

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*



# **Repositories and Research Information Management Systems**

# Repositories



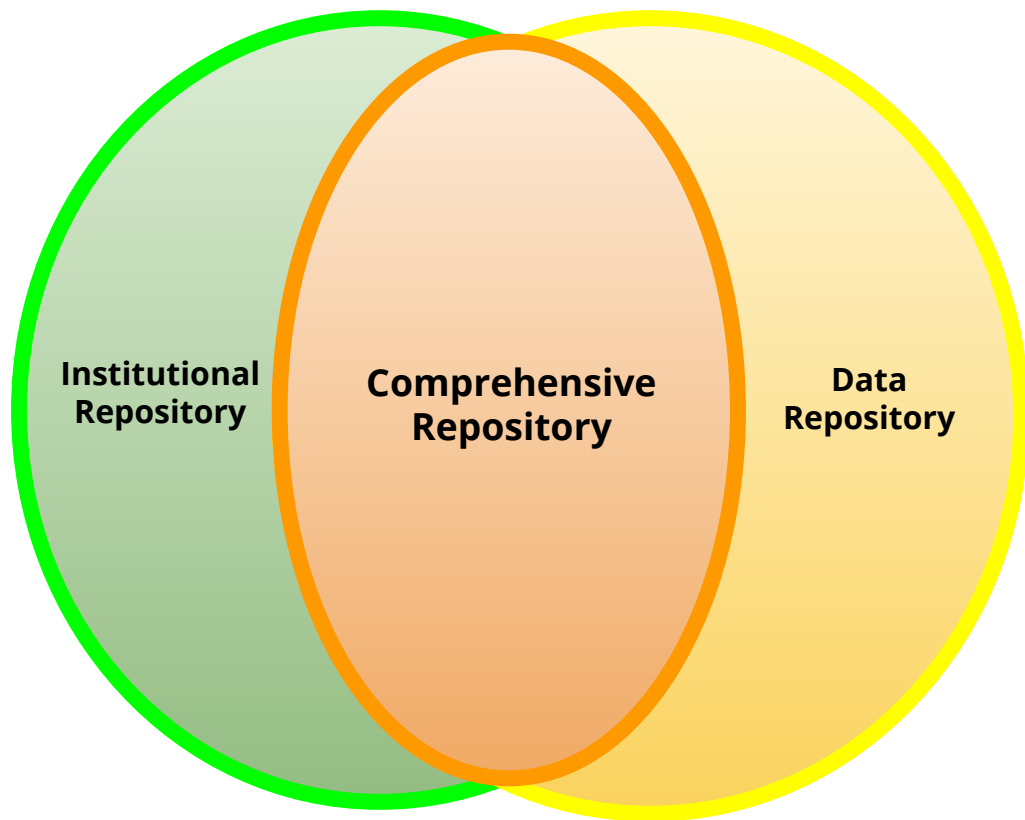
# RIMS



Elements



Carnegie  
Mellon  
University





# KiltHub

WEAVING THE FABRIC OF YOUR RESEARCH

🐦 @KILTHUB | [KILTHUB.CMU.EDU](https://kilthub.cmu.edu)

# Research Information Management Systems

Current  
Research  
Information  
Systems

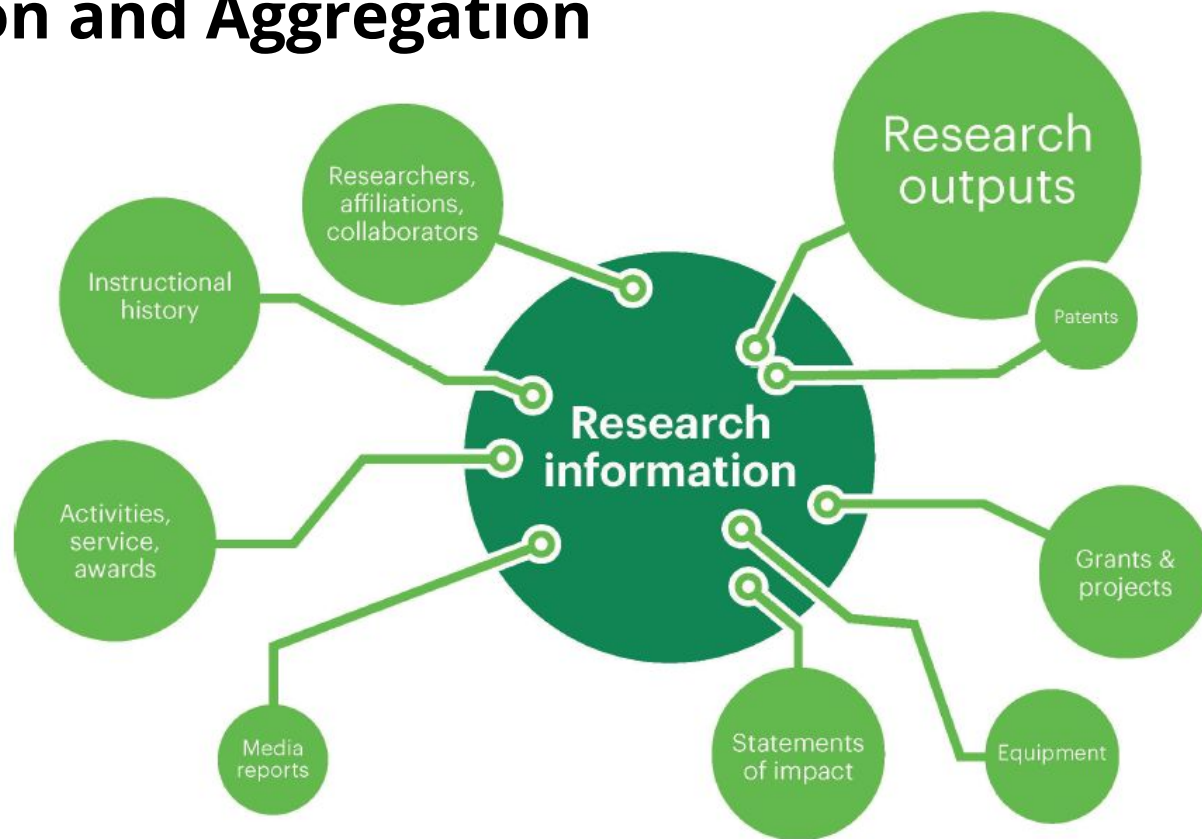
Faculty  
Information  
Systems

Scholarly  
Information  
Systems

Faculty  
Activity  
Reporting  
Systems

Carnegie  
Mellon  
University

# Curation and Aggregation



## International collaboration clusters in Africa

[+ Add a new journal article](#)

### Journal article

Adams J<sup>ORCID</sup>, Gurney K<sup>ORCID</sup>, Hook D<sup>ORCID</sup>, Leydesdorff L<sup>ORCID</sup>  
 Scientometrics **98**(1):547-556 06 Jan 2014 [DOI](#)  
[figshare for institutions](#) [doi](#)

#### Labels

0807 Library and Information Studies • Prevention •  
 Science Studies • Science & Technology [and 14 more](#)

#### Altmetrics



#### Citation Metrics

RCR	Dimensions	EMPC	Scopus	WoS
-	32	-	34	25

#### Journal Rankings

<a href="#">Scientometrics</a>		
<a href="#">SNIP</a>	<a href="#">SJR</a>	Impact Factor
1.38	1.13	2.00

### Actions

- [Mark as Favourite](#)
- [Add to Workspace](#)
- [Request a Waiver](#)

### Annual Collection

View in the Assessment page for:

[2013 Collection](#)

### Reporting Date

22 Jan 2013

### History

Showing 1 - 4 of 14 (newest first) [View all](#)

1. Claimed for Professor Daniel Hook.  
by Ms Julia Hawks  
14/11/2018
2. Claimed for Daniel Hook.  
by Ms Julia Hawks  
14/11/2018

### Data sources

[+ Add manual record](#)

 arXiv

 Crossref

 Dimensions

ID: [pub.1050166004](#) [More source info](#)

[★ Set as preferred record](#)

#### Title

International collaboration clusters in Africa

#### Authors

Adams, Jonathan<sup>ORCID</sup>; Gurney, Karen<sup>ORCID</sup>; Hook, Daniel<sup>ORCID</sup>; Leydesdorff, Loet<sup>ORCID</sup>

# An Increasing Role for Libraries

OCLC RESEARCH POSITION PAPER

## Research Information Management: Defining RIM and the Library's Role

**Rebecca Bryant**  
OCLC Research (US)

**Anna Clements**  
University of St. Andrews (UK)

**Carol Feltes**  
Rockefeller University (US)

**David Groenewegen**  
Monash University (Australia)

**Simon Huggard**  
La Trobe University (Australia)

**Holly Mercer**  
University of Tennessee-Knoxville (US)

**Roxanne Missingham**  
Australasian National University (Australia)

**Mallaca Oxnam**  
University of Arizona (US)

**Anne Rauh**  
Syracuse University (US)

**John Wright**  
University of Calgary (Canada)



OCLC RESEARCH REPORT

## Convenience and Compliance: Case Studies on Persistent Identifiers in European Research Information Management

Rebecca Bryant, Annette Dortmund,  
and Constance Malpas



OCLC RESEARCH REPORT

## Practices and Patterns in Research Information Management

### Findings from a Global Survey

Rebecca Bryant, Anna Clements, Pablo de Castro,  
Joanne Cantrell, Annette Dortmund, Jan Fransen,  
Peggy Gallagher, Michele Mennielli





## Libraries in Research Information Management

Publications &  
Scholarship Expertise

Training & Support

Discoverability, Access &  
Reputational Support

Stewardship of the  
Institutional Record

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

# What is the Scholarly Communication Ecosystem?



Universiteit Utrecht

# INNOVATIONS IN SCHOLARLY COMMUNICATION

Changing Research Workflows

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





# Scholarly Communication Tools

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

## Current List:

<https://docs.google.com/spreadsheets/d/1KUMSeqPzp4KveZ7pb5rddcssk1XBTiLHniD0d3nDqo/edit#gid=0>





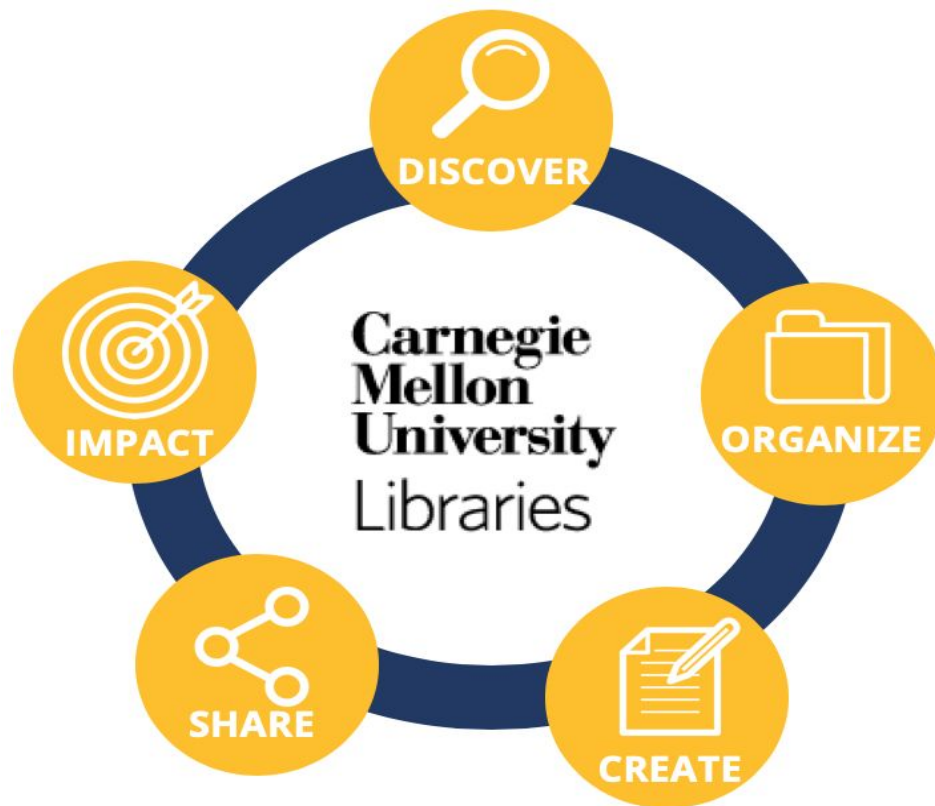
Open Science







# THE SCE at CMU

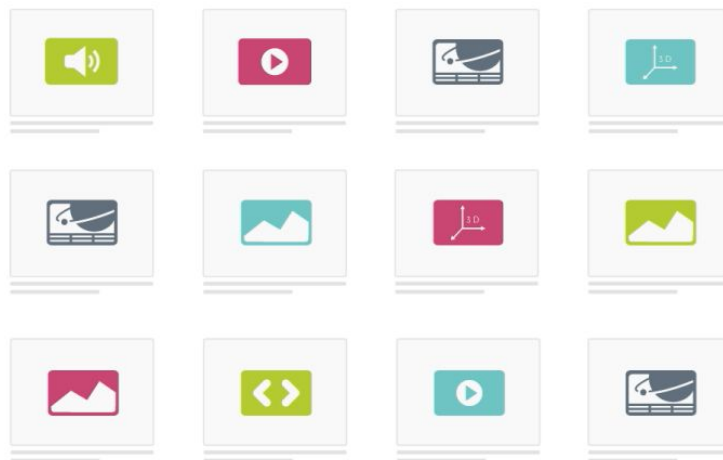


# THE SCE at CMU





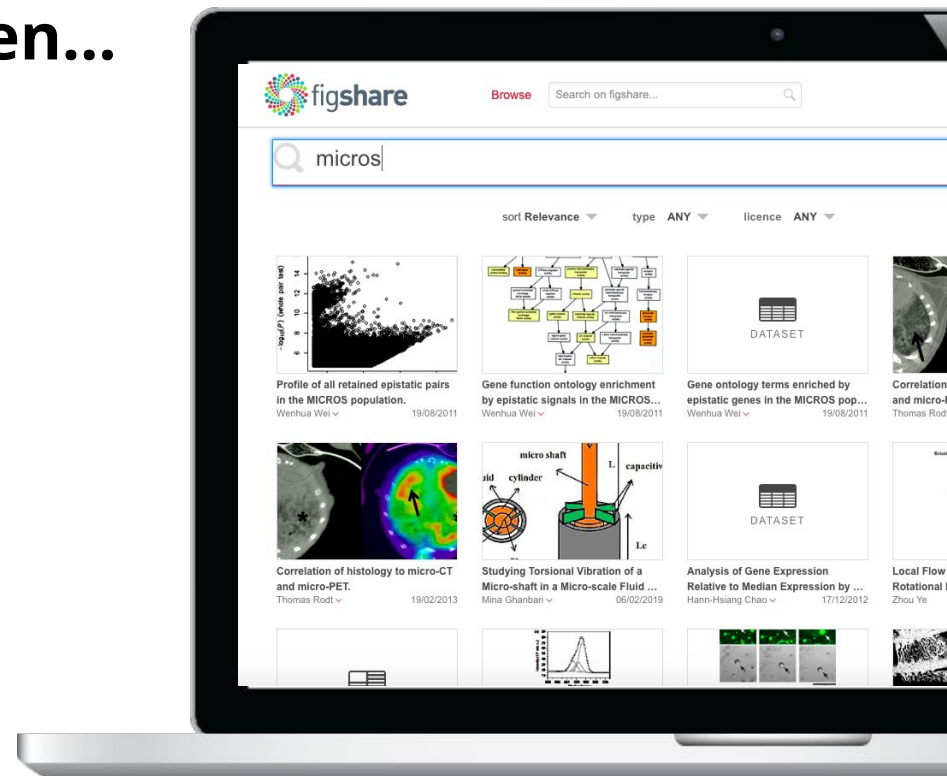
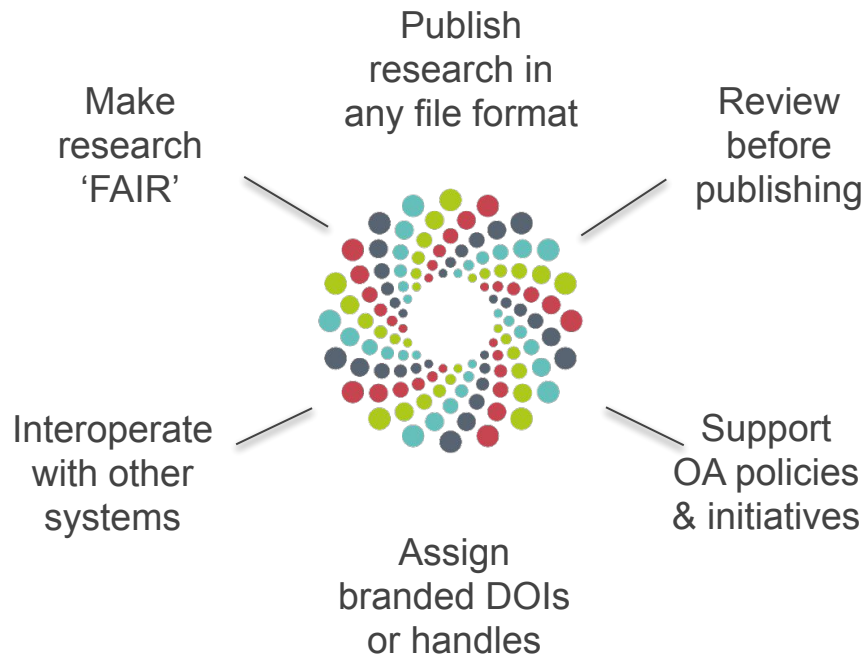
# Repositories and the SCE



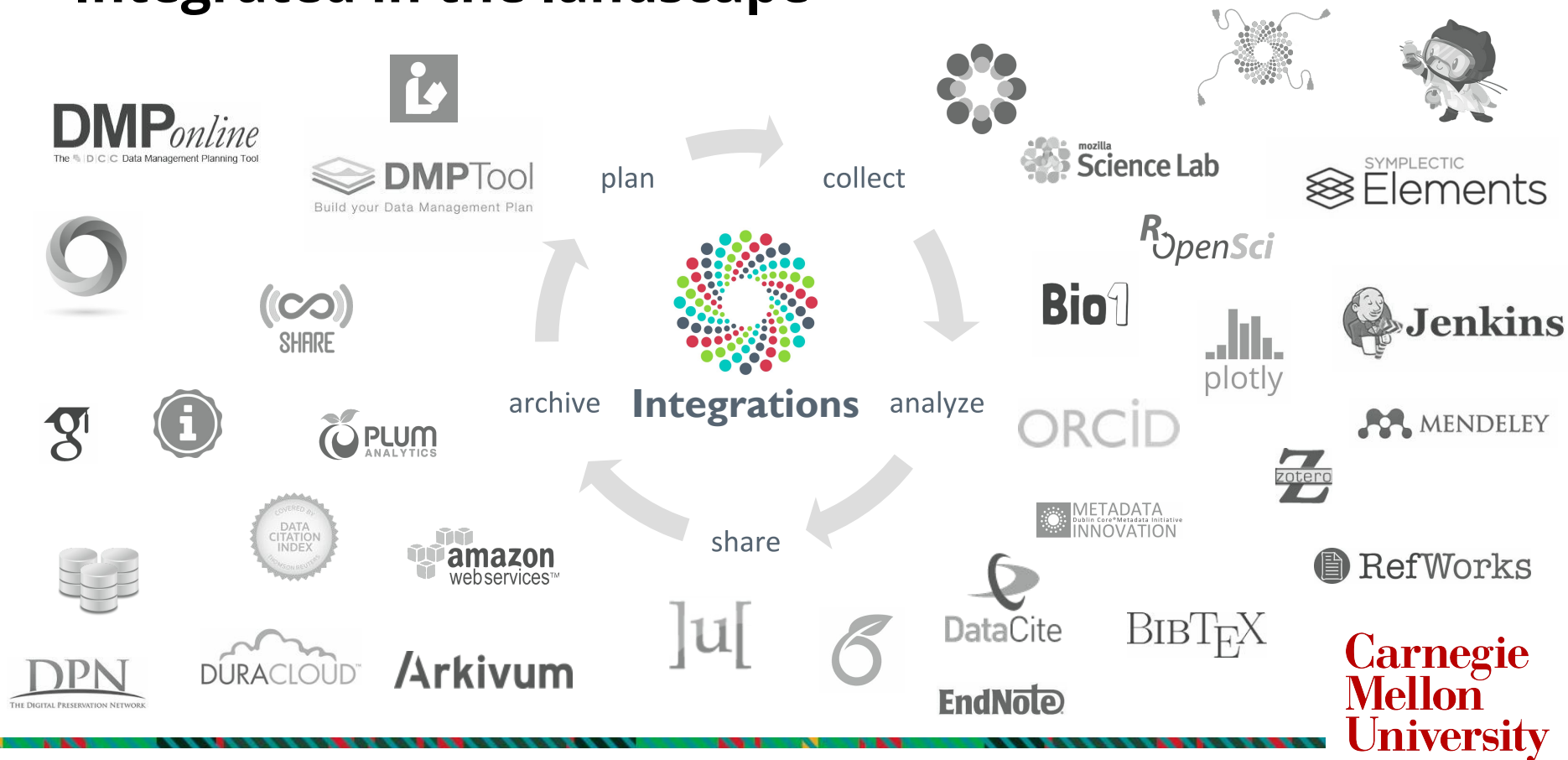
PANGAEA.

**Carnegie  
Mellon  
University**

# Helping make research open...



# Integrated in the landscape



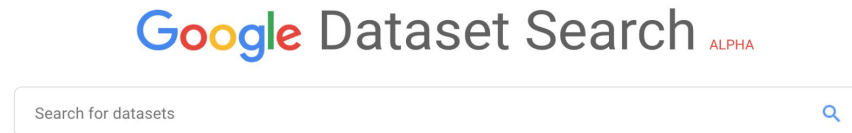
# Support for open standards and de facto standards

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

Stats  
Oai pmh  
Oai-pmh  
Base url  
Item equals article

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).



Secure | [https://figshare.com/rss/portal\\_group/plos/plosbiology](https://figshare.com/rss/portal_group/plos/plosbiology)

Apps | ALL figshare Client... | [figshare.com](#) | [Google+](#) | [https://docs.google...](#) | [figshare: LinkedIn](#) | [Tumblr](#) | [Implement](#)

```

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

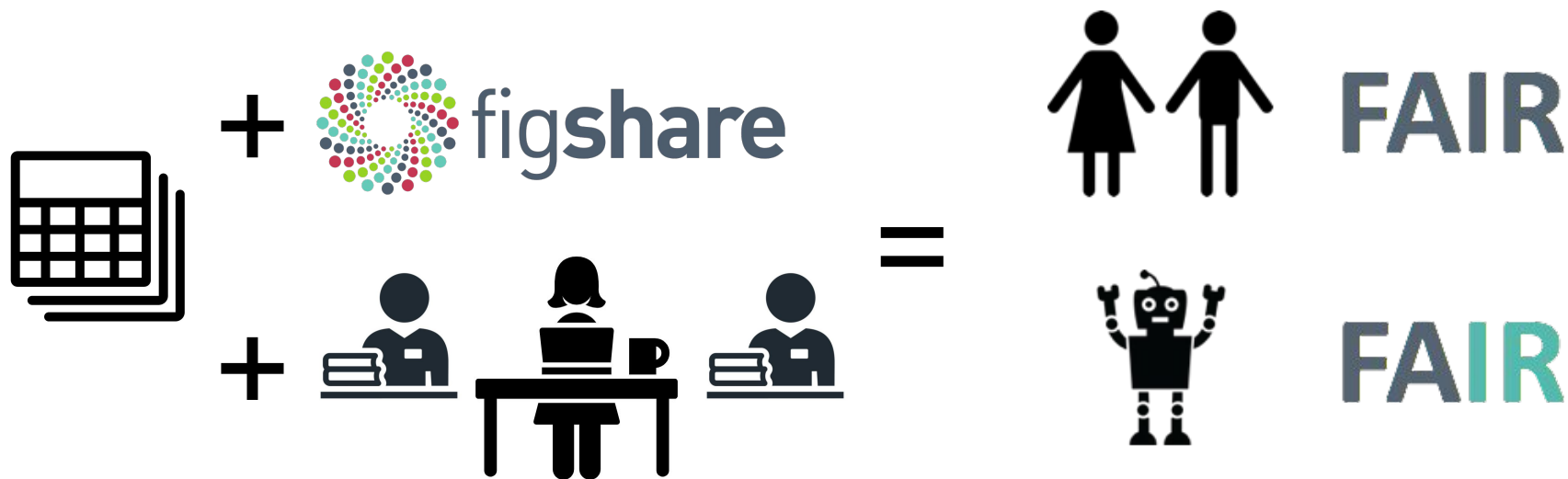
<channel>
  <title>PLOS Biology of Public Library of Science RSS Feed</title>
  <link>https://plos.figshare.com/plosbiology</link>
  <description>RSS feed for Figshare Group PLOS Biology from the Public Library of Science institution</descri

  <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>A computational model for slow motor gestures predicts the existence of three regions of th
    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>

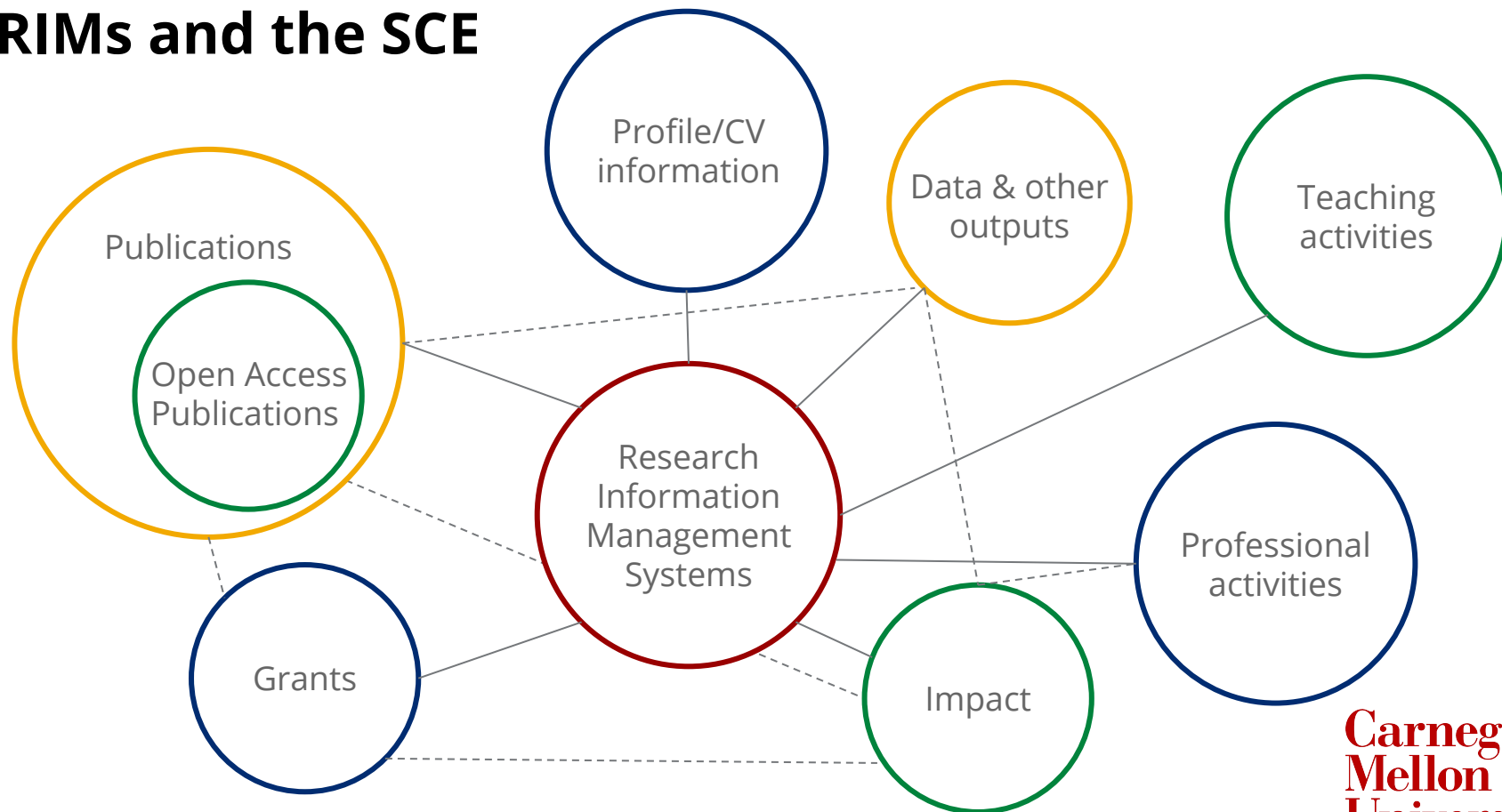
```

# Findable, Accessible, Interoperable, Reusable



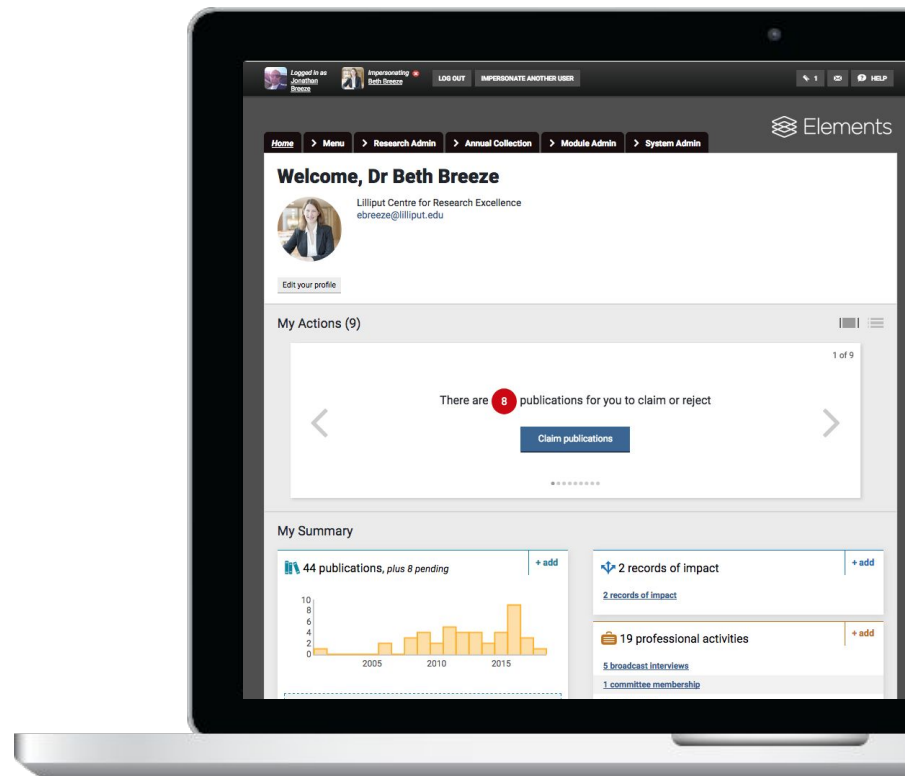
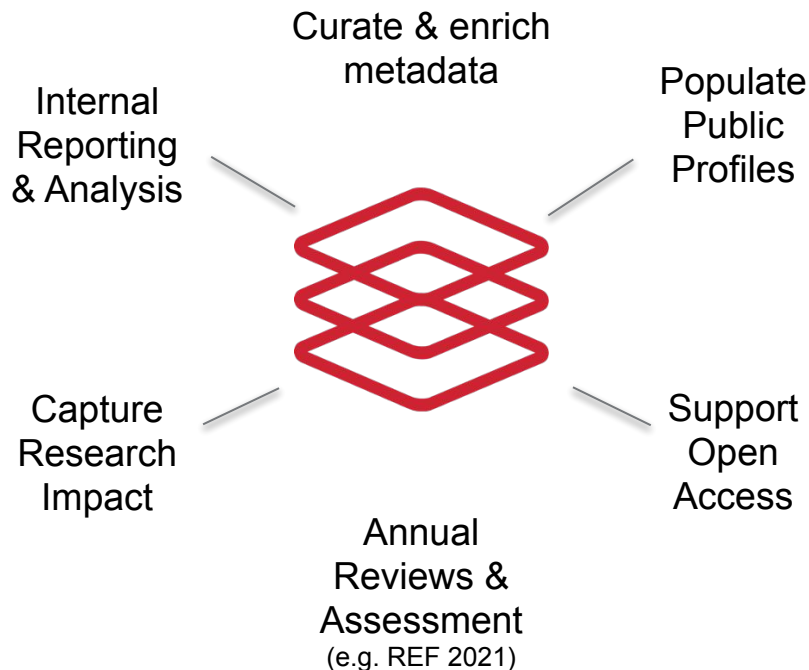
# RIMS and the SCE

# RIMs and the SCE

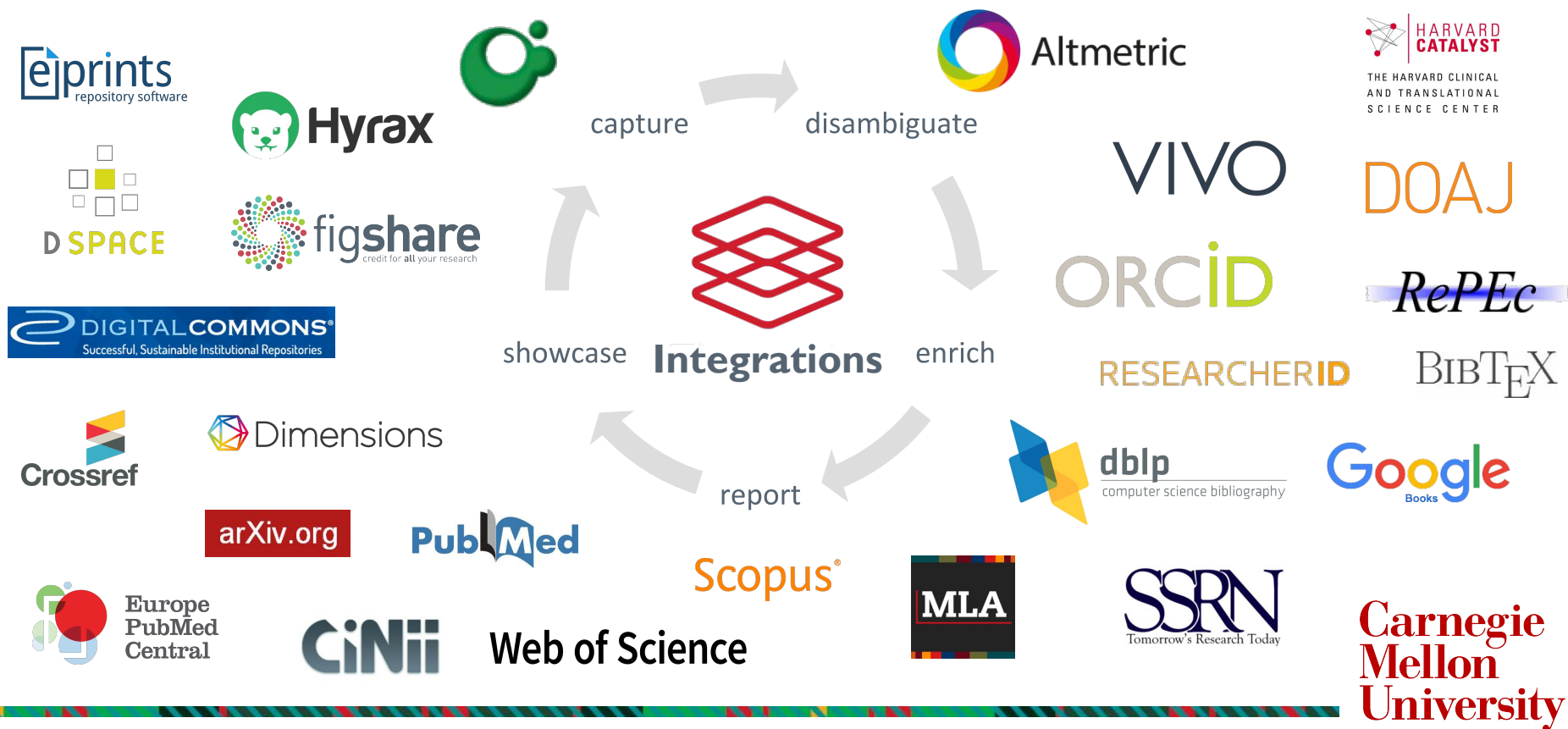




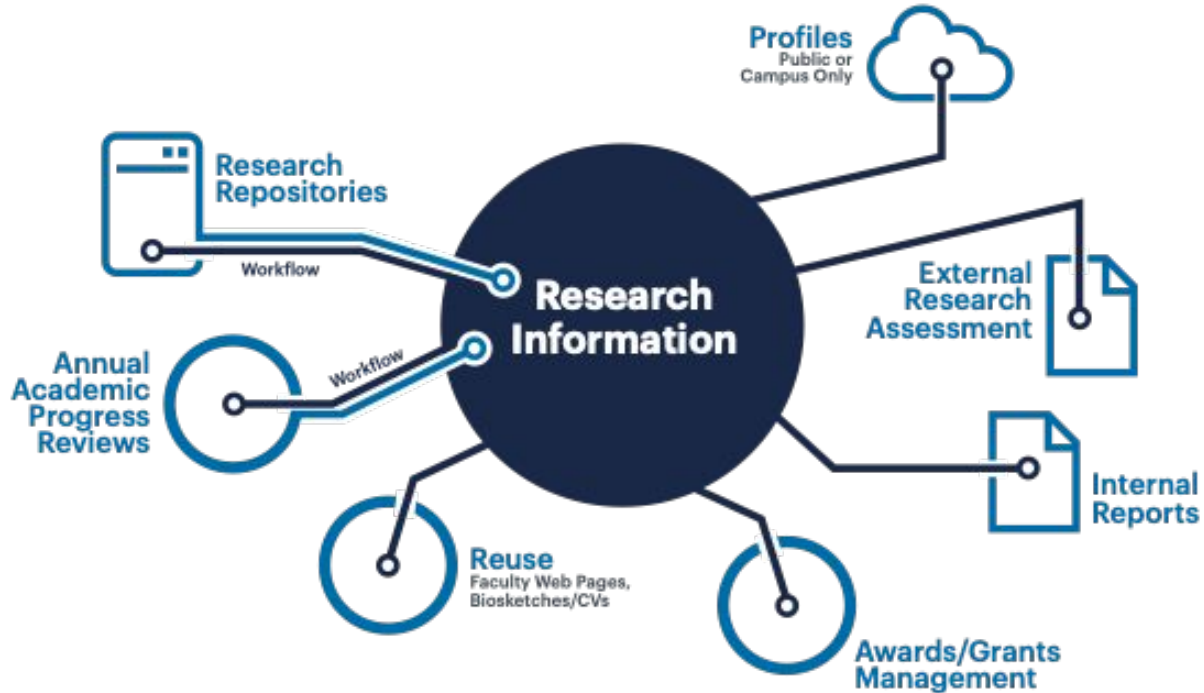
# Beyond collecting data...



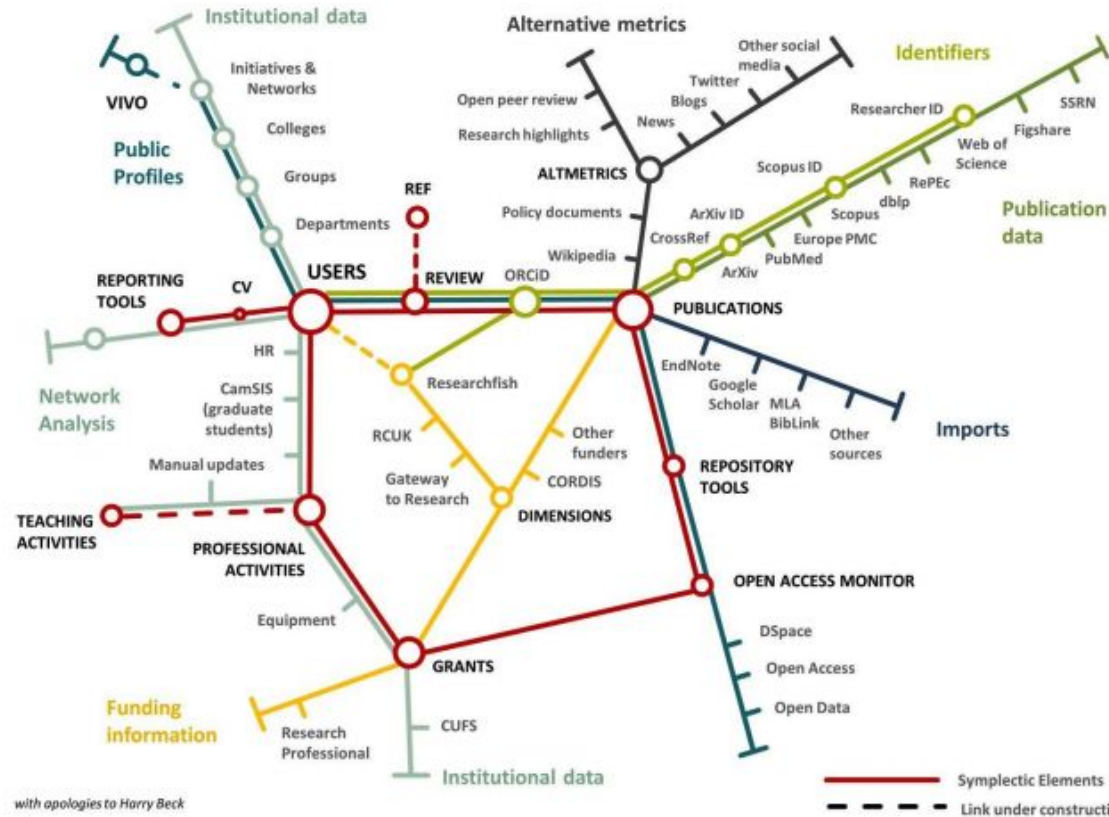
# Integrated in the landscape



# Views of the Network



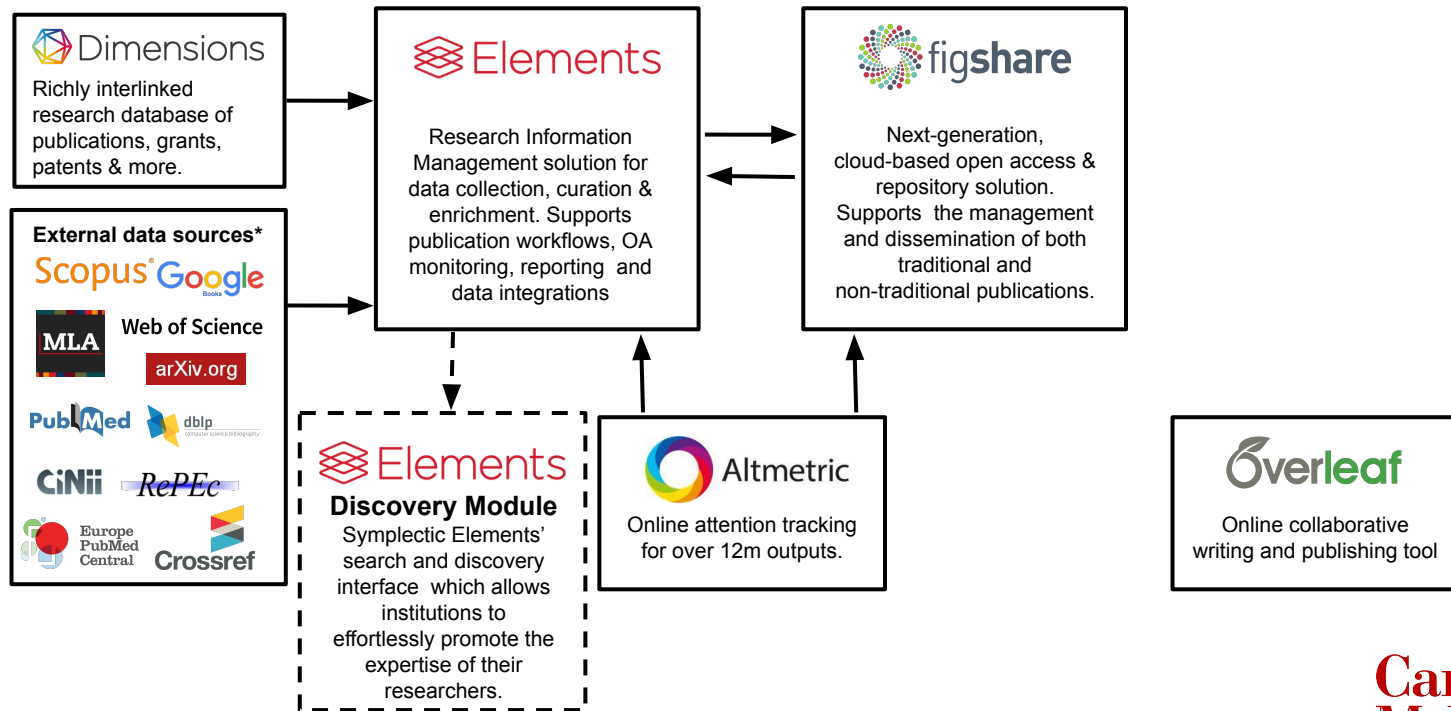
## Views of the network: Cambridge University



**Carnegie  
Mellon  
University**

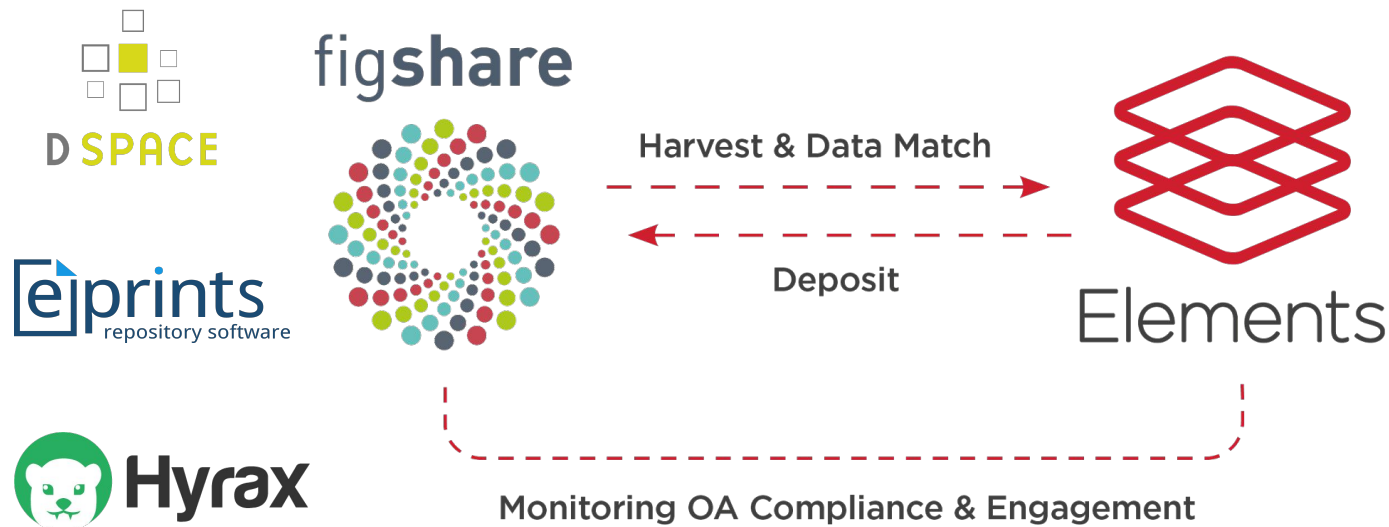
Juergen Wastl (2017). Academic Profile, Network and Collaboration Analysis : Use Cases for the University <https://doi.org/10.17863/CAM.13805>

# Views of the network: Carnegie Mellon University

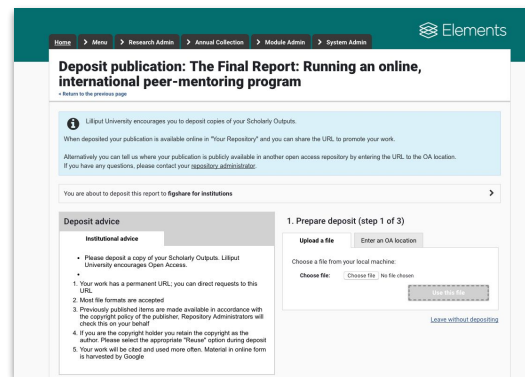
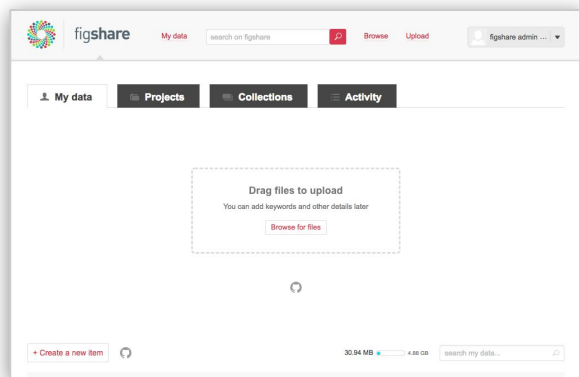


\*Data sources used by Elements. Data sources actually integrated with Primo may differ.

# Creating rich bi-directional integrations



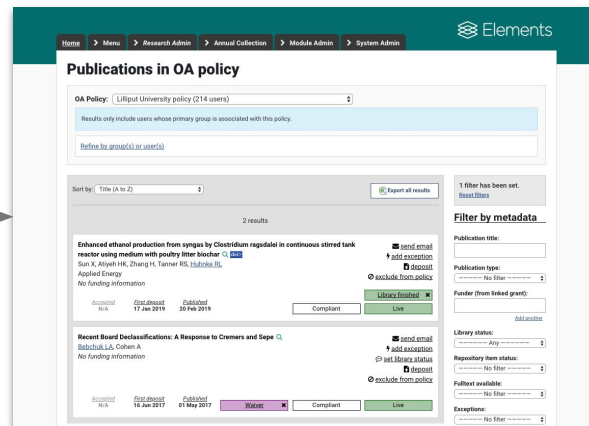
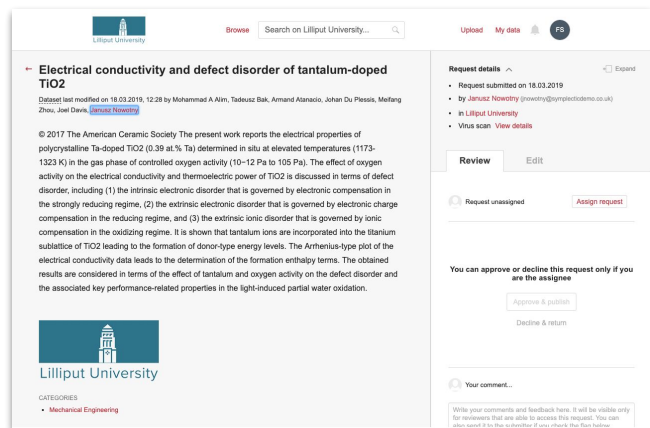
# Supports flexible workflows for researchers



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



# 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.



# Maximising reuse, minimises manual effort

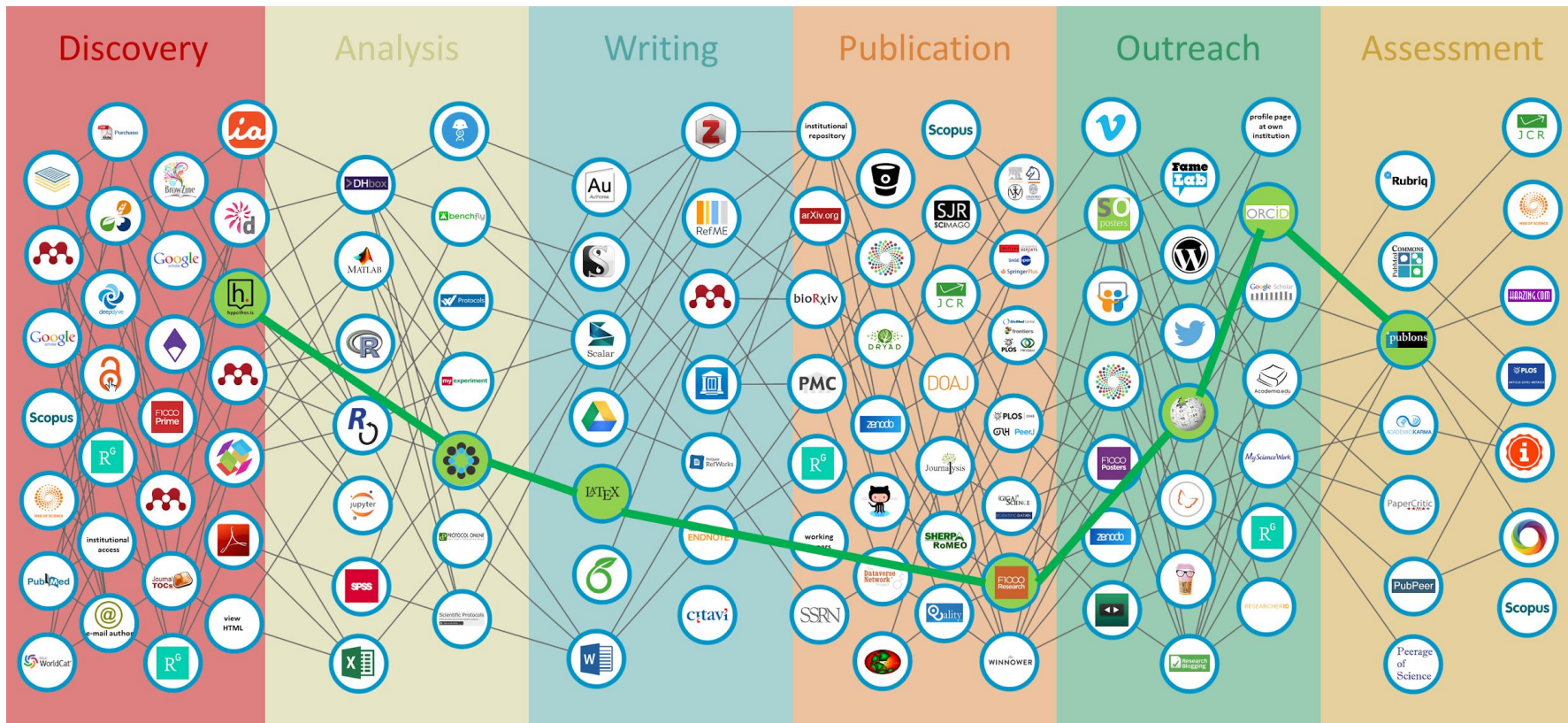


Data-focussed interoperability minimises time spent typing metadata into manual entry forms over and over, which in turn reduces the likelihood of typos and other human error.

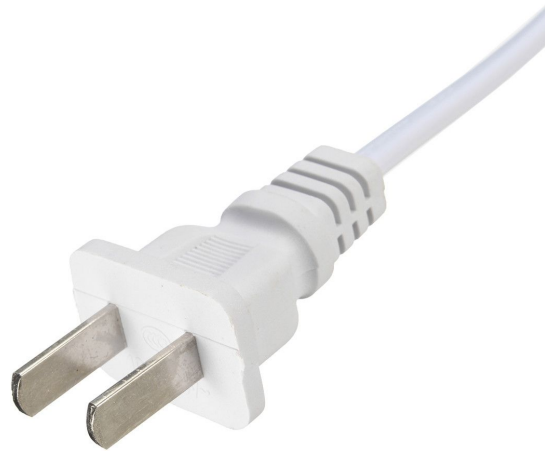
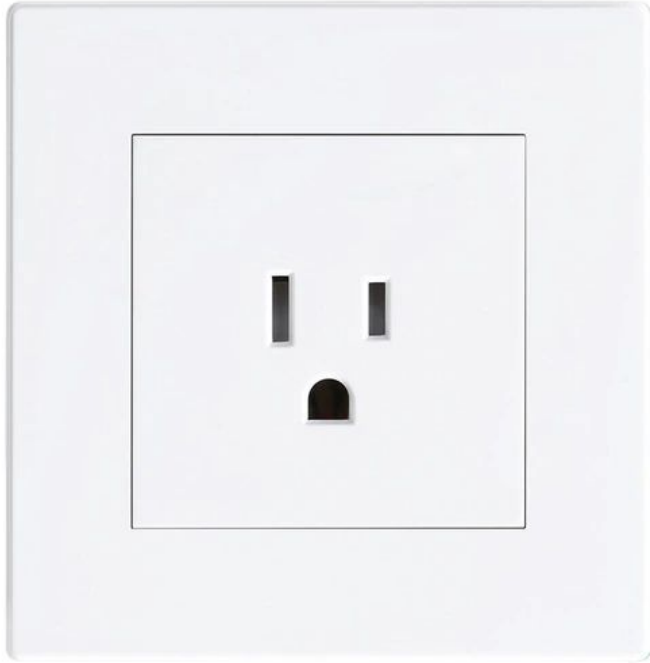
Being asked to do the same task multiple times is a common source for frustration for academics and administrators alike - reducing this frees up time for more important tasks.

# Building the Network

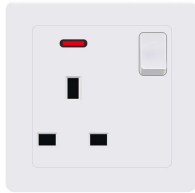
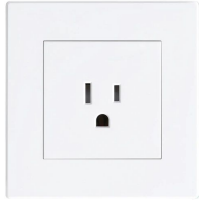
# Models for integrations



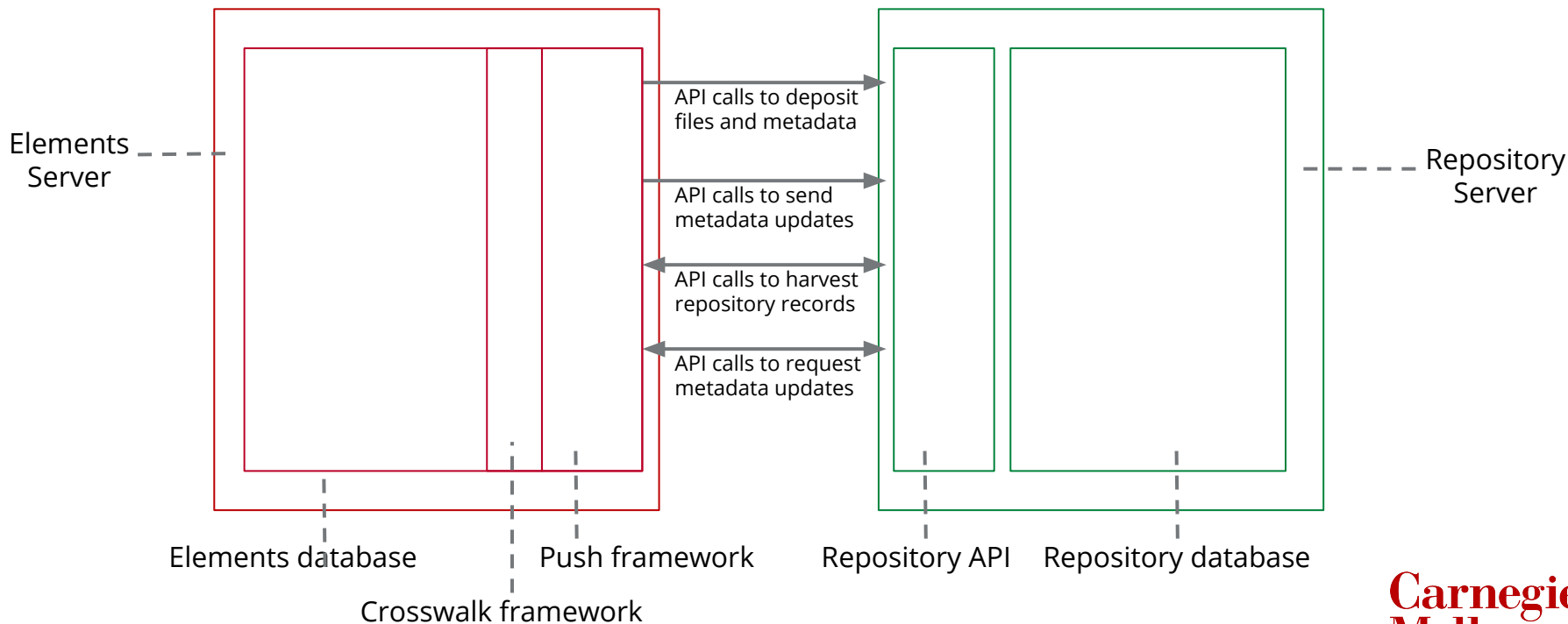
# Foundations for interconnectivity



# Foundations for interconnectivity

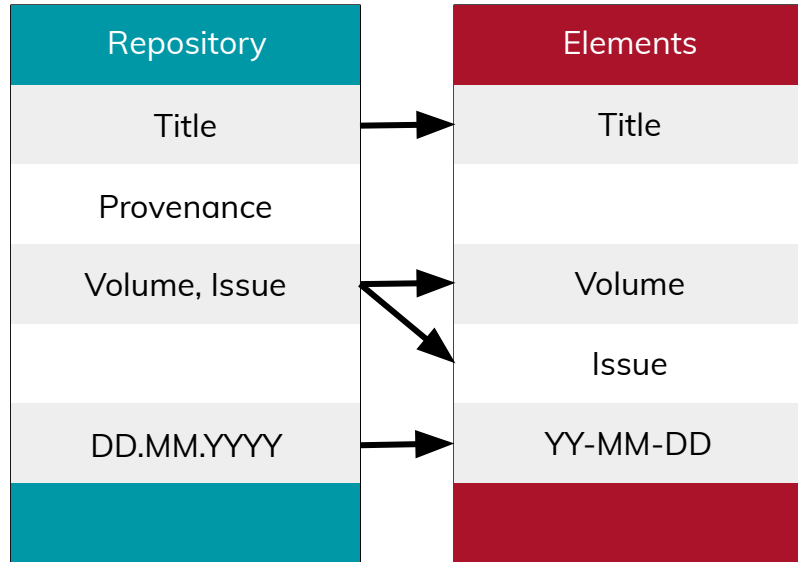


# Under the hood of our bi-directional integrations

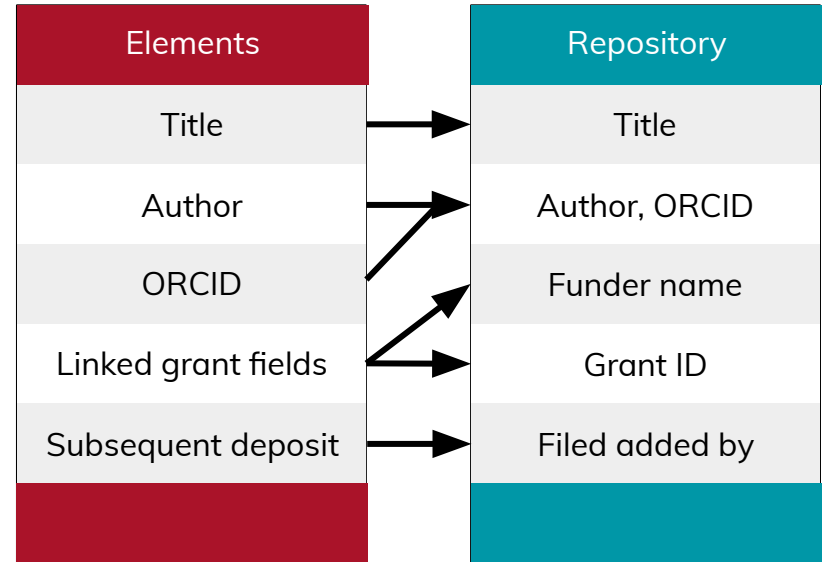


# Crosswalks & translation layers

Harvest crosswalks



Deposit crosswalks





## Edit Harvest Crosswalk for figshare for institutions

[Return to the figshare for institutions configuration page](#)

Map

Input

Output

```

122
123     <!-- BASIC BIBLIOGRAPHIC METADATA -->
124
125     <!-- defined_type_name => types -->
126     <xwalk:field-mapping to="types">
127         <xwalk:field-source from="defined_type_name" />
128     </xwalk:field-mapping>
129
130     <!-- title => title -->
131     <xwalk:field-mapping to="title">
132         <xwalk:field-source from="title" value-map="ignore-empty-string" />
133     </xwalk:field-mapping>
134
135     <!-- description => abstract -->
136     <xwalk:field-mapping to="abstract">
137         <xwalk:field-source from="description" value-map="strip-html-elements" />
138     </xwalk:field-mapping>
139
140     <!-- authors => authors -->
141     <xwalk:field-mapping to="authors">
142         <xwalk:field-source from="authors">
143             <!-- Name -->
144             <xwalk:field-source from="full_name" format="person:firstnames-lastname" />
145             <!-- Figshare for Institutions user account ID -->
146             <xwalk:field-source data-part="person:identifier">
147                 <xwalk:field-source data-part="identifier:scheme" value="figshare-for-institutions-user-account-id" />
148                 <xwalk:field-source data-part="identifier:value" from="id" value-map="ignore-empty-string" />


```

 Full screen

 Revert

 Save

Use provided input: ☐ or specify figshare for institutions ID:

 Test Crosswalk

Use this page to edit and test your figshare for institutions harvest crosswalk. Any changes you make to the harvest crosswalk will not be deployed until you save. The buttons above can be used to revert the harvest crosswalk to the currently installed version, or to save this version. The following keyboard shortcuts can be used:

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.

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

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

## Repository DOI

### 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](#)

#### FUNDING

**Center for Air, Climate, and Energy Solutions (CACES)**

Environmental Protection Agency

NLM

DC

**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 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>."

## Publisher's DOI

# Mutiple Sources

☐ **Opportunities for real-time digital reference service**




Summary

Organisation users (1)




History (2)

Data sources (3)




Links (1)




 Opportunities for Real-Time Digital Reference Service





 Opportunities for real-time digital reference service





 Opportunities for real-time digital reference service


Chapter 







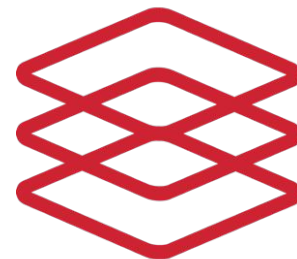




# CMU Crosswalks & Translations

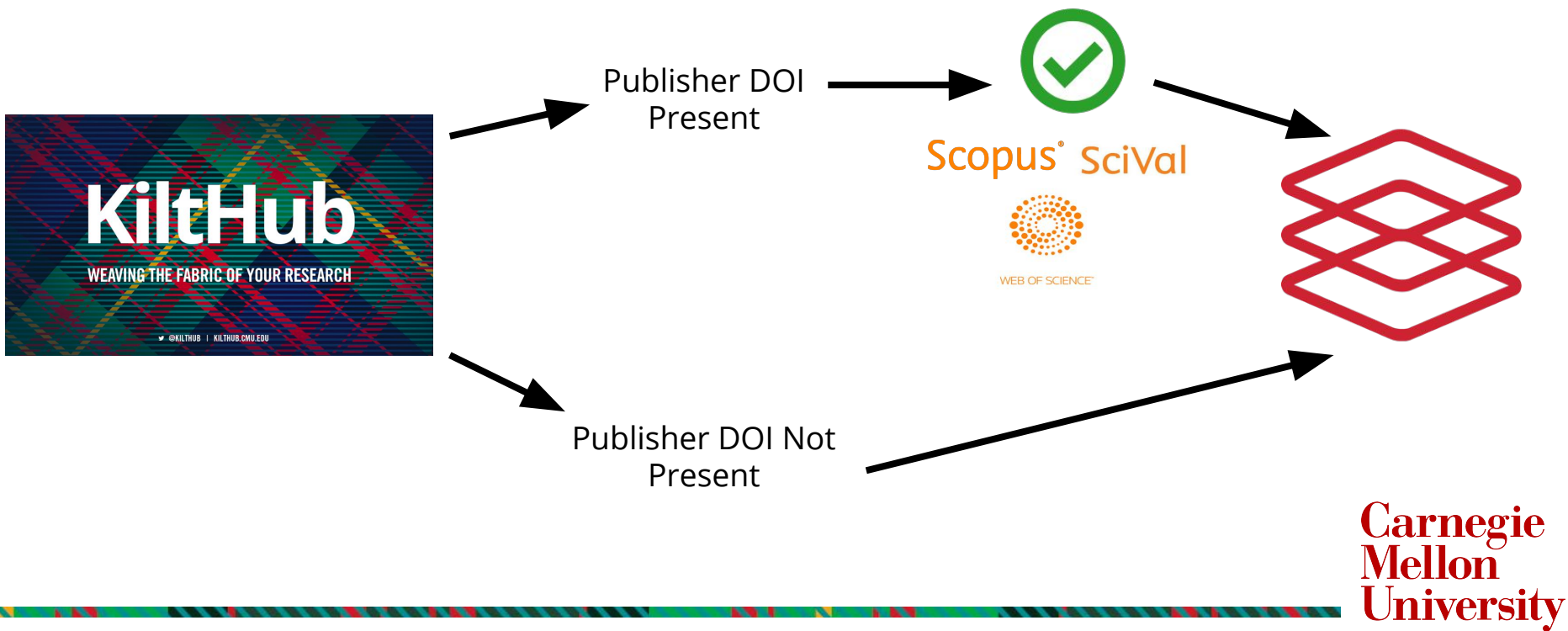


**Publisher DOI  
Present**



**Publisher DOI Not  
Present**

# CMU Crosswalks



# Carnegie Mellon University's workflows

Metadata Harvesting & Content Matching

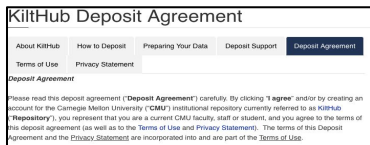


Monitoring OA Status  
and Engagement



Elements

Content Verification, Submission, and Deposit



Carnegie  
Mellon  
University

# Why this Matters



# Why this matters for institutions



# Integrations = Choices



+



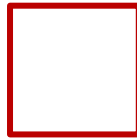
+



+



+



+



+



+



+



# 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**.

# Future of Repositories and RIMS

# 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

# 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?

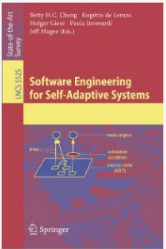
# Integrated Data Models

 Altmetric

[Access Altmetric Explorer](#) [? What is this page?](#) [Embed badge](#) [Share](#)

## Software Engineering for Self-Adaptive Systems


Overview of attention for book






**TABLE OF CONTENTS**

- Book Overview**
- Chapter 1**  
Software Engineering for Self-Adaptive Systems: A Research Roadmap
- Chapter 2**  
Modeling Dimensions of Self-Adaptive Software Systems
- Chapter 3**  
Engineering Self-Adaptive Systems through Feedback Loops
- Chapter 4**  
Improving



**Attention for Chapter 3: Engineering Self-Adaptive Systems through Feedback Loops**

 **13**

**Mentioned by**

-  **1** news outlet
-  **1** Wikipedia page
-  **160** Dimensions


**Citations**

-  **350** Mendeley
-  **9** CiteULike

**Readers on**


**SUMMARY** [News](#) [Wikipedia](#) [Dimensions citations](#)

**?** So far, Altmetric has seen **1** news story.

 **Why a robot could be the best boss you've ever had**  
The Guardian, 09 Sep 2016  
Would you like to work for a robot? Although the automation of skilled jobs is a reality, your boss is probably (still) a human.




# Integrated Data Models








[Access Altmetric Explorer](#)[What is this page?](#)[Embed badge](#)[Share](#)


## Software Engineering for Self-Adaptive Systems


Overview of attention for book




**TABLE OF CONTENTS**


-  Publisher Version
-  Repository Version 1
-  Repository Version 2
-  Repository Version 3
-  Repository


 Attention for Chapter 3: **Engineering Self-Adaptive Systems through Feedback Loops**




Aggregate of Scholarship


 1 news outlet

 1 Wikipedia page

 160 Dimensions

Readers on

 350 Mendeley


 9 CiteULike


SUMMARY

News

Wikipedia

Dimensions citations

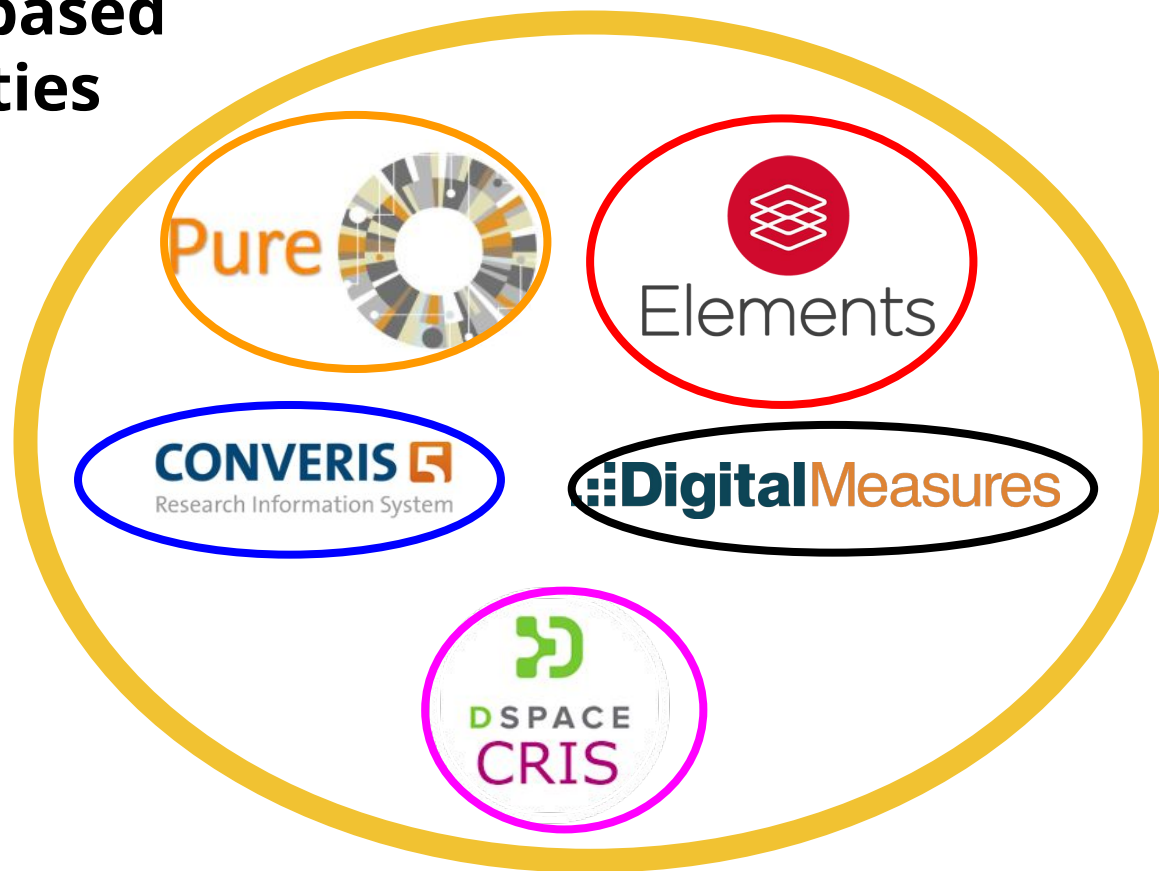
 So far, Altmetric has seen 1 news story.

 **Why a robot could be the best boss you've ever had**

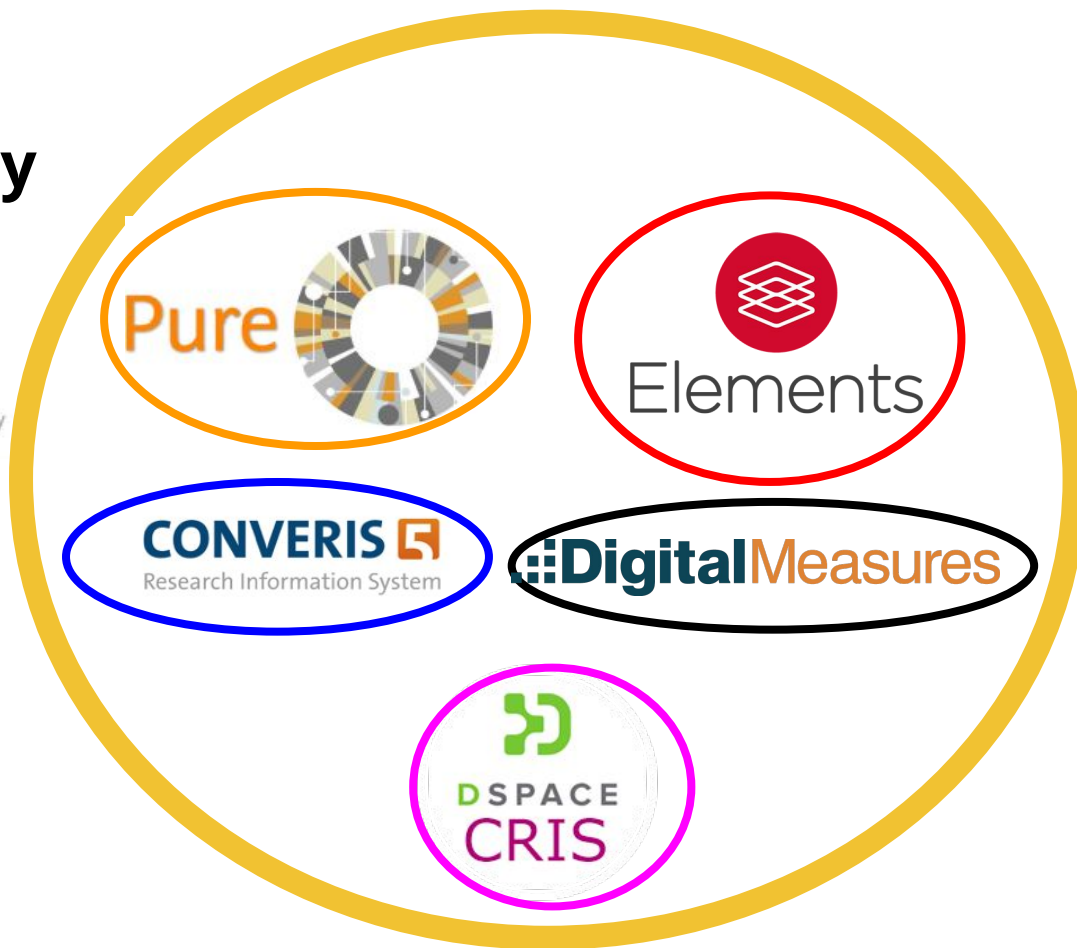
The Guardian, 09 Sep 2016

Would you like to work for a robot? Although the automation of skilled jobs is a reality, your boss is probably (still) a human.

# Platform-based Communities

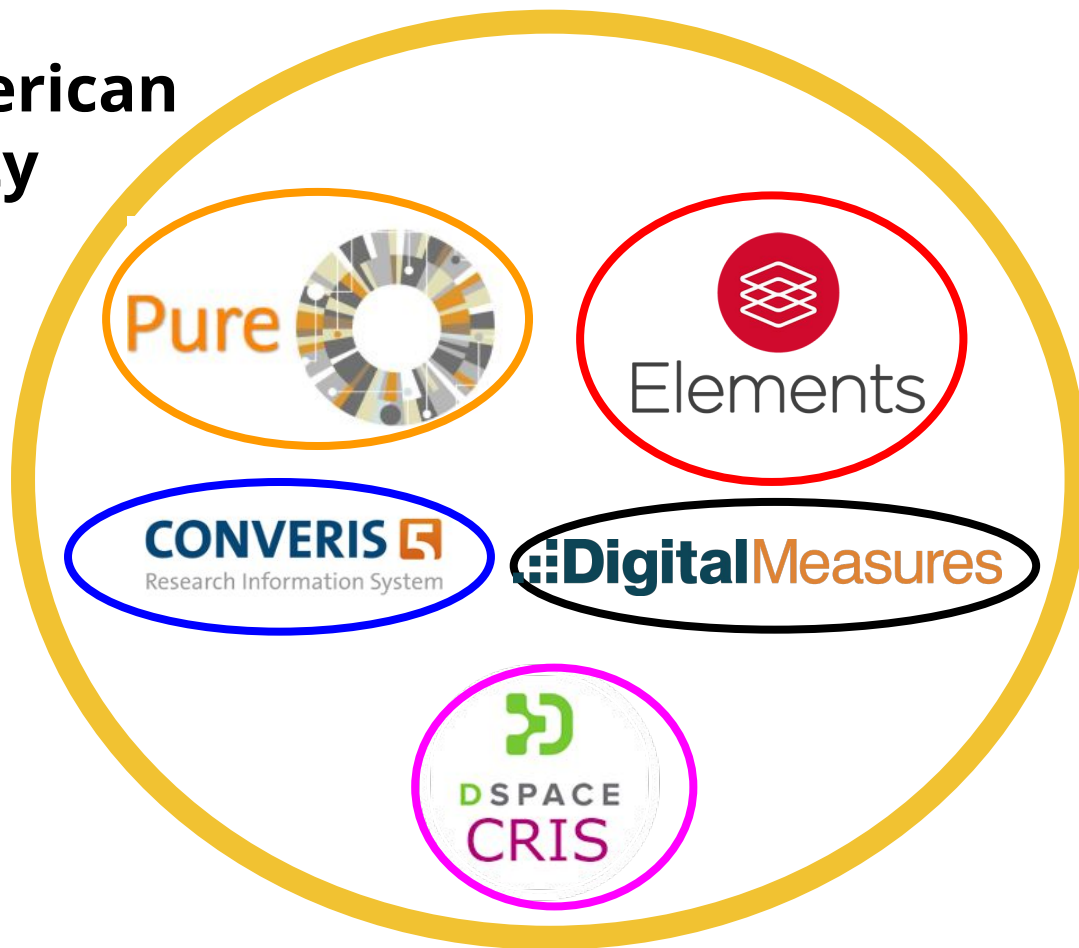


# Global Community



**Carnegie  
Mellon  
University**

# North American Community



?

# Question/Comments

David Scherer

[daschere@andrew.cmu.edu](mailto:daschere@andrew.cmu.edu)

@davidascherer

Mark Hahnel

[mark@figshare.com](mailto:mark@figshare.com)

@markhahnel

Kate Byrne

[kate@symplectic.co.uk](mailto:kate@symplectic.co.uk)

@katecbyrne

Dan Valen

[dan@figshare.com](mailto:dan@figshare.com)

@dnvln