L'IA al servizio della Scienza Aperta

Esperienze di data stewardship e monitoraggio della produzione scientifica

GenOA week 2026 – 5 novembre 2026 – **Sessione 7: Costi umani e competenze (seconda parte)**

Ugo Moschini, Lucia Corti, Marianna Laviola, Aurora Guaresi, Valentina Pasquale

Istituto Italiano di Tecnologia

This work is licensed under a Creative Commons Attribution 4.0 International License.

Scienza Aperta: nuove figure, nuovi approcci

Nuove figure professionali

Esperti Open Science, data analyst, data steward

Nuovi modi di valutare

Monitoraggio e valutazione della ricerca con parametri che valorizzino qualità e pluralità dei prodotti scientifici

Risorse limitate, IA in corsa

Carenza di risorse e difficoltà di tenere il passo con l'evoluzione dirompente dell'IA

L'opportunità: l'IA può potenziare il lavoro di supporto alla ricerca, lasciando le competenze umane alle attività di maggior valore aggiunto.

Due applicazioni concrete all'IIT

1

Dashboard R/Shiny sviluppate con l'IA

- Monitoraggio della produzione scientifica: pubblicazioni, dati e software collegati, aderenza ai principi FAIR.
- Monitoraggio dell’utilizzo del repository IIT Dataverse.

2

Schemi di metadati per i dati genomici

Mappatura dei campi su un insieme rappresentativo di standard internazionali.

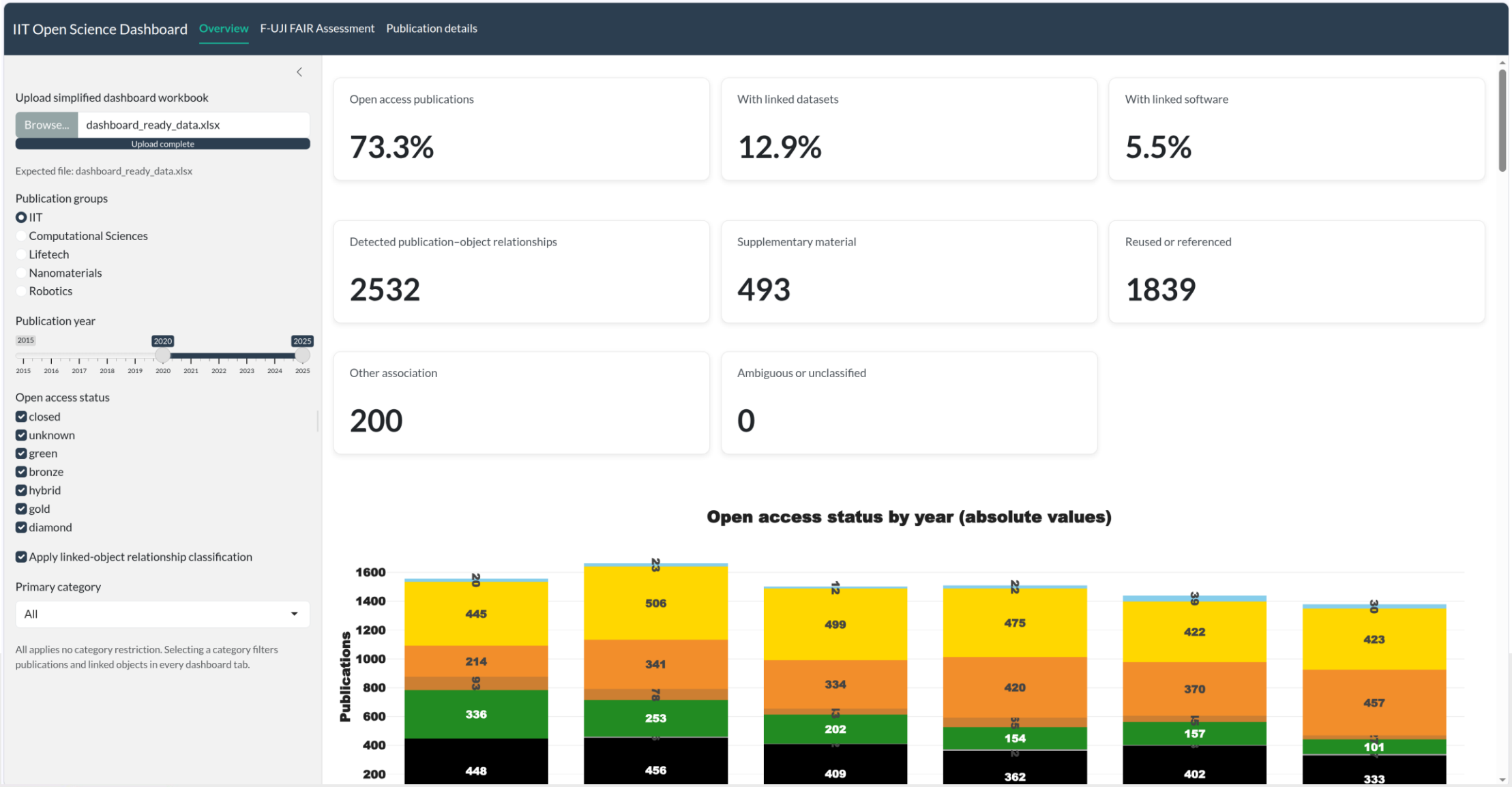
Collegare pubblicazioni, dati e software



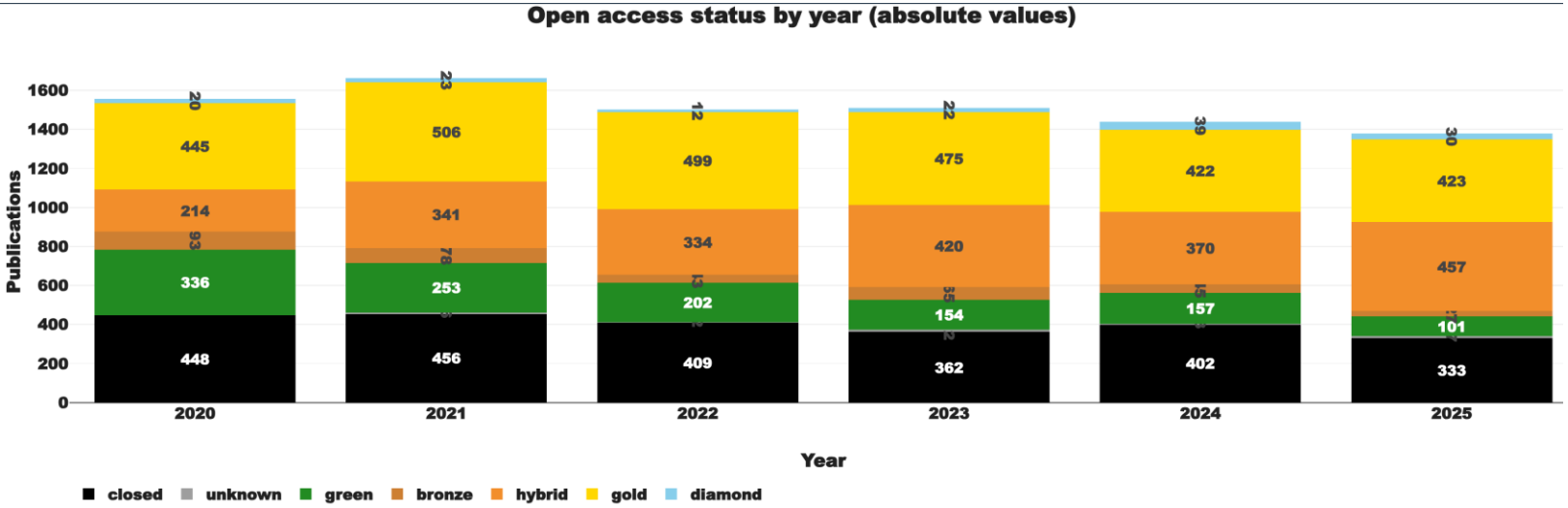
Sviluppata con assistenza dell'IA (ChatGPT/Codex 5.6)

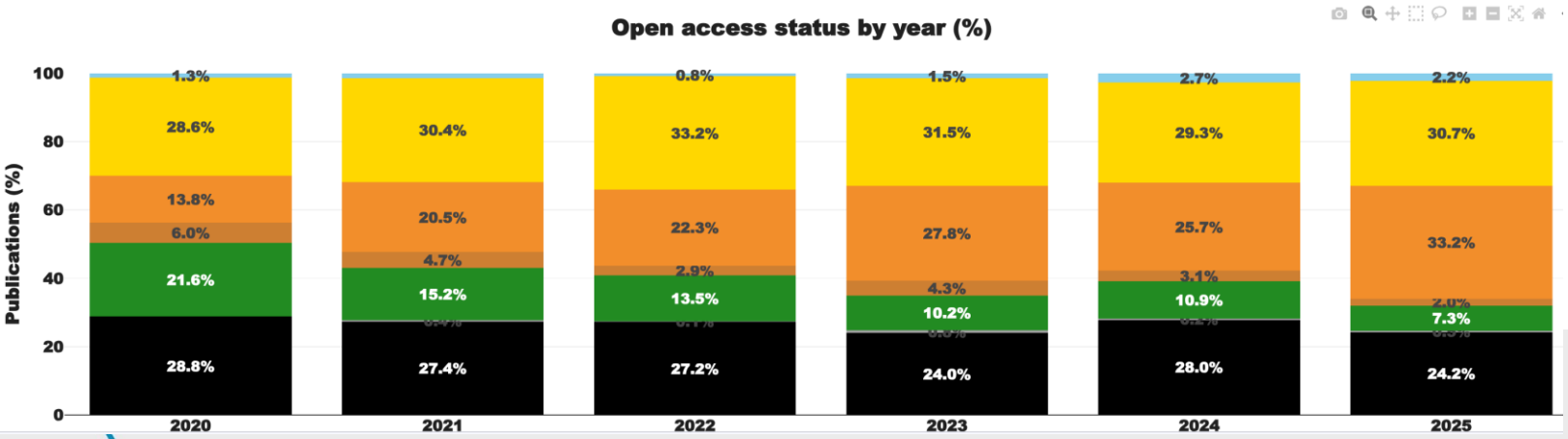
ISTITUTO ITALIANO DI TECNOLOGIA

IIT Open Science Dashboard



Open Access: l'andamento anno per anno





Stato Open Access per anno

Green, gold, hybrid, diamond, bronze, closed

Valori assoluti e percentuali

Due viste dello stesso andamento

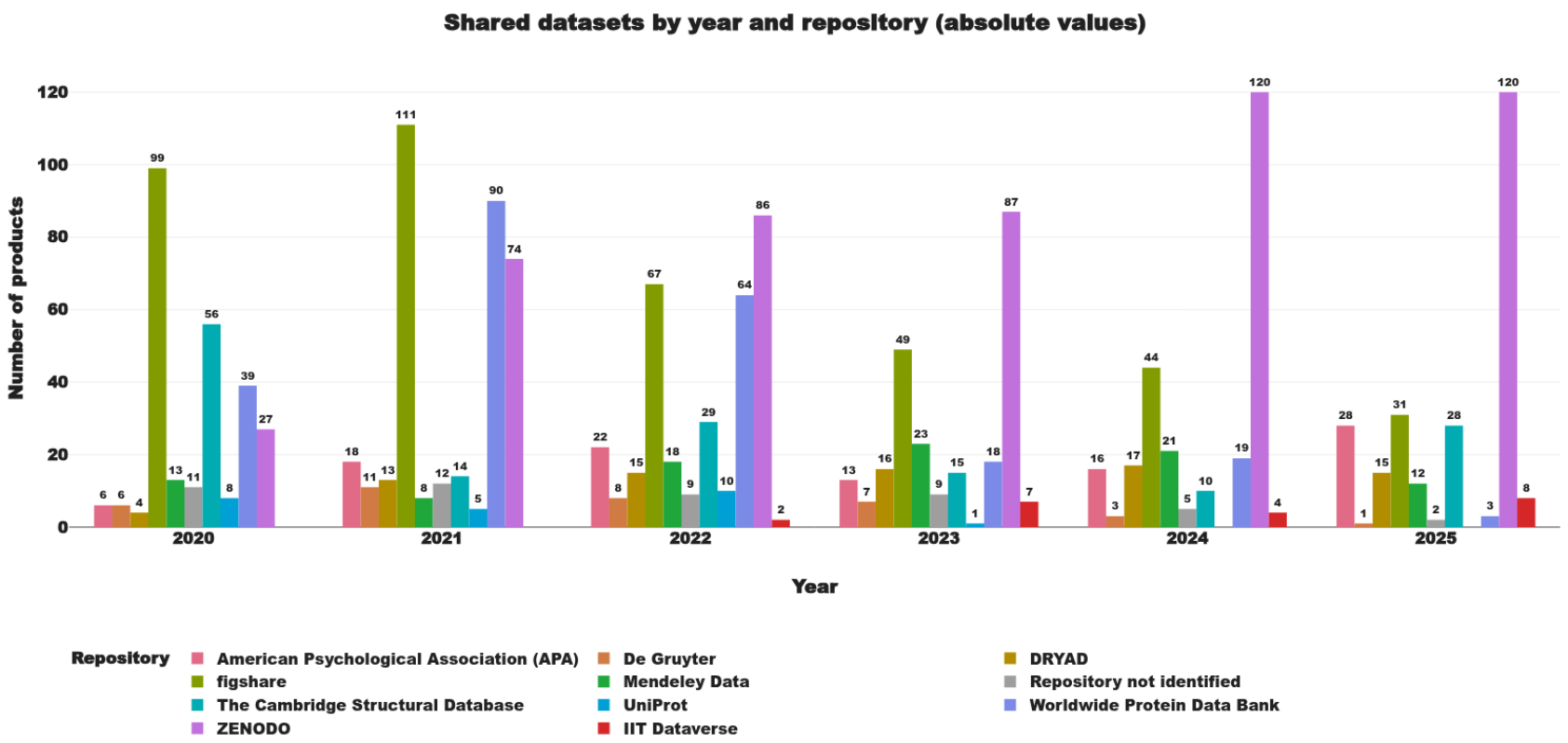
Filtri per gruppo e anno

Una barra laterale valida per tutte le viste

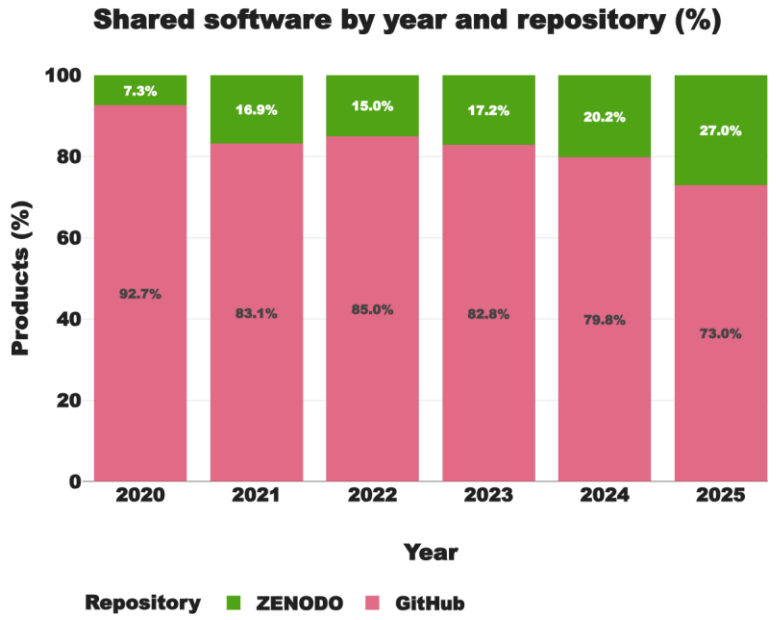
ISTITUTO ITALIANO DI TECNOLOGIA

Dove vengono condivisi dati e software?

Dataset per anno e repository



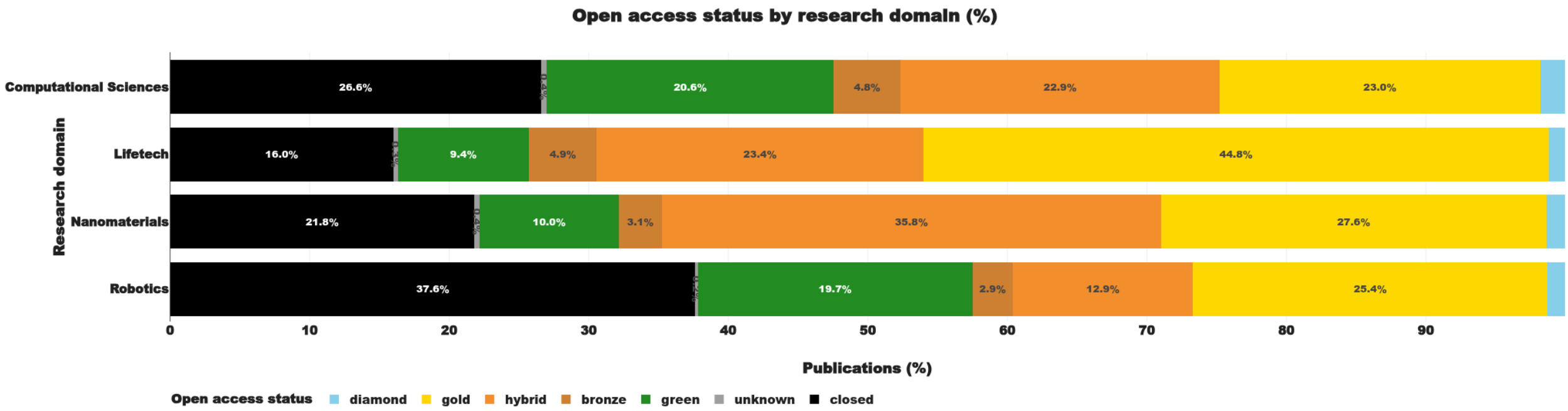
Software per anno e repository



Dati: Zenodo e figshare
Software: GitHub

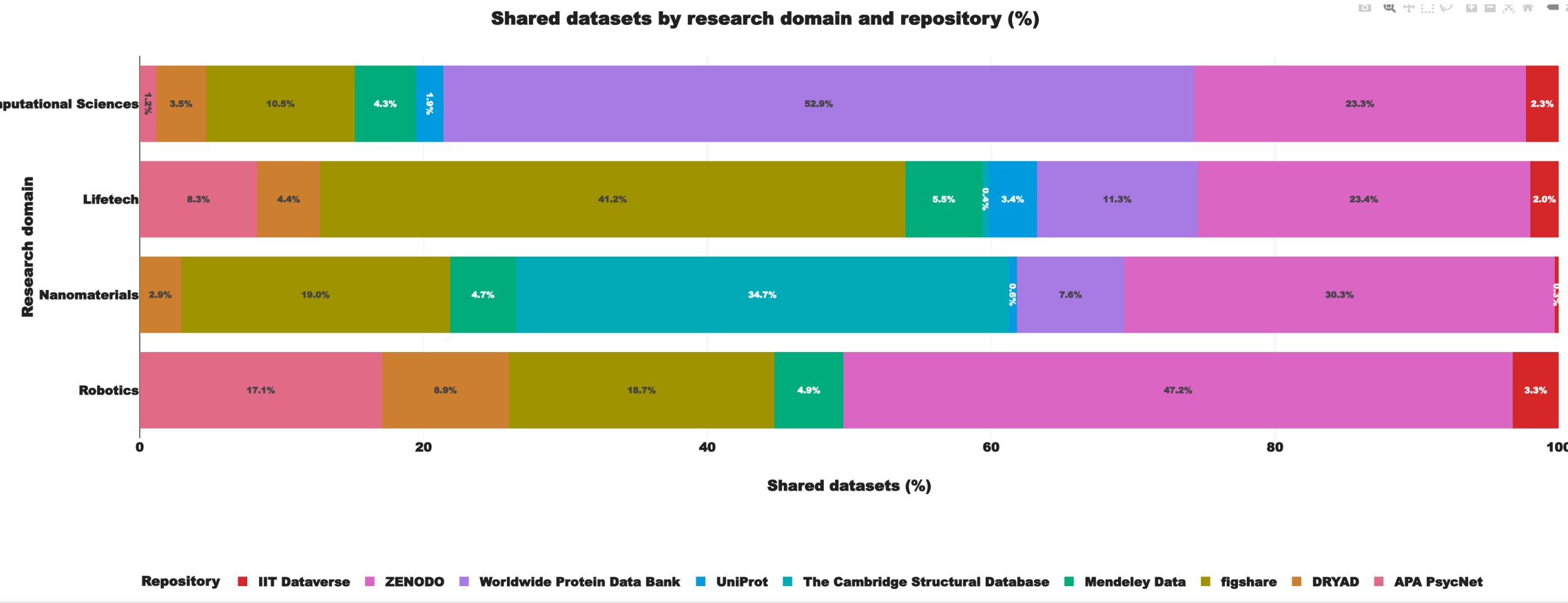
ISTITUTO ITALIANO DI TECNOLOGIA

Open Access per dominio di ricerca



ISTITUTO ITALIANO DI TECNOLOGIA

Open Data per dominio di ricerca



ISTITUTO ITALIANO DI TECNOLOGIA

F-UJI: valutazione automatica della FAIRness

Servizio open source per valutare la FAIRness dei dataset

Inserisci il DOI o l'URL del dataset

Raccoglie metadati da più fonti

Esegue test sulle metriche FAIRsFAIR

Calcola i punteggi FAIR

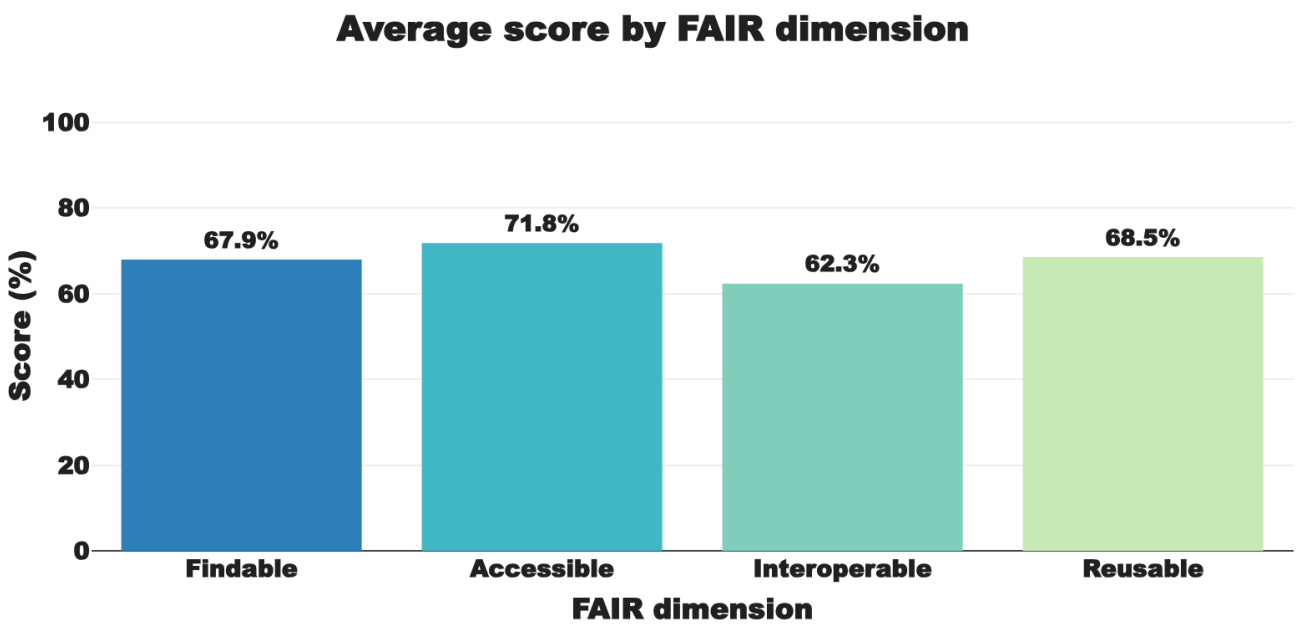
Restituisce un report in JSON o RDF

Indica dove migliorare la FAIRness del dataset

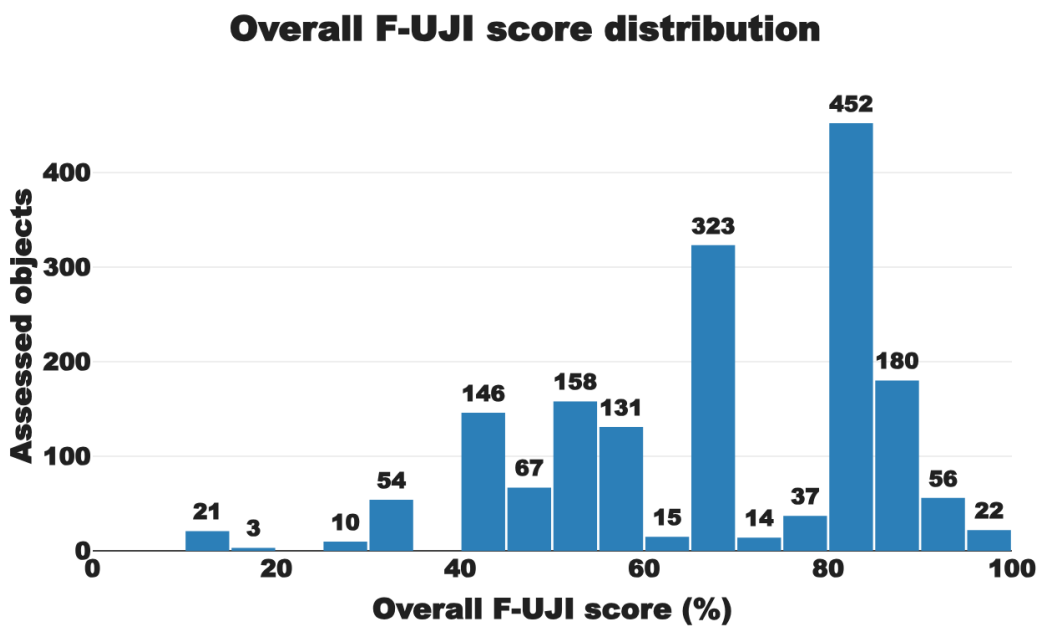
ISTITUTO ITALIANO DI TECNOLOGIA

Quanto sono FAIR i dati condivisi?

FAIR dimension scores



Overall F-UJI score distribution



1.690

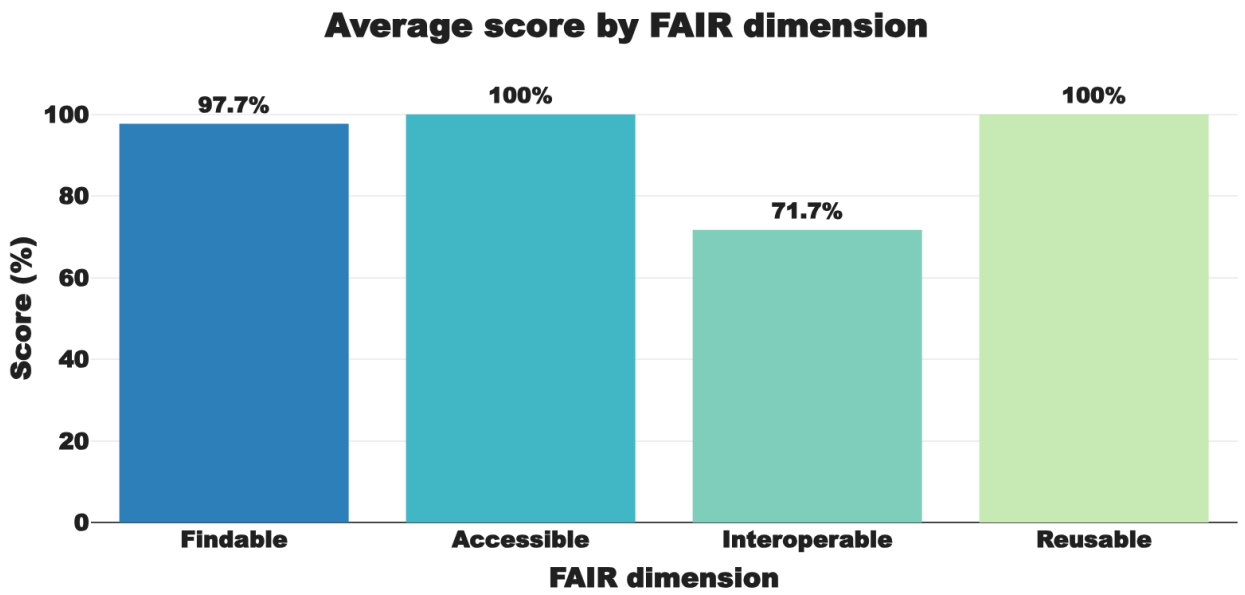
dataset individuati

69,2%

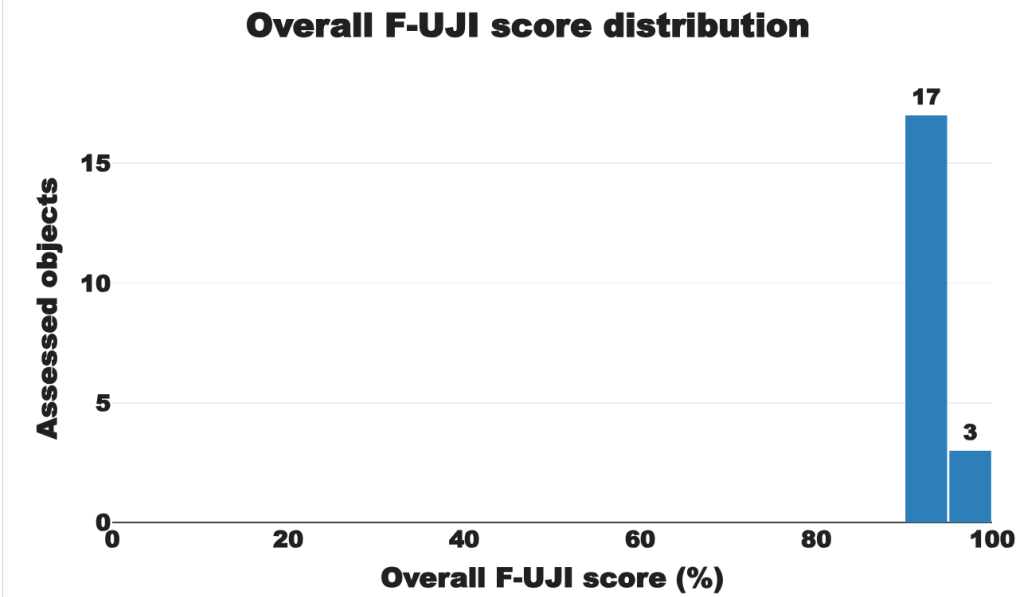
punteggio FAIR mediano

Il repository istituzionale IIT Dataverse

FAIR dimension scores

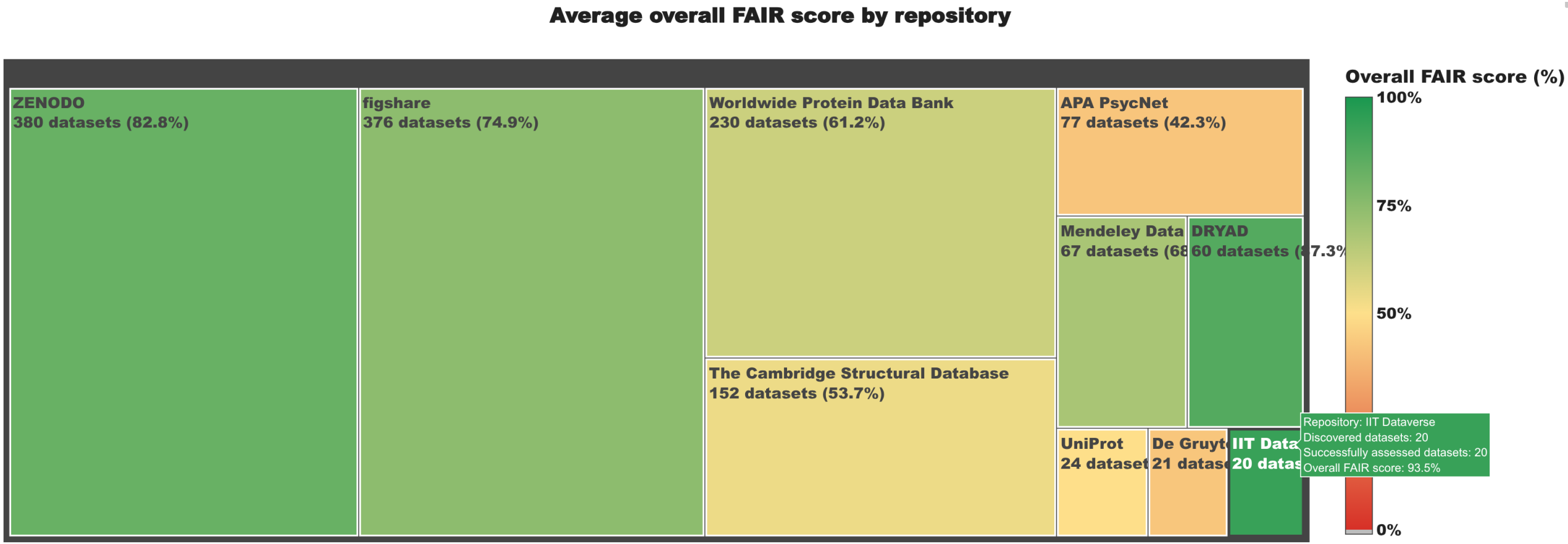


Overall F-UJI score distribution



ISTITUTO ITALIANO DI TECNOLOGIA

FAIRness by repository



ISTITUTO ITALIANO DI TECNOLOGIA

Dove l'IA ha accelerato il lavoro

Studiare API poco familiari

Sviluppare le integrazioni

Interpretare le risposte

Consolidare dati da fonti diverse

Realizzare test complessi

Tempi ridotti in modo significativo

Le competenze si spostano verso la validazione

Meno tempo su

- Studiare API e strumenti nuovi
- Sviluppare integrazioni
- Consolidare dati e scrivere test

Più impegno su

- Validazione dei risultati
- Controllo della qualità
- Attività a maggior valore aggiunto

Una dashboard per IIT Dataverse

Obiettivo: monitorare l'utilizzo del repository istituzionale dei dati di ricerca, IIT Dataverse, per orientare supporto e (in)formazione.

L'IA (ChatGPT) come strumento di supporto in tutte le fasi di sviluppo della dashboard (R/Shiny).

Pianificazione

Definizione dell'approccio e delle fasi di lavoro.

Estrazione dei dati

Interrogazione delle API e interpretazione delle risposte.

Elaborazione dei dati

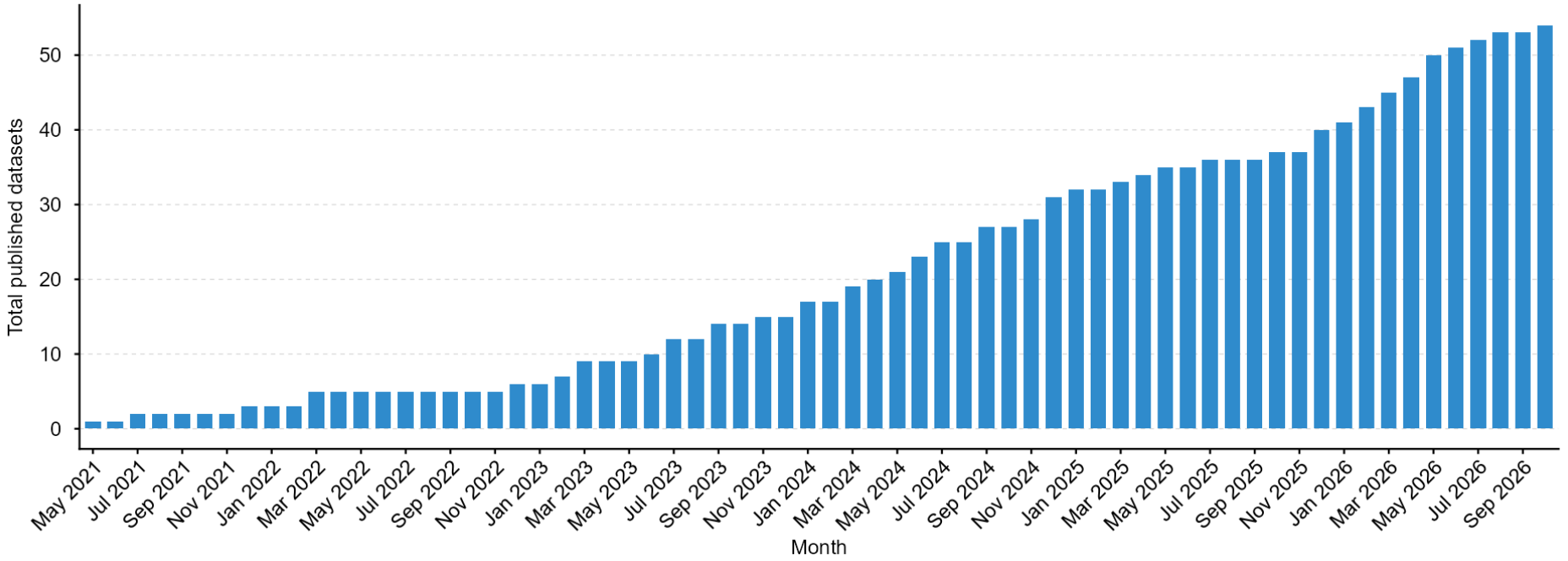
Pulizia e riorganizzazione dei dati in tabelle strutturate.

Sviluppo dashboard

Visualizzazione dei dati e implementazione dell'interfaccia grafica.

Dataverse usage across IIT

Overview published datasets over time

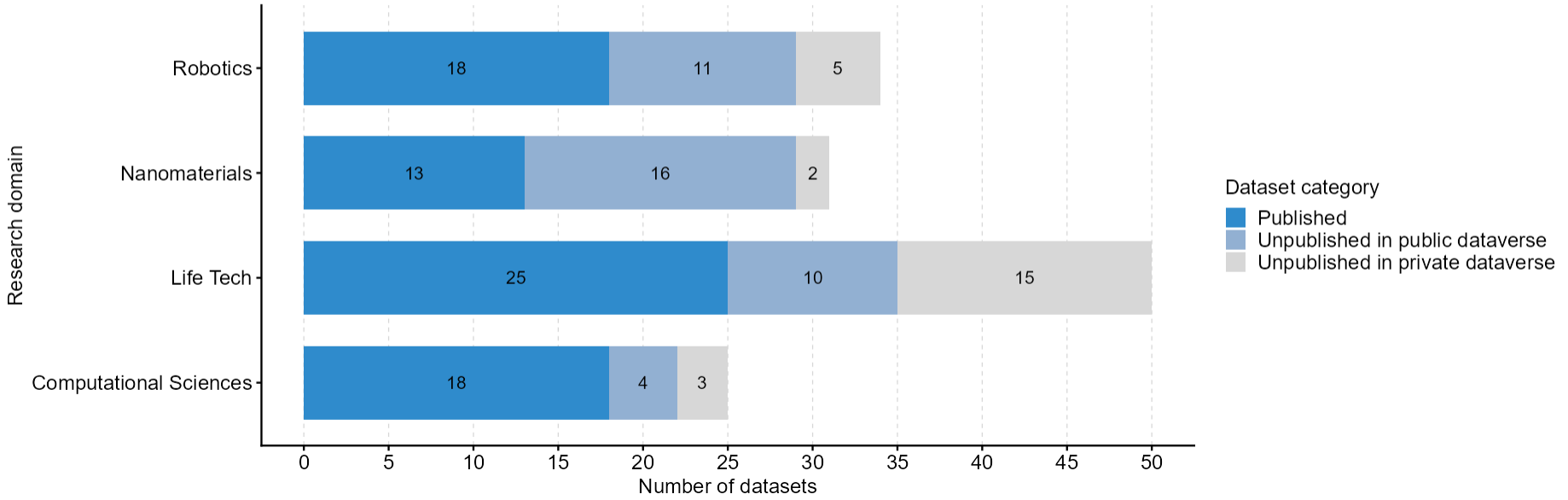


Overview of published and unpublished datasets

Published datasets	Unpublished datasets in public dataverses	Unpublished datasets in private dataverses
54	**35 (3 currently in review)**	**22**

ISTITUTO ITALIANO DI TECNOLOGIA

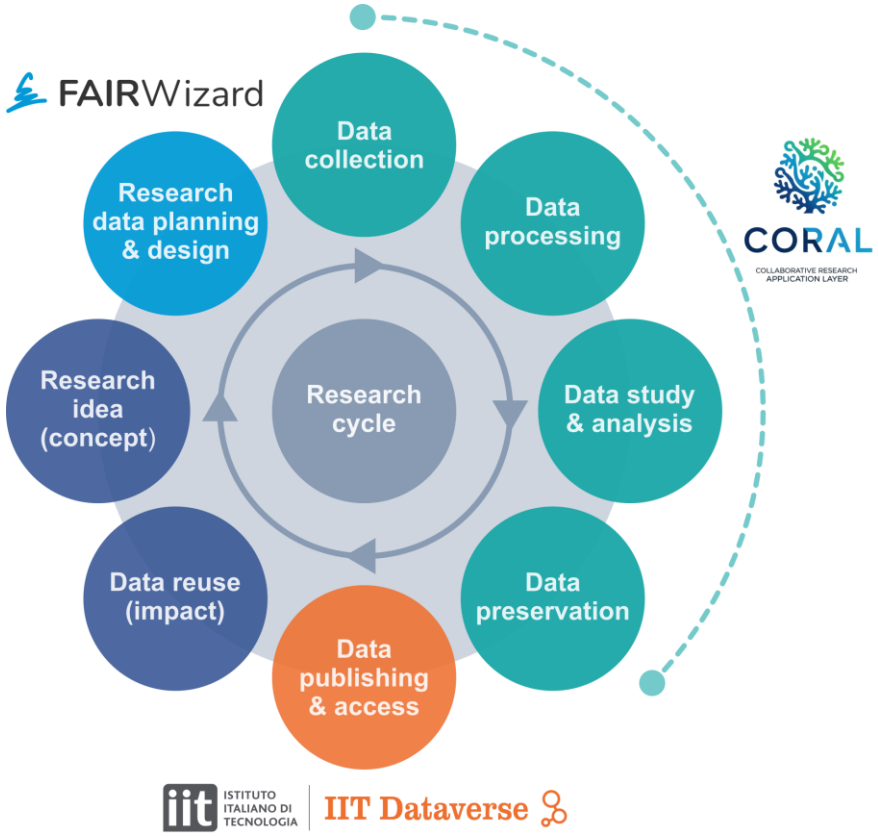
Dataverse usage across research domains

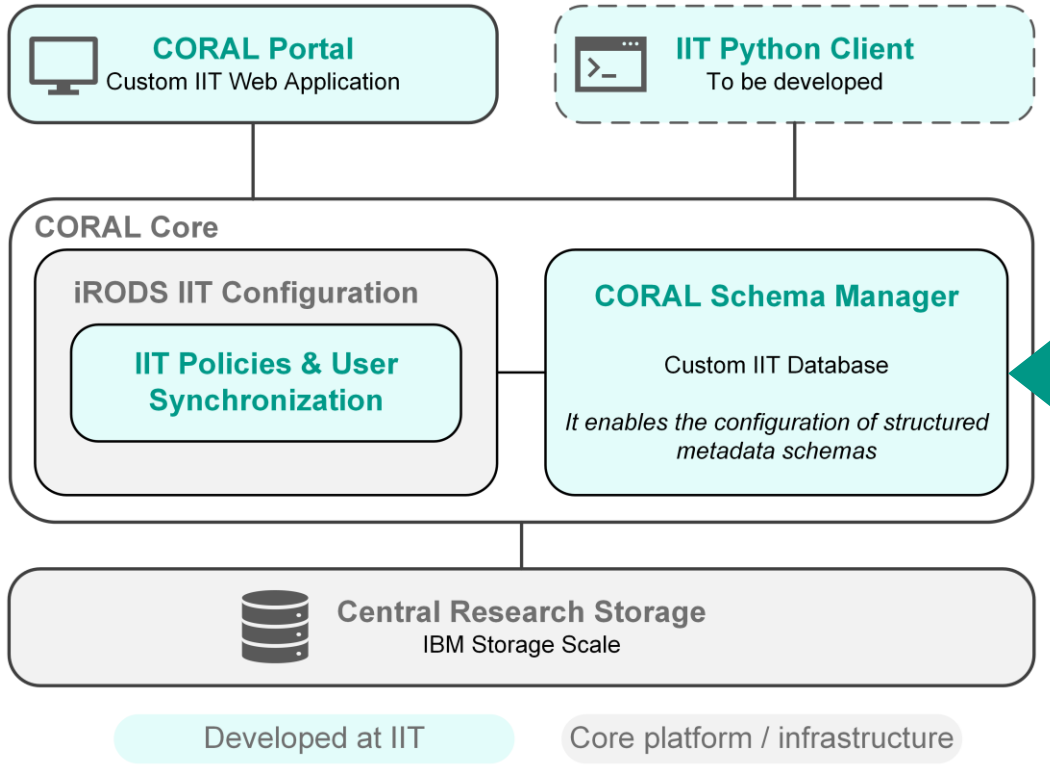


Dataverse usage across research units

- ✓ Spazio di archiviazione occupato
- ✓ Numero di dataset pubblicati
- ✓ Numero di dataset nel dataverse pubblico, ma non ancora pubblicati
- ✓ Numero di dataset nel dataverse privato

CORAL (COllaborative Research Application Layer)

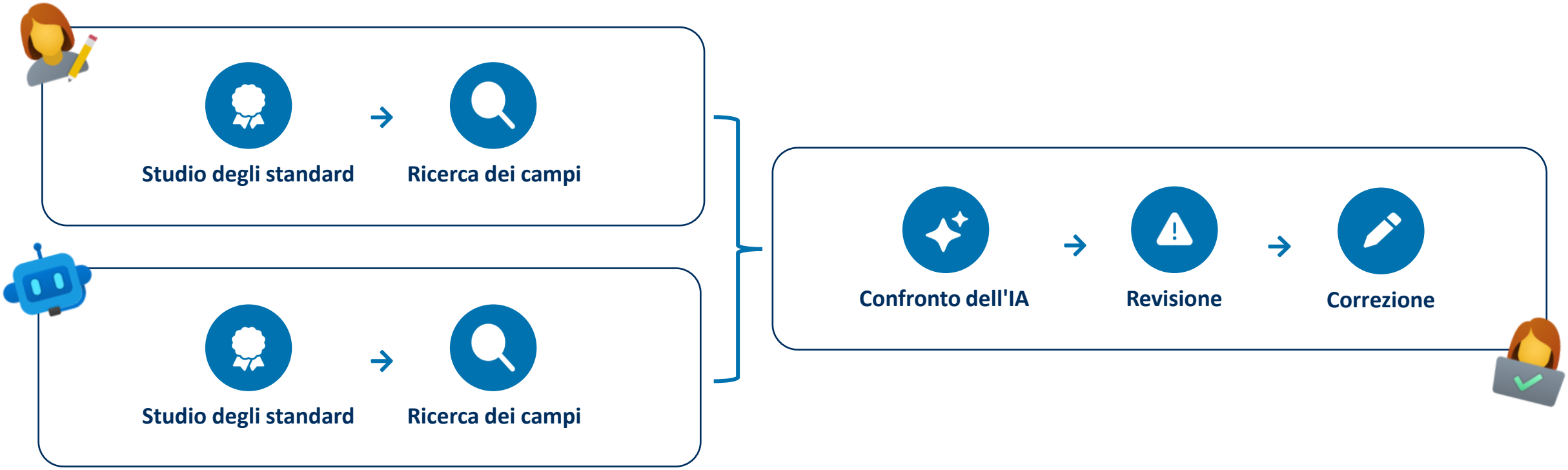




Necessità di definire schemi di metadati che consentano ai ricercatori IIT di documentare i dati in modo strutturato e uniforme.

ISTITUTO ITALIANO DI TECNOLOGIA

Doppi, tripli, quadrupli... Standard



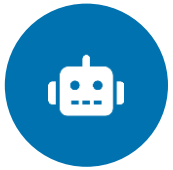
Esempio: data di nascita?

1+MG MMM

Croissant v1

...

Conclusioni



L'IA riduce i tempi

Sviluppo più rapido di dashboard, integrazioni e schemi di metadati



Le competenze umane restano

Servono per validare e controllare la qualità dei risultati



Nuove professionalità

Data steward, data analyst ed esperti Open Science sono centrali

Contatti



Data Analysis Office | Research Data Management Stewardship Office

data.analysis@iit.it | rdm@iit.it

Grazie per l'attenzione!

ISTITUTO ITALIANO DI TECNOLOGIA

ISTITUTO ITALIANO DI TECNOLOGIA
Sede Legale: Via Morego, 30 16163 Genova
Uffici di Roma: Via Guidubaldo del Monte, 54 00197 Roma
Tel. 010 2896
www.iit.it