Interoperable persistent identifiers

Agenda

- ORCID: Laura Paglione
- Datacite: Sebastian Peters
- INSPIREHEP – Implementation: Sünje Dallmeier-Tiessen
- Dryad – Implementation: Ryan Scherle
- Overarching challenges
- Q&A

ODIN

ORCID and DataCite Interoperability Network

BRITISH LIBRARY

DataCite

CERN

DRYAD

ORCID

Cornell University Library

arXiv.org

ands

AUSTRALIAN NATIONAL DATA SERVICE

SEVENTH FRAMEWORK PROGRAMME

ORCID

Connecting Research and Researchers

The ORCID Registry

Open Repositories 2014

June 12, 2014

Laura Paglione
Technical Director, ORCID
L.Paglione@ORCID.org
http://orcid.org/0000-0003-3188-6273

ORCID
Connecting Research and Researchers

What is ORCID ?

The ORCID Organization

- Non-profit, non-proprietary, open, and community-driven
- Global, interdisciplinary
- Independent
- Supported by the membership of organizations using the ORCID API
  - Funding organizations
  - Professional societies
  - Universities & research institutes
  - Publishers

The ORCID iD

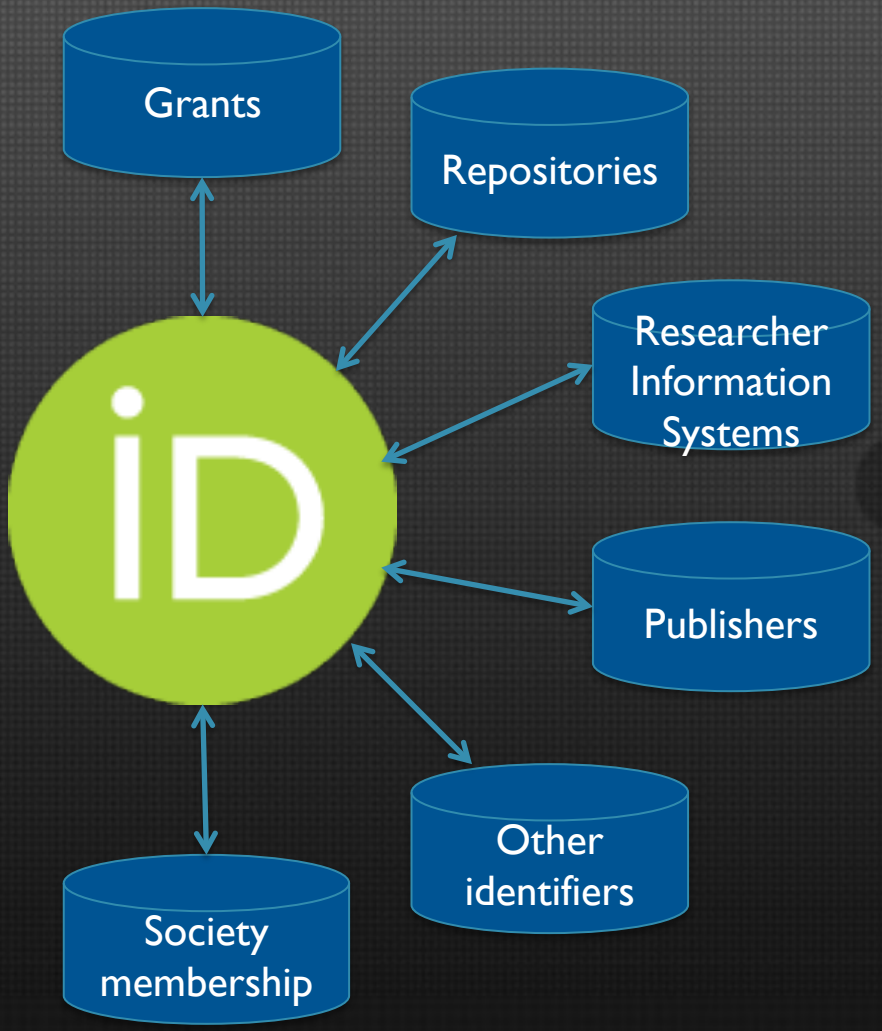
- Unique, persistent identifier for researchers, scholars & analysts
- Free to researchers
- Used throughout one's career
- Used across professional activities, affiliations, disciplines, nations & languages
- Embedded into workflows & metadata
- API enables interoperability between siloed systems

ORCID
Connecting Research and Researchers

Facilitating Info Exchange

The ORCID API enables information exchange between systems:

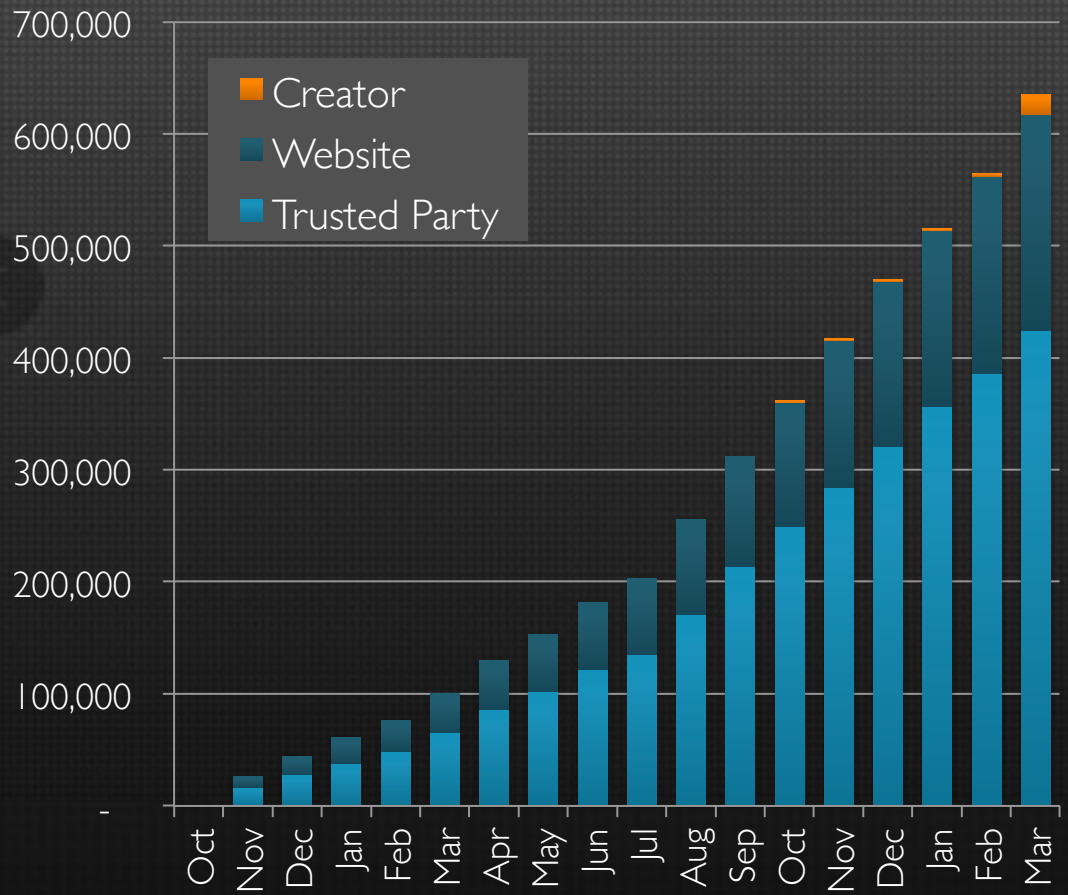
- Less time re-keying
- Improved data
- Easier maintenance
- Better sharing across systems

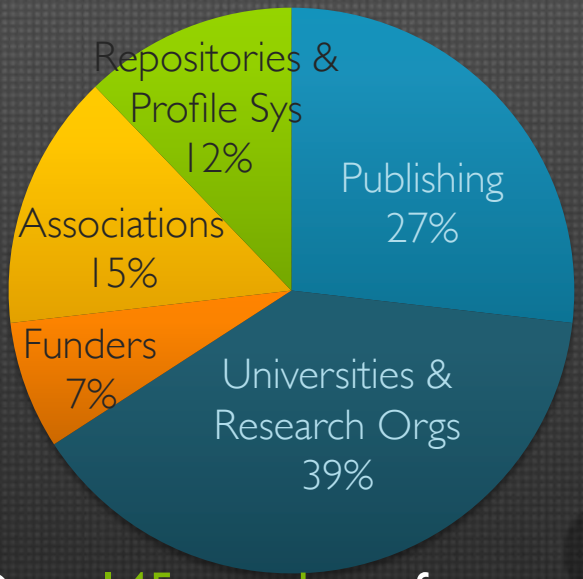


ORCID
Connecting Research and Researchers

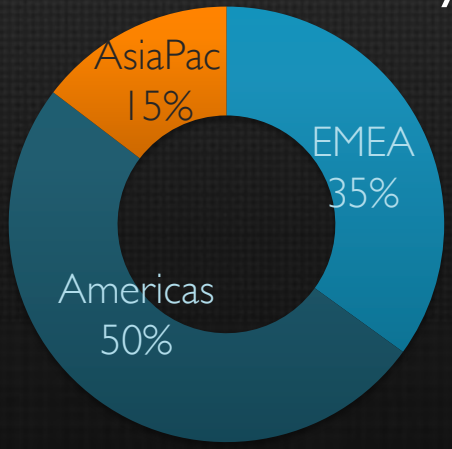
Adoption and Integration

ORCID has issued over 740,000 iDs since our launch in October 2012. Integration and use is international.



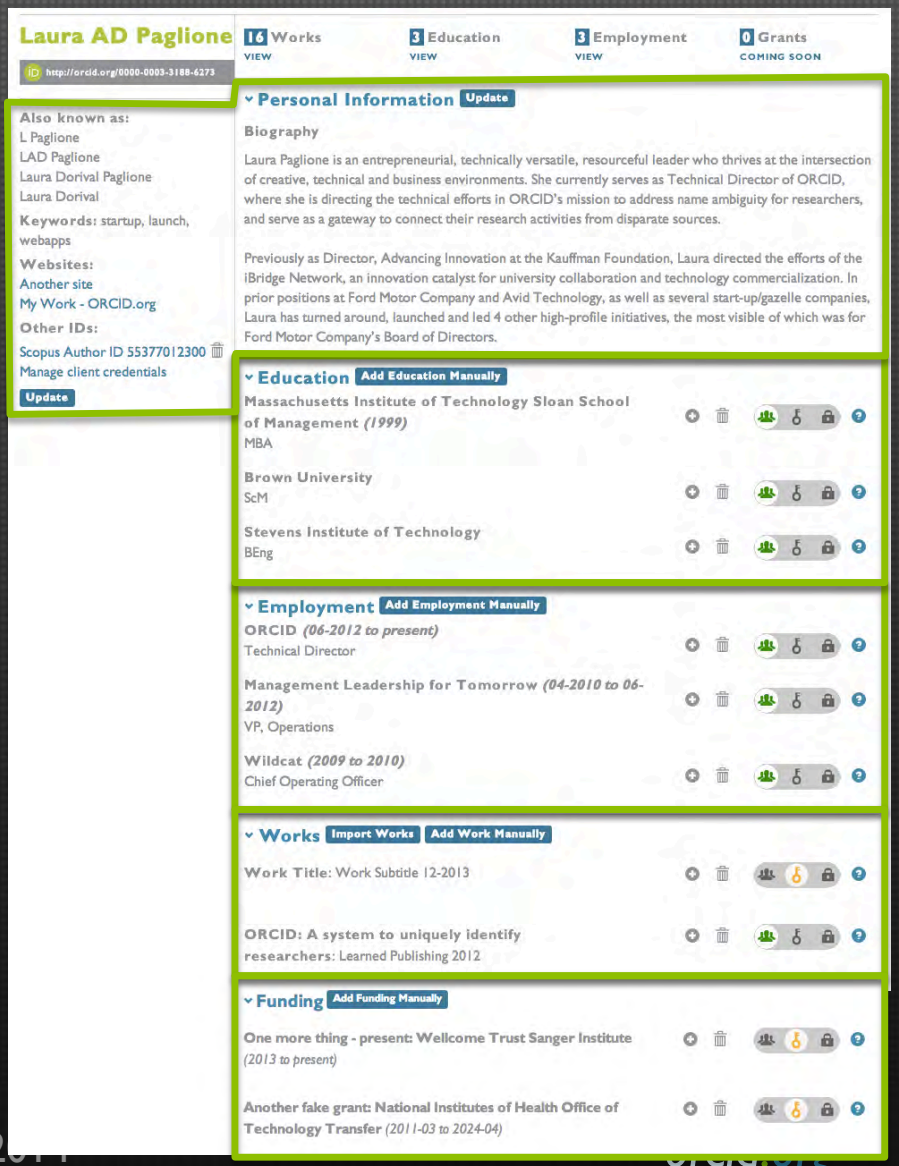


Over 145 members, from every sector of the international research community



ORCID
Connecting Research and Researchers

The ORCID Record

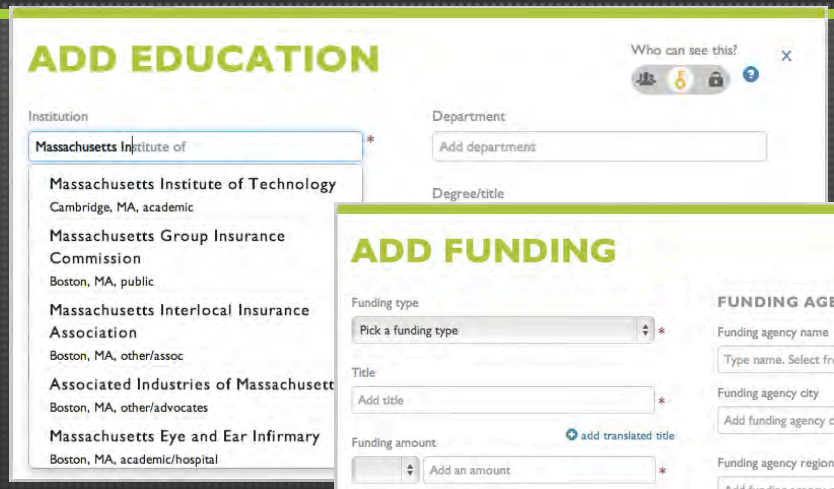


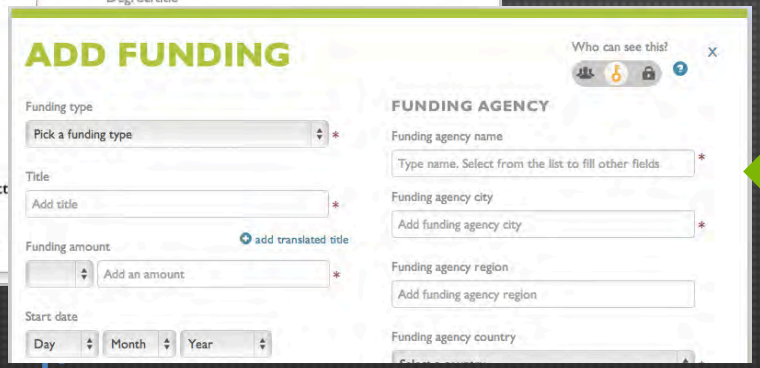
Contains:

- Biographical Data
- Education Data
- Employment Data
- Works Data
- Funding Data

ORCID
Connecting Research and Researchers

Shared IDs & Connections

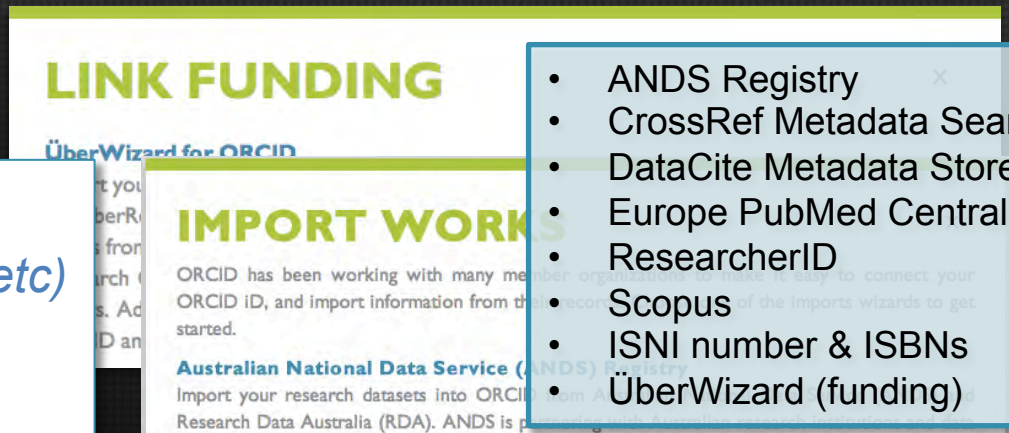




Funding agency list consistent with FundRef

↑ Organization list from Ringgold *(an ISNI Registrar)*

Search & link Wizards ↓



- ANDS Registry
- CrossRef Metadata Search
- DataCite Metadata Store
- Europe PubMed Central
- ResearcherID
- Scopus
- ISNI number & ISBNs
- ÜberWizard (funding)

External IDs
(DOIs, ISBNs, ISNIs, PubMed IDs, etc)

Other Person IDs
(Scopus ID, ResearcherID, ISNIs, Institution IDs, etc)

ORCID
Connecting Research and Researchers

Who is Integrating?

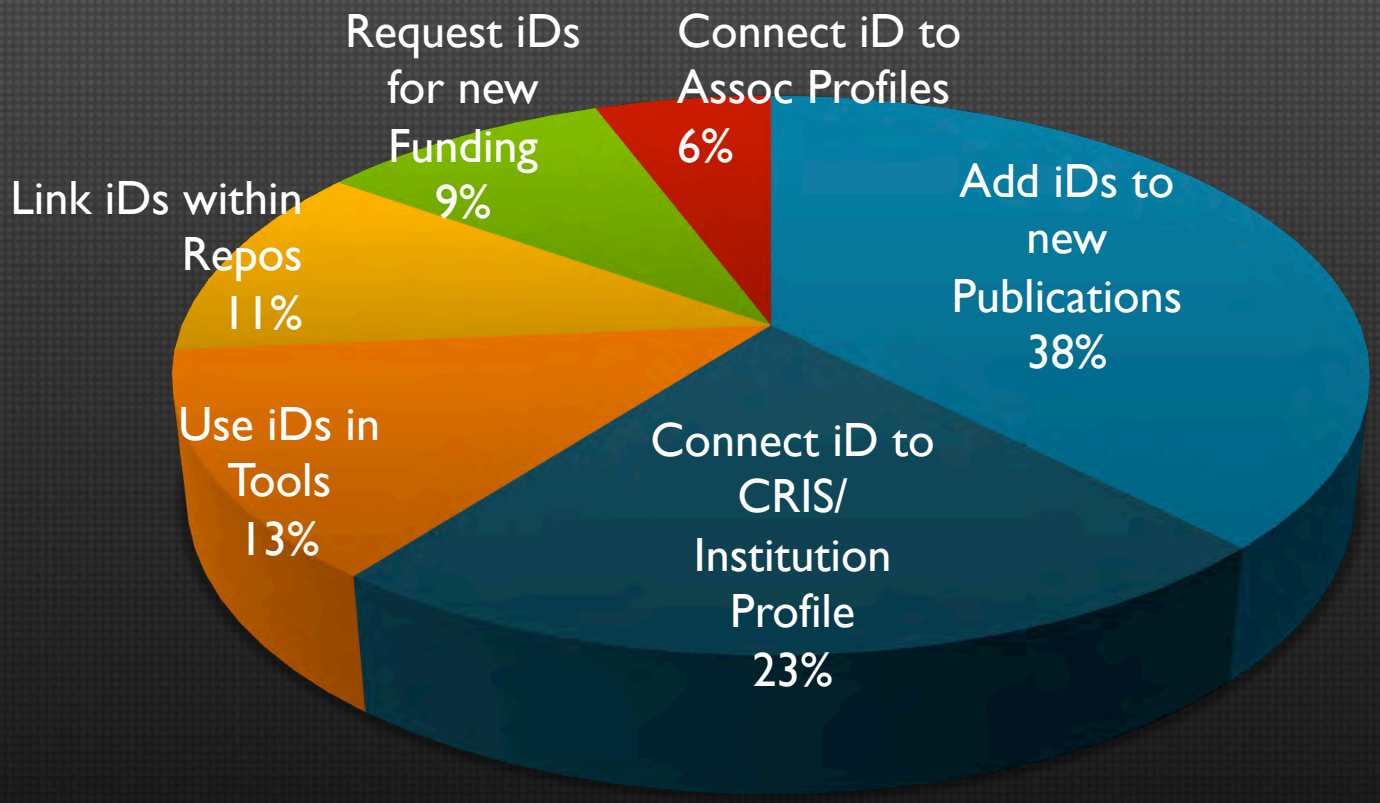
- Research Funders
- Professional Associations
- Publishers
- Repositories, CRIS Vendors, Metrics Providers
- Universities and Research Orgs

For a list of organizations and integrations see
http://orcid.org/organizations/integrators

ORCID
Connecting Research and Researchers

Types of Connections

Over 55 integrations by members and others using the public API



http://orcid.org/organizations/integrators/current

ORCID
Connecting Research and Researchers

Standard Metadata Schemas

- ORCID works collaboratively with the research community to ensure use and adoption of research information exchange standards

ORCID
Connecting Research and Researchers

Challenges

- Remembering all of those identifiers!
- Identifiers without other info are hard; they don't provide enough info (resolvable iDs help)
- Type the ORCID iD
  - Prone to errors
  - Requires researcher to remember the iD
- Search for someone to associate an iD
  - Disambiguation problems – which Smith / Wang / Johansson are you?
- iD display – long

More information

Thank you!

- Find out more at http://orcid.org
- More on membership at http://orcid.org/about/membership
- Learn about tools to embed ORCID iDs at http://support.orcid.org/knowledgebase/
- Subscribe to our blog at http://orcid.org/about/news and follow @ORCID_Org on Twitter
- Contact the ORCID Executive Director at l.haak@orcid.org

DataCite

or

DOI is more than just CrossRef

Sebastian Peters

DataCite Tech Team Lead

German National Library of
Science and Technology

tech@datacite.org / sebastian.peters@tib.uni-hannover.de

TIB | German National Library of Science and Technology

DataCite

DOI Registration Agencies

1) Airiti, Inc.
2) China National Knowledge Infrastructure (CNKI)
3) **CrossRef**
4) **DataCite**
5) Entertainment Identifier Registry (EIDR)
6) The Institute of Scientific and Technical Information of China (ISTIC)
7) Japan Link Center (JaLC)
8) Multilingual European DOI Registration Agency (mEDRA)
9) Publications Office of the European Union (OP)

What is DataCite?

- Global consortium carried by local institutions
- focused on improving the scholarly infrastructure around datasets and other non-textual information
- focused on working with datacentres and organizations that hold data
- Providing standards, workflows and best-practice
- **Make research data citable** (data publication)
- Initially, but not exclusively based on the DOI system

http://www.datacite.org

TIB | German National Library of Science and Technology

DataCite

DataCite History

- Founded December 1st 2009 in London
- Forerunner (since 2004), founder member and managing agent: German National Library of Science and Technology (TIB)
- 1st DOI for a dataset registered by TIB in 2004
- Happy 10th Anniversary: http://www.datacite.org/node/121

TIB | German National Library of Science and Technology

DataCite

Facts & Figures

- 22 full members, 9 associated; from 18 countries
- ~ 3.4 Mio. DOIs
- ~ 3.2 Mio. metadata objects
- > 300 DOI prefixes
- > 300 datacentres (publishing agents)

DOI – Digital Object Identifier

- *Digital Identifier* of an *Object*
- Maintained by the International DOI Foundation (IDF)
- Based on Handle System
- Actionable, interoperable, persistent
- ISO standard 26324:2012
- ~ 100 Mio. DOIs registered
- Example: doi:10.5524/100001, http://doi.org/10.5524/100001
- Resolve to Landing Pages

http://www.doi.org

TIB | German National Library of Science and Technology

DataCite

Challenge: Getting DOI Metadata

Problem

Every RA has a custom metadata format and API

Solution

Content Negotiation on doi.org level

Content Negotiation – CrossRef & DataCite

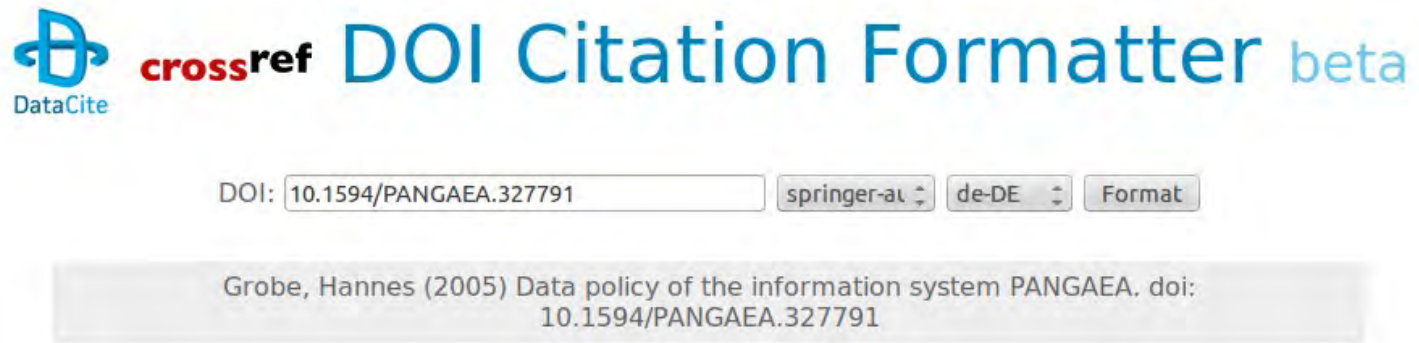
- HTTP Content Negotiation with [dx.]doi.org proxy
- You don‘t need to know the actual DOI RA
- Common Formats of CrossRef & DataCite:
  - RDF, RIS, BibTeX, Citeproc, …
- http://crosscite.org/cn/
- Example:
  - `curl -LH "Accept:application/rdf+xml" http://doi.org/10.5524/100001`

TIB | German National Library of Science and Technology

DataCite

Citation Formatter

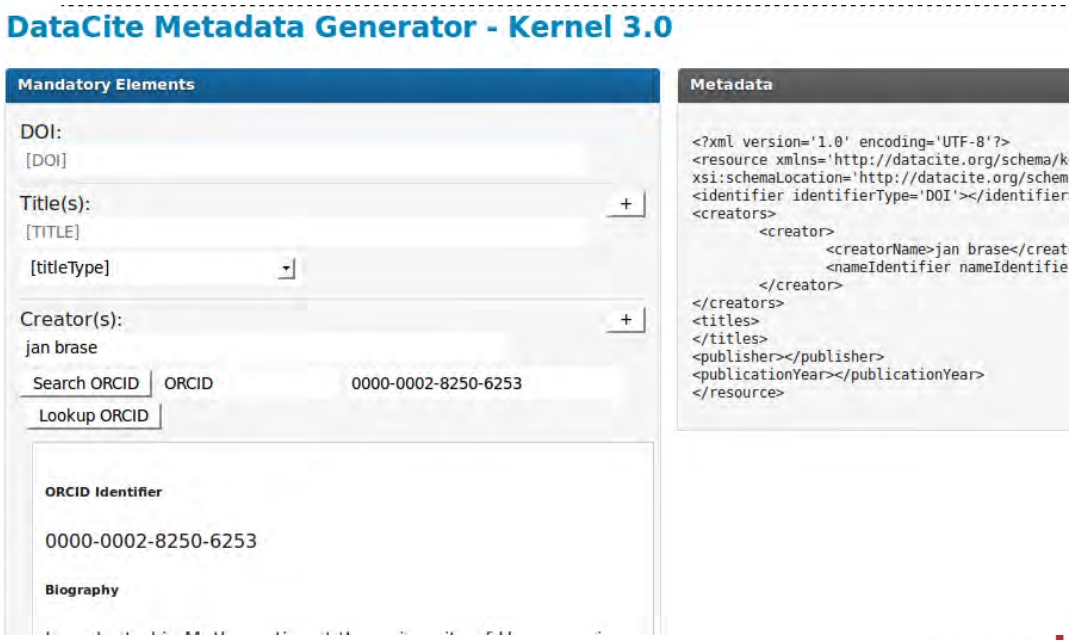
- Using Content Negotation
- More than 500 styles (http://citationstyles.org)
- http://crosscite.org/citeproc/
- https://github.com/crosscite/citeproc-doi-server



DataCite

DataCite Metadata Generator and ORCID

- Metadata formular with ORCID lookup
- Result from the last code sprint at CERN
- http://www.datacite.org/node/102
- https://github.com/koelnconcert/datacite-metadata-generator



GERMAN NATIONAL LIBRARY OF SCIENCE AND TECHNOLOGY

DataCite

Combining Identifiers for Citing Fragments

- Media Fragment Identifier (MFID), W3C
- Prototype for TIB's AV-Portal
- http://doi.org/10.5446/12780#t=00:20,00:27
- Out of the box solution
- DOI resolver agnostic
- Content negotiation still works

TIB | German National Library of Science and Technology

DataCite

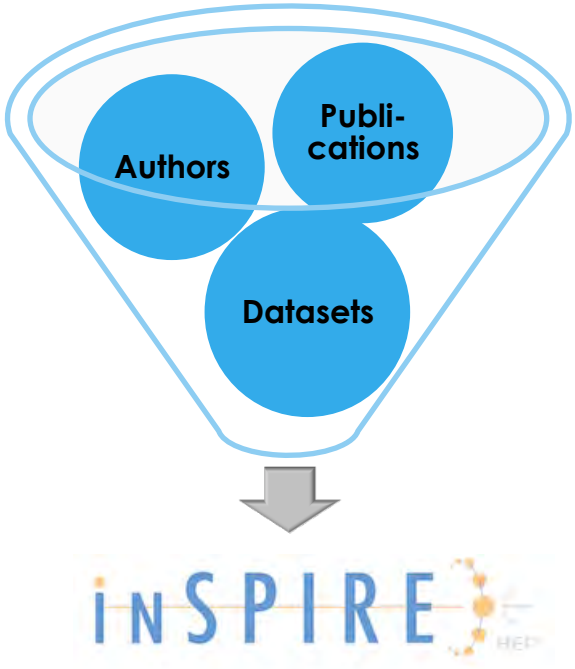
Shared identifiers in a large-scale digital library: INSPIRE-HEP

sunje.dallmeier-tiessen@cern.ch

What is INSPIRE?

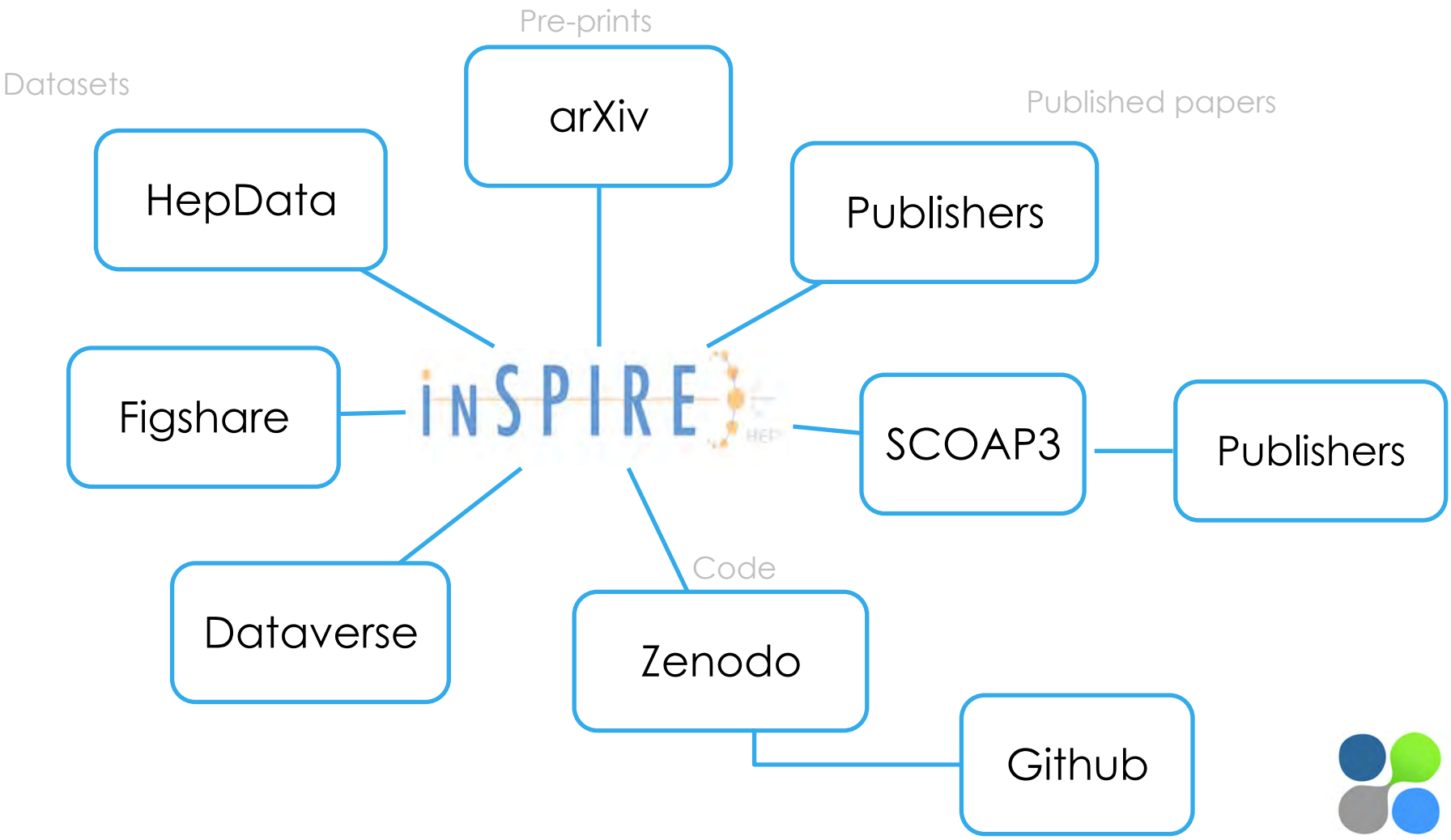
- Based on Invenio
- Digital library in High-Energy Physics
  - And an information hub:
    - Pre-prints
    - Published papers
    - Authors
    - Data

Linked!

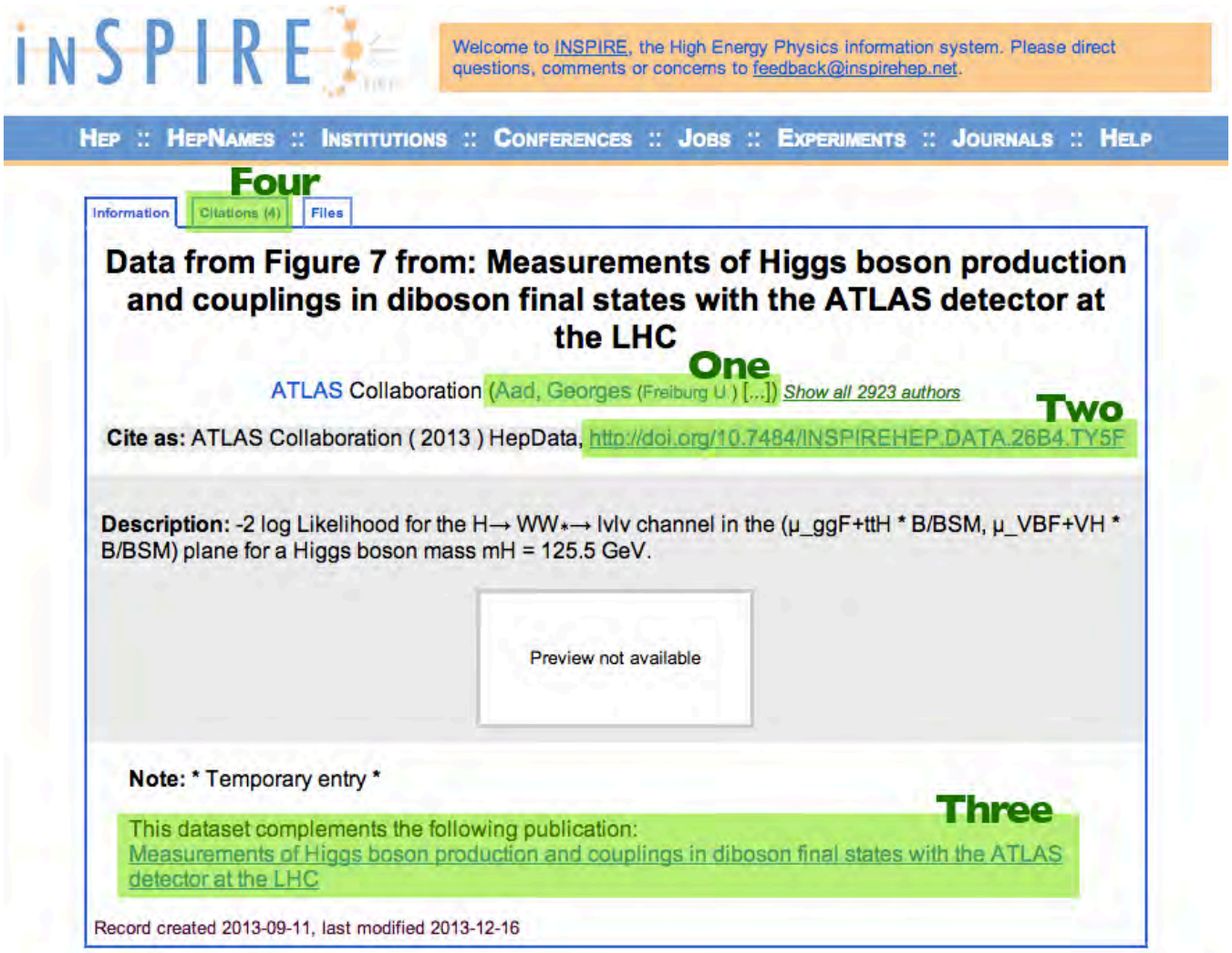


- Some numbers:
  - 1.000.000 publications
  - 100.000 authors
  - 50.000 datasets
  - And institutions, conferences, journals, job offers…

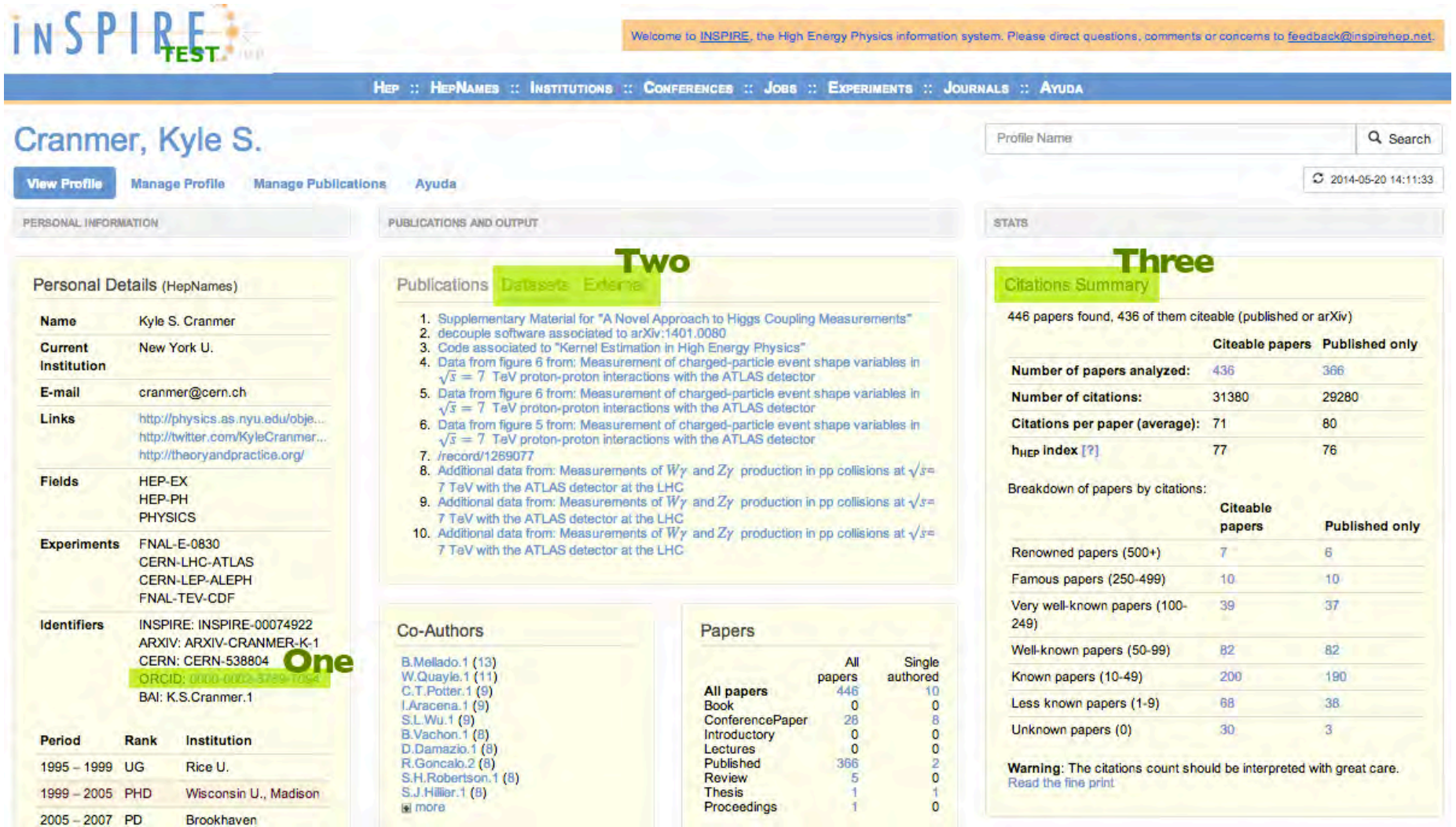
What's inside?

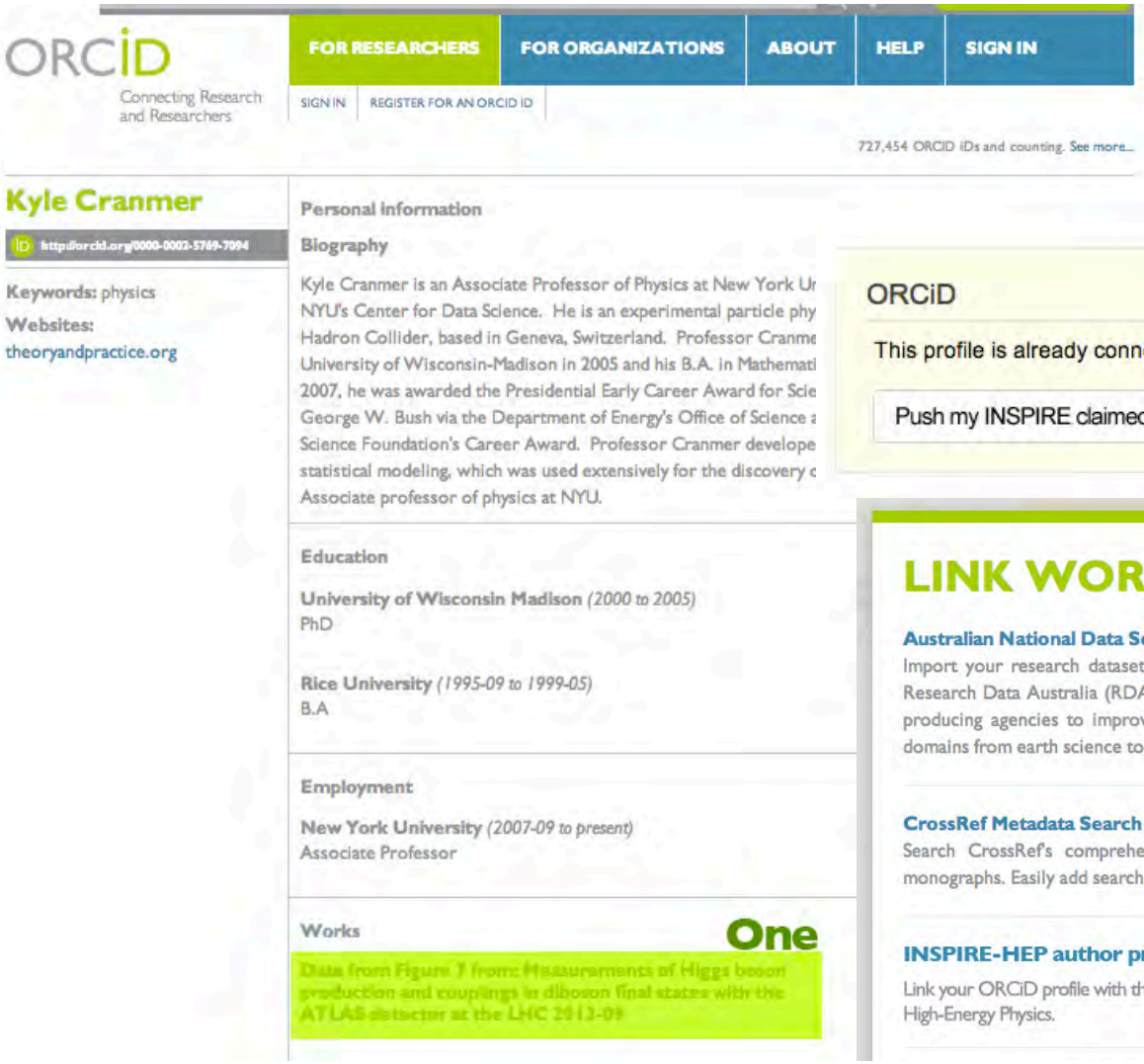


Research data (citation)



Authors on INSPIRE



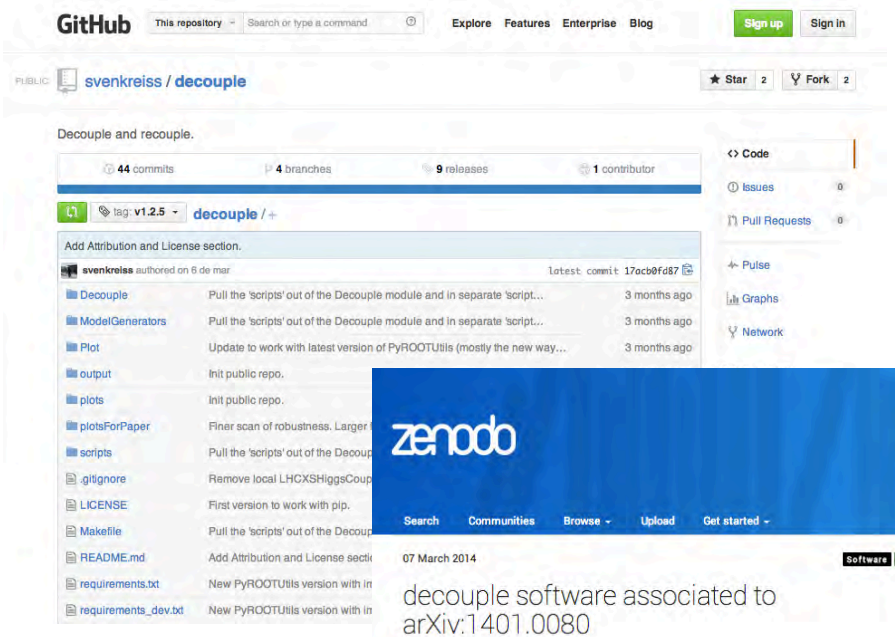

ORCID
Connecting Research and Researchers
FOR RESEARCHERS
FOR ORGANIZATIONS
ABOUT
HELP
SIGN IN
SIGN IN
REGISTER FOR AN ORCID ID
727,454 ORCID iDs and counting. See more...
Kyle Cranmer
http://orcid.org/0000-0002-5769-7094
Keywords: physics
Websites:
theoryandpractice.org
Personal information
Biography
Kyle Cranmer is an Associate Professor of Physics at New York Ur
NYU's Center for Data Science. He is an experimental particle phy
Hadron Collider, based in Geneva, Switzerland. Professor Cranme
University of Wisconsin-Madison in 2005 and his B.A. in Mathemati
2007, he was awarded the Presidential Early Career Award for Scie
George W. Bush via the Department of Energy's Office of Science a
Science Foundation's Career Award. Professor Cranmer develope
statistical modeling, which was used extensively for the discovery c
Associate professor of physics at NYU.
Education
University of Wisconsin Madison (2000 to 2005)
PhD
Rice University (1995-09 to 1999-05)
B.A
Employment
New York University (2007-09 to present)
Associate Professor
Works
One
Data from Figure 7 from: Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC 2013-09

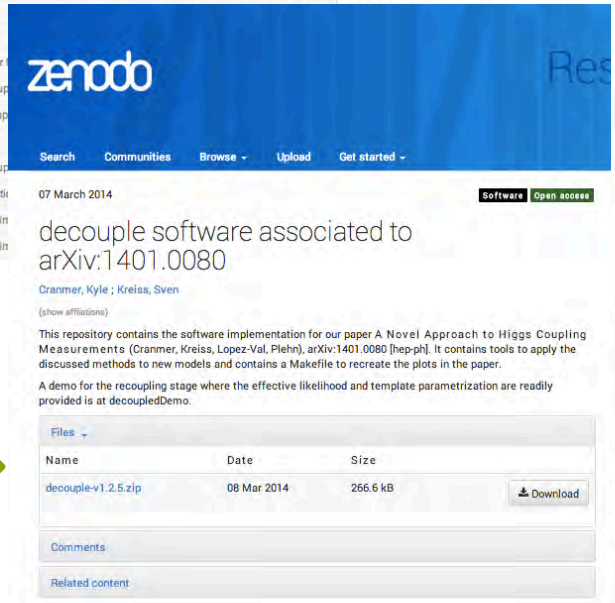
Very soon... !

ORCiD
This profile is already connected to the following ORCiD: 0000-0003-0762-2235
Push my INSPIRE claimed publications to ORCID
?

LINK WORKS
x
Australian National Data Service (ANDS) Registry
Import your research datasets into ORCID from Australian National Data Service (ANDS) and Research Data Australia (RDA). ANDS is partnering with Australian research institutions and data producing agencies to improve discovery and reusability of research data across many research domains from earth science to technology and engineering.
CrossRef Metadata Search
Search CrossRef's comprehensive metadata on journal articles, conference proceedings and monographs. Easily add search results to your ORCID profile.
INSPIRE-HEP author profiles
Link your ORCiD profile with the INSPIRE-HEP portal to import your claimed publications and datasets in High-Energy Physics.

And a bit more complex…



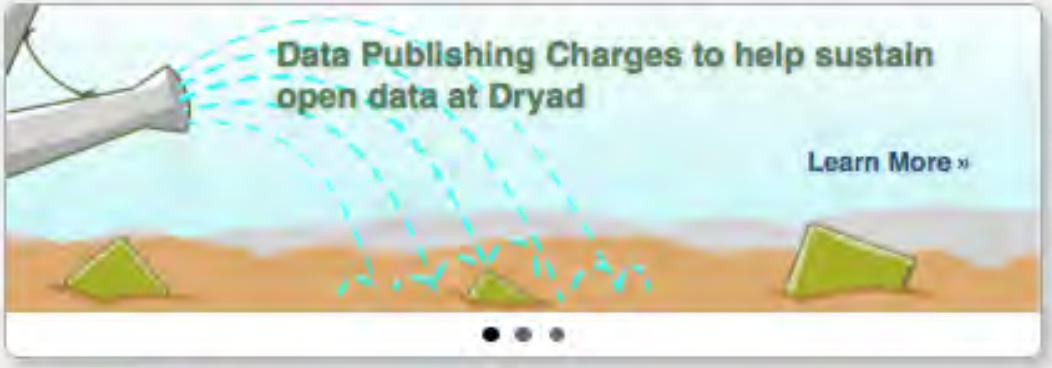


Identifiers and Interoperability: Dryad Perspective

Ryan Scherle

Dryad Digital Repository

CC BY



Submit data now

How and why?

Search for data

Enter keyword, author, title, DOI ...etc Go

Advanced search

Browse for data

Recently published | Popular | By Author | By Journal

Recently Published Data

Phadke S, Zufall R, Paixao T, Pham S, Pham T (2013) Data from: Genetic background alters dominance relationships between mat alleles in the ciliate Tetrahymena thermophila. *Journal of Heredity* doi:10.5061/dryad.qv731

Chiu C, Jost L, Chao A (2013) Data from: Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs* doi:10.5061/dryad.vr4k2

Salgado-Salazar C, Rossman AY, Chaverri P (2013) Data from: Not as ubiquitous as we thought: taxonomic crypsis, hidden diversity and cryptic speciation in the cosmopolitan fungus Thelonectria discophora (Nectriaceae, Hypocreales, Ascomycota). *PLoS ONE* doi:10.5061/dryad.q3s66

Morgan EMJ, Green BS, Murphy NP, Strugnell JM (2013) Data from: Investigation of genetic structure between deep and shallow populations of the southern rock lobster, Jasus edwardsii in Tasmania, Australia. *PLoS ONE* doi:10.5061/dryad.656qf

Robinson BW (2013) Data from: Evolution of growth by genetic accommodation in Icelandic freshwater stickleback. *Proceedings of the Royal Society B*

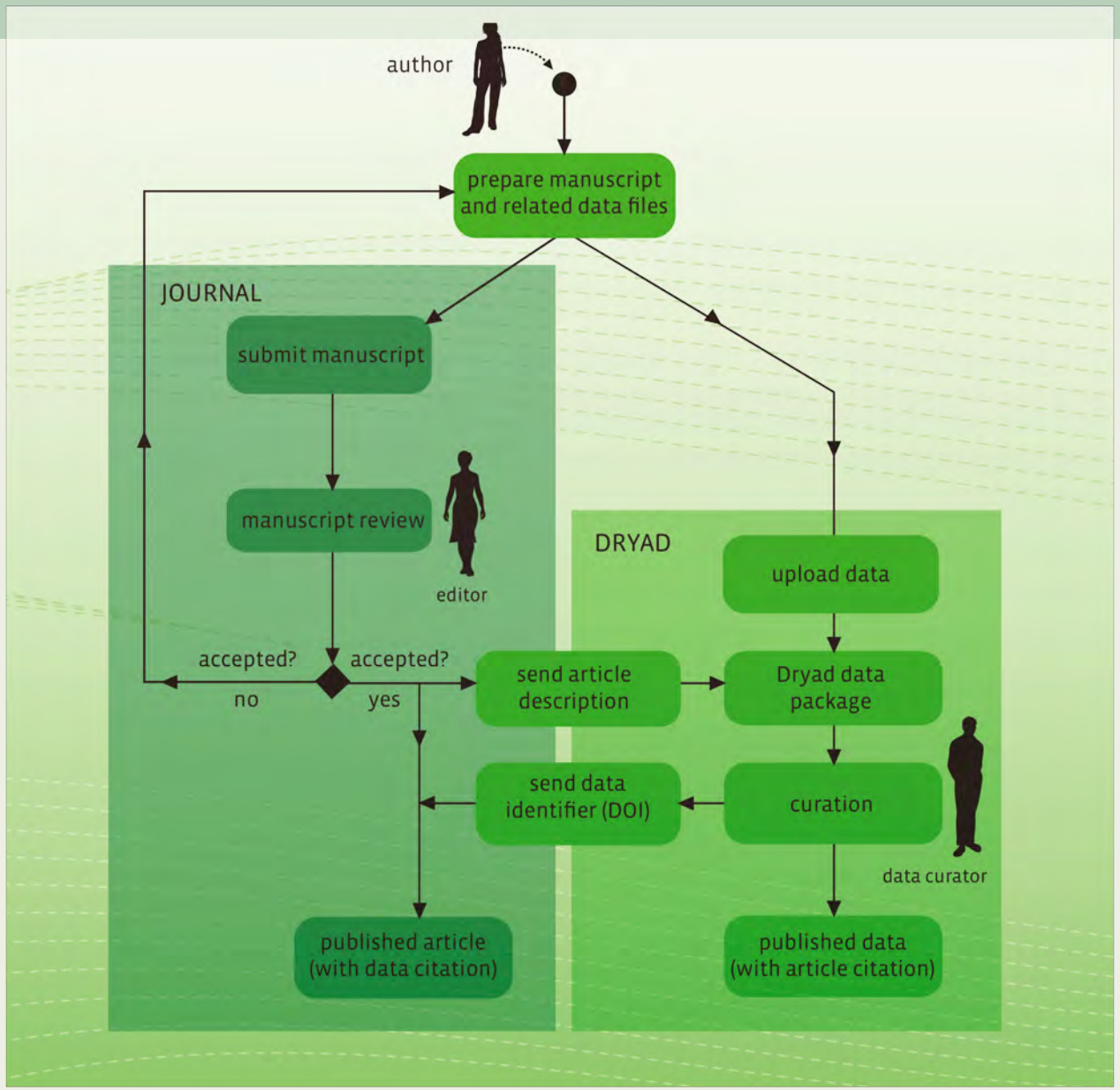
Be part of Dryad

Publishers, societies, universities, libraries, funders, and other stakeholder organizations are invited to become members. Tap into an active knowledge-sharing network, receive discounts on submission fees, and help shape Dryad's future.

Submission integration is a free service that allows publishers to coordinate manuscript and data submissions. It makes submitting data easy for researchers; makes linking articles and data easy for journals; and enables confidential review of data prior to publication.

Submission fees support the cost of keeping Dryad's content free to use. Flexible pricing plans provide volume discounts.

Integrated article/data submission



Encouraging citations

When using this data, please cite the original article:

> Levy J, Foulsham T, Kingstone A (2012) Monsters are people too. Biology Letters 9(1): 20120850.doi:10.1098/rsbl.2012.0850

Additionally, please cite the Dryad data package:

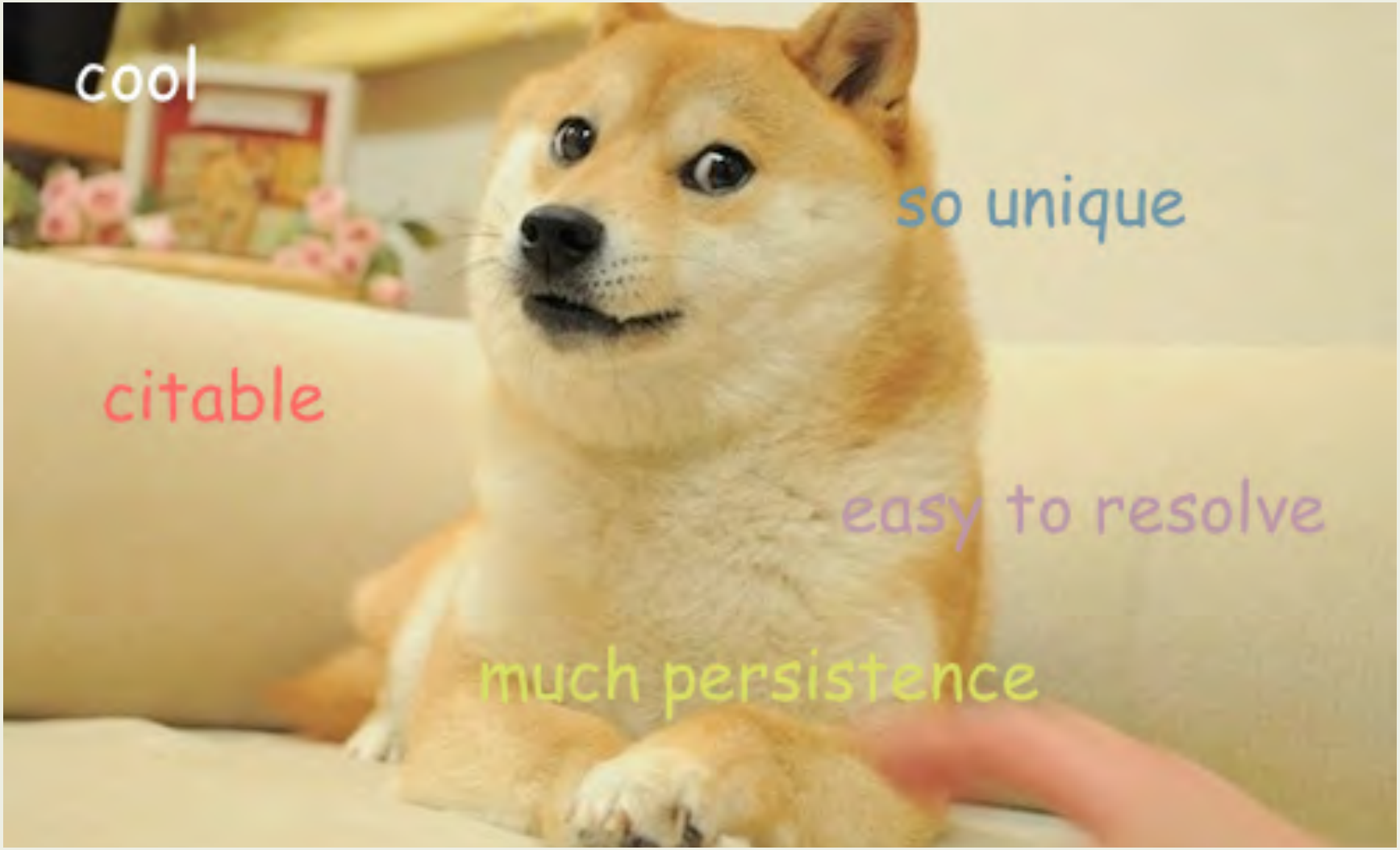
> Levy J, Foulsham T, Kingstone A (2012) Data from: Monsters are people too. Dryad Digital Repository. doi:10.5061/dryad.4rk06

Cite | Share

Principles of Citable Identifiers

- Use DOIs
- Keep identifiers simple
- Use syntax to illustrate relationships
- Create new identifiers when (and only when) the meaning of the content changes

From: Creating Citable Data Identifiers. Ryan Scherle and Mark Diggory. Open Repositories 2012.


cool
so unique
citable
easy to resolve
much persistence

Opportunities

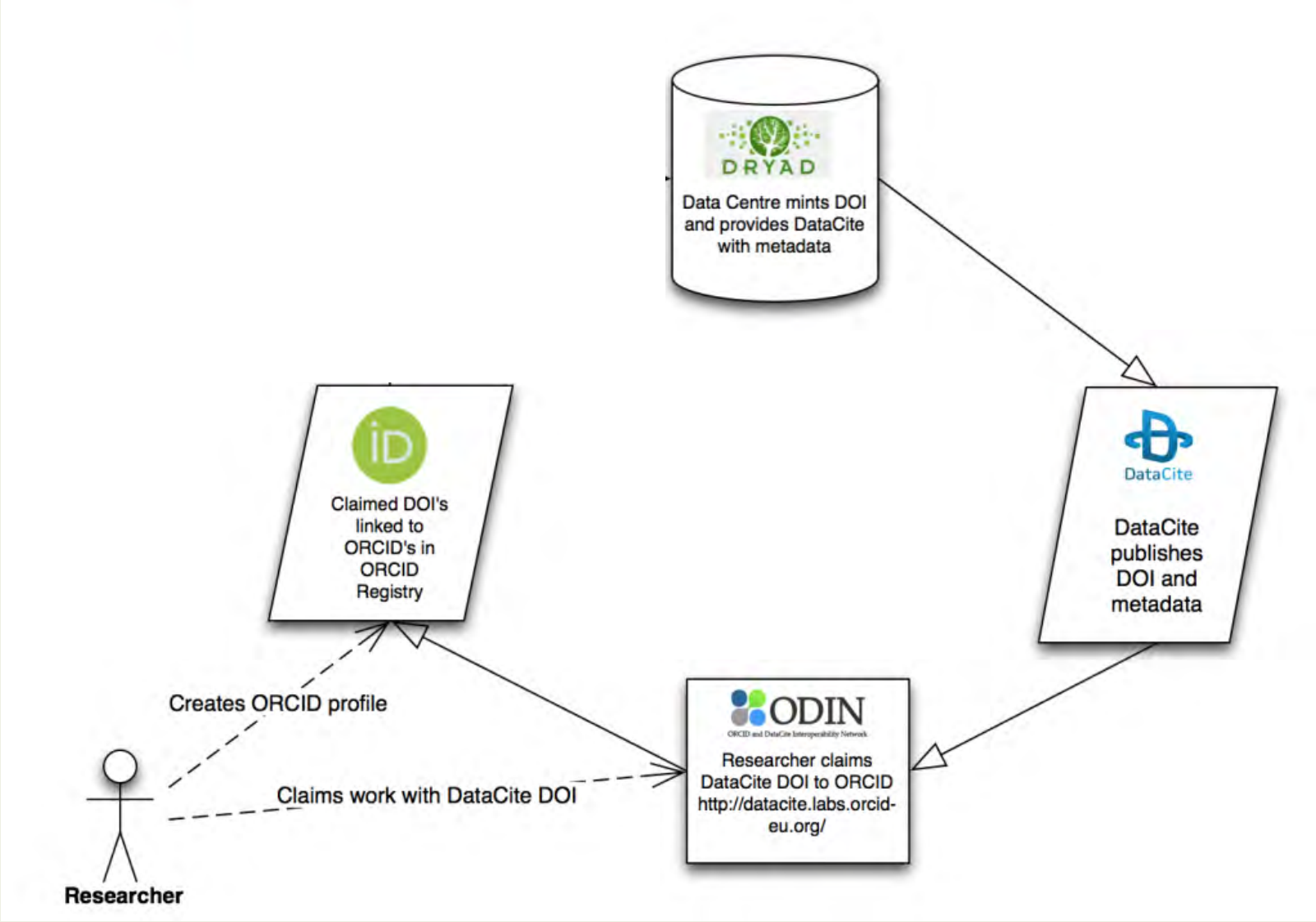
DataCite DOIs

- Leverage central infrastructure to distribute and transform metadata

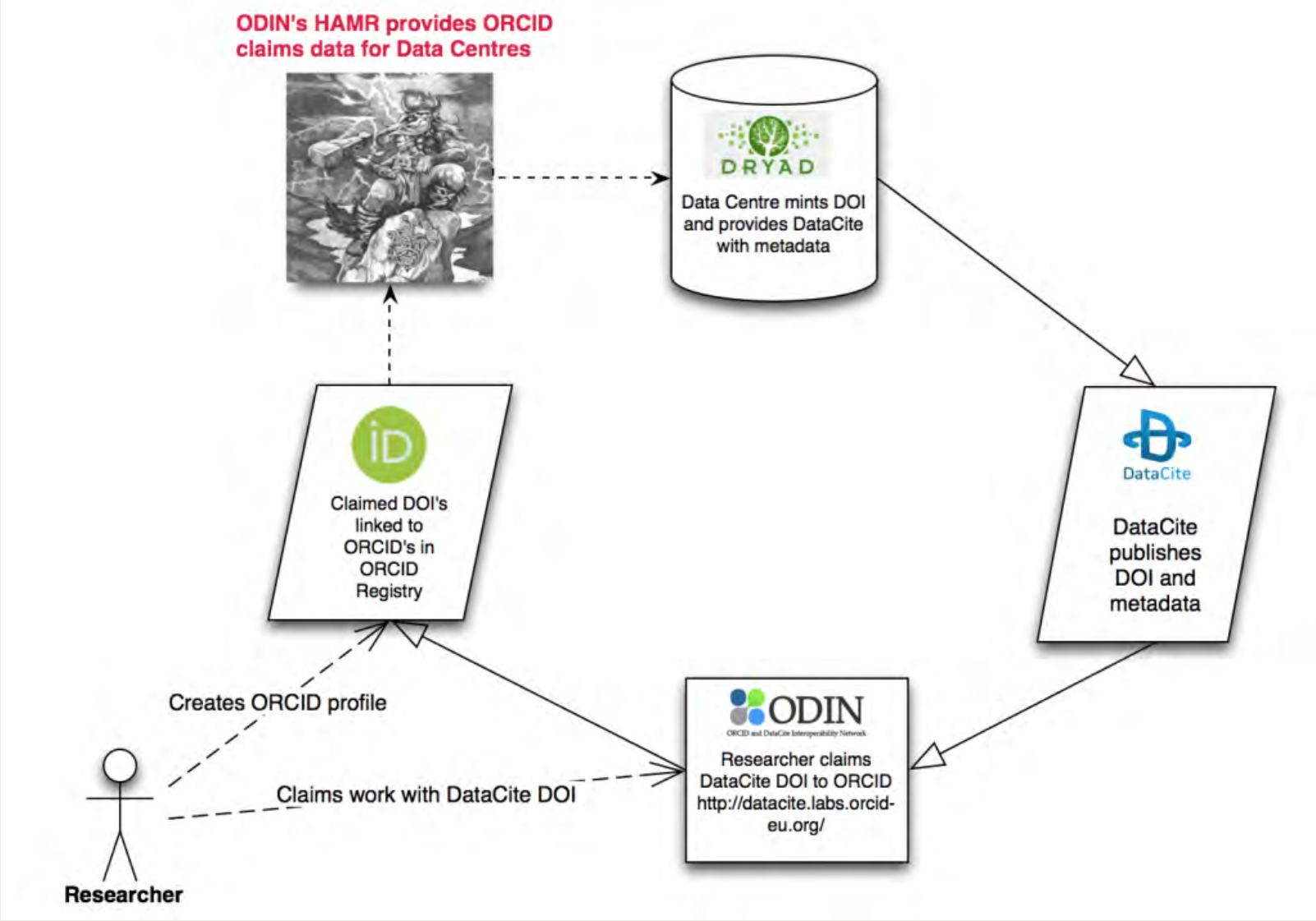
ORCIDs

- Reduce ambiguity
- Permissions
- Add “missing” metadata

ODIN's HAMR


DRYAD
Data Centre mints DOI and provides DataCite with metadata
iD
Claimed DOI's linked to ORCID's in ORCID Registry
DataCite
DataCite publishes DOI and metadata
Creates ORCID profile
Claims work with DataCite DOI
ODIN
Researcher claims DataCite DOI to ORCID http://datacite.labs.orcid-eu.org/
Researcher

ODIN's HAMR


ODIN's HAMR provides ORCID claims data for Data Centres
DRYAD
Data Centre mints DOI and provides DataCite with metadata
iD
Claimed DOI's linked to ORCID's in ORCID Registry
DataCite
DataCite publishes DOI and metadata
ODIN
ORCID and DataCite Interoperability Network
Researcher claims DataCite DOI to ORCID http://datacite.labs.orcid-eu.org/
Creates ORCID profile
Claims work with DataCite DOI
Researcher

Challenges

DOI metadata requirements vs privacy

Challenges

DOI metadata requirements vs privacy

Identifier:	doi:10.5061/DRYAD.PK85D Get link

About the identified object

Location (URL):	http://datadryad.org/publicationBlackout
Creators:	(:tba)
Title:	Dryad Item 10.5061/DRYAD.PK85D
Publisher:	Dryad Digital Repository
Publication year:	9999
Subjects:	(:tba)
Resource type:	Dataset/DataPackage
Related identifiers:	10.5061/DRYAD.PK85D/1 [DOI] [HasPart]

Challenges

Prioritizing tool development for identifiers

Challenges

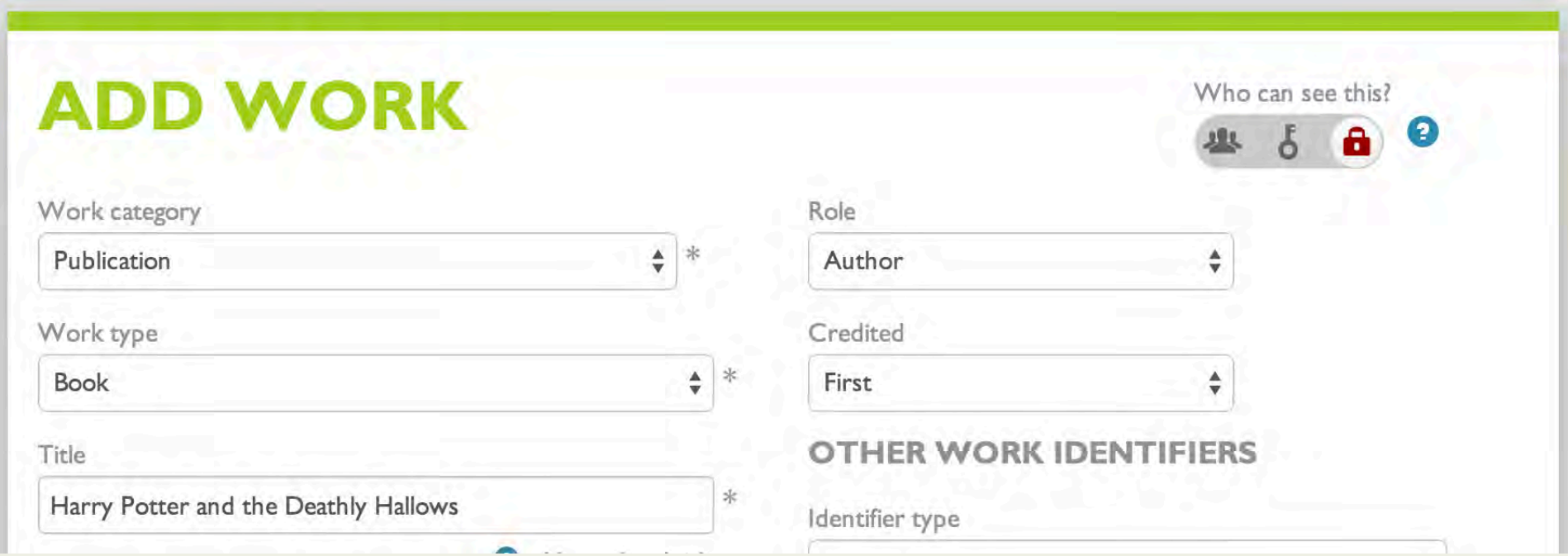
Prioritizing tool development for identifiers

Challenges

Levels of assurance – when using ORCIDs for permissions

Challenges

Levels of assurance – when using ORCIDs for permissions



Challenges

Identifiers and content mutability

Challenges

Identifiers and content mutability

Challenges

Identifiers and content mutability

Challenges & Opportunities

- Confusion – between identifiers & services
- Tension between data needs and privacy concerns
- Identifier not enough – prioritizing tool development
- Levels of assurance of identifier & relationships
- Resolvability – of identifiers; consistency in metadata & APIs