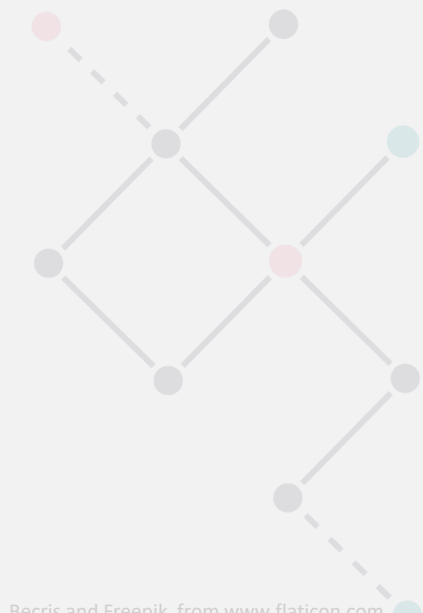
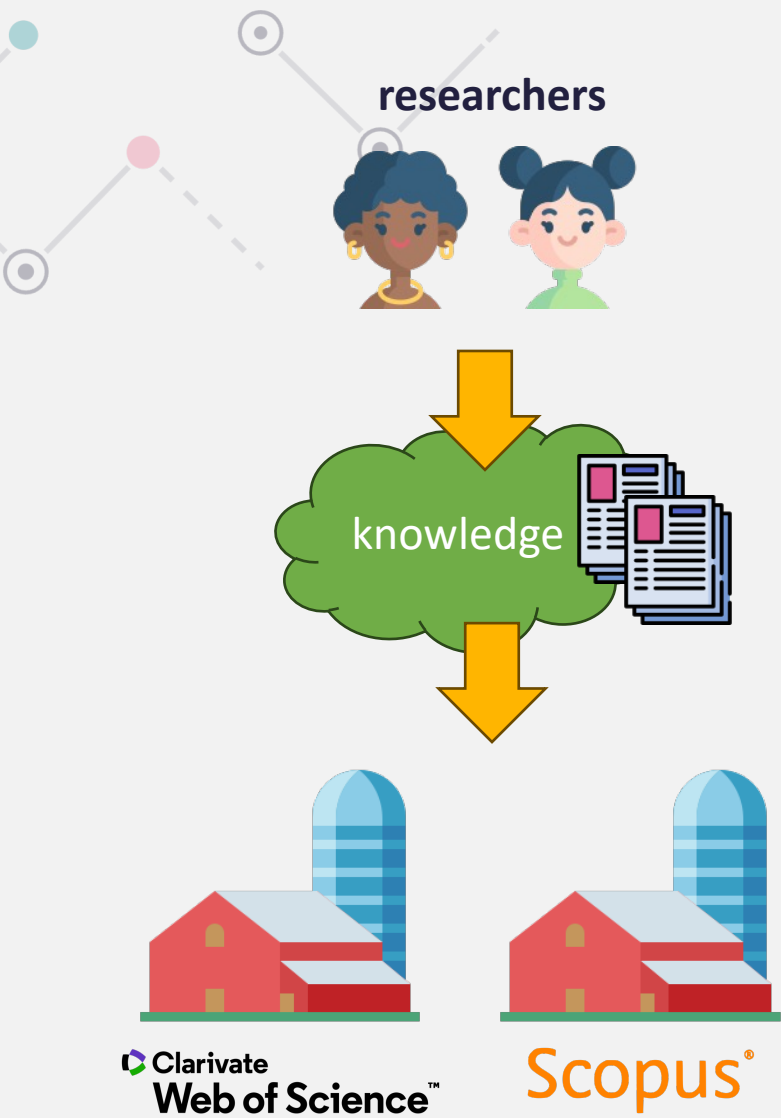




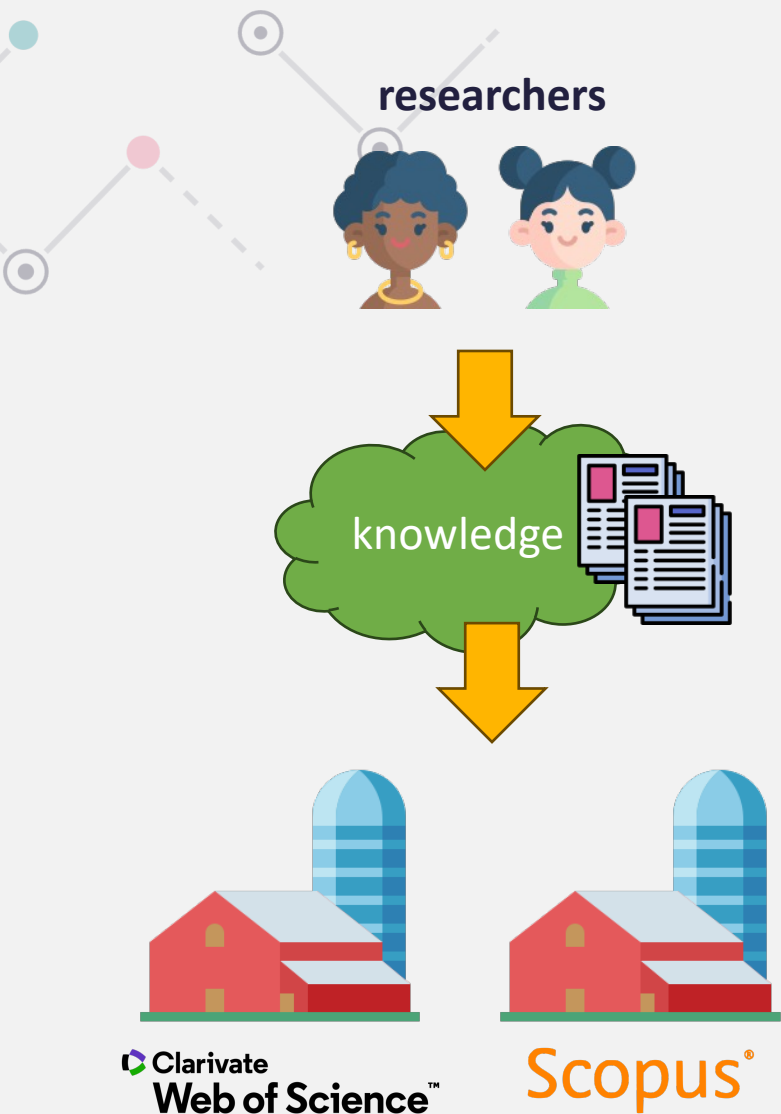
OpenAIRE Graph: status & enrichments via the SciLake project

Open Science Fair 2023, Madrid

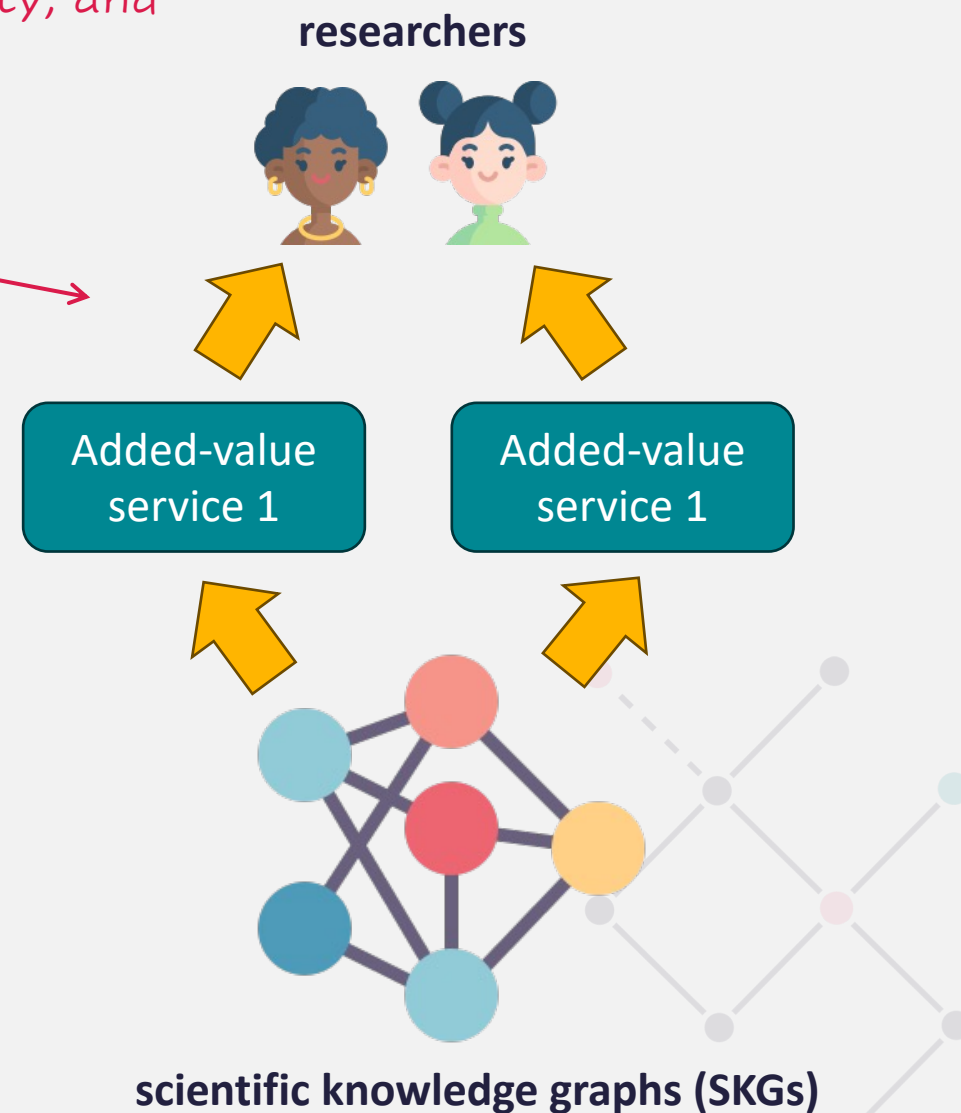
Thanasis Vergoulis, ATHENA RC



improved processes in scientific knowledge discovery, scientific credibility, and research assessment



Open Science

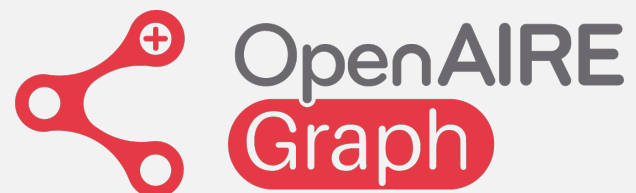




Scilake



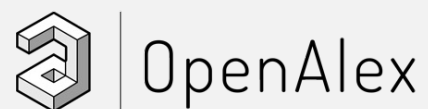
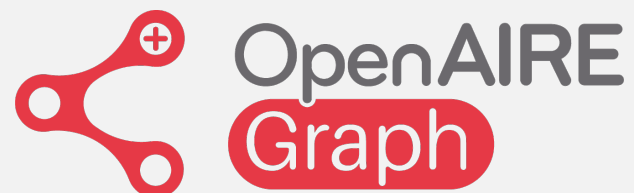
ORCID



Semantic Scholar



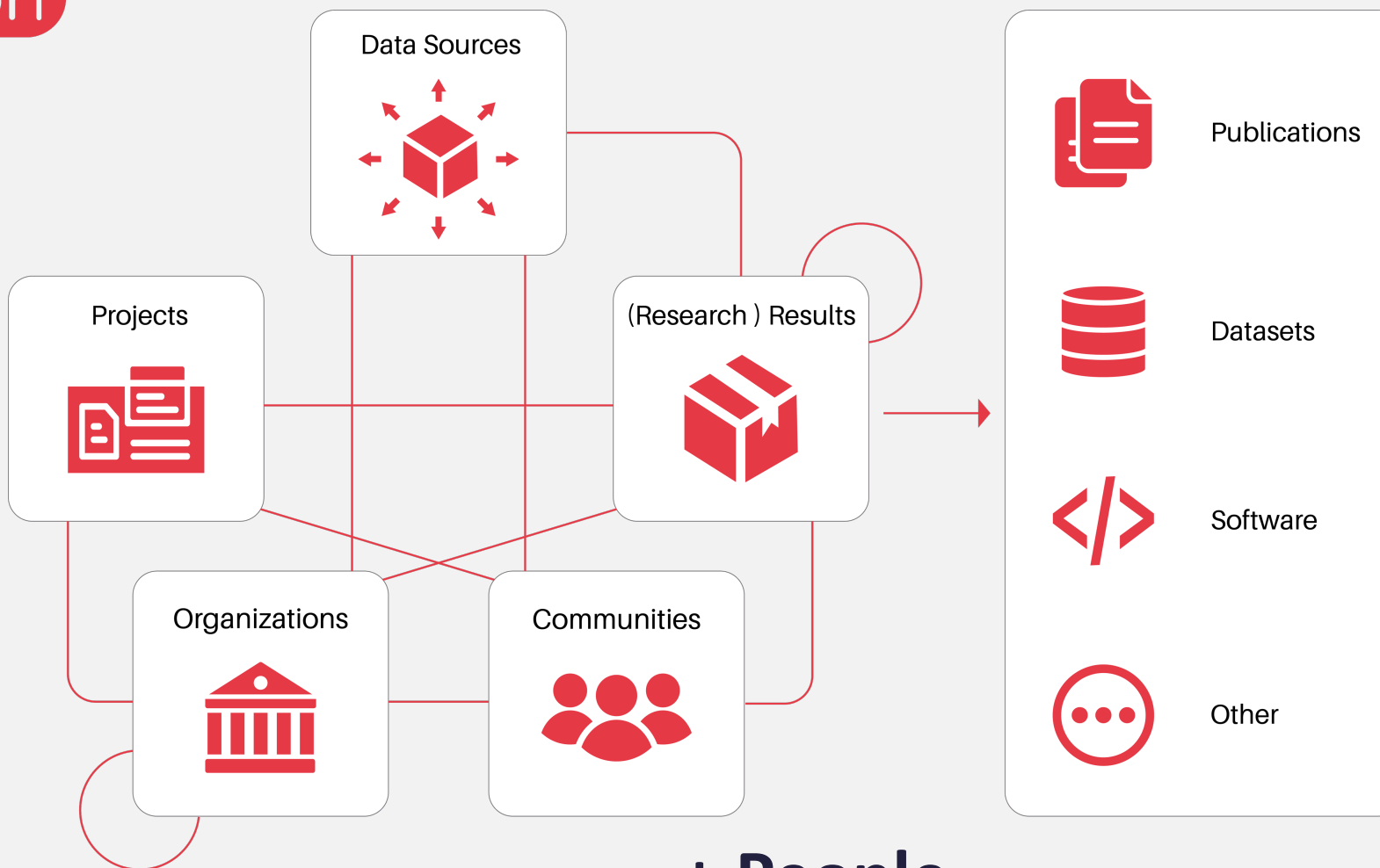
Solving The Problem Of Problem Solving™



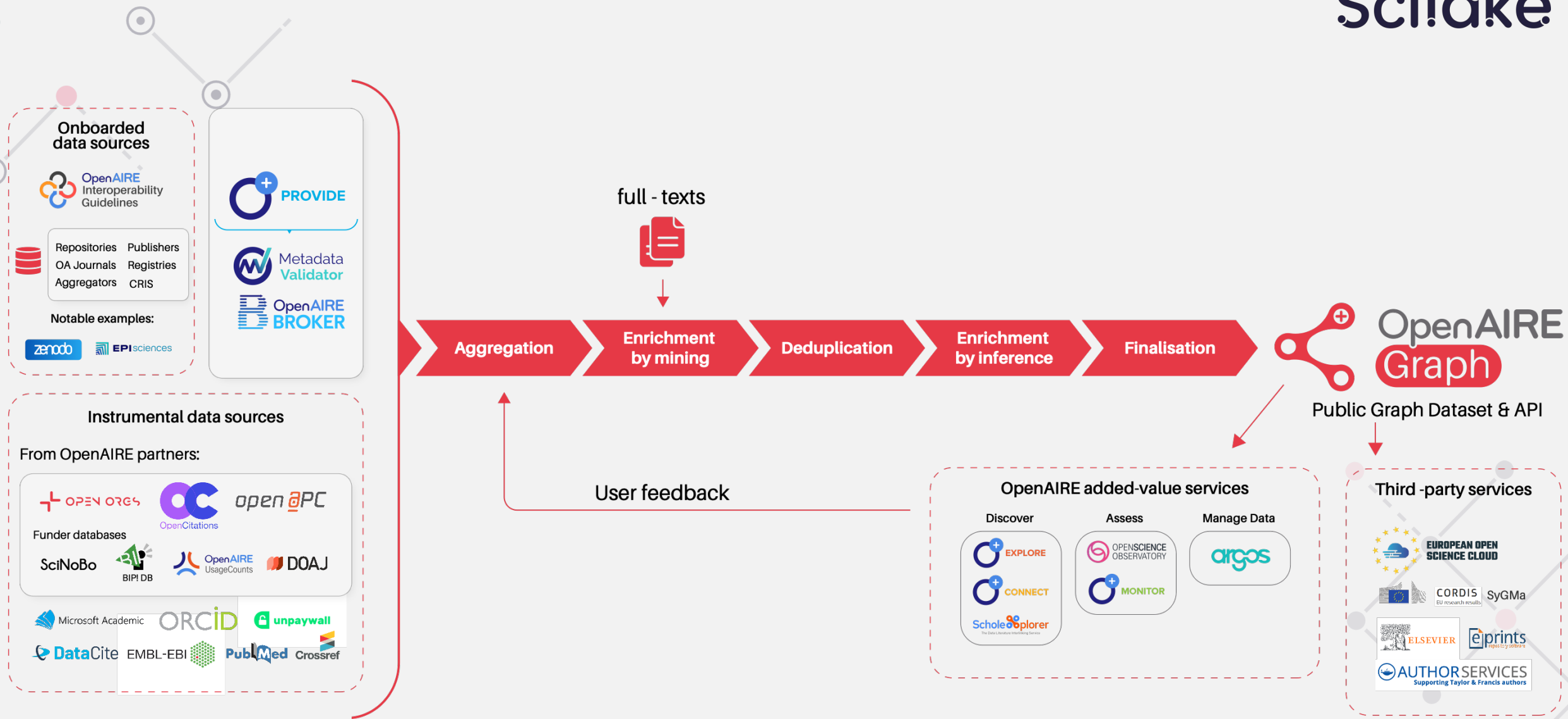
Semantic Scholar



OpenAIRE Graph



+ People
+ Venues





Challenges

- Difficult to **combine domain knowledge with domain-agnostic SKGs** (e.g., OpenAIRE Graph) & leverage them to build **domain-tailored added-value services**:
 - Knowledge fragmented in various locations.
 - Heterogeneous formats: unstructured files (e.g., textual data), databases and knowledge bases.
 - Non-English texts in some domains
 - Lack of technical expertise (e.g., graph & text mining) from domain-specific experts.
 - Not interoperable with domain-agnostic SKGs
 - Including everything in domain-agnostic SKGs is not sustainable.



At a glance

- SciLake aims to alleviate the previous problems introducing an **open, transparent & customizable Scientific-Lake-as-a-Service**
 - Data lake hosting domain-specific & domain-agnostic SKGs together with less structured domain knowledge
 - Empowers & facilitates the creation, interlinking & maintenance of community-created SKGs.
 - Offers a unified way to access & query the contained assets.
- Two discipline-tailored, **added-value services** to facilitate:
 - navigating the respective vast knowledge space (exploiting indicators of scientific impact)
 - improving research reproducibility in the respective domain
- Customise, test & demonstrate the developed services in **real-world pilots**.

Pilot 1: Neuroscience

Pilot 2: Cancer re-
search

Pilot 3: Transporta-
tion research

Pilot 4: Energy re-
search

9 countries / 13 organisations



Map created from <https://www.mapchart.net>



9 countries / 13 organisations



Map created from <https://www.mapchart.net>





Sneak peak on services

Scientific Lake service

- Main components:
 - SKG query engine
 - SKG creation assistant
 - Text & graph mining
 - Automatic translation module
- Main contents:
 - OpenAIRE Graph
 - Full-text Open Access Manuscripts
 - Domain-specific texts
 - Domain-specific SKGs

Impact-based knowledge discovery service

- Exploit citation analysis to calculate impact indicators for research products.
- Use these indicators to help researchers explore the vast knowledge space of their domain.
- Exploit domain-specific knowledge from the SKGs.
- Offer tailored functionalities to the respective communities (e.g., search presets, domain-specific annotations).

Reproducibility Assistance service

- Exploit natural language & graph processing technologies to estimate reproducibility of research works.
- Give back recommendations on how to make research more reproducible (e.g., recommend missing links to datasets & code).
- Exploit domain-specific knowledge from the SKGs.
- Offer tailored functionalities to the respective communities.

SciLake

Contact us

Email: vergoulis@athenarc.gr

Twitter/X: [@vergoulis](https://twitter.com/vergoulis)

Thank you

Mastodon: [@vergoulis@scicomm.xyz](https://mastodon.social/@vergoulis@scicomm.xyz)

