



# Scholarly Link Exchange Scholix RDA 10<sup>th</sup> plenary Montreal

research data sharing without barriers  
[rd-alliance.org](http://rd-alliance.org)

# Agenda

- Quick recap of Scholix
- Learning through real implementations  
10x examples & feedback
- Community materials 'how to'
- Finalizing the guidelines: the last 10%  
takes 90% of our time
- Joint assessment: where do we stand

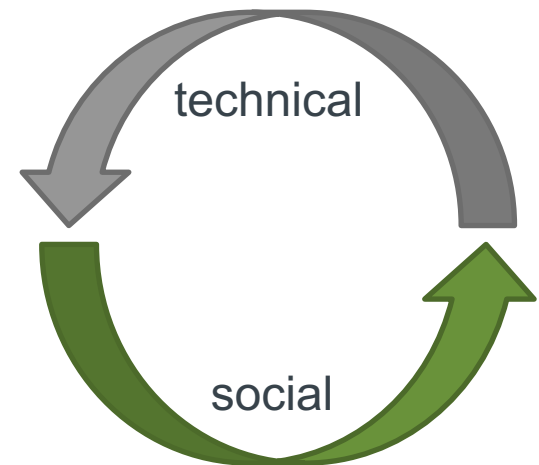
# Recap: what's the problem?

3

*Linking Research Data with the Literature* is of great value, yet current solutions are not realizing the potential

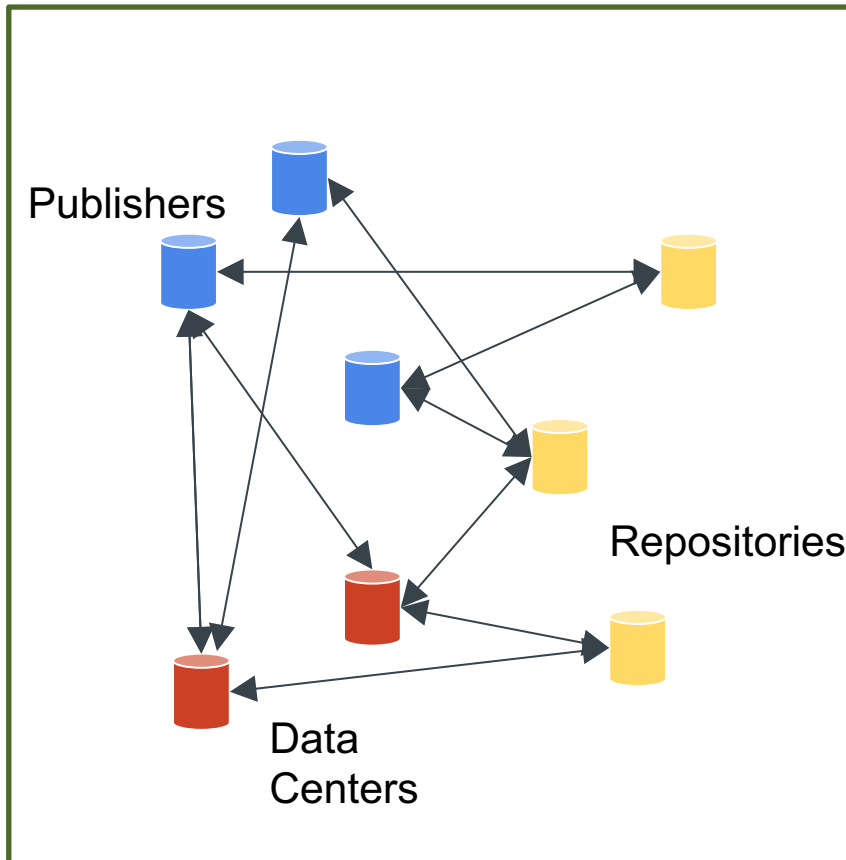
## What is the problem?

1. Many disconnected sources (publishers, data centers, repositories, infrastructure providers, ...)
2. Heterogeneity of practices, for example:
  - Different PID systems (DOI, accession numbers)
  - Different ways of referencing data (formal citations, in-text references, ...)
  - Different moments of citing data (at publication, post publication, ...)

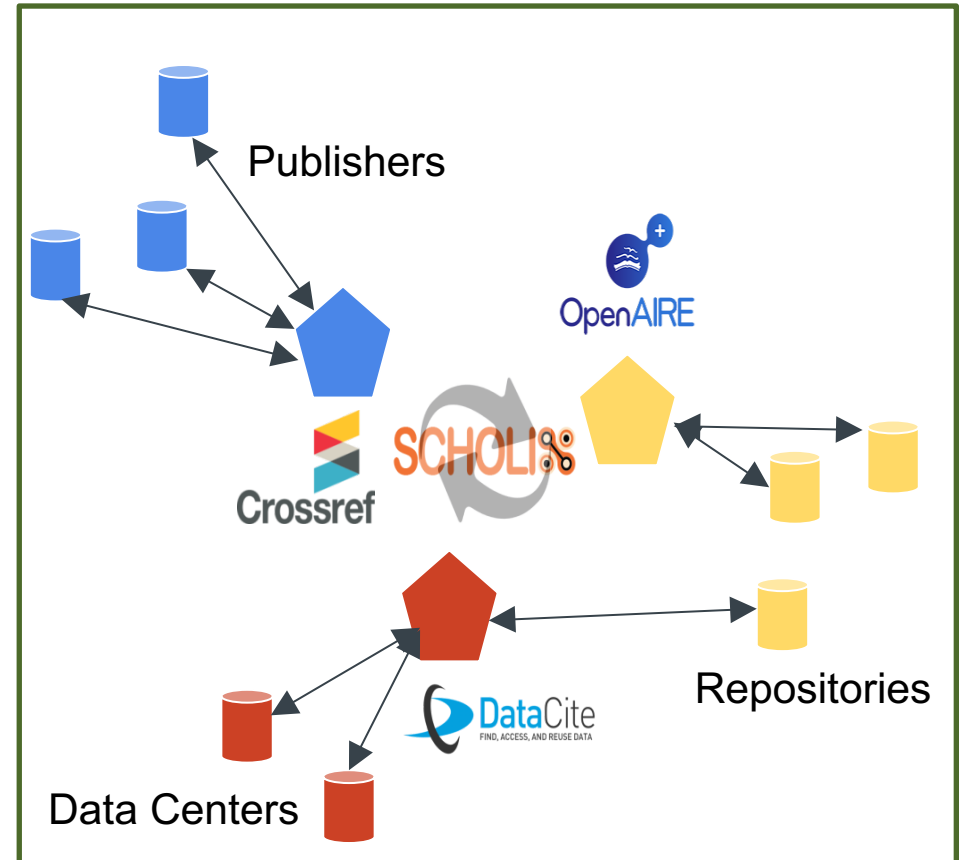


# Recap Scholix: connecting the dots

**Past:** disconnected sources using heterogeneity of practices



**Future:** standard set of guidelines for exposing and consuming links, supported by hubs



# Recap: Scholix Guidelines

- A framework for standardizing the exchange of *scholarly link* information between scholarly infrastructure providers
  - Information Model for scholarly links representation
  - Recommendation and provision of exchange formats and protocols



See also <http://www.scholix.org/guidelines>

# Recap of Scholix benefits

see: [www.Scholix.org](http://www.Scholix.org)

6



## For data repositories and journal publishers

- linking data and the literature will increase their visibility and usage
- can support additional services to improve the user experience on online platforms
- More scalable and robust due to a global standard rather than bilateral agreements

## For research institutes, bibliographic service providers, and funding bodies

- the infrastructure will enable advanced bibliographic services and productivity assessment tools
- track datasets and journal publications within a common and comprehensive framework

## For researchers:

- Easier finding and accessing relevant articles and data sets
- track long-term impact of their data (and publications)
- thereby providing additional incentives to share data.

# Example: Scholix at Europe PMC

Florian Graef

7

- Different origin Data-Literature links in various places in API and User Interface
- Consolidation into one API method providing links in Scholix format

☐ A tick salivary protein targets cathepsin G and chymase and inflammation and platelet aggregation.

(PMID:20940421 PMID:PMC3031492)

Abstract

Citations

Related Articles

Data

BioEntities

External Links

## Data behind this article



**BioStudies.** Primary data and supplemental files

<http://www.ebi.ac.uk/biostudies/studies/S-EPMC3031492>

Figures are available in the [full text of the article](#)

## Data associated with this article

[4 UniProt records that cite this article](#)

[1 PDB record that cites this article](#)

[4 ENA records that cite this article](#)

[2 OMIM records that cite this article](#)

EuropePMC obtains data-literature links in a few ways:

-DB-Crossreferences (external data records cite a publication and tell us about it (e.g. PDBe)

-Text mined accessions – extracted by our text mining pipeline in publications (PDBe, ENA)

-External links – various entities provide us with links to resources which are related to individual publications (e.g. Altmetrics, Wikipedia, Publons)

There are differences in directionality/ the way we obtain them but all are data-literature links

-> Consolidation of the API to provide all data in one response (Scholix format)

-> Will gradually replace current API methods starting with a single Tab summing up data cited/produced in a study

Internal testing is ongoing

research data sharing without barriers  
rd-alliance.org

Questions?   
[graf@ebi.ac.uk](mailto:graf@ebi.ac.uk) RESEARCH DATA ALLIANCE



# Example: Scopus.com (articles/citations) => dataset

Eleonora Presani

8

Scopus

Search

Sources

Alerts

Lists

Help

SciVal

Register

Login



## Document details

Back to results | 1 of 1

Journal Finder | Library | Copac | Order Document | Export | Download | Add to List | More...

Acta Crystallographica Section C: Structural Chemistry

Volume 71, 1 October 2015, Pages 900-902

### Halogen-bonded adduct of 1,2-dibromo-1,1,2,2-tetrafluoroethane and 1,4-diazabicyclo[2.2.2]octane (Article)

Brisdon, A.K., Muneer, A.M.T., Pritchard, R.G.

School of Chemistry, University of Manchester, Oxford Road, Manchester, United Kingdom

#### Abstract

Halogen bonding is an intermolecular interaction capable of being used to direct extended structures. Typical halogen-bonding systems involve a noncovalent interaction between a Lewis base, such as an amine, as an acceptor and a halogen atom of a halofluorocarbon as a donor. Vapour-phase diffusion of 1,4-diazabicyclo[2.2.2]octane (DABCO) with 1,2-dibromotetrafluoroethane results in crystals of the 1:1 adduct,  $C_2Br_2F_4C_8H_{12}N_2$ , which crystallizes as an infinite one-dimensional polymeric structure linked by intermolecular N...Br halogen bonds [2.829 (3) Å], which are 0.57 Å shorter than the sum of the van der Waals radii. © 2015 International Union of Crystallography.

#### Author keywords

1,4-diazabicyclo[2.2.2]octane; bromofluorocarbon trapping; crystal structure; dibromotetrafluoroethane; extended structures; halogen bonding; one-dimensional polymeric structure; phase change

#### Indexed keywords

**Engineering controlled terms:** Bromine; Fluorine; Ions; Polymers; Van der Waals forces

1,4-diazabicyclo[2.2.2]octane; bromofluorocarbon trapping; Dibromotetrafluoroethane; Extended structures; Halogen bonding; Phase Change; Polymeric structures

**Engineering main heading:** Crystal structure

ISSN: 20532296 CODEN: ACSCE Source Type: Journal Original language: English

DOI: 10.1107/S2053229615016472 Document Type: Article

Publisher: International Union of Crystallography

References (18)

- Link to external repository (in this case CCDC)
- Powered by DLI

Cited by 0 documents

Inform me when this document is cited in Scopus:

Set citation alert | Set citation feed

#### Related Research Data

##### CCDC 1422107: Experimental Crystal Structure Determination

Brisdon, A.K., Muneer, A.M.T., Pritchard, R.G.

Date of Collection: 2015

Cambridge Crystallographic Data Centre

Data linking provided by

#### Related documents

##### Crystal structure of triphenyl(vinyl)phosphonium tetraphenylborate

Bradfield, J.L., Braun, R.A., White, F. (2014) Acta Crystallographica Section E: Structure Reports Online

##### Type II halogen...halogen contacts are halogen bonds

Metrangolo, P., Resnati, G. (2014) IUCrJ

##### Raman spectra of triethylenediamine at T=298K

Sauvajol, J.L. (1980) Journal of Physics C: Solid State Physics

View all related documents based on references



## Document details

1 of 1

[Export](#) [Download](#) [Print](#) [E-mail](#) [Save to PDF](#) [Add to List](#) [More... >](#)[Full Text](#) [Copac](#) [BioRxiv](#)

Nuclear Instruments and Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms  
Volume 346, 1 March 2015, Pages 26-44

Activation cross-sections of proton induced reactions on  $^{nat}\text{Sm}$  up to 65 MeV (Article)Tárkányi, F.<sup>a</sup>, Hermanne, A.<sup>b</sup>, Takács, S.<sup>a</sup>, Ditrói, F.<sup>a</sup>, Ignatyuk, A.V.<sup>c</sup><sup>a</sup>Institute for Nuclear Research, Hungarian Academy of Sciences (ATOMKI), Debrecen, Hungary<sup>b</sup>Cyclotron Laboratory, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, Brussels, Belgium<sup>c</sup>Institute of Physics and Power Engineering (IPPE), Obninsk, Russian Federation

## Abstract

[View references \(24\)](#)

Activation cross sections for proton induced reactions on Sm are presented for the first time for  $^{nat}\text{Sm}(p,xn)^{154,152m2,152m1,152g,150m,150g,149,148,147,146,145}\text{Eu}$ ,  $^{nat}\text{Sm}(p,x)^{153,145}\text{Sm}$ ,  $^{nat}\text{Sm}(p,x)^{151,150,149,148g,148m,146,144,143}\text{Pm}$  and  $^{nat}\text{Sm}(p,x)^{141}\text{Nd}$  up to 65 MeV. The cross sections were measured via activation method by using a stacked-foil irradiation technique and high resolution gamma ray spectroscopy. The results were compared with results of the nuclear reaction codes ALICE, EMPIRE and TALYS (results taken from TENDL libraries). Integral yields of the activation products were calculated from the excitation functions. © 2015 Elsevier B.V.

## Author keywords

[Cross-section measurement](#) [Europium samarium promethium and neodymium radio-isotopes](#) [Proton irradiation](#) [Samarium target](#) [Yield calculation](#)

## Indexed keywords

Engineering controlled terms:

[Chemical activation](#)[Protons](#) [Samarium](#)[Activation method](#)[Activation products](#)[Cross-section measurement](#)[Excitation function](#)[Gamma-ray spectroscopy](#)[High resolution](#) [Integral yields](#)[Proton induced reaction](#)

Engineering main heading:

[Nuclear reactions](#)

- Example2: Scopus => Mendeley Data repository
- also powered by DLI

## Metrics ⓘ

[View all metrics >](#)

1 Citation in Scopus

0 Field-Weighted Citation Impact



PlumX Metrics

Usage, Captures, Mentions,  
Social Media and Citations  
beyond Scopus.

## Cited by 1 document

Excitation functions of proton induced reactions on  $^{nat}\text{Os}$  up to 65 MeV: Experiments and comparison with results from theoretical codes

Hermanne, A., Adam Rebeles, R., Tárkányi, F.  
(2015) Nuclear Instruments and Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms

[View details of this citation](#)

Inform me when this document is cited in Scopus:

[Set citation alert >](#)[Set citation feed >](#)

## Related research data ⓘ

Data for: Activation cross-sections of proton induced reactions on  $^{nat}\text{Sm}$  up to 65MeV

[Mendeley](#)

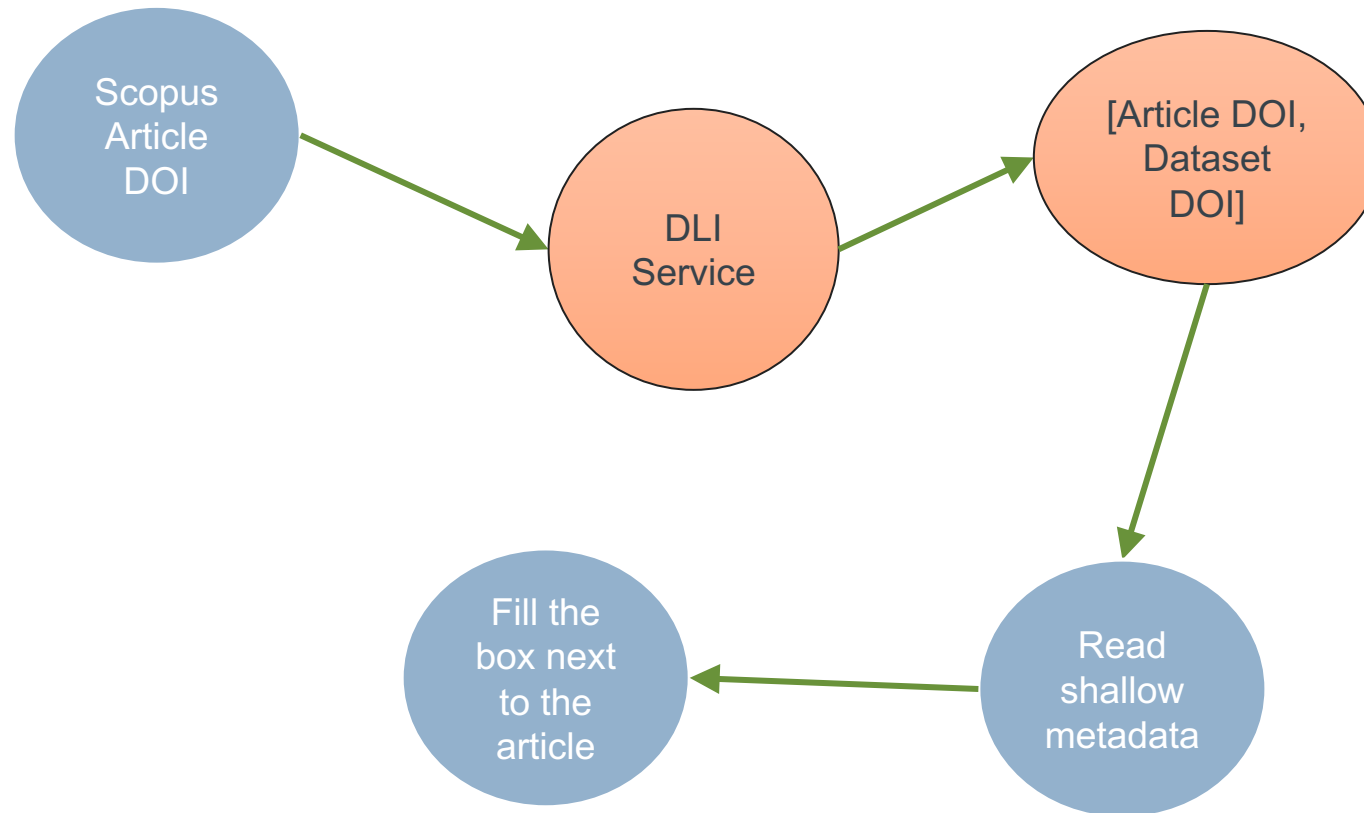
Data linking provided by

# Scopus and DLI Service within Scholix

*Eleonora Presani*

10

- For each document entry in Scopus (with DOI) DLI receives a query



***Eleonora Presani***

- 11

[illegible]

- A clearer information on where the data is actually stored (URL resolution)
- A recipe to retrieve deeper metadata from the data repository
  - Dataset description / abstract
  - Version
  - References (if any)
  - Keywords
  - Data type
  - ....
- Maybe contributing repositories can offer a key to retrieve those from them?

```
- {  
  facet: "Content Provider",  
  query_name: "provenance",  
  - values: [  
    - {  
      key: "CrossRef",  
      value: 1  
    },  
    - {  
      key: "Mendeley Data and published articles",  
      value: 1  
    }  
  ]  
}
```

# Example: Dryad participation in the Scholix initiative

Elizabeth Hull & Fiona Murphy

*Dryad's vision is to promote a world where research data is openly available, integrated with the scholarly literature, and routinely re-used to create knowledge.*

- As a repository specifically for data associated with publications, recognize the value of a common, global approach to these linkages, and pleased to be an early implementer
- Launched in July 2017 and applied to all existing data packages - **18,000 article links and counting**

Work in progress / for discussion:

- Traditional use of / preference for “IsReferencedBy” field for article DOIs over “IsSupplementTo,” but currently reporting to DataCite in both fields



research data sharing without barriers  
rd-alliance.org



<http://datadryad.org>



# ScienceDirect/CCDC Example

Ian Bruno, Helena Cousijn

14

ScienceDirect Journals Books Wouter Hark Help

Download PDF Export Search ScienceDirect Advanced search

Article outline Show full outline

Highlights Abstract Graphical abstract Keywords 1. Introduction 2. Results and discussion 3. Conclusions 4. Experimental Acknowledgments Appendix A. Supplementary data References

Figures and tables Table 1

Original article

European Journal of Medicinal Chemistry Volume 96, 26 May 2015, Pages 281–295

Synthesis, crystal structure and effect of indeno[1,2-b]indole derivatives on prostate cancer *in vitro*. Potential effect against MMP-9

Gricela Lobo<sup>a</sup>, Melina Monasterios<sup>a</sup>, Juan Rodriguez<sup>a</sup>, Neira Gamboa<sup>a</sup>, Mario V. Capparelli<sup>a</sup>, Javier Martinez-Cuevas<sup>a</sup>, Michael Lein<sup>a,\*</sup>, Klaus Jung<sup>a,\*</sup>, Claudia Abramjuk<sup>a,†</sup>, Jaime Charris<sup>a</sup>

doi: 10.1016/j.ejmech.2015.04.023 Get rights and content

Highlights

- Compounds were easily synthesized and with highly regioselectivity.
- Crystals consist of equimolar mixtures of the RR and SS diastereomers.
- All tested compounds proved to be moderately active, except one.

Data for this Article

CCDC Cambridge Crystallographic Data Centre: Crystallographic data

Currently linking between article and dataset is facilitated by bespoke Elsevier and CCDC services

## Plan is to pilot linking using the DLI service

- Concerns: timeliness of link being available relative to article being published - more systems need to update compared to current mechanisms
- Benefit: CCDC will no longer need to maintain services provided to specifically support Elsevier linking

Cambridge Crystallographic Data Centre

CSD entry: KUJVOC Search

Your query was: Doi: 10.1016/j.ejmech.2015.04.023 and returned 2 records

Results

CCDC # Refcode

1022819 KUJVOC

1022841 KUJNVV

Download

KUJVOC: 7,7-Dimethyl-5-(2,4-dimethylphenyl)-4,6,8,9-tetrahydro-4b,5,6,7,8,9b-hexahydro-indeno[1,2-b]indole-9,10-dione

Space Group: P2<sub>1</sub>/n, Cell: a 9.6297(4) Å b 18.0974(9) Å c 12.2883(5) Å α 90.00° β 93.797(10)° γ 90.00°

3D viewer

Chemical diagram

Style Labels Packing Measure

Ball and Stick No Labels Focus Rotate

View group symbols key

## Query:

<https://api-dlIService-prototype-dli.d4science.org/v1/linksFromPid?pid=10.1016/j.poly.2007.03.016&pidType=doi>

## Response:

```
{ "schema": "doi", "identifier": "http://doi.org/10.5517/ccnc0qq", "repoAcronym": "CCDC", "publisher": "Cambridge Crystallographic Data Center", "title": "CCDC 606413: Experimental Crystal Structure Determination", "repoDescr": "Crystallographic data" }, { "timestamp": 1503930494052 }
```

## Extract:

'Publisher' , 'title' , 'identifier' and combine with 'repository name' , 'descriptor' and logo from existing database

## Display:

### Research data for this article



Cambridge Crystallographic Data Center

Crystallographic data

*Data DOIs associated with this article:*

[CCDC 1488871: Experimental Crystal Structure Determination](#) ➤

[CCDC 1488870: Experimental Crystal Structure Determination](#) ➤

[CCDC 1488872: Experimental Crystal Structure Determination](#) ➤

[Show all DOIs](#) ▼





# Hub example: DataCite

*Martin Fenner*

16

- DataCite is a non-profit organization based in Germany with currently 60 member organizations across the globe
- DataCite provides persistent identifier services to its users, focussing on DOIs and associated metadata for research data
- DataCite services link research data to journal articles, software, people, funding and samples and helps locate, identify, and cite research data

# Hub example: Data Literature Interlinking Service (ScholExplorer) - Paolo Manghi

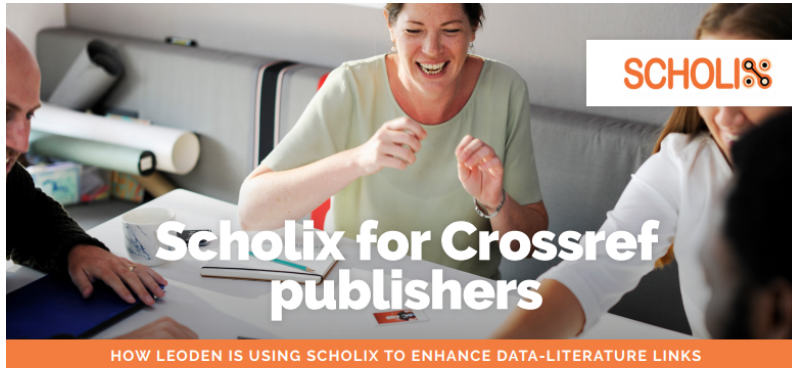
17

- **Beta system** at <https://dliservice-prototype-dli.d4science.org>
- Content from publishers and data archives
  - Datasets+Pubs: 4,200,000
  - Dataset-Pub links: 38,000,000
- In sync with DataCite every 5-6 hours
- Scholix compatible APIs to resolve DOIs
  - Scopus is a consumer
- **Production system @OpenAIRE Nov 2017**

# How to take part

- Community materials 'how to'
- [Scholix for repository managers](#) (Natasha Simons)
- [Scholix for DataCite members](#) (Catherine Brady)
- [Scholix for CrossRef members](#) (Joe Wass)
- [Frequently Asked Questions \(FAQs\)](#) (Natasha Simons)
- [Scholix Scopus story](#) (Natasha Simons & Eleonora Presani)

# How to take part



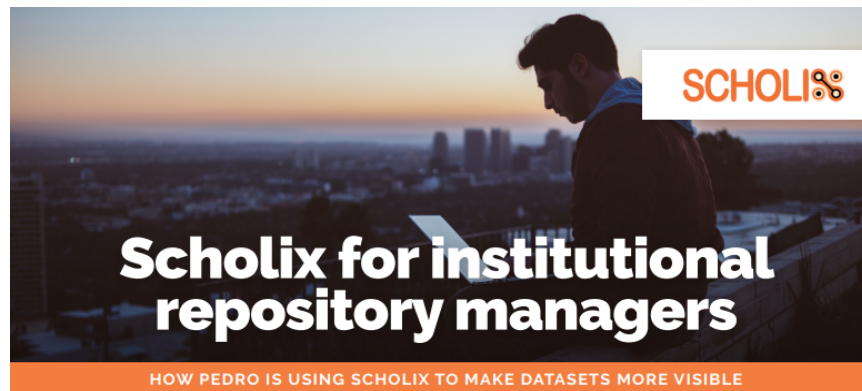
Leoden is a publisher member of Crossref. The member publishes journals and registers their content with Crossref.

Crossref, as one of the Scholix hubs, assists in the global aggregation of literature-data link information for their publisher members. Leoden already provides a wealth of metadata through the DOI registration process. Upon hearing about the Scholix initiative, she decides to include links to data as part of the citation metadata records she provides to Crossref. In doing this, Leoden makes sure that globally unique persistent identifiers - DOIs wherever possible - are used for both the dataset and piece of literature. As a result, the data-literature links provided by Leoden are made available via



In doing this, Anne makes sure that globally unique persistent identifiers - DOIs wherever possible - are used for both the dataset and piece of literature. As a result, the data-literature links provided by Anne are made available via Scholix aggregators such as the DLI service.

Publishers and service providers such as Scopus can query the DLI Scholix aggregation to find datasets related to journal articles they hold. Articles on the journal website can now display a link back to the datasets held in Anne's data centre (via the DOI link), driving web traffic to the data centre and facilitating research discovery.



Pedro manages the institutional research data repository at a university. The repository holds and provides access to datasets from his institution. Information about journal publications related to each dataset are included in the repository metadata records. However, when Pedro visits the journal websites there are no links back to the datasets held in his repository.

Pedro decides to take advantage of the Scholix interoperability framework to get the message out about the literature that is related to data in his repository.

Pedro already provides descriptions of his data through to the international OpenAIRE aggregation. He now simply includes any links to literature in the metadata records harvested

by OpenAIRE (making sure that persistent identifiers - DOIs wherever possible - are used for both the dataset and piece of literature). OpenAIRE, as one of the Scholix hubs, assists in the global aggregation of data-literature link information. The data-literature links from Pedro's repository are now available in Scholix aggregators such as the DLI service.

Publishers and service providers such as Scopus can then query the DLI Scholix aggregation to find datasets related to journal articles they hold. Articles on the journal website now display a link back to the datasets held in Pedro's repository (for example via the DOI link), driving web traffic to the repository and facilitating research discovery.

# Next steps - guidelines

- Guidelines. Almost final, but there are still loose ends. Focus was on testing it (e.g. DLI, Scopus, CCDC, Pangaea).
- We will now put it into a final version. Uncertainties remaining:
  - Scope: focus on data-article right now. Software, labs, and other entities more in the future. Agreed to stay open for these kind of relations
  - “Cited by” and “Supplement to” are enough as relations. We use the subset of the DataCite list for now. Out of scope right now is the standardized terms.
  - “Link Provider” = place where you got it (and -tbd- not the original source)
  - “Publisher” = mandatory (but discussion still open)
  - Does a PID have to be URL? (e.g. does the http part need to be part of it; does the means of resolving it be separate from the PID) => ask advice from the PID group
- Work on documenting schema, support materials and validation tool will kick off after the Barcelona meeting

# Timeline: where do we stand



## Deliverables

- Inventory of interlinking, cross-referencing, and other tools and processes relevant to data publication currently in place.
- An analysis of pros and cons, with an emphasis on scalability and doability.
- Gap analysis, including an analysis of needs & use cases for key stakeholders (data repositories, journal publishers, providers of bibliographic services, funding bodies, research institutions, researchers)
- Recommendations for a one-to-all cross-resolving service. These recommendations will include technical, organizational, governance, and cost aspects
- An operational and publicly available service for cross-referencing datasets and articles

## Deliverables:

A critical mass of Scholix conformant hubs providing the enabling infrastructure for a global view of data-literature links

Pathfinder services providing aggregations, query services, and analyses

Beneficiaries of these services accessing data-literature link information to add value to scholarly journal sites, data centre portals, research impact services, research discovery services, research management software, etc.

Operational workflows to populate the infrastructure with data-literature links

Better understanding of current data-literature interlinking landscape viewed from the perspective of e.g. disciplines, publishers, repositories etc.

# Thank you!

22

## Questions?

Email [info@scholix.org](mailto:info@scholix.org)

RDA WG site & Scholix web site: <http://www.scholix.org/>

Co-chairs:

- Adrian Burton [adrian.burton@ands.org.au](mailto:adrian.burton@ands.org.au)
- Wouter Haak [W.Haak@Elsevier.com](mailto:W.Haak@Elsevier.com)
- Paolo Manghi [paolo.manghi@isti.cnr.it](mailto:paolo.manghi@isti.cnr.it)
- Martin Fenner [Fenner@datacite.org](mailto:Fenner@datacite.org)