June 07, 2019

Collaborative Approaches to Integrate Repositories within the Research Information Ecosystem

Creating Bridges for Common Goals

David Scherer, *Carnegie Mellon University*
Kate Byrne, *Symplectic*
Mark Hahnel, *Figshare*
Dan Valen, *Figshare*

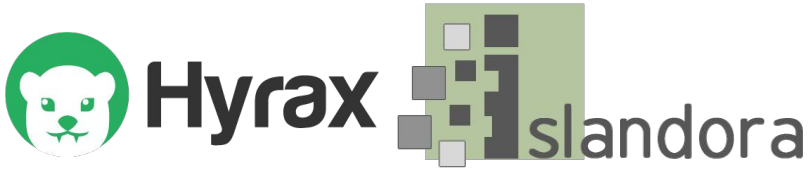
Carnegie Mellon University

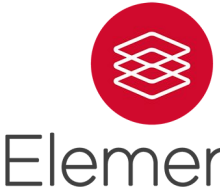
@davidascherer @katebyrne @dnvln @MarkHahnel #nasig19

Repositories and Research Information Management Systems

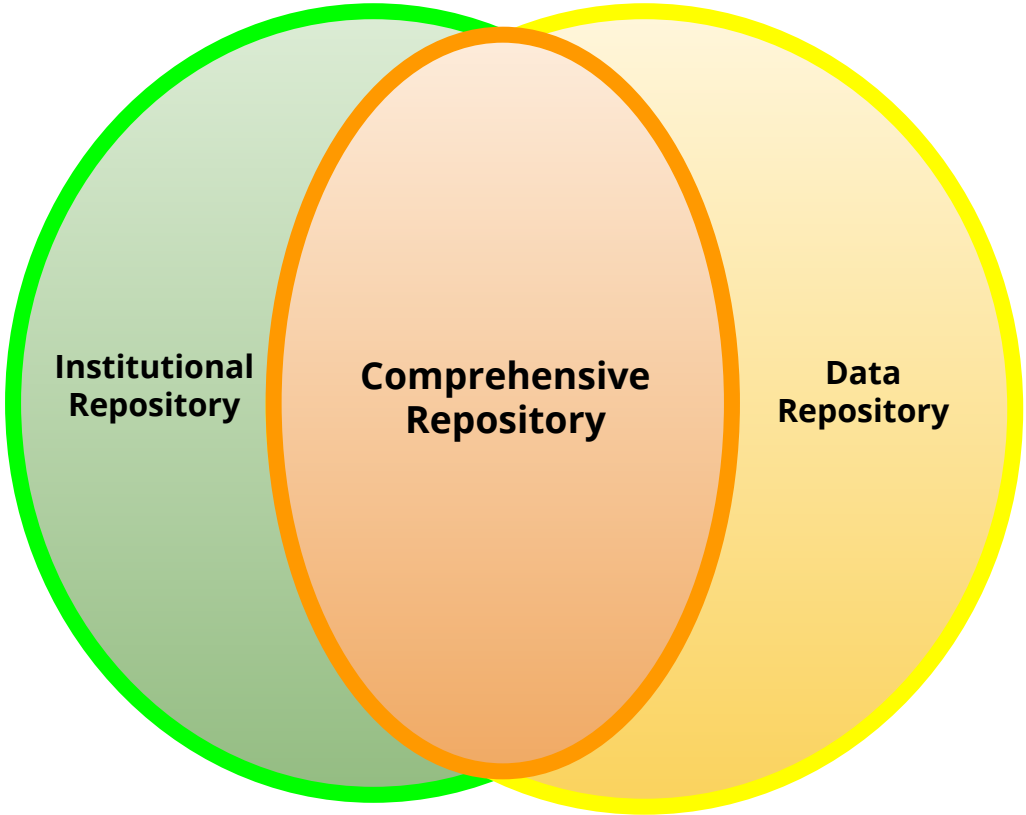
Repositories

RIMS





Carnegie Mellon University



Carnegie Mellon University

KiltHub
WEAVING THE FABRIC OF YOUR RESEARCH
@KILTHUB | KILTHUB.CMU.EDU

Research Information Management Systems

Current Research Information Systems

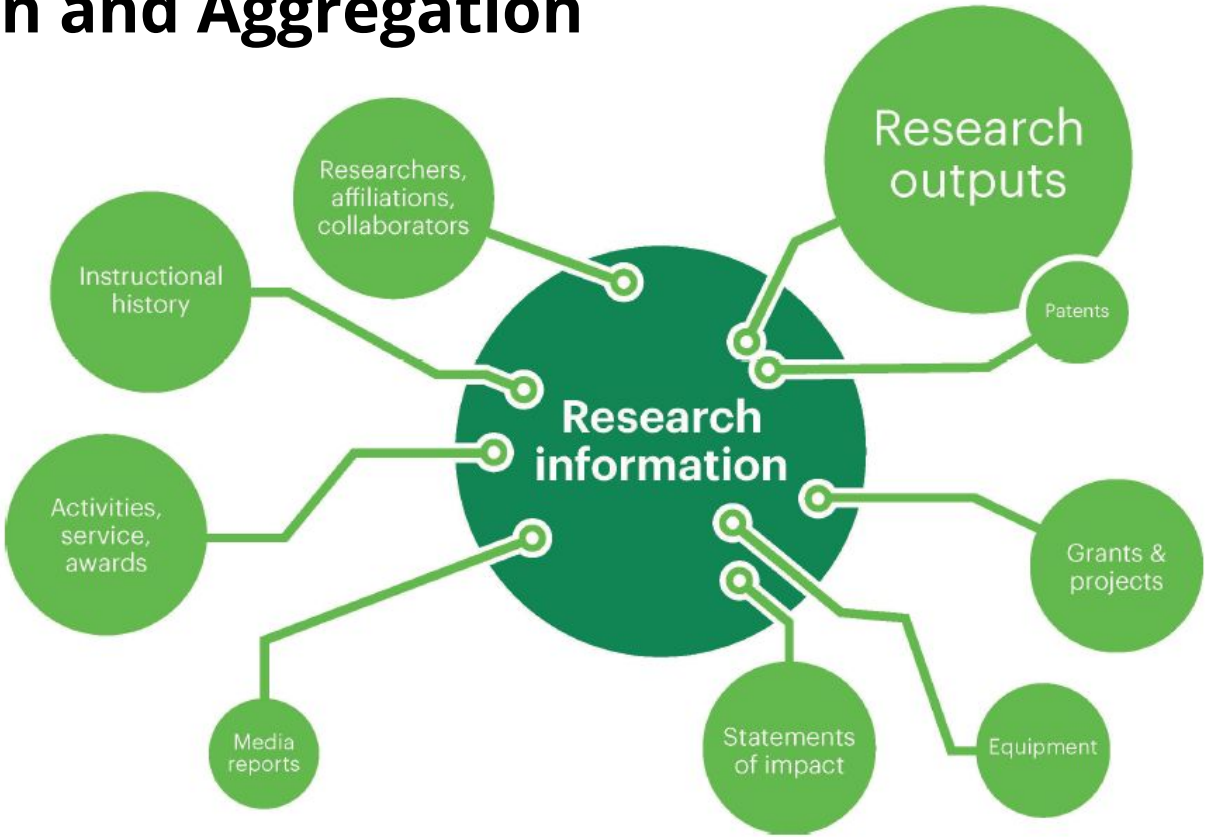
Faculty Information Systems

Scholarly Information Systems

Faculty Activity Reporting Systems

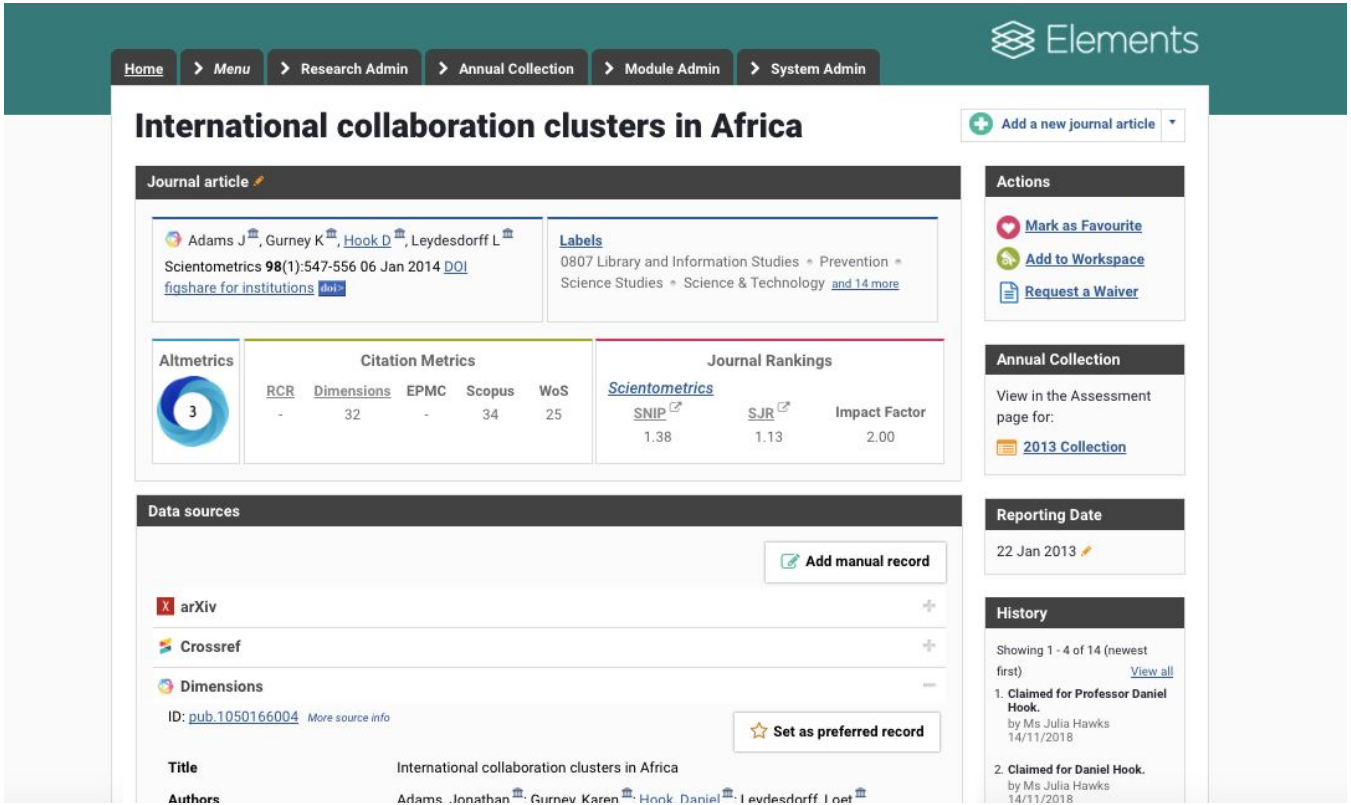
Carnegie Mellon University

Curation and Aggregation



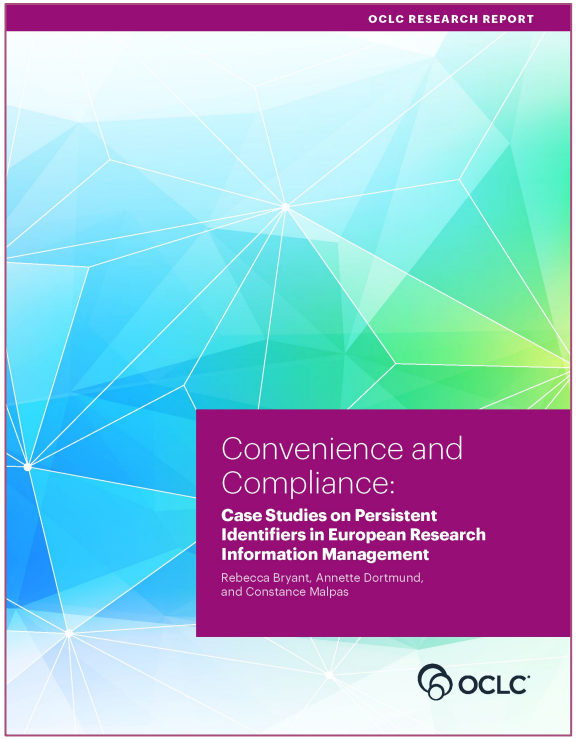
"RIM Metadata" by OCLC Research, from *Research Information Management: Defining RIM and the Library's Role* (doi.org/10.25333/C3NK88), CC BY 4.0

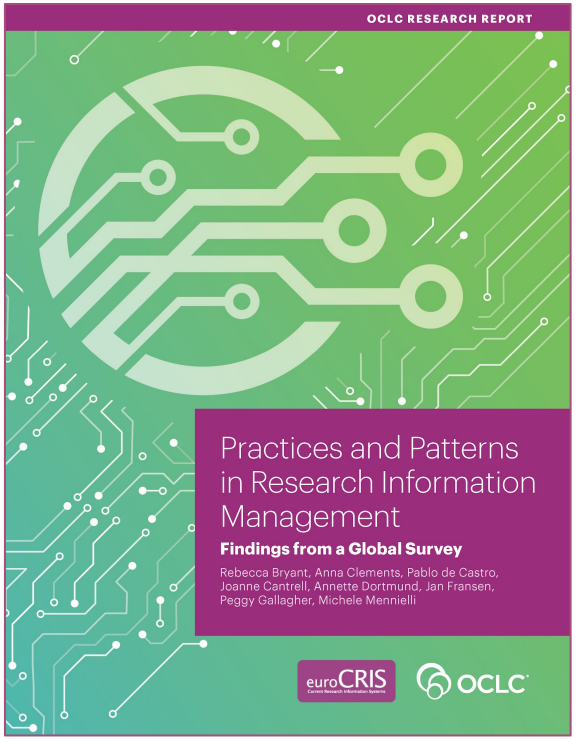
Carnegie Mellon University



Carnegie Mellon University

An Increasing Role for Libraries





"Libraries in Research Information Management" by OCLC Research, from *Research Information Management: Defining RIM and the Library's Role* (doi.org/10.25333/C3NK88), CC BY 4.0)

Carnegie Mellon University

What is the Scholarly Communication Ecosystem?

Carnegie Mellon University

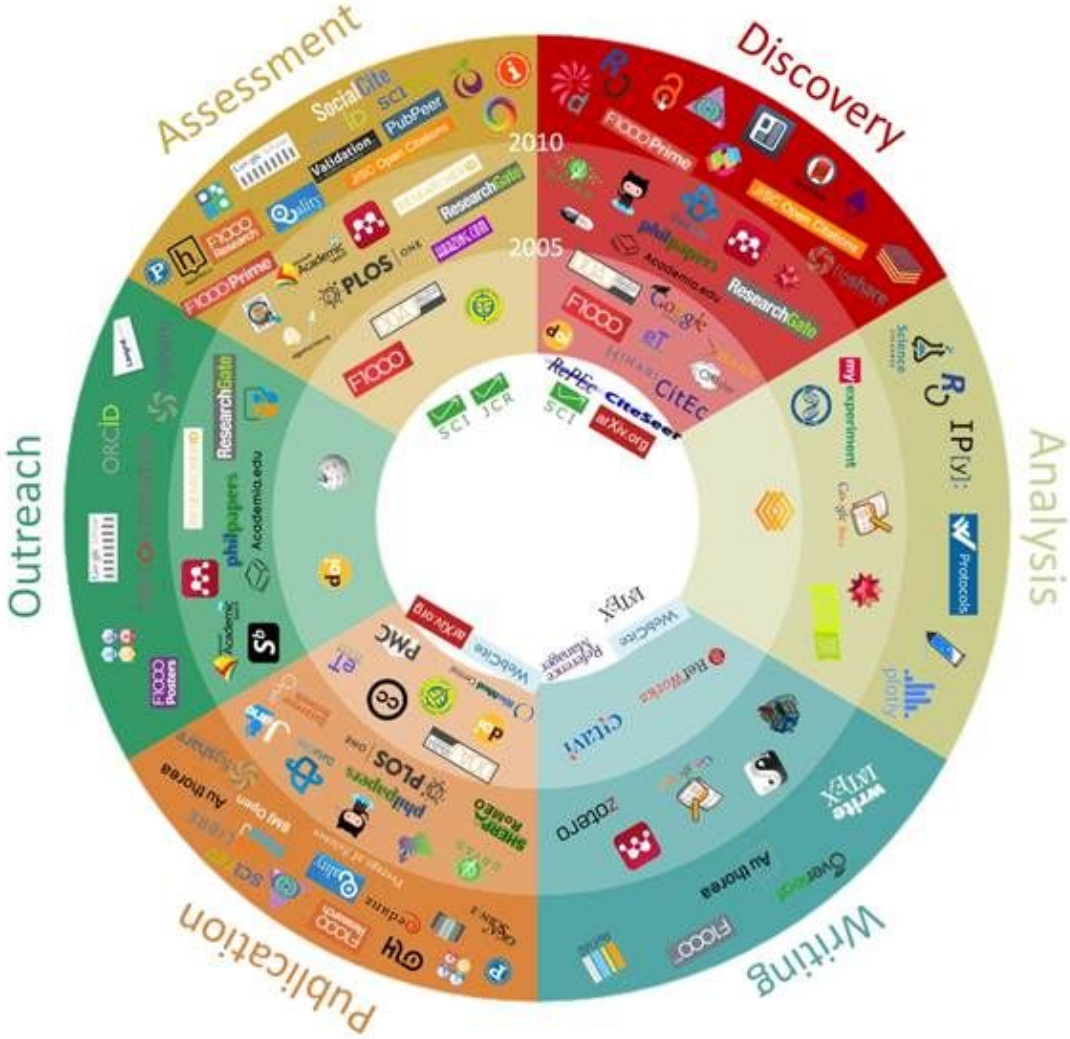
Universiteit Utrecht

Kramer, B., Bosman, J. (2015) *101 Innovations in Scholarly Communication - the Changing Research Workflow.* https://doi.org/10.6084/m9.figshare.1286826.v1

Carnegie Mellon University

Research Lifecycle/Workflow

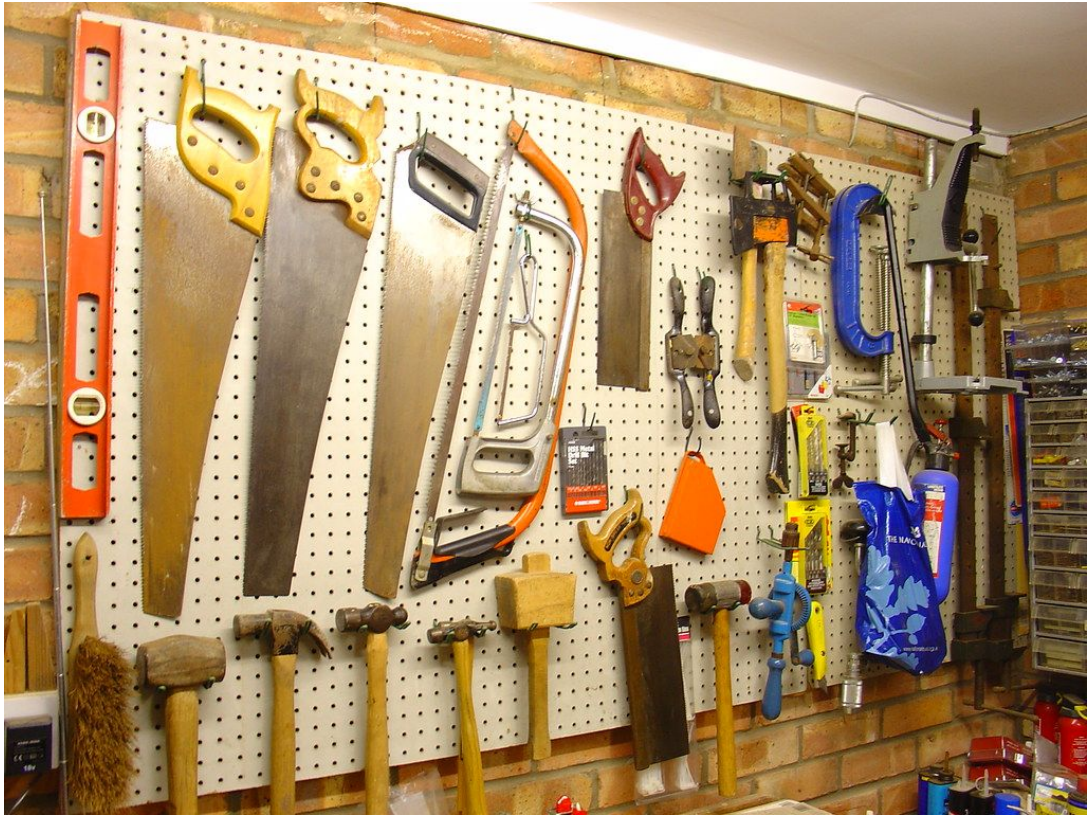
Carnegie Mellon University


Discovery
Analysis
Writing
Publication
Outreach
Assessment
2010
2005

Scholarly Communication Tools

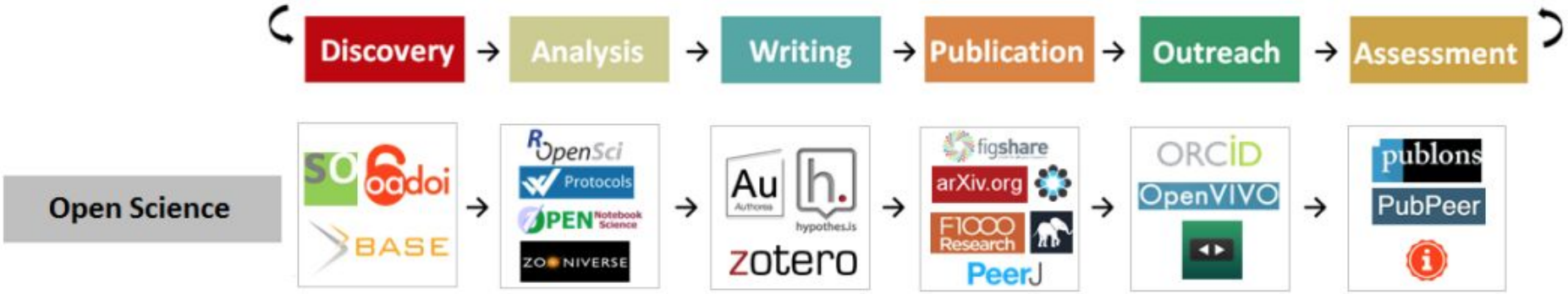
- 2015 - 101 Tools
- 2015 - 400+ Tools
- 2019 - 680+ Tools!

Current List:
https://docs.google.com/spreadsheets/d/1KUMSeq_Pzp4KveZ7pb5rddcssk1XBTiLHniD0d3nDqo/edit#gid=0



Tools by Andrew Fogg CC-BY 2.0

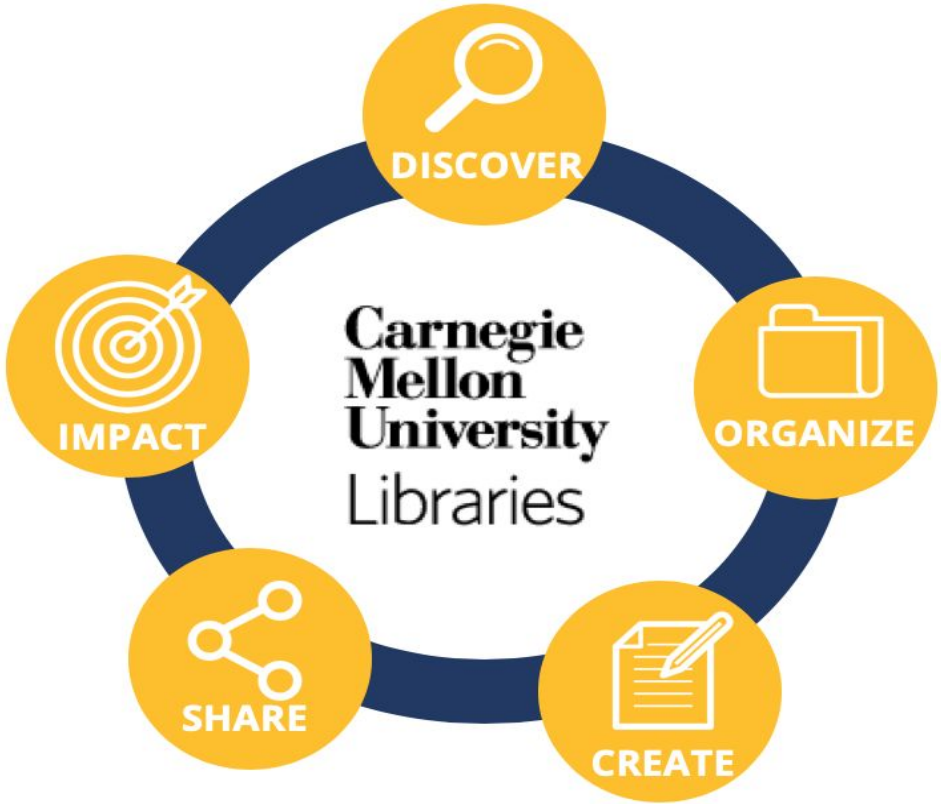
Carnegie Mellon University



Carnegie Mellon University

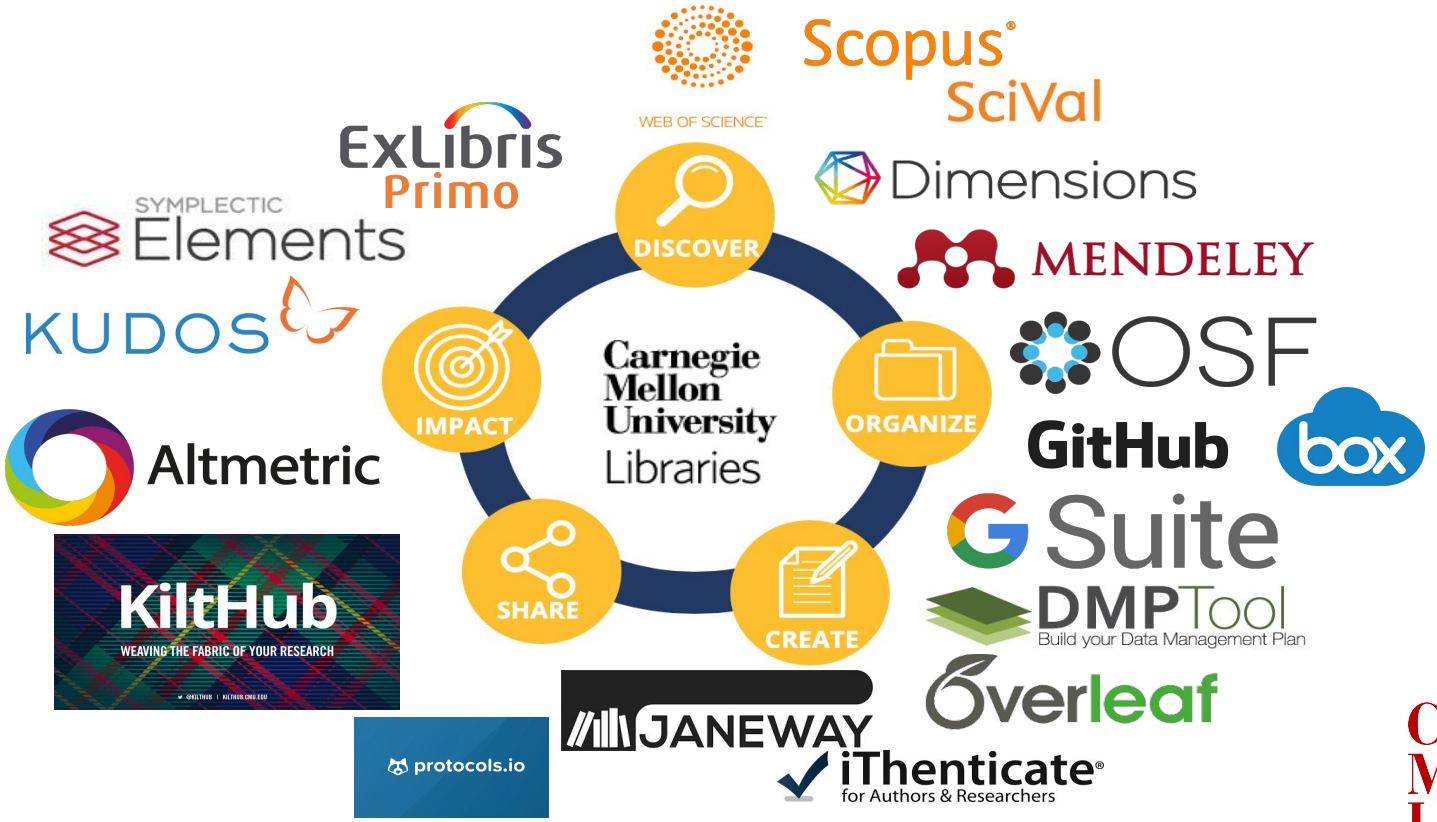
Carnegie Mellon University

THE SCE at CMU



Carnegie Mellon University

THE SCE at CMU



Carnegie Mellon University

Repositories and the SCE

Carnegie Mellon University

DSPACE

figshare

eprints
repository software

zenodo

bepress™

DRYAD

Fedora™

PANGAEA.

Carnegie Mellon University

Helping make research open…

Make research ‘FAIR’

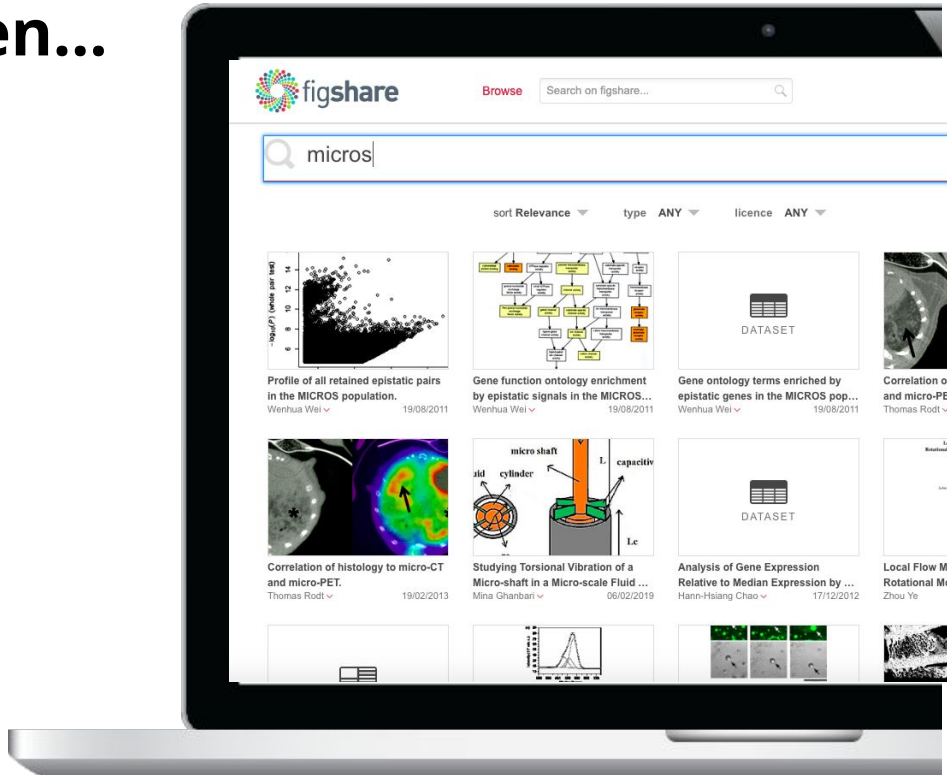
Publish research in any file format

Review before publishing

Interoperate with other systems

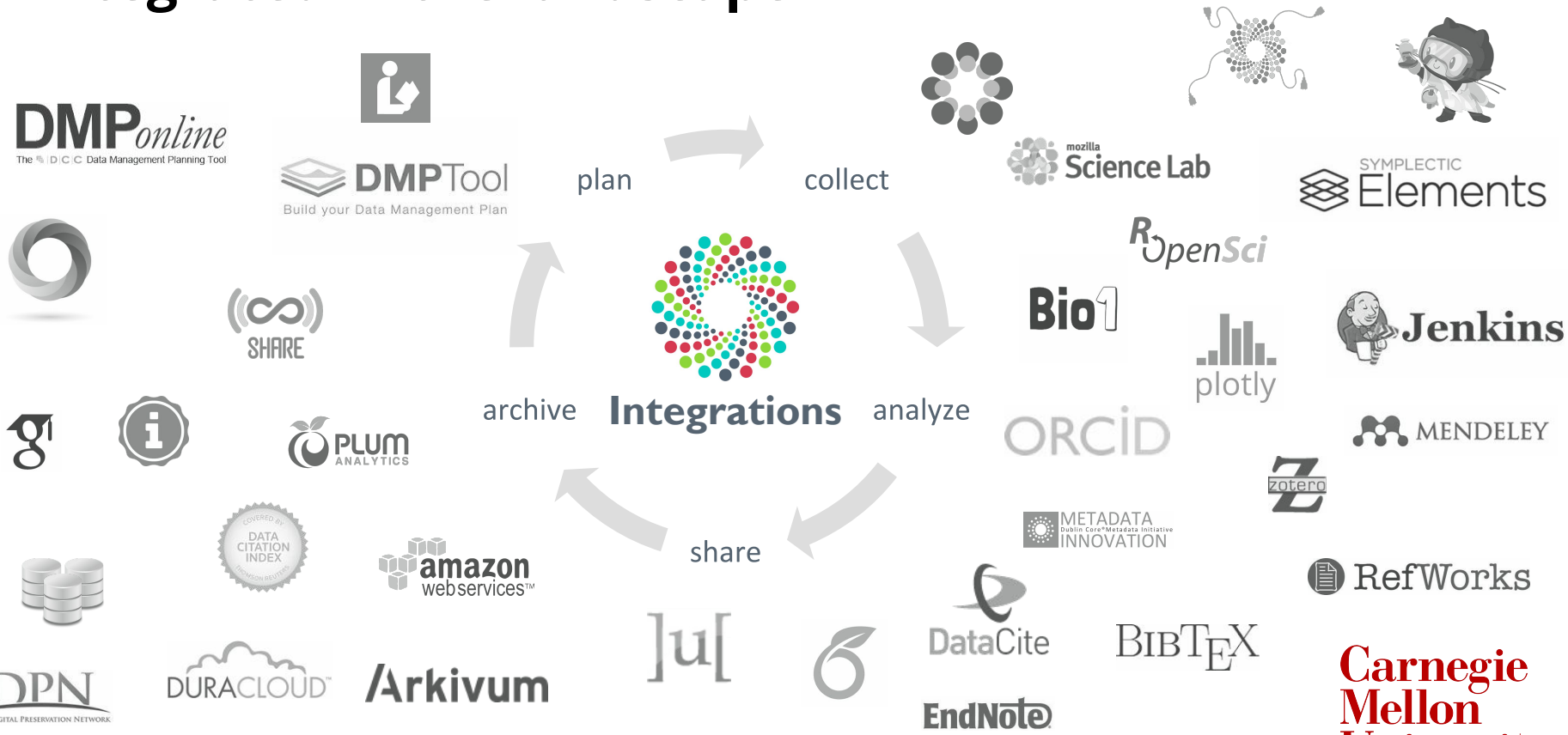
Support OA policies & initiatives

Assign branded DOIs or handles



Carnegie Mellon University

Integrated in the landscape



Carnegie Mellon University

Support for open standards and de facto standards

Type	Indexed in
Datasets	Google Dataset Search Datacite DataONE Dimensions OpenAire SHARE Clarivate Data Citation Index PubMed National metadata registries eg. Research Data Australia
Papers	Google Scholar
Preprints	Google Scholar Dimensions
Theses	Google Scholar Networked Digital Library of Theses and Dissertations

Metadata formats

Currently, the supported formats are: **Dublin Core** (*oai_dc*), **Datacite** (*oai_datacite*), **RDF** (*rdf*), **CERIF XML** (*cerif*), and **Qualified Dublin Core** (*qdc*) (hasPart support).

Google Dataset Search ALPHA

Search for datasets

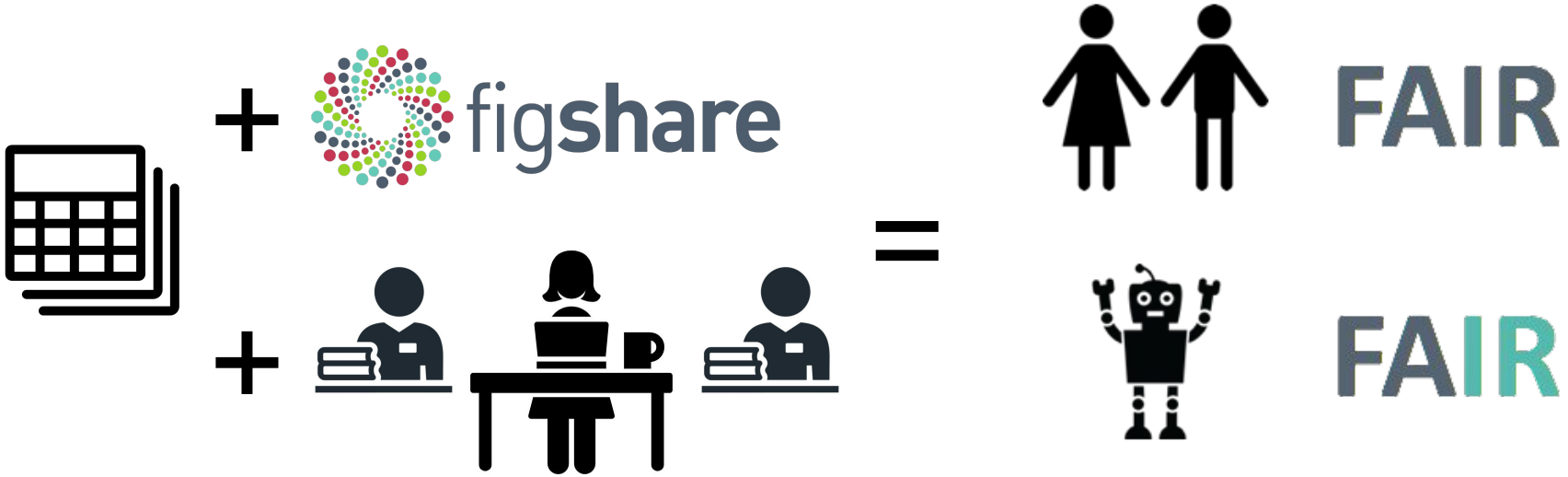
Secure https://figshare.com/rss/portal_group/plos/plosbiology

Apps ALL figshare Client... figshare.com Google+ https://docs.google... figshare: LinkedIn Tumblr Implementa

```
<?xml version="1.0" encoding="UTF-8" ?>
<rss version="2.0">

<channel>
  <title>PLOS Biology of Public Library of Science RSS Feed</title>
  <link>https://plos.figshare.com/plosbiology</link>
  <description>RSS feed for Figshare Group PLOS Biology from the Public Library of Science institution</descri

    <item>
      <title>Towards an integrated view of vocal development</title>
      <link>https://figshare.com/collections/Towards_an_integrated_view_of_vocal_development/4039298</link>
      <description>Vocal development is usually studied from the perspective of neuroscience. In this issue, Z
growth might condition the process. They study the vocalizations of marmoset infants with a wide range of tech
experiments that mimic growth reversal. Their results suggest that the qualitative changes that occur during d
between the nervous system and the biomechanics involved in respiration. This work illustrates how an integrat
</description>
      <category>Ecology, Developmental Biology, Science Policy, Mental Health, Environmental Sciences not else
classified</category>
      <pubDate>2018-03-22 17:27:36</pubDate>
    </item>

    <item>
      <title>The schematics of a bifurcation diagram and its use in experimental design.</title>
      <link>https://figshare.com/articles/The_schematics_of_a_bifurcation_diagram_and_its_use_in_experimental_
      <description>A computational model for slow motor gestures predicts the existence of three regions of th
region, qualitatively different solutions (different behaviors) are expected. One of the parameters is related
varied, different solutions can be found at early stages of development (light grey arrow), and only one solut
Placing marmoset infants in a heliox atmosphere, Zhang and Ghazanfar mimic the reversal of a parameter that co
behaviors (green arrow).</description>
      <category>Ecology, Developmental Biology, Science Policy, Mental Health, Environmental Sciences not else
classified</category>
      <pubDate>2018-03-22 17:27:36</pubDate>
    </item>
```

negie
on
University

Findable, Accessible, Interoperable, Reusable

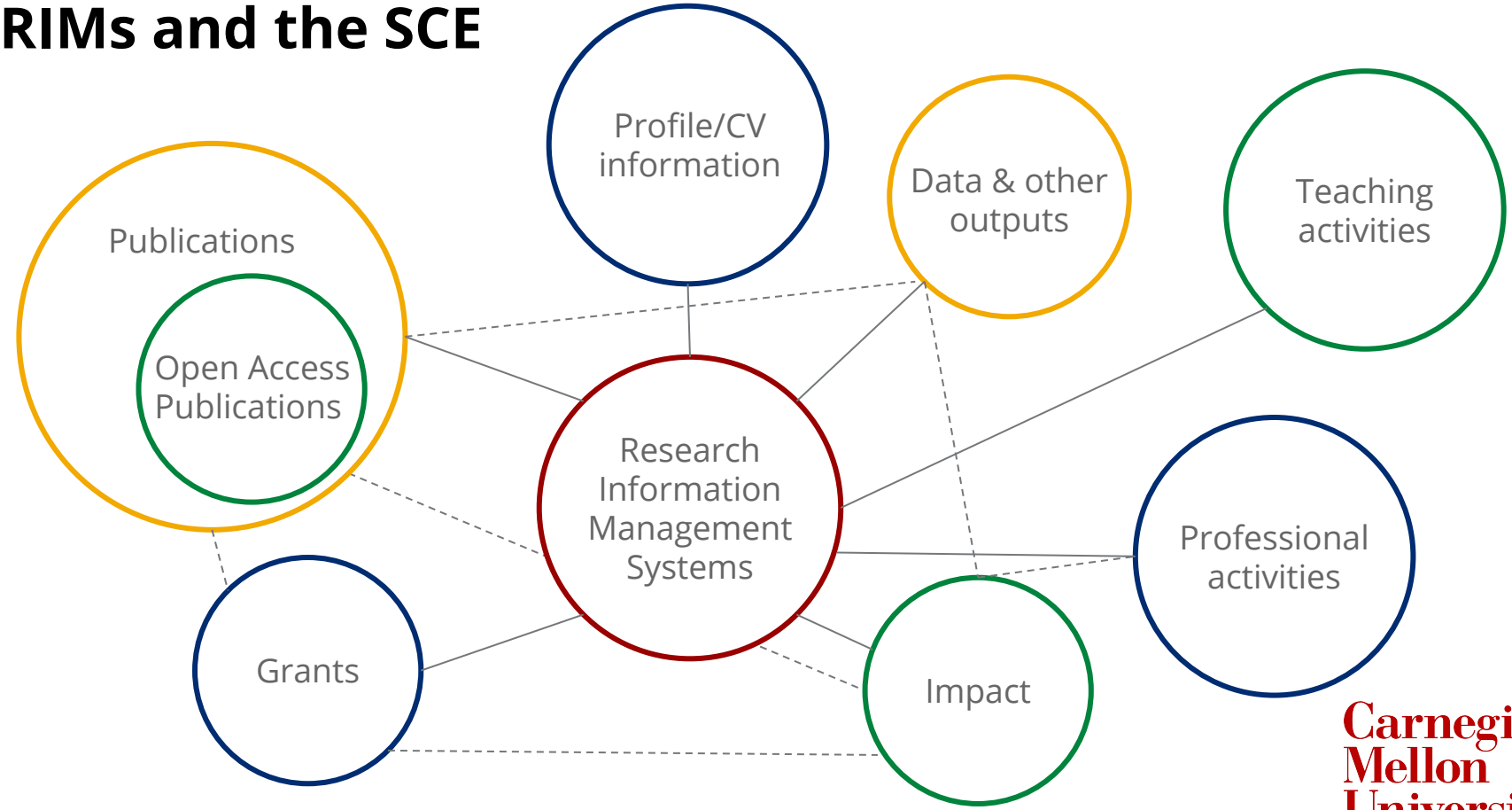


Carnegie Mellon University

RIMS and the SCE

Carnegie Mellon University

RIMs and the SCE



Carnegie Mellon University

Beyond collecting data…

Curate & enrich metadata

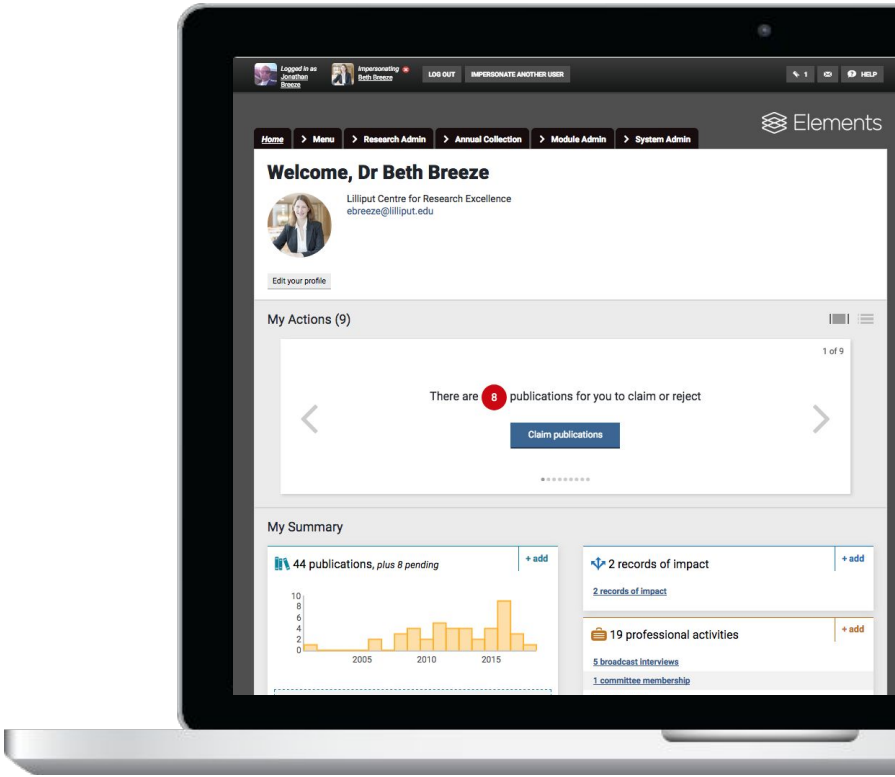
Internal Reporting & Analysis

Populate Public Profiles

Capture Research Impact

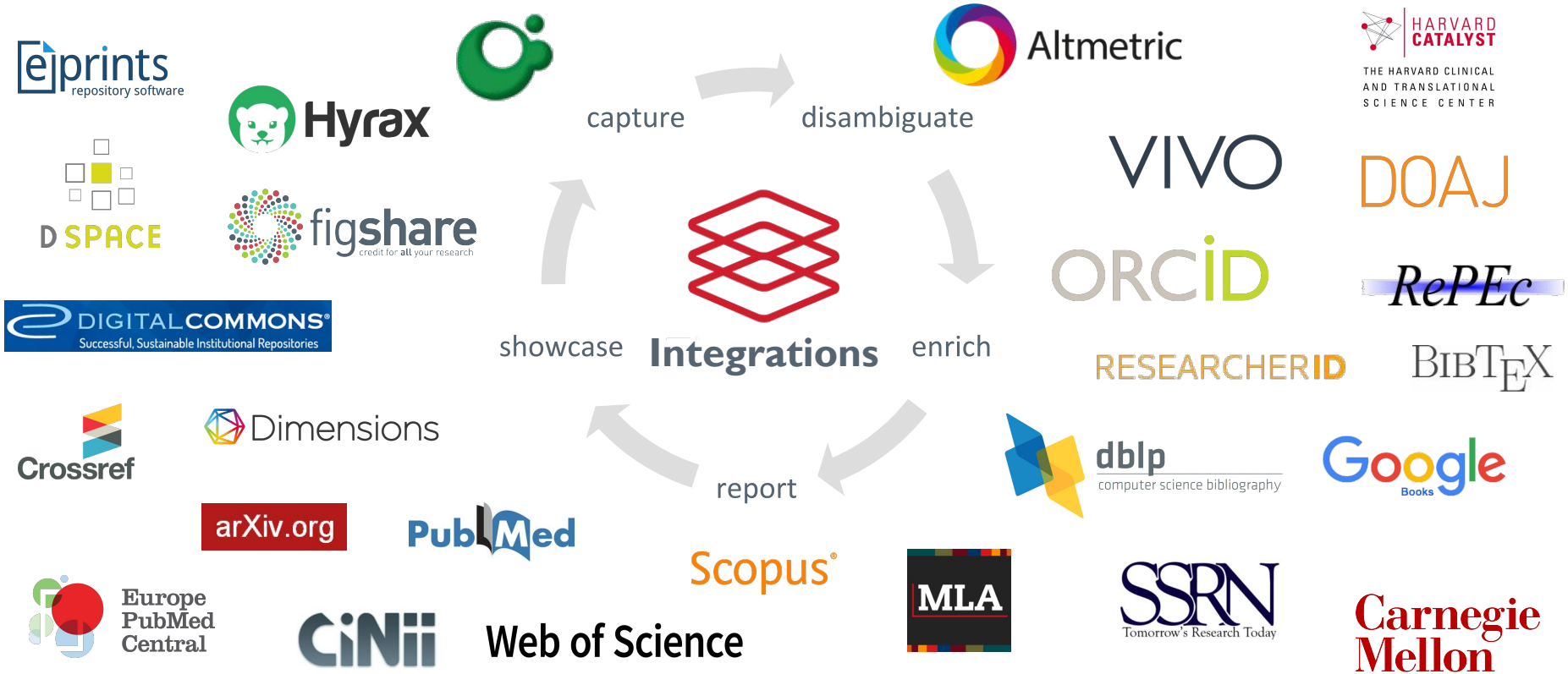
Support Open Access

Annual Reviews & Assessment
(e.g. REF 2021)



Carnegie Mellon University

Integrated in the landscape

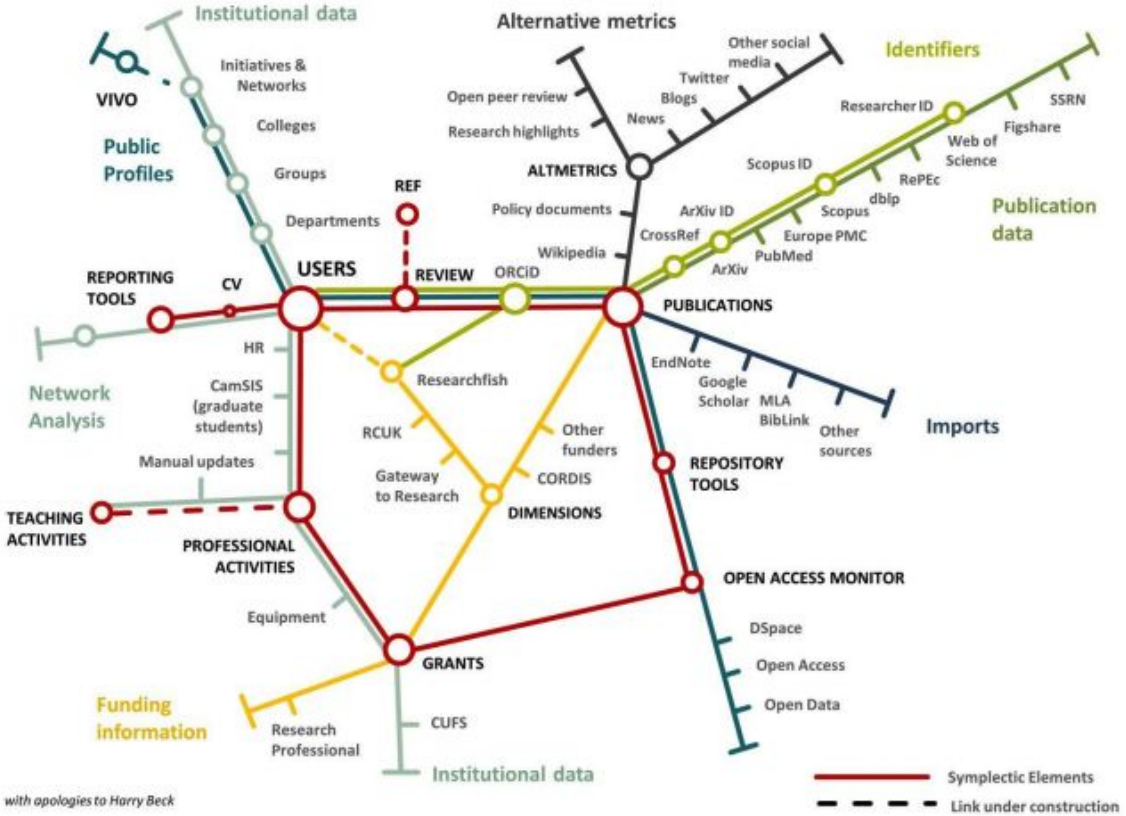


Views of the Network

"RIM Uses" by OCLC Research, from *Research Information Management: Defining RIM and the Library's Role* (doi.org/10.25333/C3NK88), CC BY 4.0)

Carnegie Mellon University

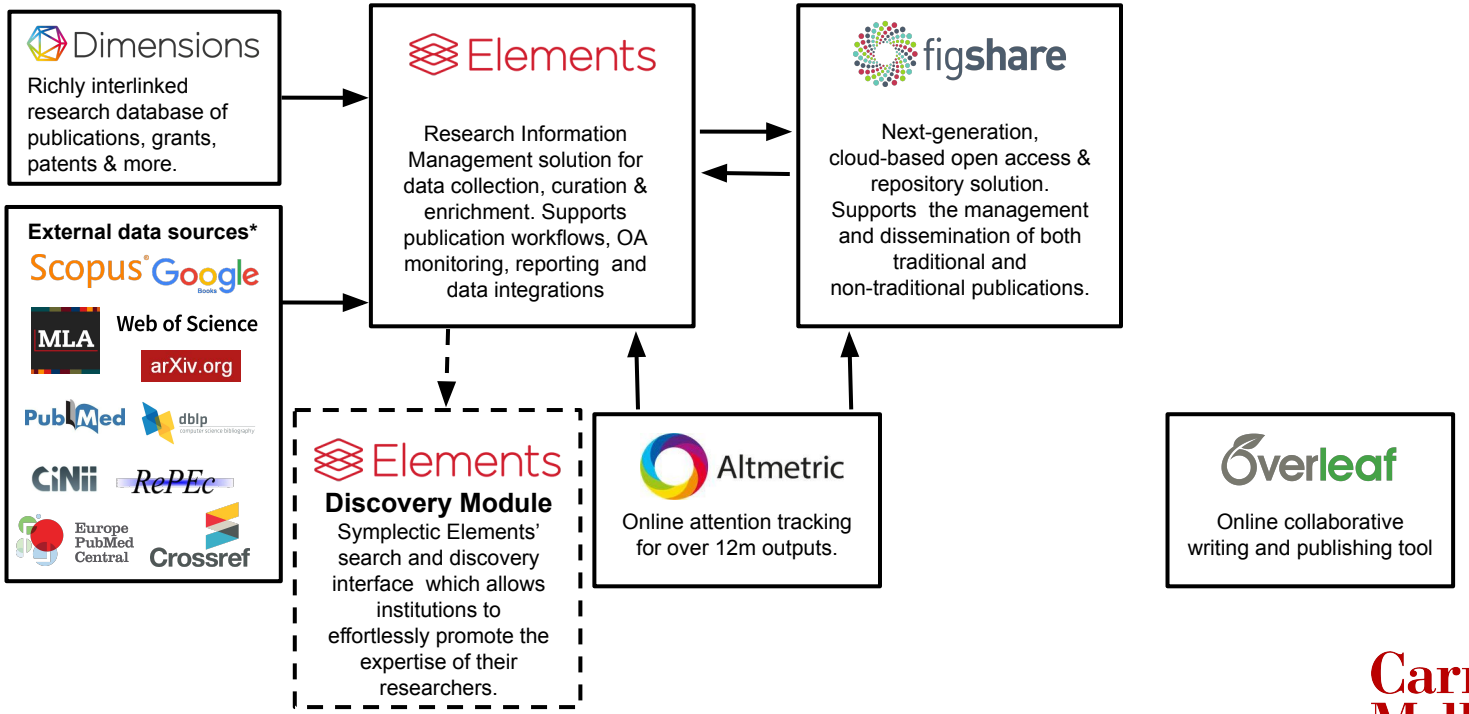
Views of the network: Cambridge University



Juergen Wastl (2017). Academic Profile, Network and Collaboration Analysis : Use Cases for the University https://doi.org/10.17863/CAM.13805

Carnegie Mellon University

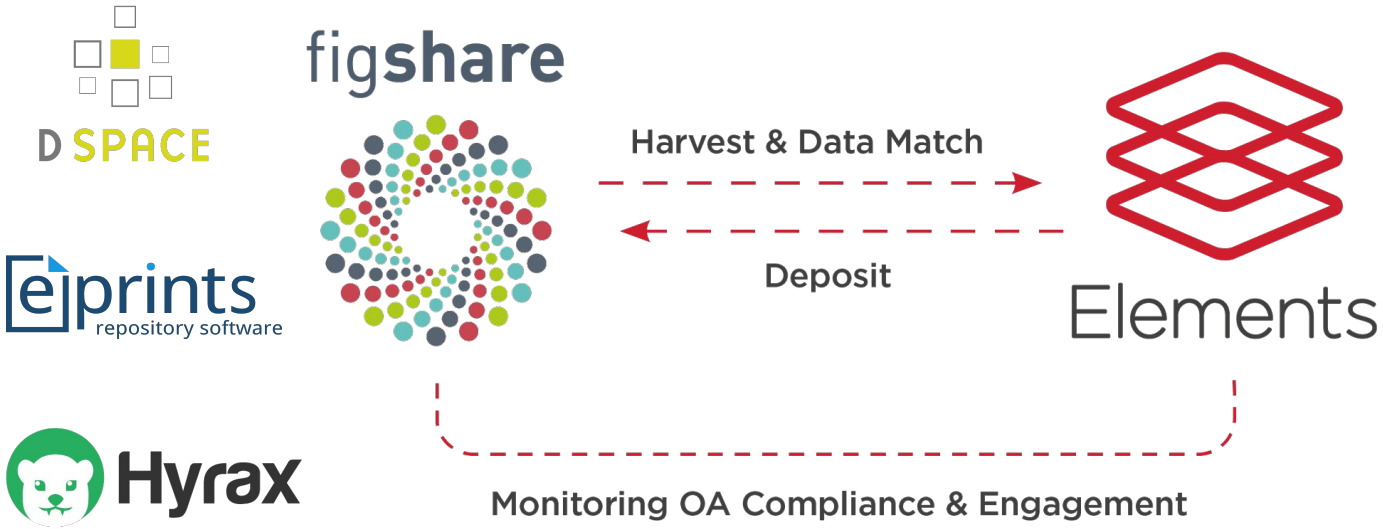
Views of the network: Carnegie Mellon University



*Data sources used by Elements. Data sources actually integrated with Primo may differ.

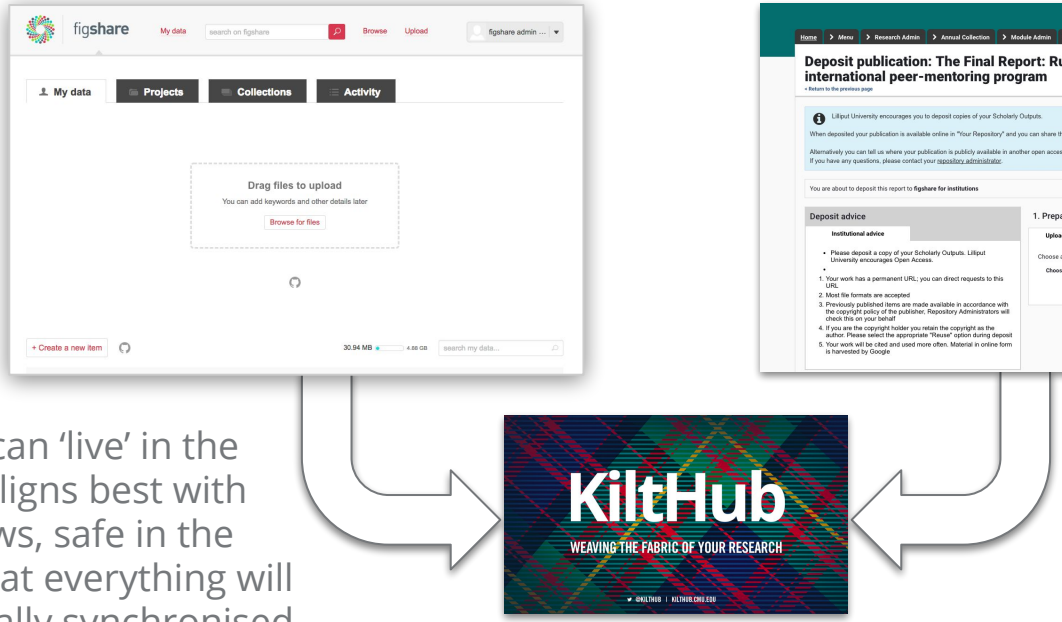
Carnegie Mellon University

Creating rich bi-directional integrations



Carnegie Mellon University

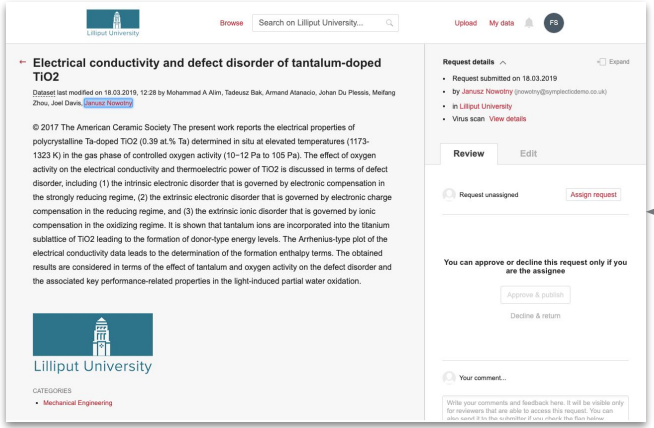
Supports flexible workflows for researchers

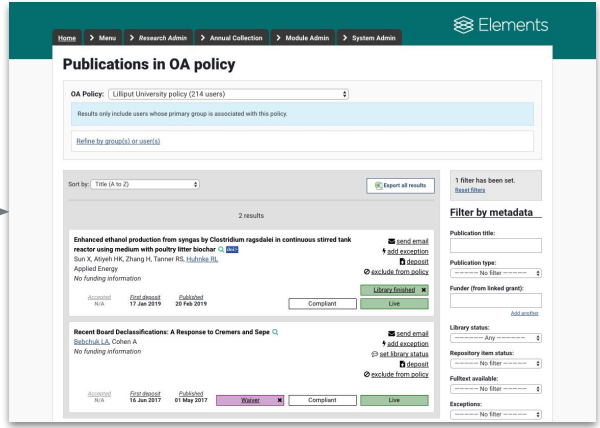


Researchers can ‘live’ in the system that aligns best with their workflows, safe in the knowledge that everything will be automatically synchronised between the systems

Carnegie Mellon University

Consolidates activities for administrators





Manage repository collections including:

- Review & verify deposits
- Categorise materials
- Manage embargos
- Manage reuse licenses
- Store & preserve files

Support OA engagement & compliance including:

- Identify which outputs are ‘in policy’ (aka targeted for deposit)
- Monitor proportion of outputs that have been deposited.
- Tools to encourage deposits or deposit on researcher’s behalf.

Carnegie Mellon University

Maximising reuse, minimises manual effort

Data-focussed interoperability minimises time spent typing metadata into manual entry forms over and over, which in turn reduces the likelihood of typos and other human error.

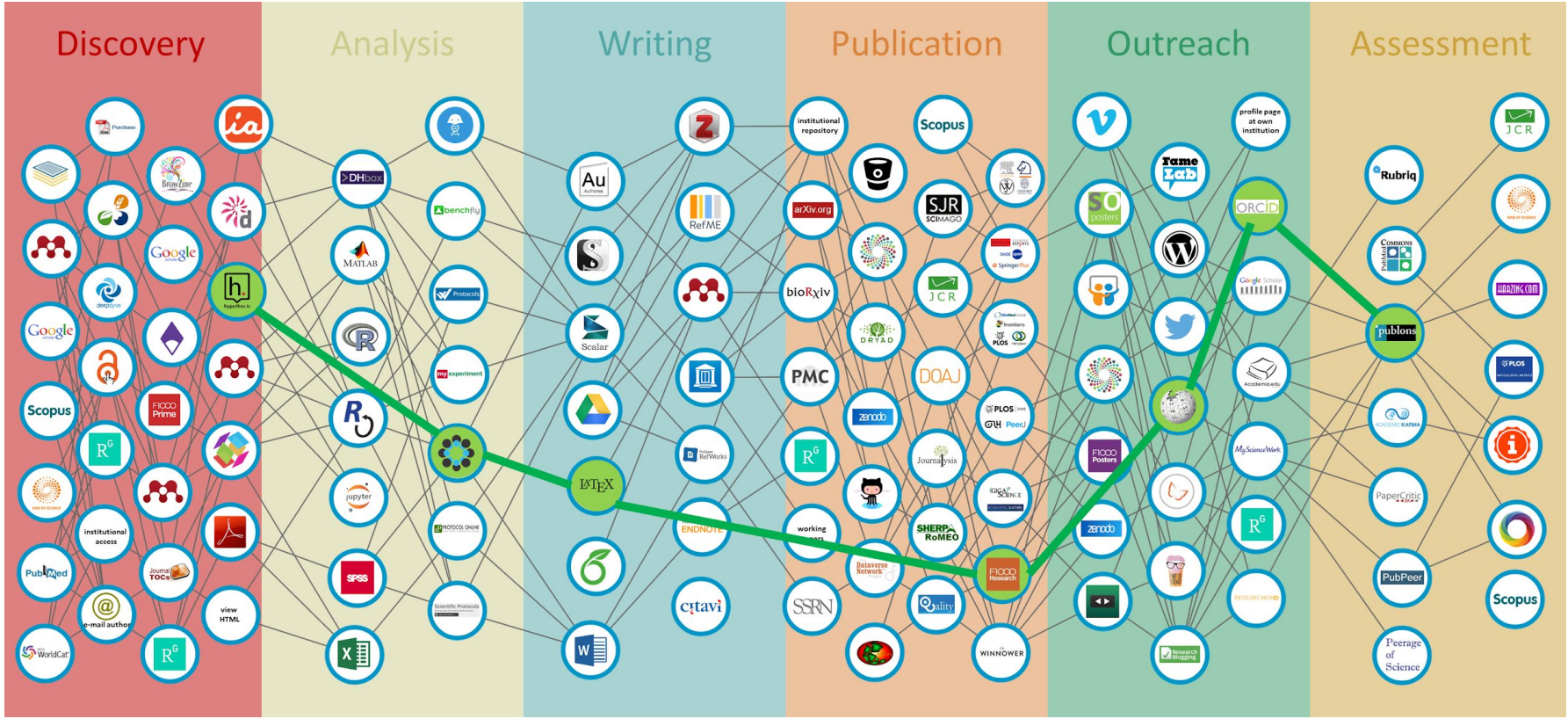
Being asked to do the same task multiple times is a common source for frustration for academics and administrators alike - reducing this frees up time for more important tasks.

Carnegie Mellon University

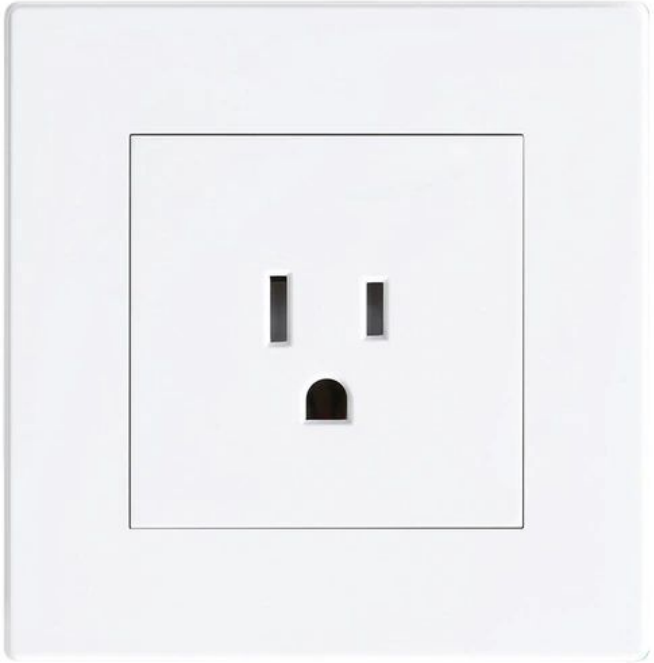
Building the Network

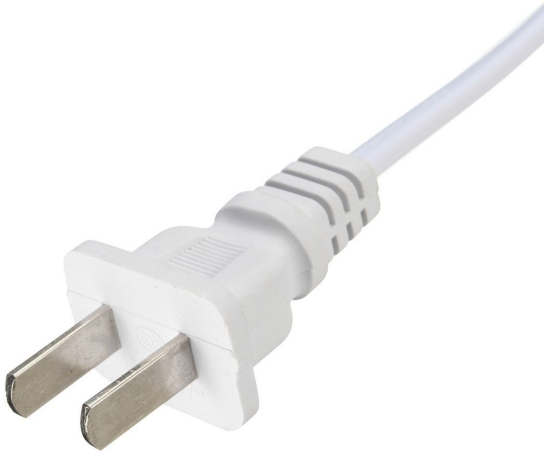
Carnegie Mellon University

Models for integrations



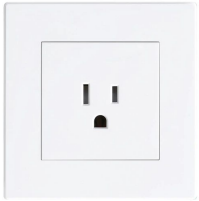
Foundations for interconnectivity

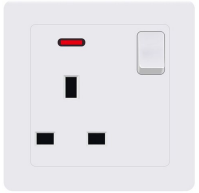




Carnegie Mellon University

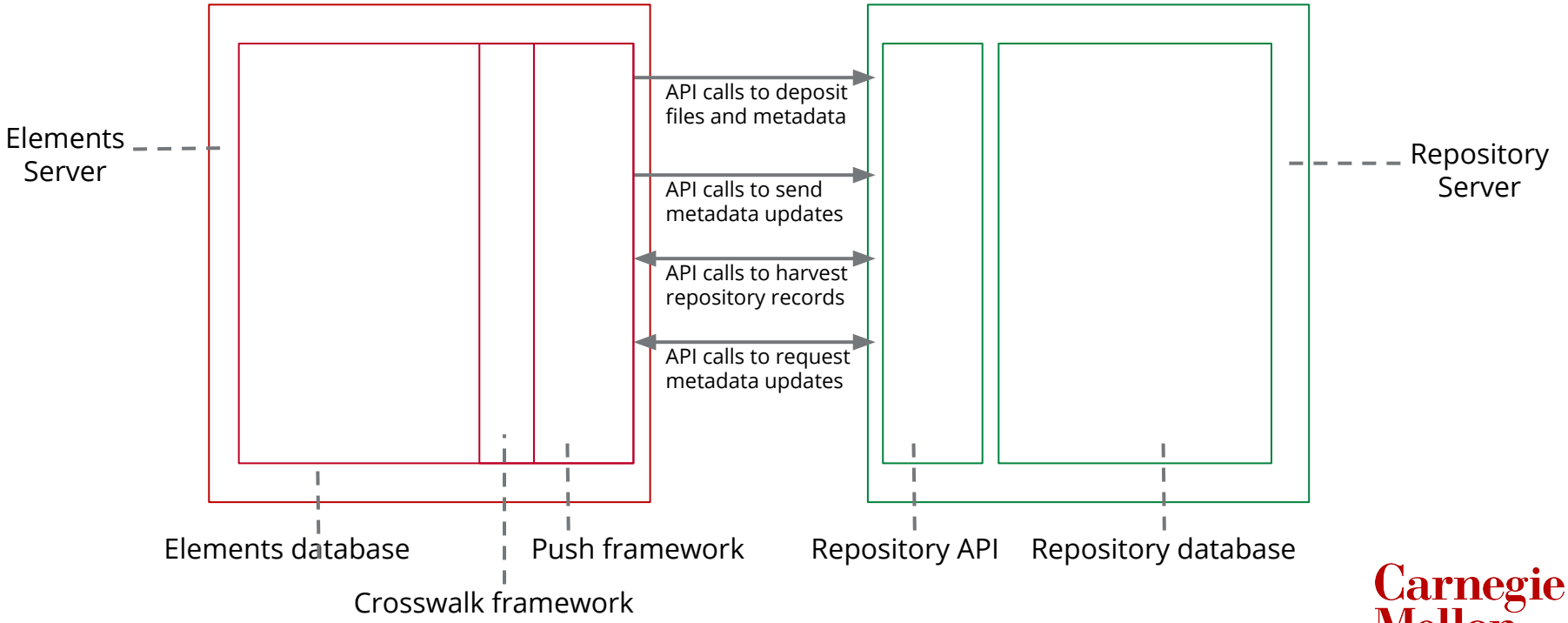
Foundations for interconnectivity





Carnegie Mellon University

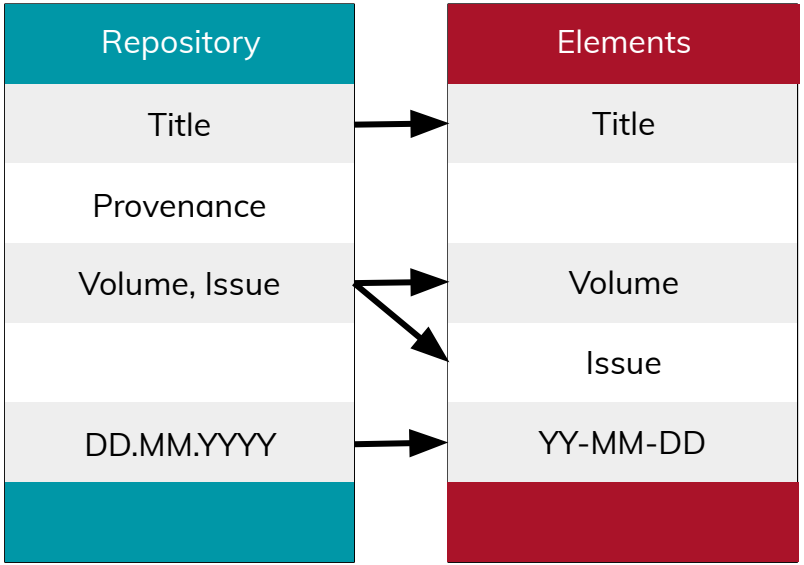
Under the hood of our bi-directional integrations



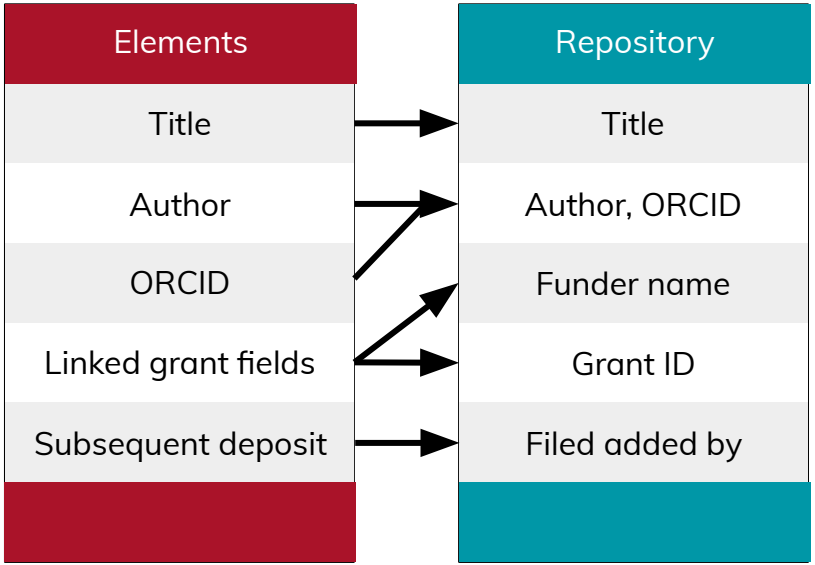
Carnegie Mellon University

Crosswalks & translation layers

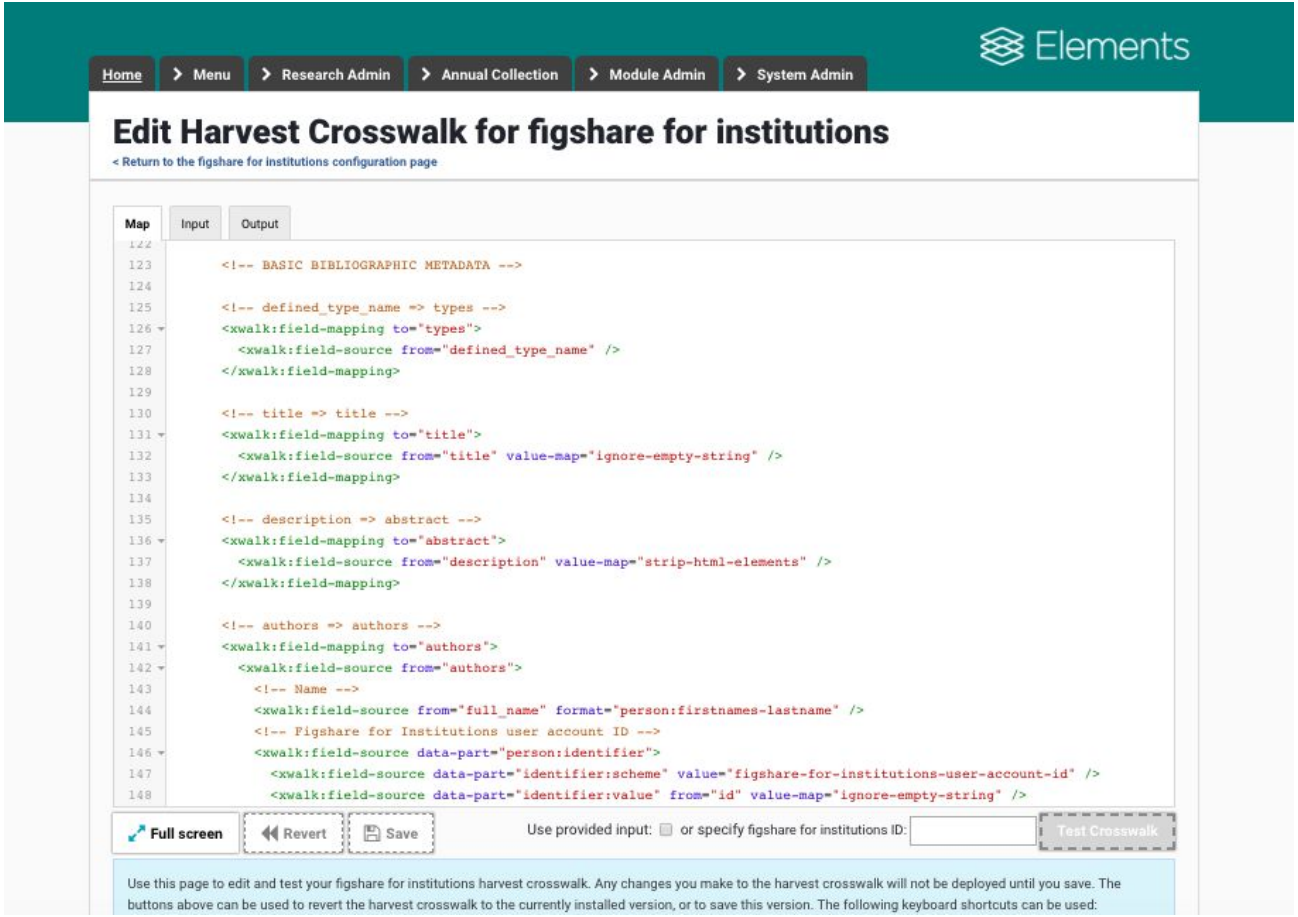
Harvest crosswalks



Deposit crosswalks



Carnegie Mellon University



Carnegie Mellon University

Cite Share + Collect (you need to log in first)

DataCite

Gu, Peishi; Li, Hugh Z.; Ye, Qing; Robinson, Ellis; Apte, Joshua S.; Robinson, Allen; et al. (2019): Intracity Variability of Particulate Matter Exposure Is Driven by Carbonaceous Sources and Correlated with Land-Use Variables. figshare. Journal contribution.

Select your citation style and t
your mouse over the citation t
select it.

Repository DOI

https://doi.org/10.1184/R1/7568579.v1

Intracity Variability of Particulate Matter Exposure Is Driven by Carbonaceous Sources and Correlated with Land-Use Variables

305 views | 0 downloads

Journal Contribution posted on 23.09.2018, 20:00 by Peishi Gu, Hugh Z. Li, Qing Ye, Ellis Robinson, Joshua S. Apte, Allen Robinson, Albert Presto

CATEGORIES

- Atmospheric Sciences
- Atmospheric Aerosols

DataCite

NLM

DC

FUNDING

Center for Air, Climate, and Energy Solutions (CACES)
Environmental Protection Agency

Spatial Variability in Organic Aerosol: Relationships between Sources, Processing, and Composition
Directorate for Geosciences

TIMELINE

Posted date
24.09.2018

First online date
Revision date
Submission date
09.01.2019

REFERENCES

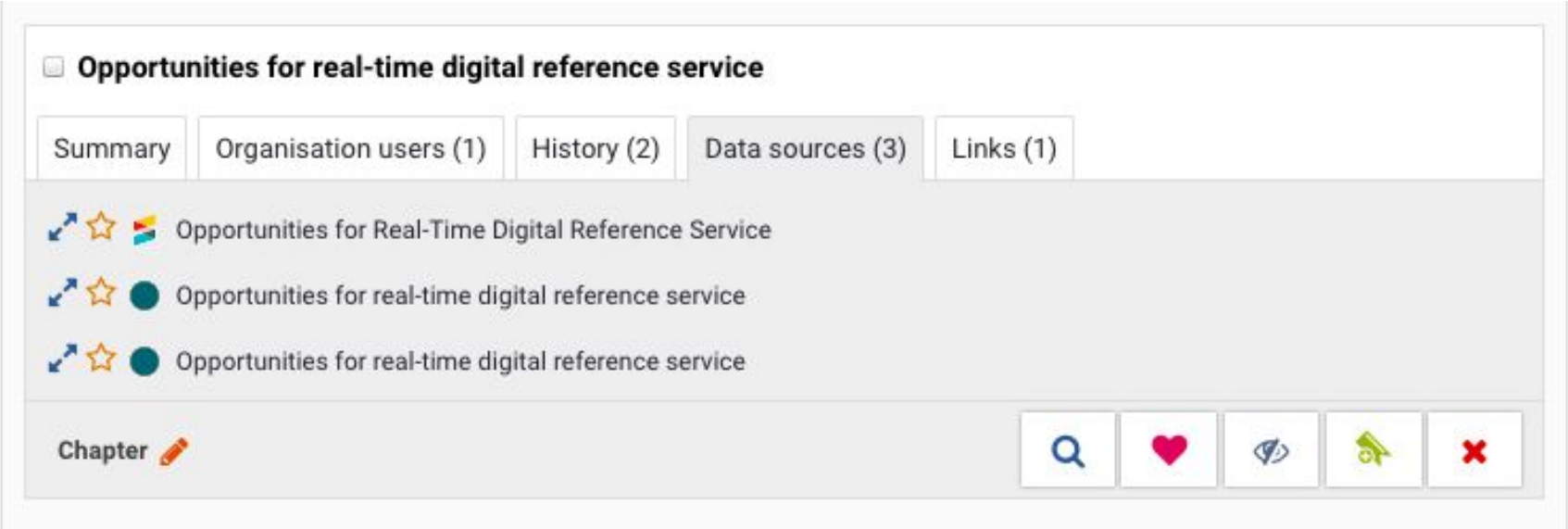
- https://pubs.acs.org/doi/10.1021/acs.est.8b03833
- https://doi.org/10.1184/R1/7353287

Publisher's DOI

PUBLISHER STATEMENT

This is the author's accepted manuscript version of, " Gu, P., Li, H., Ye, Q., Robinson, E., Apte, J., Robinson, A., & Presto, A. (2018). Intracity Variability of Particulate Matter Exposure Is Driven by Carbonaceous Sources and Correlated with Land-Use Variables. Environmental Science & Technology, 52(20), 11545–11554. https://doi.org/10.1021/acs.est.8b03833."

Mutiple Sources



Carnegie Mellon University

CMU Crosswalks & Translations

Publisher DOI Present

Publisher DOI Not Present

Carnegie Mellon University

CMU Crosswalks

Publisher DOI Present

Scopus® SciVal

WEB OF SCIENCE™

Publisher DOI Not Present

Carnegie Mellon University

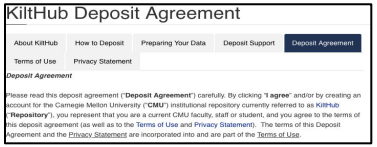
Carnegie Mellon University's workflows

Metadata Harvesting & Content Matching

Monitoring OA Status and Engagement

Elements

Content Verification, Submission, and Deposit



SHERPA RoMEO

Carnegie Mellon University

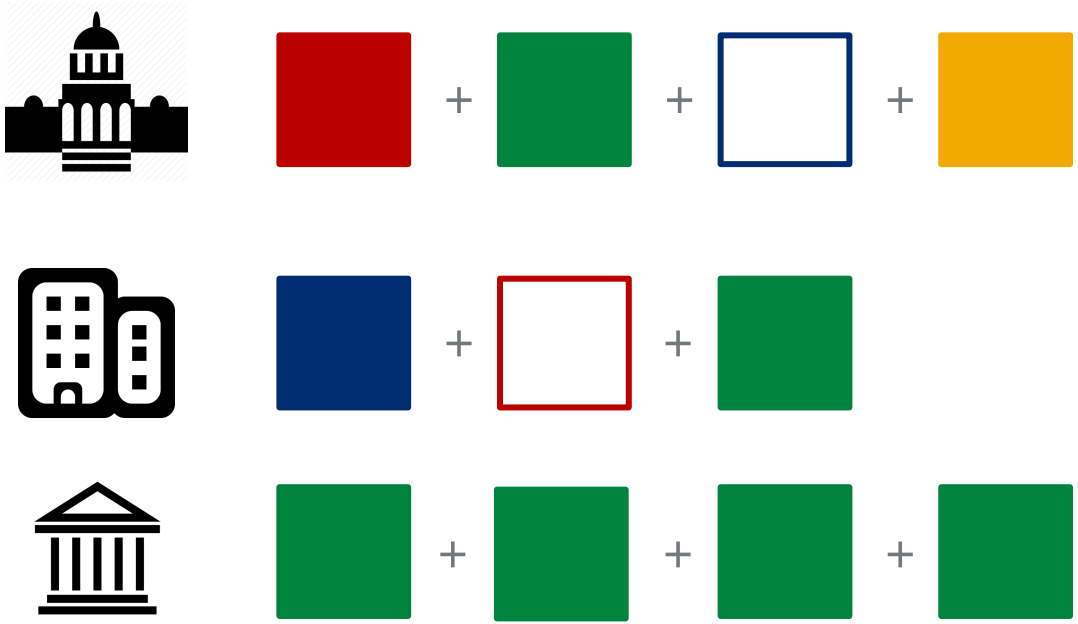
Why this Matters

Why this matters for institutions

"Institutional Collaborators" by OCLC Research, from *Research Information Management: Defining RIM and the Library's Role* (doi.org/10.25333/C3NK88), CC BY 4.0

Carnegie Mellon University

Integrations = Choices



Carnegie Mellon University

Why it matters to do it the right way?

We need to make integrations **sustainable and scalable.**

Symplectic supports over 80 repository integrations for over 70 institutions - we need to ensure those integrations meet our clients needs but also are **easy to implement, use, manage and maintain.**

Carnegie Mellon University

Future of Repositories and RIMS

Carnegie Mellon University

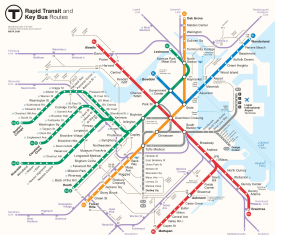
Continuing to build the network

Interoperability is paramount

- APIs should be standard
- Also translation layers & admin toolsets
- Power to swap out any one system

Wider range of integrations between systems

- Connections in line with the research lifecycle
- Consider opportunities for chaining together integrations



Flexible workflows & configurability

- Consistent but still configurable.
- Support choice eg. multiple deposit options.

Socket Wrench and Sockets by Kae. CC-BY 3.0

Carnegie Mellon University

Enhanced interoperability starts with asking questions

Creating connections

- Does the platform offer APIs or other kinds of interfaces for feeding & receiving data from other systems? Automated/manual
- What kinds of operations do the APIs support? Read/Write/Update etc.

Metadata Quality

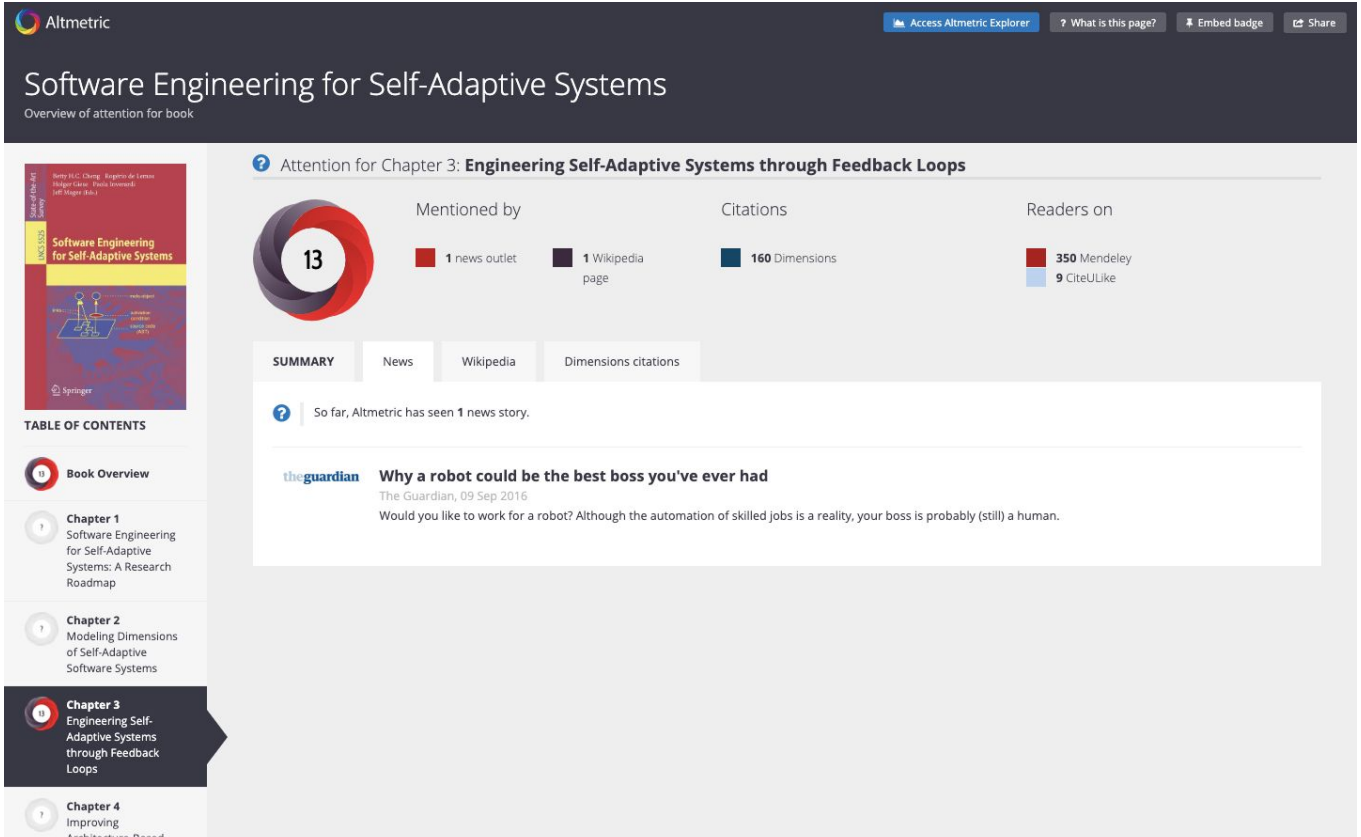
- Which metadata standards does the system use?
- Does the system offer crosswalks/translation layers or other kinds of metadata mapping tools?
- How does the system handle metadata updates?

Sustainability

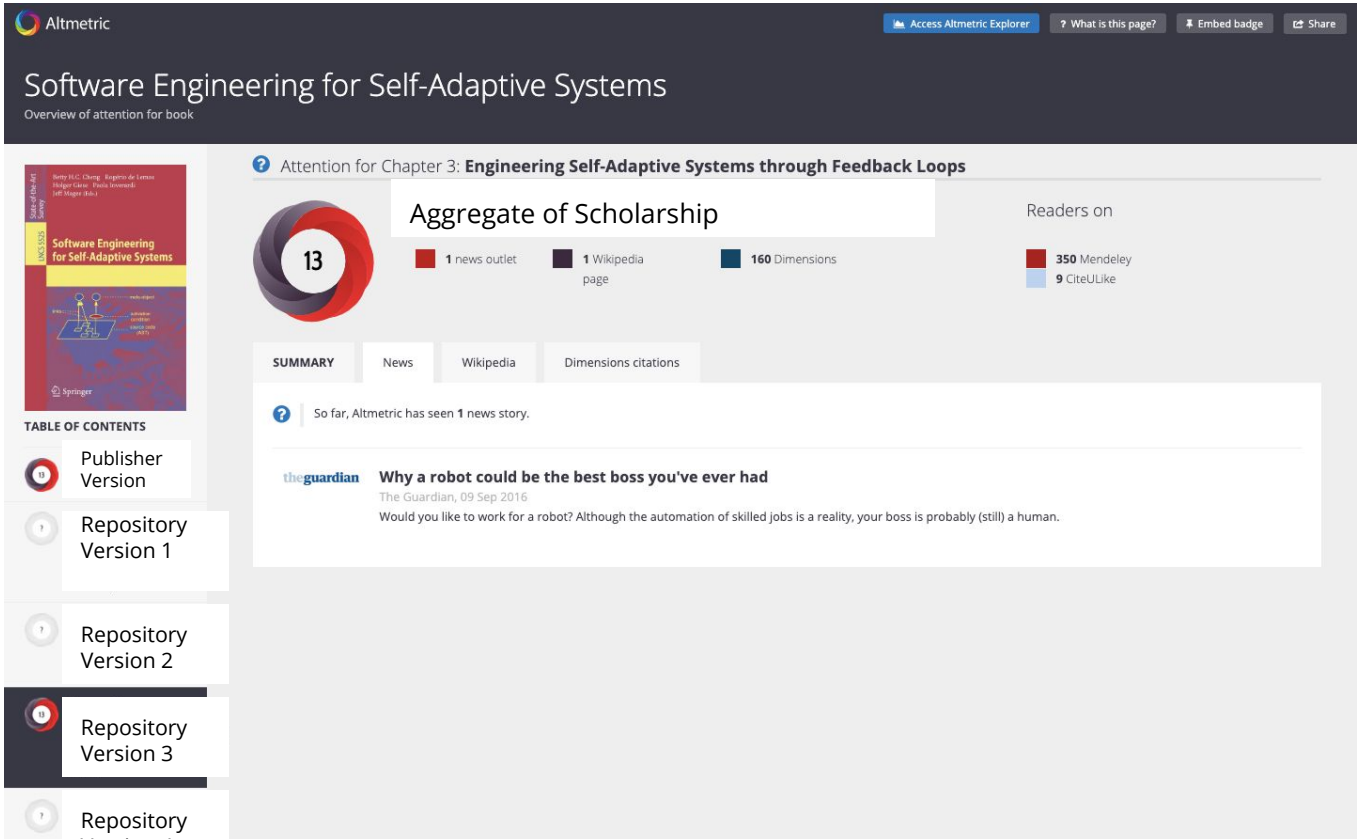
- What tools does the system offer to help maintain the integration?
- How does the system manage data at volume?
- How can we configure the integration to meet our needs?

Carnegie Mellon University

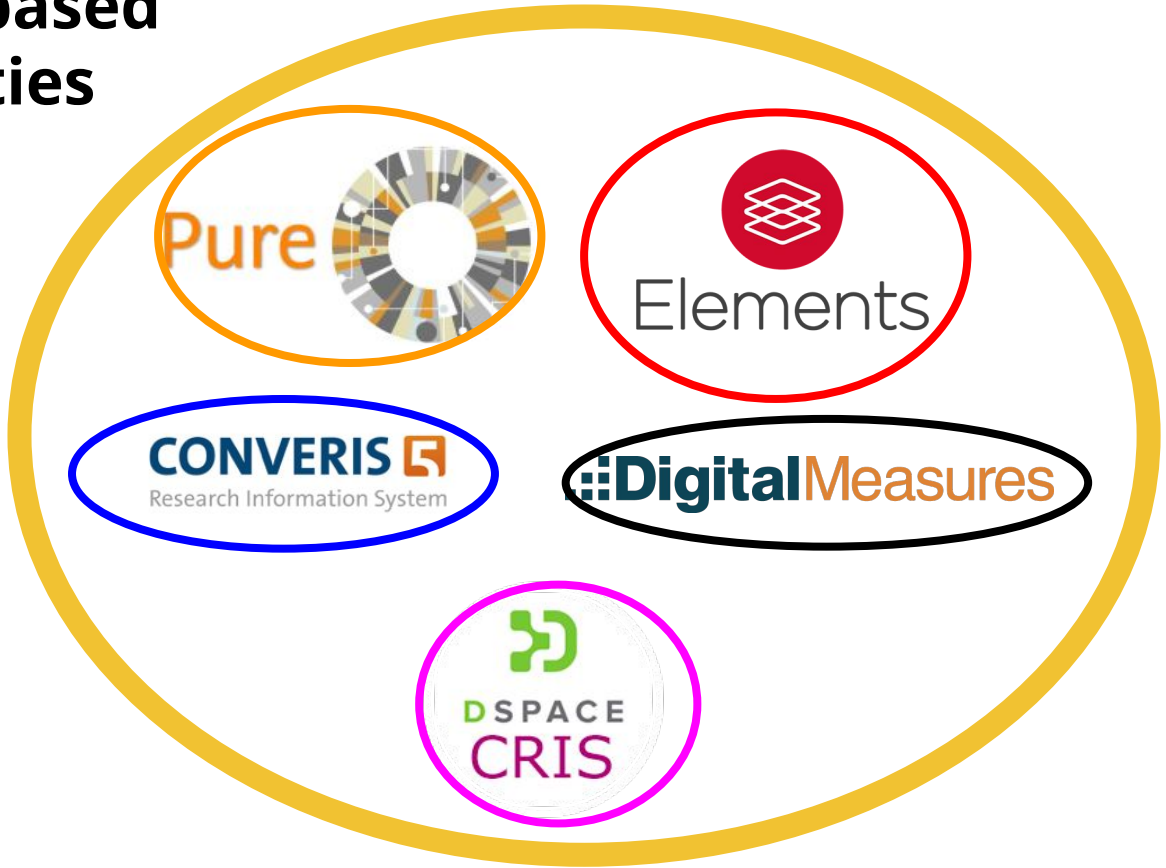
Integrated Data Models



Integrated Data Models

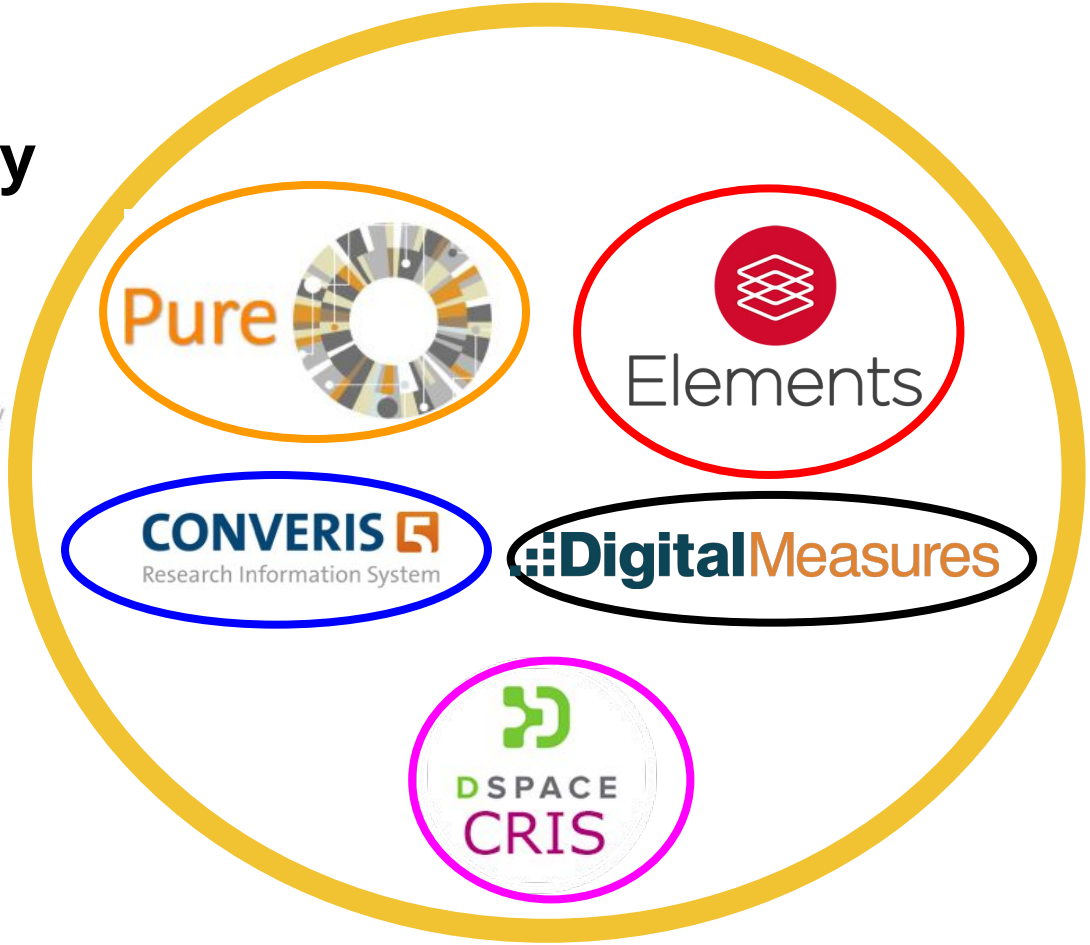


Platform-based Communities



Carnegie Mellon University

Global Community



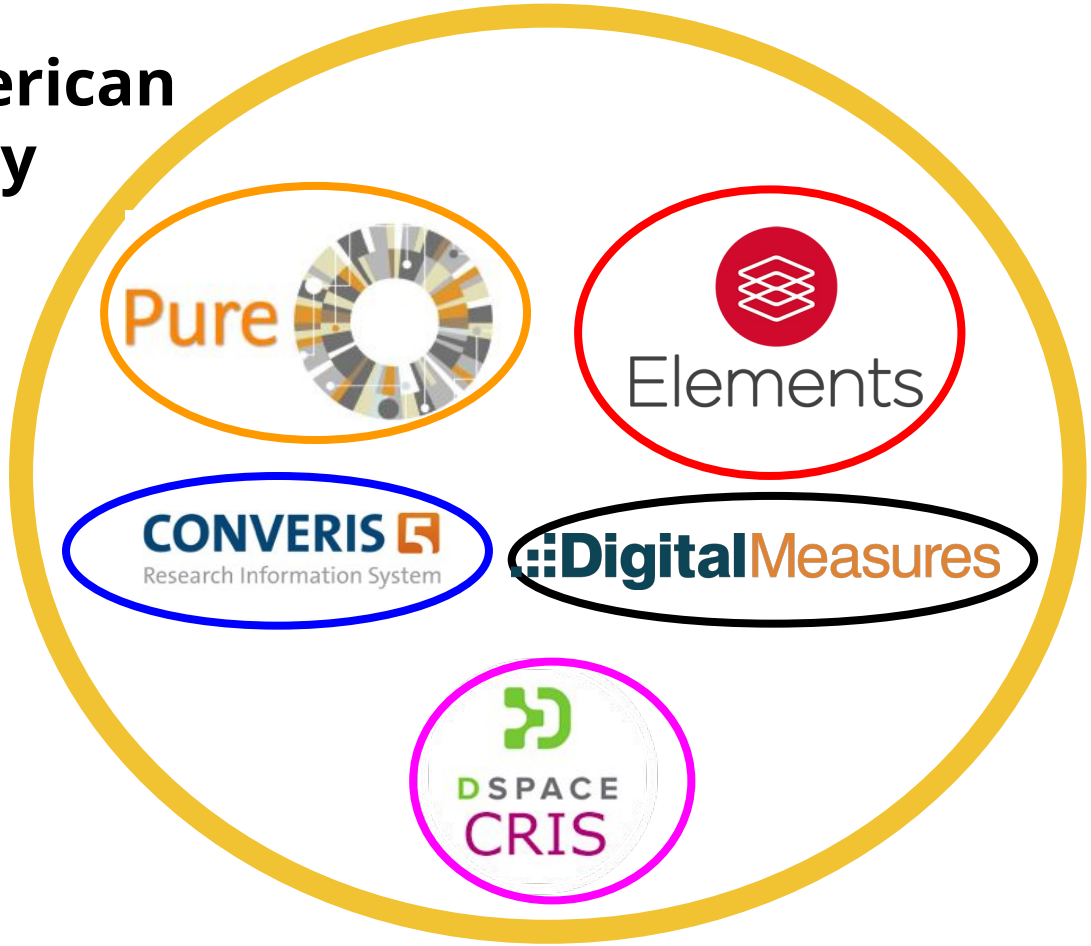
ARMS
Australasian Research Management Society

uk
arma
Association of Research Managers and Administrators

euroCRIS
Current Research Information Systems

Carnegie Mellon University

North American Community



OCLC

?

Carnegie Mellon University

Question/Comments

David Scherer
daschere@andrew.cmu.edu
@davidascherer

Mark Hahnel
mark@figshare.com
@markhahnel

Kate Byrne
kate@symplectic.co.uk
@katecbyrne

Dan Valen
dan@figshare.com
@dnvln

Carnegie Mellon University