

FAIR-дані та FAIR workflow

FAIR-дані

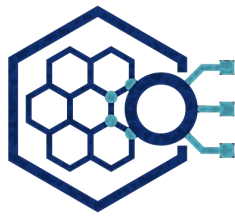
це **дані**, властивості яких забезпечують їхню знаходжуваність, доступність, інтеоперабельність і повторне використання

FAIR workflow

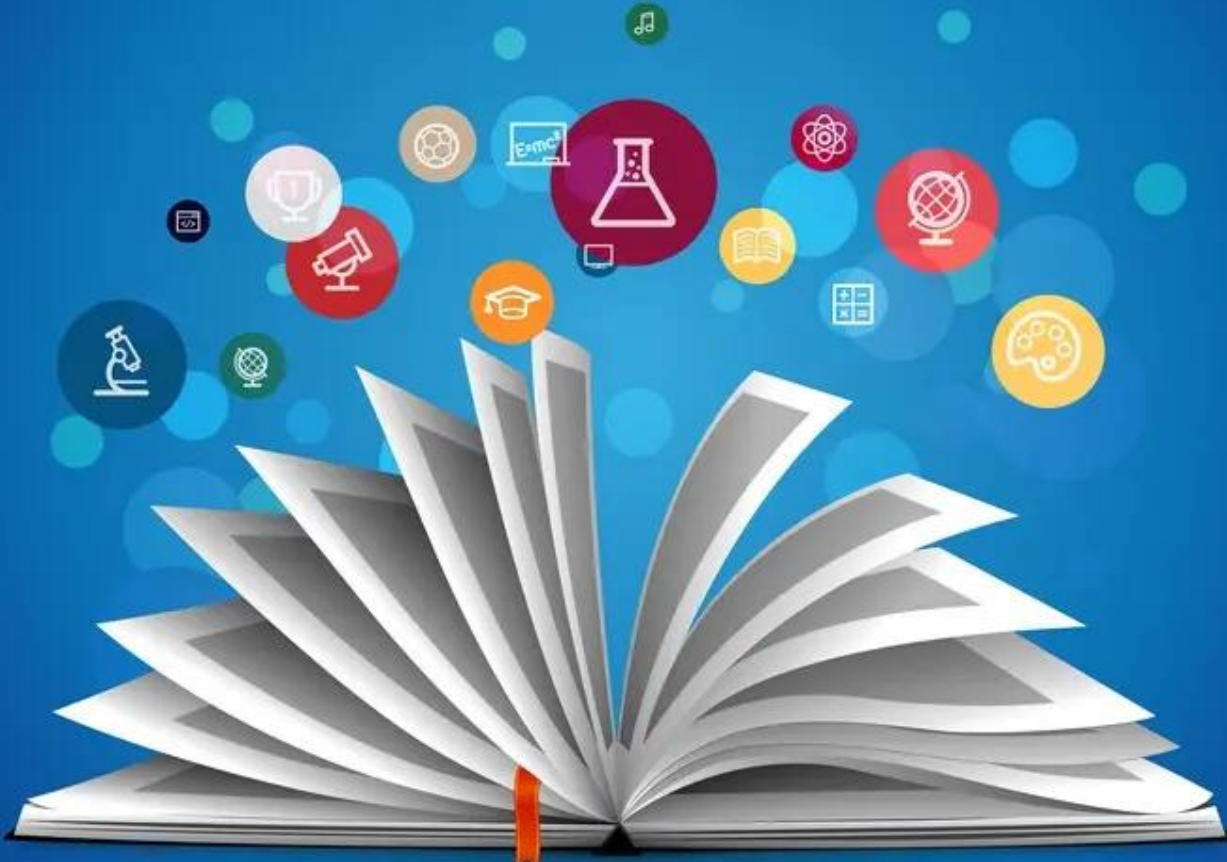
це **процес і сукупність практик роботи з даними**, спрямованих на забезпечення цих властивостей – знаходжуваність, доступність, інетроперабельність і повторне використання.

FAIR-дані та FAIR workflow

Питання	FAIR data	FAIR workflow
Що є об'єктом?	Набір даних / Колекція даних	Послідовність обчислювальних або аналітичних етапів та їхніх залежностей
Що описується?	Файли, змінні, формати, ліцензія, автор, ідентифікатори	вхідні та вихідні дані, програмне забезпечення, версії, параметри, середовище, порядок виконання та взаємозв'язки між етапами
Чим забезпечено повторне використання?	Метадані + доступні файли + чітко визначена ліцензія	Виконуваний або відтворюваний процес + метадані + походження даних + пов'язані дані/програмне забезпечення
Що показує провенанс?	Звідки походять ці дані?	Як було отримано цей результат: на основі яких вхідних даних, за допомогою яких етапів і за яких умов?



УПРАВЛІННЯ ДАНИМИ ДОСЛІДЖЕНЬ



Управління даними досліджень (Research Data Management, RDM)

Управління даними досліджень (Research Data Management, RDM)

– це система заходів для ефективного управління науковими даними протягом усього їхнього життєвого циклу: від планування збору даних, їх організації та зберігання під час дослідження до довгострокового збереження, обміну та повторного використання після його завершення.

Головна мета RDM

– забезпечити, щоб дані були безпечними, стійкими, доступними для пошуку, розуміння та використання .

ІСТОРИЧНІ ВИТОКИ ПЛАНУ УПРАВЛІННЯ ДАНИМИ

- ❑ Ідея **Data Management Plan** (плану управління даними) сформувалась у США наприкінці 2000-х.
- ❑ **2006 рік** – перші рекомендації з боку деяких наукових бібліотек і центрів досліджень, які зіштовхнулися з проблемою збереження великих обсягів наукових даних.
- ❑ **2011 рік** – **Національний науковий фонд США (NSF)** зробив DMP обов’язковою частиною всіх заявок на фінансування.
Це вважають офіційною відправною точкою практики у світовому масштабі.
- ❑ Пізніше практика набула поширення у Великій Британії (EPSRC, UKRI), Нідерландах, Австралії, Канаді, а після 2016 року — і в ЄС через програму **Horizon 2020**, нині — **Horizon Europe**.

Життєвий цикл наукових (дослідницьких) даних

Етап	Зміст
1. Планування	Визначення видів даних та процесу управління ними
2. Створення / Збір	Отримання первинних даних шляхом експериментів, вимірювань, опитувань, моделювання
3. Обробка	Первинна очистка, структурування, перетворення даних у придатний формат, кодування змінних
4. Аналіз	Використання методів та інструментів для інтерпретації даних (статистичний аналіз, моделювання тощо)
5. Документування / Опис (Метадані)	Опис даних за стандартами метаданих - документування контексту, методів, походження, щоб інші могли зрозуміти і повторити
6. Зберігання	Організація даних у захищених і стабільних системах (репозитаріях)
7. Поширення / Публікація	Надання доступу до даних для інших дослідників і суспільства (через відкриті репозитарії, журнали з відкритим доступом)
8. Повторне використання	Дані застосовуються у нових дослідженнях чи валідації результатів, в освіті
9. Архівування / Довготривале збереження	Забезпечення доступності та цілісності даних у перспективі (через цифрові архіви, національні дата-центри)