Introducing the Data Publishing Services WG

How to move from a plethora of (mostly) bilateral arrangements to a one-for-all service model infrastructure for the research data publication landscape?

- Increase interoperability
- Decrease systemic inefficiencies
- Power new tools and functionalities to the benefit of researchers
Main deliverable: an open, universal cross-linking service

Primary Focus:
Universal literature – data cross-linking service

- Given article A, what relevant data D exists – and vice versa
- Additional metadata about the nature of the relationship, e.g. supplementary data, related data, formal citation.
- Additional metadata for article and/or data set
Linking data and the literature

Why? Improve visibility, discoverability, re-use and reproducibility

Many organizations are already doing this – for example:

• **Data repositories** keep track of articles that cite, or refer to, their data
• **Publishers** devise applications to link the articles they publish with data hosted externally
• **Providers of bibliographic information** (i.e. repositories) are increasingly looking at data alongside the traditional article output
• **Organizations such as CrossRef, DataCite and OpenAIRE** are developing systems to track or infer relationships between data and the literature

But it’s all very scattered, which limits the value!
Deliverable: We want to bring existing article/data links together, normalize them using a common schema, and expose the full set as an open service.

Philosophy:
- Open, collaborative & inclusive, cross-stakeholder
- Domain-agnostic, (aim for) comprehensiveness
- Quality and provenance is key
- Flexible, test & learn, hands-on
High-level architecture

Search APIs  OAI-PMH  Web Portal  OAI-PMH intersection

Information Space

PID resolving  De-duplicating  Harmonizing

Links collection

Core Data Model

Over 1M links!

Data Sources (authoritative)

Examples:
- Pairs of DOIs
- DataCite records
- PANGAEA records
Data-Linking Service – prototype

Give it a spin: [http://dliservice.research-infrastructures.eu](http://dliservice.research-infrastructures.eu)

Powered by OpenAIRE D-NET software and PANGAEA search engine

Connected to RD-switchboard
Next steps

- Develop tailored API’s and services to meet use cases from prospective users: data centers, publishers, etc.
- Embed service in real-life situation
- Additional article-data link contributors
- De-duplication

Hope to see you at the WG session, today @ 15:30 pm