

ДОСЛІДНИЦЬКІ ДАНІ ЯК КЛЮЧОВИЙ СКЛАДНИК ВІДКРИТОЇ НАУКИ

Тетяна Павленко
2025

Відкрита наука

Обмін
дослідницькими
даними

Відкритість,
цілісність,
відтворюваність

Прискорення
наукового
прогресу

Обмін дослідницькими даними

- Поширена практика сучасного дослідницького середовища
- Вимога організацій-спонсорів наукових досліджень
- Політика рецензованих журналів

Відкриті дослідницькі дані

Визначення

- Цифрові та аналогові
- Необроблені і оброблені
- Метадані
- Числові дані
- Текстові записи
- Зображення
- Аудіозаписи
- Протоколи
- Коди аналізу
- Робочі процеси

Дозволи

- Відкритість
- Багаторазове використання
- Збереження
- Повторне поширення третіми особами за умови вказівки джерела

Керівні принципи (FAIR)

- Зручність пошуку
- Доступність
- Функціональна сумісність
- Можливість повторного використання

Дослідницькі дані

- Будь-яка інформація, отримана в ході наукових досліджень для підтвердження оригінальних результатів
- Набір даних, метод збору, формат збереження залежить від галузі та предметної області, правових аспектів, етичних норм тощо
- Біомедичні дослідження вимагають особливої уваги
- Систематизація даних за принциповими ознаками обумовлює продуктивність їх організації та управління

Біомедичні дослідницькі дані

За походженням

- **Первинні дані (необроблені)** – отримані безпосередньо в ході клінічних випробувань, лабораторних досліджень, медичних обстежень, спостережень за пацієнтами або анкетування
- **Вторинні дані (оброблені)** – отримані шляхом аналізу або агрегування вже наявних медичних даних (метааналіз клінічних випробувань, статистичні звіти про захворюваність, епідеміологічні дослідження тощо)

Біомедичні дослідницькі дані

За формою представлення

- **Кількісні дані** – виражені у числовій формі, що дозволяє проводити статистичний аналіз (рівень глюкози в крові, артеріальний тиск, частота серцевих скорочень, доза медикаментів тощо)
- **Якісні дані** – текстові, графічні або аудіовізуальні дані, які не можуть бути виражені простою числовою шкалою (результати інтерв'ю з пацієнтами, опис клінічних випадків, відеозаписи медичних процедур тощо)

Біомедичні дослідницькі дані

За структурою

- **Структуровані дані** – організовані у фіксованому форматі, наприклад, у базах медичних записів (електронні медичні карти, фармакологічні БД, реєстри пацієнтів тощо)
- **Напівструктуровані дані** – містять певні метадані без суворої схеми організації (записи електронної медичної документації у форматі XML або JSON тощо)
- **Неструктуровані дані** – представлені у вільному форматі (рукописні медичні записи, аудіофайли консультацій, медичні зображення без метаданих тощо)

Біомедичні дослідницькі дані

За методом збору

- **Дані спостережень** – отримані шляхом моніторингу стану пацієнтів або проведення довготривалих досліджень
- **Експериментальні дані** – отримані в результаті клінічних випробувань, лабораторних тестів або медичних експериментів
- **Дані моделювання** – створені шляхом комп'ютерного моделювання біологічних процесів (математичні моделі поширення вірусів, прогнозування реакції організму на медикаменти тощо)
- **Похідні дані** – отримані шляхом аналізу і компіляції існуючих даних

Біомедичні дослідницькі дані

За типом

- **Клінічні дані** – інформація, отримана під час лікування пацієнтів
- **Епідеміологічні дані** – статистика поширеності, причин і наслідків захворювань у популяціях
- **Лабораторні дані** – результати аналізів крові, генетичних досліджень, мікробіологічних тестів
- **Фармакологічні дані** – дані про клінічні випробування ліків, побічні ефекти, ефективність терапії
- **Фізіологічні та біомедичні дані** – інформація з датчиків та портативних пристроїв
- **Соціально-демографічні дані пацієнтів**

Біомедичні дослідницькі дані

За форматом – цифрові та аналогові дані в різних комбінаціях

- Текстові документи, табличні та графічні форми
- Медичні зображення та візуалізації
- Геопросторові дані та мапи
- Відео- та аудіозаписи
- Біомедичні дані, зразки та проби
- Протоколи лікування та алгоритми діагностики
- Комп'ютерні моделі
- Лабораторні журнали, щоденники спостережень, анкетні дані та результати опитувань

Біомедичні дослідницькі дані

За рівнем доступності

Відкрито, наскільки це можливо; закрито, наскільки це необхідно ([Barcelona Declaration on Open Research Information](#))

- **Відкриті дані** – доступні для публічного використання
- **Обмежено відкриті** – доступні лише для певних груп користувачів
- **Конфіденційні (закриті) дані** – доступні лише власникам або уповноваженим особам (медичні архіви, приватні фармацевтичні дослідження, військова медична інформація тощо)

Управління дослідницькими даними

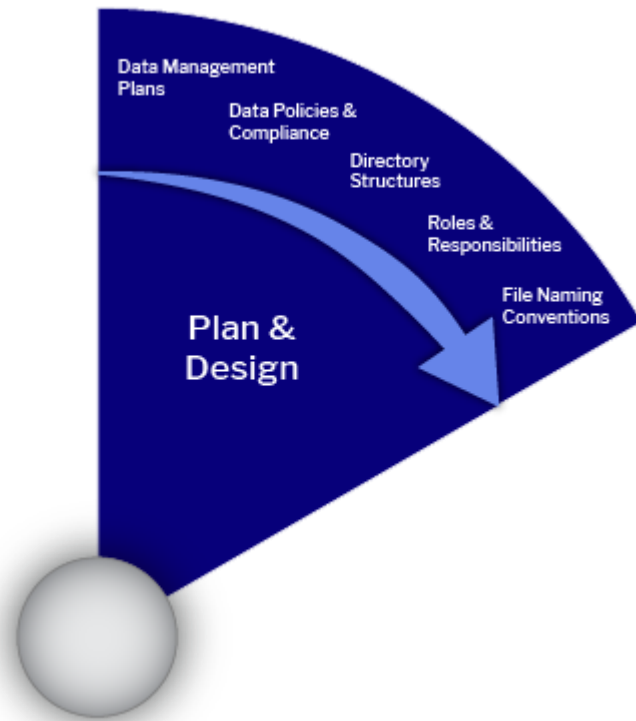
- Як збирати дослідницькі дані?
- Як зберегти?
- Як стандартизувати та опрацьовувати?
- За якими метаданими ідентифікувати?
- Як зробити доступними?
- Де та як поширювати?

Harvard Biomedical Data Management



1. Планування та проєктування
2. Збір і створення
3. Аналіз та співпраця
4. Оцінка та архівування
5. Обмін та поширення
6. Публікація та повторне використання

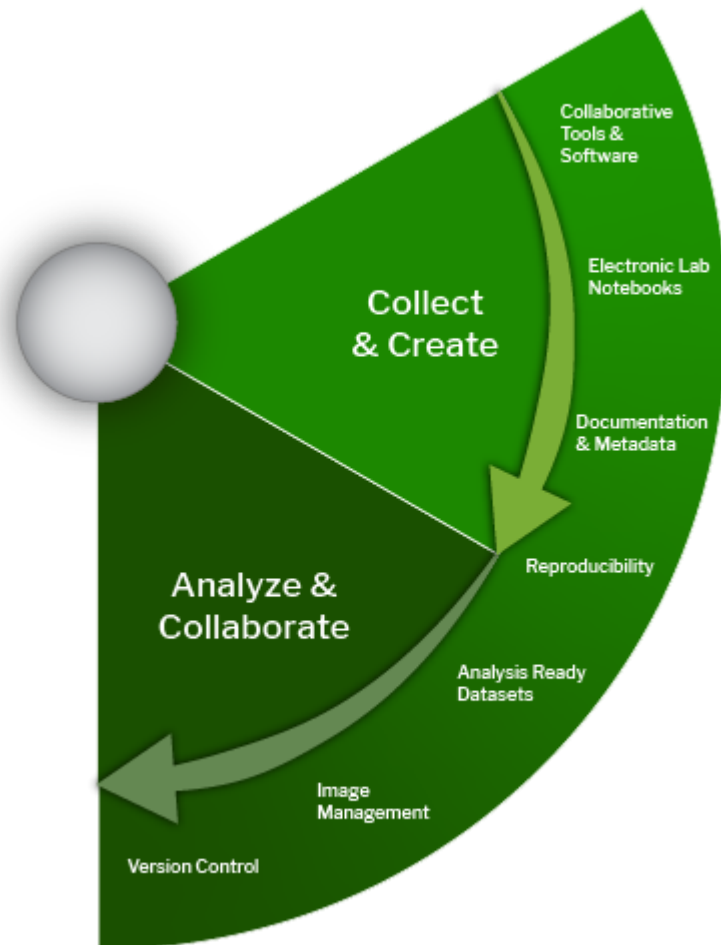
Harvard Biomedical Data Management



1. Планування та проєктування

- складання плану та бюджету дослідницьких цілей
- розробка чітких нормативів, етичних правил та протоколів, які регламентують роботу з даними
- розподіл ролей та обов'язків серед учасників

Harvard Biomedical Data Management



2. Збір і створення – організація та інтеграція наборів даних і процесів збору; підготовка переліку типів та форматів даних; опис методів їх збору та систематизації з використанням стандартизованих процедур та пов'язаних метаданих; створення супровідної документації

3. Аналіз та співпраця – перевірка відповідності зібраних даних; їх категоризація і угруповання; очищення від шумів і помилок; цифрова обробка; складання протоколів щодо інтерпретації та аналізу даних; командна співпраця у формулюванні висновків та прийнятті рішень

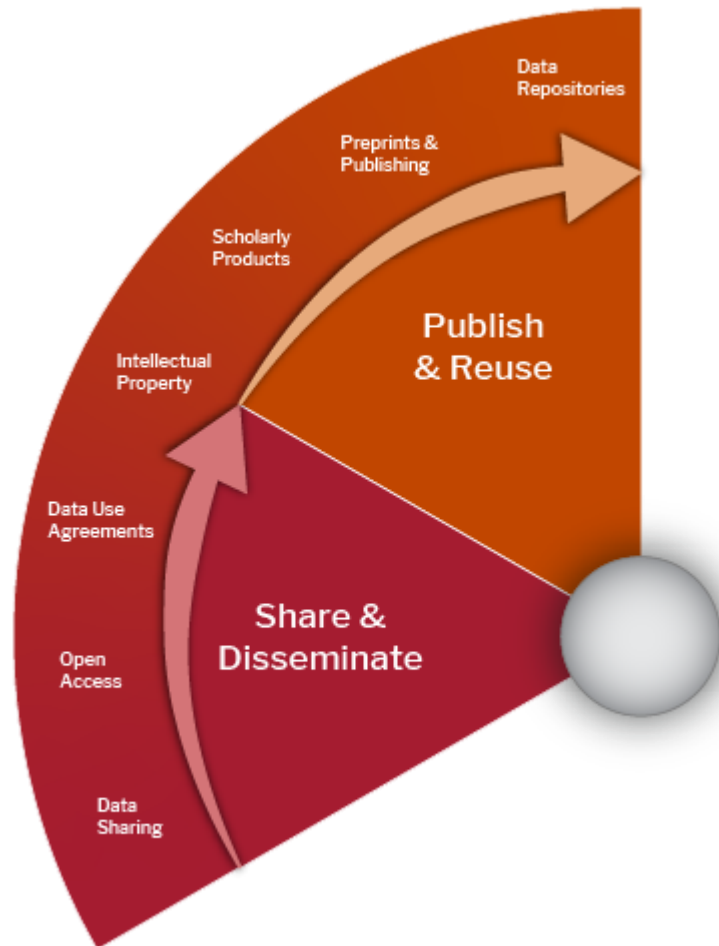
Harvard Biomedical Data Management



4. Оцінка та архівування

- визначення політик захисту і цілісності даних
- обрання надійної платформи для архівування і резервного копіювання
- визначення рівня важливості даних щодо вибору стратегії їх зберігання і доступу до записів з урахуванням мінливості технологічного середовища

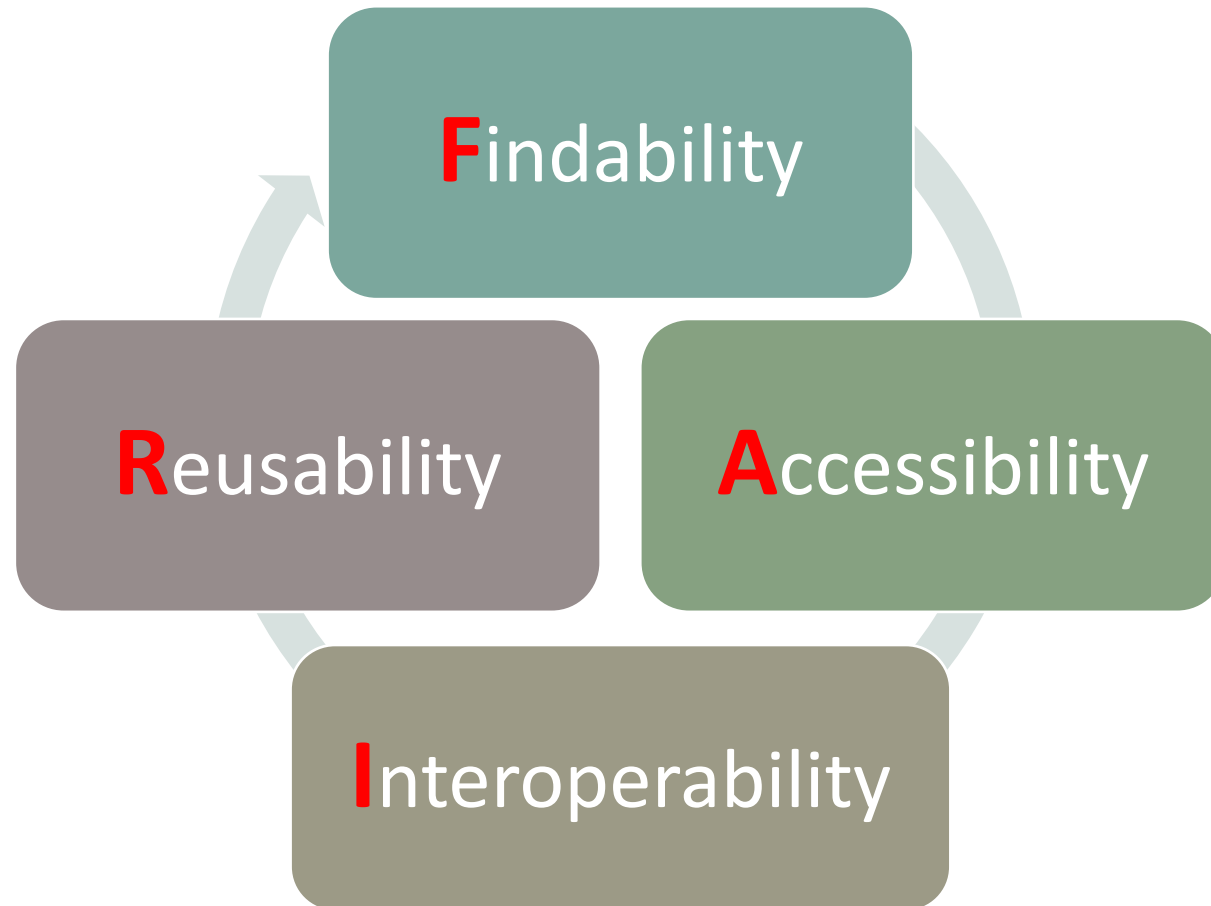
Harvard Biomedical Data Management



5. Обмін та поширення – застосування практик відкритої науки для обміну усіма компонентами дослідницьких даних, методів та протоколів; визначення регуляторних політик щодо передачі даних, їх повторного використання, зберігання, повторного поширення та утилізації чи припинення дії з урахуванням етичних та правових норм

6. Публікація та повторне використання – промоція дослідницьких даних шляхом надання відкритих ліцензій; забезпечення постійних ідентифікаторів для їх цитування; застосування системи посилань між публікаціями та даними; обрання репозитарію для розміщення дослідницьких даних

Керівні принципи FAIR



Керівні принципи FAIR

Findability (*відшукуваність*)

- Переконайтеся, що ваші дані можуть бути знайдені як людьми, так і пошуковими системами.
- Використовуйте репозитарії
- Надайте детально структуровані описи метаданих
- Вказуйте унікальний постійний ідентифікатор цифрового об'єкта

Керівні принципи FAIR

Accessibility (*доступність*)

- Переконайтеся, що ваші дані відкриті і зрозумілі як людям, так і машинам
- Доступні для використання
- Якщо дані мають певні обмеження, надайте в публічний доступ їх метадані

Керівні принципи FAIR

Interoperability (*сумісність*)

- Переконайтеся, що ваші дані функціонально здатні до взаємодії та інтеграції з іншими даними, інструментами, програмами
- Використовуйте відкриті формати, що відповідають міжнародним стандартам
- Використовуйте контрольовані словники та чітко визначені моделі як гарантію структурованості даних

Керівні принципи FAIR

Reusability *(повторне використання)*

- Переконайтеся, що дані та пов'язані метадані супроводжуються детальним описом з можливістю тиражувати та/або комбінувати в різних налаштуваннях
- Повторне використання обумовлено зрозумілою та доступною ліцензією (ліцензіями)
- Надайте посилання для забезпечення зручності цитування

Відкриті дослідницькі практики

Переваги для дослідників

- Захист автентичності і цілісності даних
- Зменшення ризику втрати даних і убезпечення їх від «цифрового старіння»
- Оптимізація їх використання в процесі дослідження та співпраці з колегами
- Демонстрація доброчесності дослідницького процесу і відповідності необхідним вимогам
- Зручність звітності
- Зміцнення наукової репутації

Відкриті дослідницькі практики

Переваги для суспільства

- Підвищення достовірності, надійності та відтворюваності наукових результатів
- Полегшення інтерпретації та повторного використання даних
- Можливість незалежного підтвердження результатів, розвитку нових підходів та вдосконалення наукових методів
- Стимулювання наукових інновацій
- Покращення якості наукової інформації
- Зростання довіри до науки

Використані джерела

- [Barcelona Declaration on Open Research Information \(2024\)](#)
- [Biomedical Data Types Explorer](#)
- Das, G., Ghosh, I. (2024). Biomedical Data Types, Sources, Content, and Retrieval. In: Sharan, A., Malik, N., Imran, H., Ghosh, I. (eds) Text Mining Approaches for Biomedical Data. Transactions on Computer Systems and Networks. Springer, Singapore. https://doi.org/10.1007/978-981-97-3962-2_2
- [Harvard Biomedical Data Management](#)
- [Human clinical and health data](#)
- Myneni S, Patel VL. Organization of Biomedical Data for Collaborative Scientific Research: A Research Information Management System. Int J Inf Manage. 2010 Jun 1;30(3):256-264. <https://doi.org/10.1016/j.ijinfomgt.2009.09.005>

Використані джерела

- [Practical guide to the international alignment of research data management](#)
- [Types of Healthcare Data: A Comprehensive Overview](#)
- [UNESCO Recommendation on Open Science \(2021\)](#)
- [Відкрита наука : вступний посібник \(2023\)](#)
- [Відкриті наукові практики : навчальний посібник \(2023\)](#)
- [Методичні рекомендації щодо управління науковими даними](#) для закладів вищої освіти та наукових установ у частині визначення механізмів збереження та повторного використання наукових даних (2024)
- [Принципи FAIR \(Findability, Accessibility, Interoperability, Reusability\)](#)

Тетяна Павленко

заступник директора Наукової бібліотеки ХНМУ

tb.pavlenko@kntu.edu.ua