-4af1f417-011.

[275] Publication Manual of the American Psychological Association. 7th ed. Washington, DC: American Psychological Association, 2020. 428 p.

[276] R Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing, 2025. URL: https://www.R-project.org/.

[277] Radicchi F., Fortunato S., Castellano C. Universality of citation distributions: toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences*. 2008. Vol. 105, no. 45. P. 17268–17272. DOI: 10.1073/pnas.0806977105.

[278] Ralph P., Baltes S., Bianculli D., Dittrich Y., Ernst M., Felderer M. et al. ACM SIGSOFT Empirical Standards for Software Engineering Research. 2021. arXiv:2010.03525. URL: https://arxiv.org/abs/2010.03525.

[279] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union*. 2016. L 119. P. 1–88. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[280] Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe — the Framework Programme for Research and Innovation. *Official Journal of the European Union*. 2021. L 170. P. 1–68. URL: https://eur-lex.europa.eu/eli/reg/2021/695/oj.

[281] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*. 2024. L series, 12.07.2024. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

[282] Research Organization Registry (ROR): persistent identifiers for research organizations. URL: https://ror.org/.

[283] Resnik D. B., Hosseini M. Hallucinated citations produced by generative artificial intelligence may constitute research misconduct when citations function as data in scholarly papers. *Accountability in Research*. 2026. 2645390. DOI: 10.1080/08989621.2026.2645390.

[284] Rossow C., Dietrich C. J., Grier C., Kreibich C., Paxson V., Pohlmann N., Bos H., van Steen M. Prudent practices for designing malware experiments: status quo and outlook. *2012 IEEE Symposium on Security and Privacy*. San Francisco: IEEE, 2012. P. 65–79. DOI: 10.1109/SP.2012.14.

[285] Runeson P., Höst M. Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*. 2009. Vol. 14, no. 2. P. 131–164. DOI: 10.1007/s10664-008-9102-8.

[286] Runeson P., Höst M., Rainer A., Regnell B. Case Study Research in Software Engineering: Guidelines and Examples. Hoboken: John Wiley & Sons, 2012. 237 p. DOI: 10.1002/9781118181034.

[287] Saaty T. L. A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*. 1977. Vol. 15, no. 3. P. 234–281. DOI: 10.1016/0022-2496(77)90033-5.

[288] Saaty T. L. Decision making with the analytic hierarchy process. *International Journal of Services Sciences*. 2008. Vol. 1, no. 1. P. 83–98. DOI: 10.1504/IJSSCI.2008.017590.

[289] Saaty T. L. The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. New York: McGraw-Hill, 1980. 287 p. ISBN 978-0-07-054371-3.

[290] Saito T., Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*. 2015. Vol. 10, no. 3. e0118432. DOI: 10.1371/journal.pone.0118432.

[291] San Francisco Declaration on Research Assessment (DORA). 2013. URL: https://sfdora.org/read/.

[292] Sandve G. K., Nekrutenko A., Taylor J., Hovig E. Ten simple rules for reproducible computational research. *PLoS Computational Biology*. 2013. Vol. 9, no. 10. e1003285. DOI: 10.1371/journal.pcbi.1003285.

[293] Sargent R. G. Verification and validation of simulation models. *Journal of Simulation*. 2013. Vol. 7, no. 1. P. 12–24. DOI: 10.1057/jos.2012.20.

[294] Saving Ukrainian Cultural Heritage Online (SUCHO): volunteer initiative for web archiving of Ukrainian digital heritage. URL: https://www.sucho.org/.

[295] Science Europe. Practical Guide to the International Alignment of Research Data Management — Extended Edition. Brussels: Science Europe, 2021. 52 p. DOI: 10.5281/zenodo.4915862.

[296] SCImago Journal & Country Rank: SJR indicator and journal quartiles. Scimago Lab. URL: https://www.scimagojr.com/.

[297] Seabold S., Perktold J. Statsmodels: econometric and statistical modeling with Python. *Proceedings of the 9th Python in Science Conference (SciPy 2010)*. Austin, 2010. P. 92–96. DOI: 10.25080/Majora-92bf1922-011.

[298] Seglen P. O. Why the impact factor of journals should not be used for evaluating research. *BMJ*. 1997. Vol. 314, no. 7079. P. 498–502. DOI: 10.1136/bmj.314.7079.497.

[299] Shadish W. R., Cook T. D., Campbell D. T. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin, 2002. 623 p. ISBN 978-0-395-61556-0.

[300] Shapiro S. S., Wilk M. B. An analysis of variance test for normality (complete samples). *Biometrika*. 1965. Vol. 52, no. 3–4. P. 591–611. DOI: 10.1093/biomet/52.3-4.591.

[301] Sharafaldin I., Habibi Lashkari A., Ghorbani A. A. Toward generating a new intrusion detection dataset and intrusion traffic characterization. *Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP 2018)*. Funchal: SciTePress, 2018. P. 108–116. DOI: 10.5220/0006639801080116.

[302] Sharp H., Dittrich Y., de Souza C. R. B. The role of ethnographic studies in empirical software engineering. *IEEE Transactions on Software Engineering*. 2016. Vol. 42, no. 8. P. 786–804. DOI: 10.1109/TSE.2016.2519887.

[303] Shaw M. Writing good software engineering research papers. *Proceedings of the 25th International Conference on Software Engineering (ICSE 2003)*. Portland: IEEE, 2003. P. 726–736. DOI: 10.1109/ICSE.2003.1201262.

[304] Simmons J. P., Nelson L. D., Simonsohn U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as

significant. *Psychological Science*. 2011. Vol. 22, no. 11. P. 1359–1366. DOI: 10.1177/0956797611417632.

[305] Simpson E. H. The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society: Series B*. 1951. Vol. 13, no. 2. P. 238–241. DOI: 10.1111/j.2517-6161.1951.tb00088.x.

[306] Sjøberg D. I. K., Hannay J. E., Hansen O., Kampenes V. B., Karahasanović A., Liborg N.-K., Rekdal A. C. A survey of controlled experiments in software engineering. *IEEE Transactions on Software Engineering*. 2005. Vol. 31, no. 9. P. 733–753. DOI: 10.1109/TSE.2005.97.

[307] Smith A. M., Katz D. S., Niemeyer K. E. Software citation principles. *PeerJ Computer Science*. 2016. Vol. 2. e86. DOI: 10.7717/peerj-cs.86.

[308] Smith A. M., Niemeyer K. E., Katz D. S., Barba L. A., Githinji G., Gymrek M. et al. Journal of Open Source Software (JOSS): design and first-year review. *PeerJ Computer Science*. 2018. Vol. 4. e147. DOI: 10.7717/peerj-cs.147.

[309] Sollaci L. B., Pereira M. G. The introduction, methods, results, and discussion (IMRAD) structure: a fifty-year survey. *Journal of the Medical Library Association*. 2004. Vol. 92, no. 3. P. 364–371. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC442179/.

[310] Sommer R., Paxson V. Outside the closed world: on using machine learning for network intrusion detection. *Proceedings of the 2010 IEEE Symposium on Security and Privacy*. Oakland: IEEE, 2010. P. 305–316. DOI: 10.1109/SP.2010.25.

[311] Springer Nature. Artificial intelligence (AI) policy for authors. URL: https://www.springernature.com/gp/authors/book-authors-guidelines/ai-policy.

[312] Stevens S. S. On the theory of scales of measurement. *Science*. 1946. Vol. 103, no. 2684. P. 677–680. DOI: 10.1126/science.103.2684.677.

[313] Stol K.-J., Fitzgerald B. The ABC of software engineering research. *ACM Transactions on Software Engineering and Methodology*. 2018. Vol. 27, no. 3. Article 11. P. 1–51. DOI: 10.1145/3241743.

[314] StrikePlagiarism.com: сервіс перевірки на плагіат у складі Національного репозитарію академічних текстів / УкрІНТЕІ. URL: https://nrat.ukrintei.ua/strikeplagiarism-servis-perevirky-na-plagiat/.

[315] Strom B. E., Applebaum A., Miller D. P., Nickels K. C., Pennington A. G., Thomas C. B. MITRE ATT&CK: Design and Philosophy. Rev. March 2020. McLean: The MITRE Corporation, 2020. 37 p. URL: https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf.

[316] Susman G. I., Evered R. D. An assessment of the scientific merits of action research. *Administrative Science Quarterly*. 1978. Vol. 23, no. 4. P. 582–603. DOI: 10.2307/2392581.

[317] Swales J. M. Genre Analysis: English in Academic and Research Settings. Cambridge: Cambridge University Press, 1990. 260 p. ISBN 978-0-521-33813-4.

[318] Sweeney L. *k*-anonymity: a model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. 2002. Vol. 10, no. 5. P. 557–570. DOI: 10.1142/S0218488502001648.

[319] Tahamtan I., Safipour Afshar A., Ahamdzadeh K. Factors affecting number of citations: a comprehensive review of the literature. *Scientometrics*. 2016. Vol. 107, no. 3. P. 1195–1225. DOI: 10.1007/s11192-016-1889-2.

[320] Think. Check. Submit.: a checklist for choosing trusted journals and conferences. URL: https://thinkchecksubmit.org/.

[321] Thorp H. H. ChatGPT is fun, but not an author. *Science*. 2023. Vol. 379, no. 6630. P. 313. DOI: 10.1126/science.adg7879.

[322] Tichy W. F. Should computer scientists experiment more? *Computer*. 1998. Vol. 31, no. 5. P. 32–40. DOI: 10.1109/2.675631.

[323] Tricco A. C., Lillie E., Zarin W. et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of Internal Medicine*. 2018. Vol. 169, no. 7. P. 467–473. DOI: 10.7326/M18-0850.

[324] Tufte E. R. The Visual Display of Quantitative Information. 2nd ed. Cheshire: Graphics Press, 2001. 197 p. ISBN 978-0-9613921-4-7.

[325] Tukey J. W. Exploratory Data Analysis. Reading: Addison-Wesley, 1977. 688 p. ISBN 978-0-201-07616-5.

[326] Turnitin. Unicheck end of life: офіційне повідомлення про припинення роботи сервісу Unicheck з 1 січня 2025 року. URL: https://www.turnitin.com/products/unicheck. Дата звернення: 22.09.2026.

[327] UNESCO Recommendation on Open Science. Paris: UNESCO, 2021. 36 p. DOI: 10.54677/MNMH8546.

[328] Unicheck: сервіс перевірки академічних текстів на збіги; умови безкоштовного доступу для закладів вищої освіти України. URL: https://mon.gov.ua/news/unicheck-timchasovo-zaprovadiv-bezkoshtovnu-perevirku-na-plagiat-dlya-ukrainskikh-zakladiv-vishchoi-osviti.

[329] van Eck N. J., Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*. 2010. Vol. 84, no. 2. P. 523–538. DOI: 10.1007/s11192-009-0146-3.

[330] van Raan A. F. J. Sleeping beauties in science. *Scientometrics*. 2004. Vol. 59, no. 3. P. 467–472. DOI: 10.1023/B:SCIE.0000018543.82441.f1.

[331] Vanhoef M., Piessens F. Key reinstallation attacks: forcing nonce reuse in WPA2. *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS '17)*. New York: ACM, 2017. P. 1313–1328. DOI: 10.1145/3133956.3134027.

[332] Vanhoef M., Ronen E. Dragonblood: analyzing the Dragonfly handshake of WPA3 and EAP-pwd. *2020 IEEE Symposium on Security and Privacy*. San Francisco: IEEE, 2020. P. 517–533. DOI: 10.1109/SP40000.2020.00031.

[333] Venable J., Pries-Heje J., Baskerville R. FEDS: a framework for evaluation in design science research. *European Journal of Information Systems*. 2016. Vol. 25, no. 1. P. 77–89. DOI: 10.1057/ejis.2014.36.

[334] Virtanen P., Gommers R., Oliphant T. E., Haberland M., Reddy T., Cournapeau D. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*. 2020. Vol. 17, no. 3. P. 261–272. DOI: 10.1038/s41592-019-0686-2.

[335] Visser M., van Eck N. J., Waltman L. Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quantitative Science Studies*. 2021. Vol. 2, no. 1. P. 20–41. DOI: 10.1162/qss_a_00112.

[336] Wagenmakers E.-J., Marsman M., Jamil T., Ly A., Verhagen J., Love J. et al. Bayesian inference for psychology. Part I: theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*. 2018. Vol. 25, no. 1. P. 35–57. DOI: 10.3758/s13423-017-1343-2.

[337] Waltman L. A review of the literature on citation impact indicators. *Journal of Informetrics*. 2016. Vol. 10, no. 2. P. 365–391. DOI: 10.1016/j.joi.2016.02.007.

[338] Waltman L., van Eck N. J. Field-normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics*. 2015. Vol. 9, no. 4. P. 872–894. DOI: 10.1016/j.joi.2015.08.001.

[339] Wasserstein R. L., Lazar N. A. The ASA statement on p-values: context, process, and purpose. *The American Statistician*. 2016. Vol. 70, no. 2. P. 129–133. DOI: 10.1080/00031305.2016.1154108.

[340] Wasserstein R. L., Schirm A. L., Lazar N. A. Moving to a world beyond «$p < 0.05$». *The American Statistician*. 2019. Vol. 73, sup. 1. P. 1–19. DOI: 10.1080/00031305.2019.1583913.

[341] Webster J., Watson R. T. Analyzing the past to prepare for the future: writing a literature review. *MIS Quarterly*. 2002. Vol. 26, no. 2. P. xiii–xxiii. URL: https://www.jstor.org/stable/4132319.

[342] Welch B. L. The generalization of «Student's» problem when several different population variances are involved. *Biometrika*. 1947. Vol. 34, no. 1–2. P. 28–35. DOI: 10.1093/biomet/34.1-2.28.

[343] Wieringa R. J. Design Science Methodology for Information Systems and Software Engineering. Berlin; Heidelberg: Springer, 2014. 332 p. DOI: 10.1007/978-3-662-43839-8.

[344] Wieringa R., Maiden N., Mead N., Rolland C. Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. *Requirements Engineering*. 2006. Vol. 11, no. 1. P. 102–107. DOI: 10.1007/s00766-005-0021-6.

[345] Wilcoxon F. Individual comparisons by ranking methods. *Biometrics Bulletin*. 1945. Vol. 1, no. 6. P. 80–83. DOI: 10.2307/3001968.

[346] Wilkinson M. D., Dumontier M., Aalbersberg I. J. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*. 2016. Vol. 3. 160018. DOI: 10.1038/sdata.2016.18.

[347] Wilsdon J., Allen L., Belfiore E. et al. The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. Bristol: HEFCE, 2015. 176 p. DOI: 10.13140/RG.2.1.4929.1363.

[348] Wohlin C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE 2014)*. New York: ACM, 2014. Article 38. P. 1–10. DOI: 10.1145/2601248.2601268.

[349] Wohlin C., Kalinowski M., Romero Felizardo K., Mendes E. Successful combination of database search and snowballing for identification of primary studies in systematic literature studies. *Information and Software Technology*. 2022. Vol. 147. 106908. DOI: 10.1016/j.infsof.2022.106908.

[350] Wohlin C., Runeson P., Höst M., Ohlsson M. C., Regnell B., Wesslén A. Experimentation in Software Engineering. Berlin; Heidelberg: Springer, 2012. 236 p. DOI: 10.1007/978-3-642-29044-2.

[351] World Medical Association. WMA Declaration of Helsinki — Ethical Principles for Medical Research Involving Human Subjects. Ferney-Voltaire: WMA, 2024. URL: ht tps://www.wma.net/policies-post/wma-declaration-of-helsinki/.

[352] Wu Q., Lu K. Withdrawal letter: «On the feasibility of stealthily introducing vulnerabilities in open-source software via hypocrite commits». University of Minnesota, 26 April 2021. URL: https://www-users.cse.umn.edu/~kjlu/papers /withdrawal-letter.pdf.

[353] Zenodo: general-purpose open repository operated by CERN. URL: https://zenodo .org/.

[354] Zhang H., Babar M. A., Tell P. Identifying relevant studies in software engineering. *Information and Software Technology*. 2011. Vol. 53, no. 6. P. 625–637. DOI: 10.1016/j.infsof.2010.12.010.

[355] Zipf G. K. Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology. Cambridge, MA: Addison-Wesley Press, 1949. 573 p.

[356] Zotero: a free, easy-to-use tool to help you collect, organize, cite, and share research. Corporation for Digital Scholarship. URL: https://www.zotero.org/.

Додаток А. Контрольні списки дослідника

Списки призначено для самоперевірки перед ключовими рішеннями. Пункт, на який немає ствердної відповіді з доказом, а не з наміром, вважається невиконаним.

А.1. Перед затвердженням теми

1. Сформульовано протиріччя, а не просто галузь інтересу: указано, що саме відоме і чого бракує.
2. Знайдено щонайменше п'ять робіт за останні три роки, у яких прогалина визнана авторами явно (розділ «future work»).
3. Об'єкт і предмет розмежовано: предмет є властивістю або аспектом об'єкта, а не його синонімом і не назвою технології.
4. Мета одна, вона досяжна й перевірна; завдання в сумі забезпечують мету й не виходять за її межі.
5. Кожне завдання має вимірний результат, за яким можна сказати «виконано» чи «не виконано».
6. Є доступ до даних або спосіб їх отримати законно; правові підстави оброблення з'ясовано.
7. Оцінено обчислювальні ресурси; для найважчого експерименту відомий порядок потрібного часу.
8. У закладі є науковий керівник за тематикою та потенційні рецензенти з профільними публікаціями.
9. Тема лишається актуальною через чотири роки, а не втрачає сенс із виходом наступної версії технології.
10. Складено реєстр ризиків: для кожної позиції обчислено $R = p \cdot I$ і визначено рівень за шкалою розділу 3; для всіх позицій рівня «високий» ($R \geq 12$) передбачено захід пом'якшення з виконавцем і терміном.

А.2. Протокол систематичного огляду

1. Дослідницькі питання огляду сформульовано *до* пошуку.
2. Рядок пошуку записано повністю, окремо для кожної бази, із зазначенням полів і фільтрів.
3. Чутливість рядка перевірено на контрольному наборі відомих релевантних робіт.

4. Критерії включення та виключення зафіксовано письмово, кожен має код.
5. Скринінг двоетапний, і на *обох* етапах усі записи перевіряють два незалежні скринери.
6. Окремо до основного скринінгу проведено *пілотну* перевірку узгодженості: на частці записів (5–10 %) обидва скринери працюють незалежно, узгодженість оцінено коефіцієнтом κ, і правило продовження за його значенням застосовано згідно з розділом 6.
7. Процедуру вирішення розбіжностей визначено наперед.
8. Застосовано зворотний і прямий снігові коми.
9. Якість первинних досліджень оцінено за явною анкетою.
10. Форму вилучення даних випробувано на п'яти роботах до основного етапу.
11. Потокову діаграму відбору заповнено числами на кожному кроці; причини виключення на етапі повного тексту наведено за кодами.

А.3. Перед статистичним аналізом і після нього

1. Шкалу кожної змінної визначено; застосовані статистики допустимі для цієї шкали.
2. Обсяг вибірки обґрунтовано розрахунком потужності, а не доступністю даних.
3. Гіпотези й план аналізу зафіксовано до перегляду результатів.
4. Передумови обраного критерію перевірено, перевірку описано.
5. Викиди не видалено мовчки: правило їх опрацювання встановлено наперед.
6. Поряд із p-значенням наведено розмір ефекту й довірчий інтервал.
7. Множинність порівнянь враховано; указано, який показник контролюється — сімейна помилка чи частота хибних відкриттів.
8. Для моделей класифікації наведено метрики, стійкі до незбалансованості, а не лише частку правильних відповідей.
9. Невизначеність метрик оцінено; порівняння моделей супроводжено відповідним критерієм.
10. Формулювання висновку відповідає отриманим числам: відсутність значущої різниці не є доведенням однаковості.

А.4. План керування даними та відтворюваність

1. Описано, які дані створюються або збираються, у якому форматі та якого обсягу.

2. Визначено правову підставу оброблення персональних або чутливих даних; передбачено мінімізацію й знеособлення.
3. Обрано репозитарій; з'ясовано, чи присвоює він постійний ідентифікатор.
4. Метадані відповідають схемі репозитарію; наведено ліцензію.
5. Репозитарій артефакту містить `README`, `LICENSE`, `CITATION.cff` і опис середовища.
6. Залежності зафіксовано з точними версіями; випадкові зерна зафіксовано або недетермінізм описано.
7. Наведено команду, що відтворює головний результат від початку до кінця.
8. Перевірено, що артефакт розгортається на чистій системі сторонньою особою.
9. Передбачено резервування й довготривале зберігання.
10. Указано, як цитувати датасет і програмний код окремо від статті.

А.5. Перед поданням тексту

1. Усі запозичення позначено; перефразування супроводжено посиланням, дослівні фрагменти — лапками.
2. Вторинне цитування позначено явно або замінено зверненням до першоджерела.
3. Кожне джерело перевірено за ідентифікатором: воно існує, автори й рік збігаються, зміст відповідає твердженню, заради якого його наведено.
4. Перевірено, чи не відкликано цитовані роботи.
5. Склад авторів відповідає критеріям внеску; внески описано.
6. Конфлікт інтересів і джерела фінансування задекларовано.
7. Використання генеративного штучного інтелекту розкрито у формі, якої вимагає видання; ШІ не зазначено співавтором.
8. Результат не публікувався раніше; подання до кількох видань одночасно виключено.
9. Обмеження дослідження й загрози валідності описано явно.
10. Дані та код доступні або пояснено, чому доступ обмежено.

Додаток Б. Шаблони формулювань

Шаблони подано як каркаси, у яких кутові дужки позначають місця, що їх заповнює здобувач. Вони не звільняють від змістовної роботи, а лише запобігають типовим композиційним помилкам.

Б.1. Актуальність

Суспільний рівень. Зростання ⟨явище, підтверджене джерелом⟩ зумовлює потребу ⟨потреба практики⟩.

Галузевий рівень. Наявні рішення ⟨клас рішень⟩ забезпечують ⟨що саме⟩, проте не враховують ⟨обмеження⟩, що в умовах ⟨умови⟩ призводить до ⟨наслідок⟩.

Науковий рівень. Відомі підходи ⟨посилання⟩ розв'язують задачу для ⟨окремий випадок⟩; питання ⟨невирішене питання⟩ лишається відкритим, що й визначає актуальність дослідження.

Б.2. Об'єкт, предмет, мета, завдання

Об'єкт дослідження — процес ⟨назва процесу⟩ у ⟨клас систем⟩.

Предмет дослідження — ⟨моделі, методи, засоби⟩ ⟨чого саме⟩, що забезпечують ⟨властивість⟩ за умов ⟨умови⟩.

Мета — підвищення ⟨вимірна характеристика⟩ ⟨об'єкт⟩ шляхом розроблення ⟨артефакт⟩.

Завдання: 1) проаналізувати ⟨…⟩ і визначити ⟨…⟩; 2) побудувати ⟨модель⟩; 3) розробити ⟨метод⟩; 4) реалізувати ⟨засіб⟩; 5) експериментально оцінити ⟨що саме й за якими показниками⟩; 6) упровадити результати.

Б.3. Гіпотеза

Якщо ⟨втручання⟩, то ⟨вимірний показник⟩ ⟨зросте або зменшиться⟩ щонайменше на ⟨величина⟩ порівняно з ⟨базовий підхід⟩ за умов ⟨умови⟩.

Умова спростування: гіпотезу буде відхилено, якщо ⟨конкретний спостережуваний результат⟩.

Б.4. Наукова новизна

Уперше ⟨дія: запропоновано, встановлено, обґрунтовано⟩ ⟨що саме⟩, яке, на відміну від ⟨відомі підходи⟩, ⟨відмітна ознака⟩, що дало змогу ⟨вимірний ефект⟩.

> *Удосконалено* ⟨що саме⟩ за рахунок ⟨зміна⟩, що, на відміну від відомого, забезпечує ⟨ефект із числом⟩.
>
> *Набули подальшого розвитку* ⟨що саме⟩ у частині ⟨уточнення⟩, що дає змогу ⟨наслідок⟩.

Формулювання без відповіді на питання «на відміну від чого» і «з яким вимірним ефектом» новизни не описує.

Б.5. Обмеження та загрози валідності

> Результати отримано за умов ⟨умови⟩ на даних ⟨джерело й обсяг⟩. Перенесення висновків на ⟨інші умови⟩ потребує окремої перевірки, оскільки ⟨причина⟩. Внутрішню валідність обмежує ⟨чинник⟩, вплив якого зменшено ⟨захід⟩.

Додаток В. Приклади бібліографічних описів за ДСТУ 8302:2015

Нижче наведено зразки опису типових для інформаційних технологій видів документів. Пунктуація відповідає ДСТУ 8302:2015 [7]; прізвища подаються перед ініціалами, між ініціалами ставиться нерозривний пробіл, цифровий ідентифікатор наводиться наприкінці. Структуру самої роботи оформлюють за ДСТУ 3008:2015 [8], а мовні норми — за чинним «Українським правописом» [45]. Якщо видання вимагає іншого стилю, застосовують ISO 690 [189], APA 7 [275] або IEEE Reference Guide [178] — але *один* стиль на весь список.

Щодо кількості авторів. ДСТУ 8302:2015 допускає і скорочення «та ін.» після трьох перших прізвищ, і повний перелік. У списку джерел цього посібника послідовно застосовано *повний перелік*, бо він потрібен для однозначної ідентифікації праці та для звірки з метаданими DOI. Зразок зі «та ін.» наведено як допустиму альтернативу: обрану форму застосовують до всього списку, не змішуючи їх.

Стаття в журналі (один–три автори)
Прізвище І. Б., Прізвище І. Б. Назва статті. *Назва журналу*. 2025. Т. 12, № 3. С. 45–58. DOI: 10.xxxx/yyyyy.

Стаття в журналі (чотири й більше авторів)
Прізвище І. Б., Прізвище І. Б., Прізвище І. Б. та ін. Назва статті. *Назва журналу*. 2024. Vol. 8, no. 2. P. 101–117. DOI: 10.xxxx/yyyyy.

Матеріали конференції
Прізвище І. Б. Назва доповіді. *Назва конференції: матеріали міжнар. наук.-практ. конф., м. Київ, 12–14 травня 2025 р.* Київ: Видавництво, 2025. С. 77–80.

Праці конференції англійською
Surname N. Title of the paper. *Proceedings of the 32nd ACM Conference on Computer and Communications Security*. New York: ACM, 2025. P. 1210–1225. DOI: 10.1145/xxxxxxx.xxxxxxx.

Книга одного автора
Прізвище І. Б. Назва книги: підручник. 2-ге вид., перероб. і допов. Київ: Видавництво, 2024. 352 с.

Книга за редакцією
Назва книги / за ред. І. Б. Прізвище. Київ: Видавництво, 2023. 410 с.

Розділ у книзі
Прізвище І. Б. Назва розділу. *Назва книги* / за ред. І. Б. Прізвище. Berlin: Springer, 2022. С. 130–158. DOI: 10.1007/xxx-x-xxx-xxxxx-x_7.

Дисертація
Прізвище І. Б. Назва дисертації: дис. … д-ра філософії: 125 / Київський столичний університет імені Бориса Грінченка. Київ, 2025. 198 с.

Автореферат
Прізвище І. Б. Назва: автореф. дис. … д-ра техн. наук: 05.13.06. Київ, 2021. 40 с.

Закон України
Про наукову і науково-технічну діяльність: Закон України від 26.11.2015 № 848-VIII. URL: https://zakon.rada.gov.ua/laws/show/848-19 (дата звернення: 15.09.2026).

Постанова Кабінету Міністрів України
Про затвердження Порядку присудження ступеня доктора філософії: постанова Кабінету Міністрів України від 12.01.2022 № 44. URL: https://zakon.rada.gov.ua/laws/show/44-2022-%D0%BF (дата звернення: 15.09.2026).

Національний стандарт
ДСТУ 3008:2015. Інформація та документація. Звіти у сфері науки і техніки. Структура та правила оформлювання. Київ: ДП «УкрНДНЦ», 2016. 26 с.

Міжнародний стандарт
ISO/IEC 29147:2018. Information technology — Security techniques — Vulnerability disclosure. Geneva: ISO, 2018. URL: https://www.iso.org/standard/72311.html.

Технічний звіт
Surname N., Surname N. Title of the report: Technical Report TR-2024-07. Keele: Keele University, 2024. 48 p.

Препринт
Surname N., Surname N. Title of the preprint. *arXiv*. 2025. arXiv:2501.01234. URL: https://arxiv.org/abs/2501.01234.

Датасет
Прізвище І. Б. Назва набору даних: dataset. Version 1.2. Zenodo, 2025. DOI: 10.5281/zenodo.xxxxxxx.

Програмний код
Прізвище І. Б. Назва програмного забезпечення: software. Version 2.0. Zenodo, 2025. DOI: 10.5281/zenodo.xxxxxxx.

Патент України
Спосіб …: пат. 123456 Україна: МПК H04L 9/00. № a202301234; заявл. 10.03.2023; опубл. 15.11.2024, Бюл. № 46. 6 с.

Електронний ресурс без автора
Назва ресурсу. *Назва сайту*. URL: https://example.org/page (дата звернення: 15.09.2026).

Про дату звернення. Її наводять для електронних ресурсів, зміст яких може змінитися: вебсторінок, політик, реєстрів, онлайнових баз. Для запису з DOI або з іншим постійним ідентифікатором версії (стаття, книга, стандарт, препринт із номером версії, депозит у Zenodo) дата звернення не обов'язкова — ідентифікатор і так фіксує конкретну версію. Допустимо також оголосити єдину дату звернення один раз на початку списку й не повторювати її в кожному записі; саме так зроблено в списку джерел цього посібника. Неприпустимо лише одне — застосовувати різні правила в межах одного списку без пояснення.

Типові помилки: відсутність дати звернення для вебсторінки, що змінюється, за відсутності загального застереження на початку списку; заміна тире дефісом у діапазонах сторінок; скорочення назви журналу без потреби; подання DOI у вигляді гіперпосилання замість ідентифікатора; змішування в одному списку кількох стилів оформлення.