

Dal perché al come: l'esperienza UNIMI con Dataverse per la gestione FAIR e la condivisione di dati di qualità

19 novembre 2025 , Dataverse in Italia: esperienze a
confronto – GenOA week

Dario Basset & Carolina Manfredini
Direzione Performance, Assicurazione qualità, valutazione e politiche di
Open Science - Università degli Studi di Milano

Il contesto

2014: elaborazione della [Policy Open Access](#) (aggiornata nel 2020)

2016: UNIMI entra in [Open APC](#)

2017: istituzione della Commissione d'Ateneo per l'Accesso Aperto (dal 2020 diventata **Commissione Open Science**)

2017: elaborazione della [Policy sulla gestione dei dati della ricerca](#) (aggiornata nel 2022)

2017/2018: definizione template istituzionale DMP

2017/2018: ricerca di un repository open source e FAIR



Perché Dataverse: la scelta strategica e i punti di forza





❑ Sfide legate alla **community**:

- **Reattività** della community non sempre immediata
- La forza di Dataverse sta nella sua comunità open source - se cala la **partecipazione** o mancano istituzioni attive, il progetto rischia di indebolirsi

❑ Sfide legate alla **sostenibilità** a lungo termine:

- **Obsolescenza** tecnologica
- **Costi** infrastrutturali crescenti (soprattutto con dataset di grandi dimensioni)
- Garantire la presenza e i fondi per le **risorse umane** coinvolte nella sua gestione e implementazione

Implementazione in UNIMI

- **2018**: sperimentazioni su Dataverse
- **2019**: lancio di **Dataverse UNIMI** e primi progetti pilota di deposito e condivisione dei dati
- **2020**: definizione di **linee guida** & avvio di servizio di supporto per la **redazione di DMP**
- **2022**: creazione di un **sito web** per la gestione dei dati della ricerca, avvio di **corsi** dedicati & arruolamento di un **repository manager**
- **2023**: arruolamento di due **data stewards**
- **Primavera 2024**: introduzione di un nuovo **workflow** con sistema di **validazione prima della pubblicazione** & definizione di nuove linee guida
- **Autunno 2024**: arruolamento di una terza data stewards & ottenimento del **CoreTrustSeal**
- **2025**: rinnovamento dei corsi dedicati a RDM, avvio di seminari pratici e dipartimentali, lancio del nuovo sito web

Ottenimento della certificazione CoreTrustSeal

Il percorso verso la certificazione CoreTrustSeal, iniziato nel 2022, si è concluso con successo a **ottobre 2024**. Sono state necessarie **4 interazioni** con il gruppo certificatore per chiarire alcuni aspetti su cui è stata chiesta documentazione specifica.

CoreTrustSeal: 16 requisiti (+1)

Context

Digital Object
Management

Provenance and authenticity (R07)
Deposit & Appraisal (R08)
Preservation plan (R09)
Quality Assurance (R10)
9
Workflows (R11)
Discovery and Identification (R12)
Reuse (R13)

Organizational
Infrastructure

Mission & Scope (R01)
Rights Management (R02)
Continuity of Service (R03)
Legal & Ethical (R04)
Governance & Resources (R05)
Expertise & Guidance (R06)

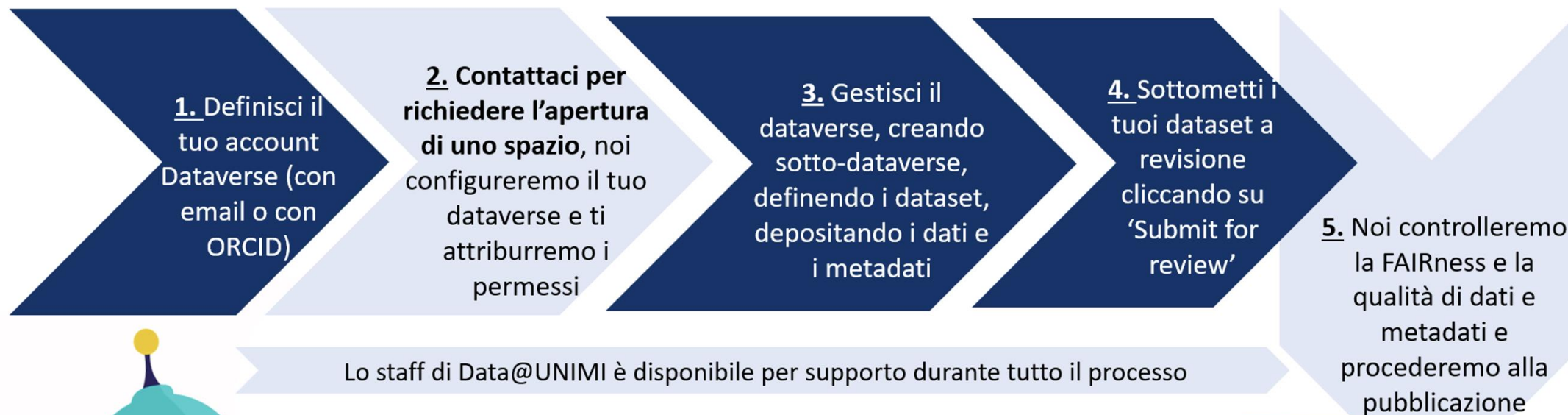
Information
Technology &
Security

Storage & Integrity (R14)
Technical Infrastructure (R15)
Security (R16)

Alcune osservazioni sul processo di certificazione:

1. La certificazione richiede una serie di documenti che dimostrino come si svolge il lavoro presso l'istituzione, quali sono i **processi, i sistemi e il flusso di lavoro adottato**. Tali documenti vanno preparati almeno in lingua inglese e servono anche come guida dell'utente.
2. Occorre avere ben chiaro il concetto di **user population**, cioè chi sono gli utenti reali e potenziali del sistema.
3. Occorre avere **un'organizzazione strutturata di risorse umane** in grado di lavorare e di far lavorare sul sistema, in maniera che possano essere erogate tutte le attività relative alla promozione, al supporto ed alla formazione, a vari livelli.
4. Gli **aspetti formali** sono importanti, se si chiama una parte degli utenti in una maniera, in tutto il documento lo stesso nome va usato in maniera appropriata.
5. Gli **aspetti tecnici** rimangono essenziali, soprattutto per quello che riguarda l'attenzione agli aspetti di **sicurezza e conservazione dei dati**.

Il workflow di Dataverse UNIMI



Dataverse UNIMI: chi siamo?

3 data stewards

- Gestione, mantenimento, sviluppo e **operazioni quotidiane** nel repository
- Definizione di **linee guida** e documentazione
- Gestione e aggiornamento del [sito web RDM](#)
- Controllo quotidiano dei contenuti caricati sul repository
- **Contatti quotidiani** con gli utenti via e-mail e tramite incontri online
- Curatela dei dati e **controlli di qualità** specifici
- **Corsi e formazione**
- **Supporto disciplinare**

1 repository manager

- **Supervisione** alla gestione, mantenimento, sviluppo e operazioni quotidiane nel repository
- Supporto per la definizione linee guida e documentazione
- **Supporto tecnico** per il sito RDM
- Configurazione e implementazione di **nuove funzionalità**
- Statistiche di utilizzo
- **Aspetti tecnico-informatici** e programmi dedicati
- Supervisione alla curatela dei dati
- **Supporto tecnico specifico**



Incontri periodici con la **dirigente della Direzione Performance, Assicurazione Qualità, Valutazione e Politiche di Open Science & il responsabile dell'Ufficio Qualità dei Dati sulla Ricerca** per aggiornamenti, accordi, implementazione delle pratiche, questioni specifiche sulla gestione e qualità dei dati

Il controllo della FAIRness: integrazione tra strumenti automatici e verifica delle data steward

Dataset pubblicati

- Programma automatizzato '**Completeness Check**'
- Eseguito periodicamente (**due volte all'anno**) e i suoi risultati sono verificati dalle data stewards che possono richiedere agli utenti di apportare modifiche o **integrazioni per migliorare la qualità dei dataset**

Dataset non pubblicati - sottomessi

- Programma automatizzato '**Admin Check**'
- Esegue verifiche **quotidiane** per garantire che le autorizzazioni di pubblicazione siano riservate solo al gruppo di supporto
- Dati e metadati sono sottoposti alla **validazione** delle data stewards **prima della pubblicazione**, generalmente entro un giorno dal ricevimento della richiesta

Dataset non pubblicati – in progress

- Programma automatizzato '**Draft Check**'
- Eseguito **mensilmente** per individuare i dataset che sono rimasti **in stato di bozza per un periodo prolungato**. Notifica il gruppo di supporto che valuta se gli utenti abbiano dimenticato di pubblicarli, se stiano ancora lavorando attivamente su di essi o se abbiano creato una versione in bozza per modificare un dataset già pubblicato

La validazione dei dataset: cosa controlliamo?

UNIMI Dataverse > Paola test > Paola Test Dataverse Child > Customized dataverse >


Info – This draft version needs to be published. When ready for sharing, please **publish** it so that others can see these changes.

Unpublished Dataset Private URL – Privately share this dataset before it is published: <https://dataverse.unimi.it/privateurl.xhtml?token=df9c50ab-d121-4186-9e8b-929f7d75292a>

- ☐ Struttura del dataverse
- ☐ Naming convention e coerenza dei titoli

ingestion test data

Draft **Unpublished**



---, 2025, "ingestion test data", https://doi.org/10.13130/RD_UNIMI/9YJ0SJ, UNIMI Dataverse, DRAFT VERSION, UNF:6:VijUV/SmasFppLMBRLvHHQ== [fileUNF] ?

[Cite Dataset](#) [Learn about Data Citation Standards.](#)

Access Dataset ▾

[Publish Dataset](#)

Edit Dataset ▾

[Contact Owner](#) [Share](#)

Dataset Metrics ?


0 Downloads ?

Description ?

Subject ?

Other

License/Data Use Agreement

 **CC0 1.0**

- ☐ Licenza assegnata e termini di accesso

[Files](#) [Metadata](#) [Terms](#) [Versions](#)

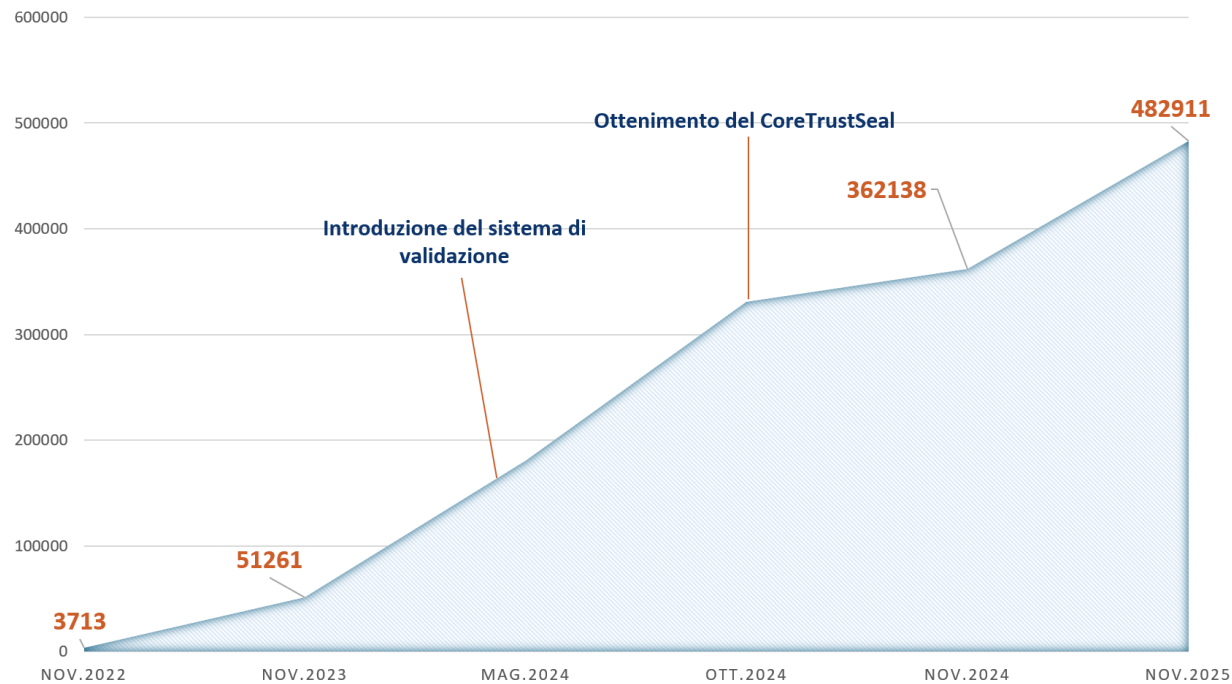
- ☐ Formato dei file
- ☐ Presenza del README file

+ Upload Files

Sort ▾

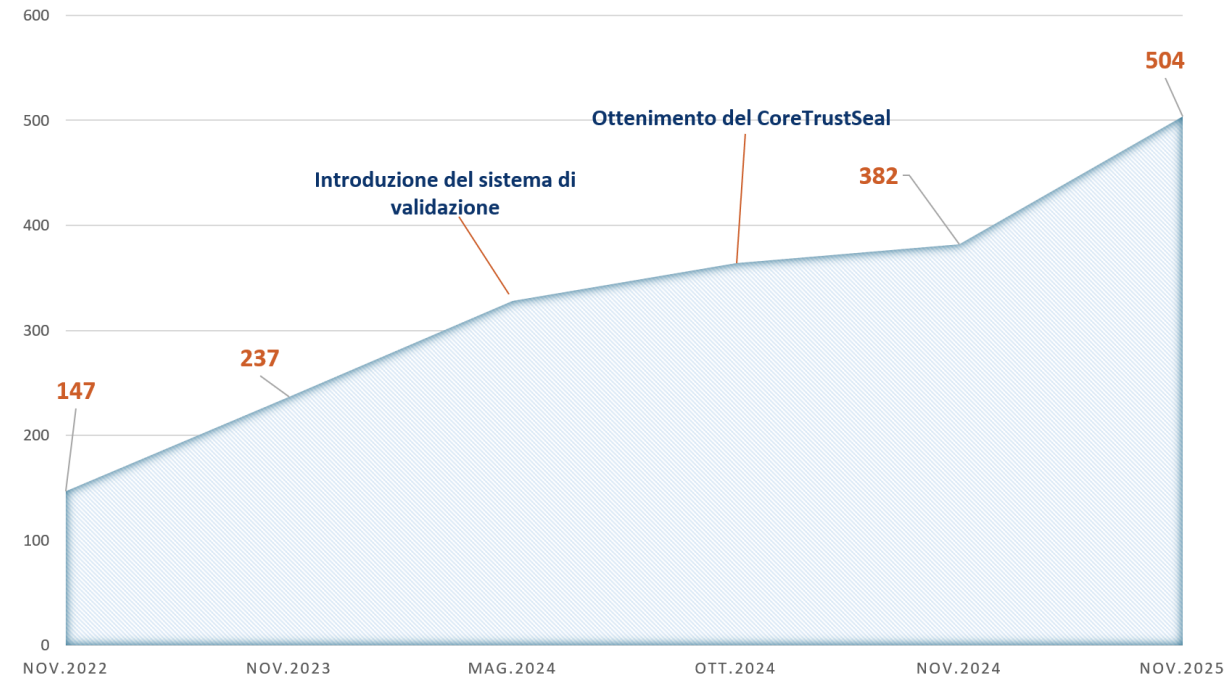
Risultati concreti: Dataverse

TOTALE DOWNLOAD



Tra 2022 e 2025, **aumento percentuale** dei download pari a **+14610%**

TOTALE DATASET PUBBLICATI



Tra 2022 e 2025, **aumento percentuale** dei dataset pubblicati pari a **+385%**
- in parallelo, crollo delle pubblicazioni su Zenodo, [v. poster](#)

Formazione e supporto: promuovere una cultura dell'apertura e della qualità dei dati

Sito web
dedicato e
aggiornato
periodicament
e

Linee guida,
template e
strumenti utili
specifici per
Dataverse,
DMP e
gestione FAIR
dei dati

Reportistica su
community
Zenodo

Calendario di
ricevimento
online

Webinar
mensili
introduttivi e
avanzati su
Dataverse,
DMP, dati FAIR
e gestione dei
dati sensibili

Workshop
pratici e
disciplinari per
early-career
researchers

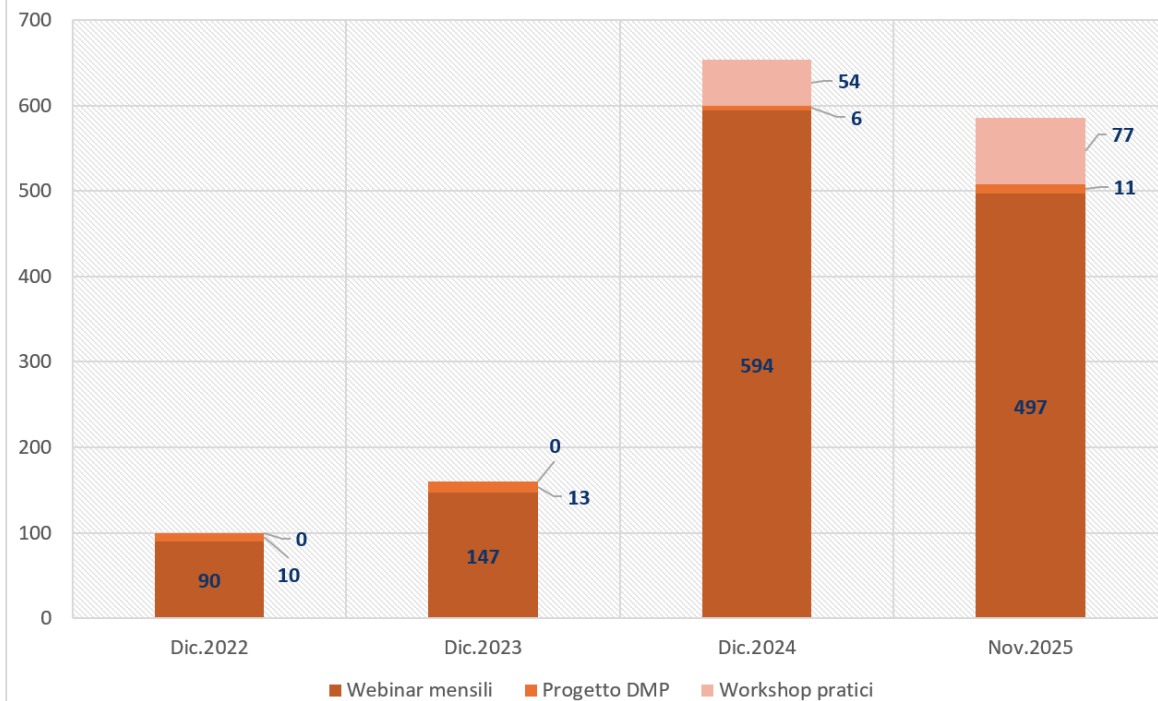
Incontri
personalizzati
per
dipartimenti e
progetti di
ricerca

Progetto DMP
per PhD (qui
una
testimonianza)

Supporto quotidiano e capillare tramite e-mail e incontri MS Teams per Dataverse, DMP e pratiche di gestione dei dati

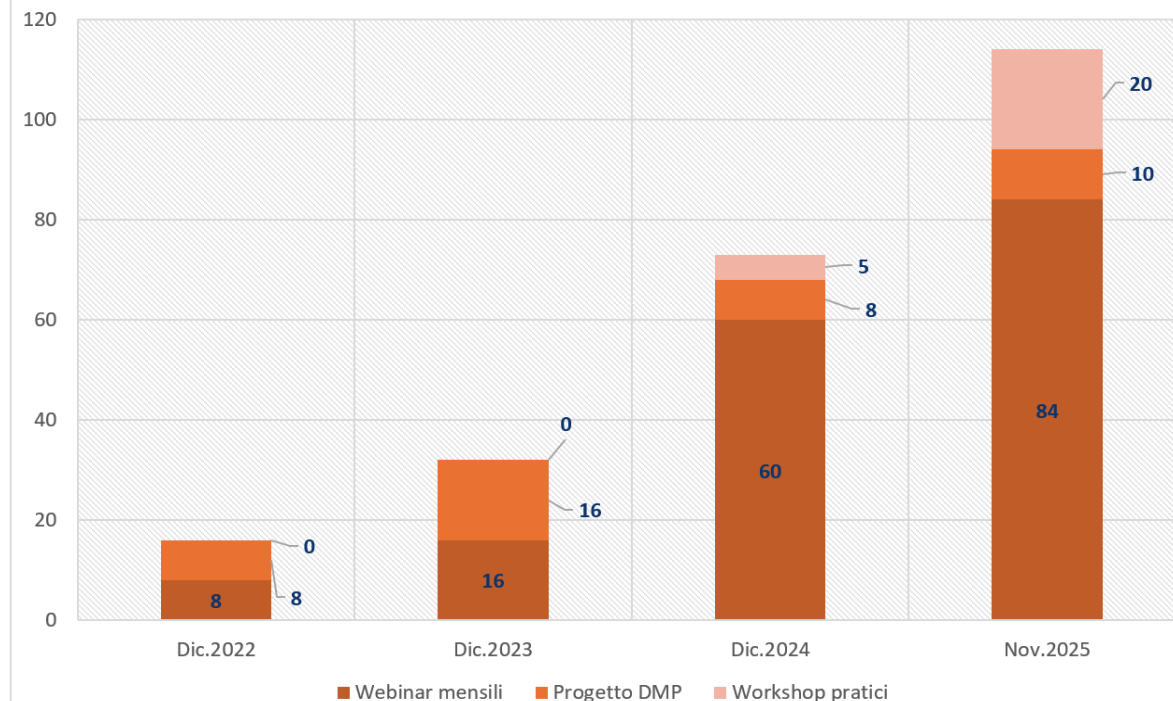
Risultati concreti: formazione

TOTALE ISCRITTI ALLA FORMAZIONE RDM



Tra 2022 e 2025, **aumento percentuale** degli iscritti alla formazione RDM pari a **+452%**

TOTALE ORE ANNUE DI FORMAZIONE RDM



Tra 2022 e 2025, **aumento percentuale** delle ore di formazione RDM pari a **+721%**

Messaggi chiave e prospettive future

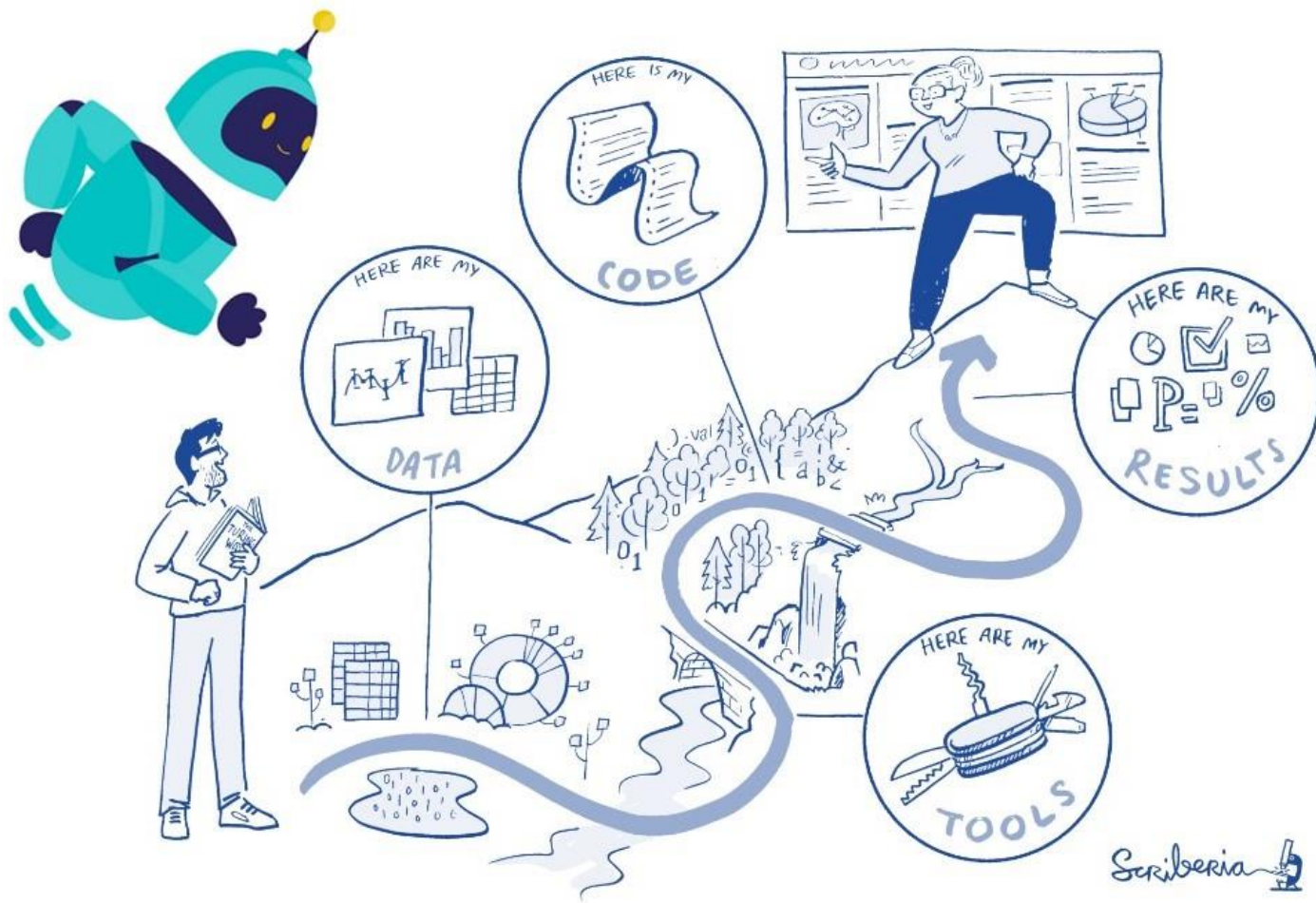


Illustration by Scriberia with The Turing Way community, used under a CC-BY 4.0 licence. DOI: [10.5281/zenodo.3332807](https://doi.org/10.5281/zenodo.3332807)

- ❑ Necessità di **implementazioni tecniche** nella visualizzazione dei file di dati, per stare al passo e **garantire interoperabilità e FAIRness**
- ❑ CoreTrustSeal, percentuali e scelte dei ricercatori come specchio del **lavoro capillare di supporto**
- ❑ **Non c'è qualità senza formazione e dialogo** con la comunità accademica
- ❑ Dataverse come **infrastruttura viva** e strumento di cultura di condivisione di dati di qualità
- ❑ Investire oggi per la sostenibilità e la qualità dei dati di domani



UNIMI Dataverse

Grazie per l'attenzione!

dataverse@unimi.it