Why should more publishers come to PIDapalooza: connecting PIDs and Open Infrastructure to publisher workflows

Hindawi as a use case

Created by: Alessandra Auddino

Supervisor: Catriona Maccallum

Special thanks to...

Crossref

Ginny Hendricks
Rachael Lammey
Paul Davis

ROR

Maria Gould

Jisc

Josh Brown

NISO

Alice Meadows

DataCite
FIND, ACCESS, AND REUSE DATA

Martin Fenner

Silvio Peroni

OAS

Yvonne Campfens

Sally Wynn-Jones
Angelina Ilievska
Catherine Glover   Andrew Smeall   Alex Pricop
Old publishing team
Julie Young   Toby Prior   Jennifer Leech
Raul Popovici   Ioana Onofrei
Andrea Magnus   Mohamed Hakeem   Alexandru Pricop   Naomi Froude
Adrian Trefas   Catriona Maccallum   Matt Green
Sarah Greaves   Matt Hodgkinson   Mathias Astell   Kelly Duoos
Finance team
Rares Iurea   Eric Liao   Nicko Goncharoff
Nisha Meghani   Iulian Nechifor   Paul Flood   Luigi Mantu
Thomas Faust   Rebecca Windless   Kamila Tyran
Aurelia Apetrei
Marketing team

Hindawi

Alessandra Auddino
alessandra.auddino@hindawi.com

"We aim to **maximize the impact** of scientific research through **openness** and global **collaboration** as we truly believe that science works best when **research is open**"

hindawi.com

The vision of the research communication world that Hindawi is contributing to the creation of:

A research more open, global, collaborative, creative and closer to society.
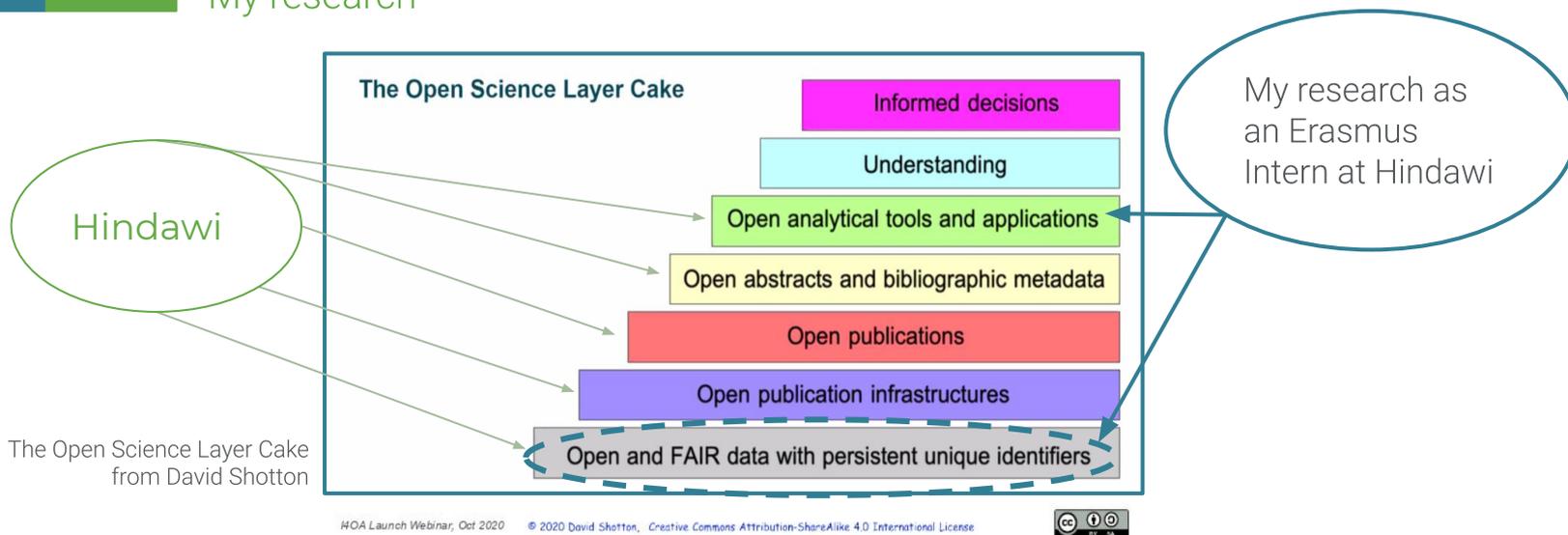
Hindawi supports Open Science.

Gold OA publisher

Open publishing infrastructure

My research

The Open Science Layer Cake from David Shotton

Hindawi

My research as an Erasmus Intern at Hindawi

Hindawi is a service-provider: we help authors publish OA content that aims to be as fair as possible (improving our content findability, accessibility, interoperability, and reusability) on an open publishing infrastructure and we share our metadata and abstracts with the community so that other can build additional services on them.

Paul Peters. 'An Open Approach to Developing Infrastructure for Open Science'. Hindawi (blog), 2017. https://www.hindawi.com/post/a-radically-open-approach-to-developing-infrastructure-for-open-science/.
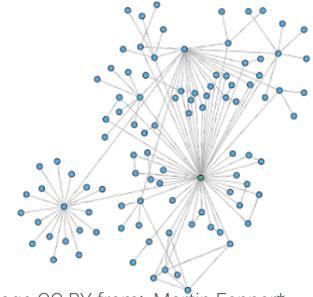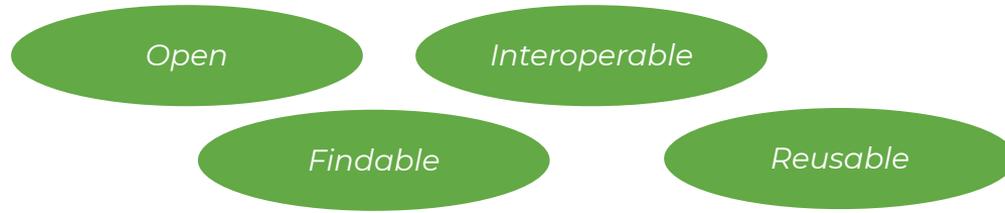
FAIR data

Open

Interoperable

Findable

Reusable

Image CC BY from: Martin Fenner*

For articles to persist in a **global research network of connections** across disciplines, borders, and time and to improve the **discoverability** of our content and the cost-effectiveness of our workflows, we need:

- Richer **deposits of PIDs and metadata**
  - DOIs & ORCID IDs but also ROR, RAID and others
- Using common community standards and following industry best practices
- Providing Open metadata and Open APIs

Ludo Waltman. 'Publications Should Be FAIR'. Leiden Matrics (blog), 2020. https://www.leidenmadtrics.nl/articles/publications-should-be-fair.
Cousijn, Helena, Ginny Hendricks, and Alice Meadows. 'Why Openness Makes Research Infrastructure Resilient'. Learned Publishing 34, no. 1 (2021): 71–75. https://doi.org/10.1002/leap.1361.

Why does Hindawi need ROR (Research Organization Registry)?

To contribute to the network effects of widespread adoption by publishers

To make our workflows more machine readable and cost effective

To be part of this OA community project

ROR

To improve the quality of our reports

To enrich and improve our data

To take part in the OA Switchboard pilot test

To provide a better service for our authors and publishing partners

To implement innovative and useful features in Phenom

ROR user case

- *To improve the quality of our reports*

Defining a **target by institute** for our marketing communication

From which **institute/funder** do the authors get their capital for the publication invoice?

Can we **map the institutes** from which our authors/editors/reviewers are from?

How many time articles from a certain institute have been **read/cited**?

Are some institutions are **more profitable** than others?

Can we automatically share our articles in the **institutes/funders' archives and databases**?

ROR user case

- *To improve the quality of our **reports***

Monitor and increase the quality of the editorial process and guarantee the quality preventing ethical issues

From which institutes there are more **rejected manuscripts?** Why?

Can we discover and manage institutes that are in country with **international sanction?**

What institutes are more active in a certain **field of research?**

Is there a trend of **suspicious behavior** (retracted papers, corrections etc) from authors from the same institute?

Can we **automatically prevent any conflict of interest** involving authors/reviewers/editors from the same institute?

ROR user case

- *To implement **new features in Phenom***

Autocompletion of data at submission	Easier data retrieval and checks on the authors' affiliation institutes	Include ROR IDs in the metadata we share with the community
Automatic prevention of conflict of interest	OA Switchboard integration	Automatically deposition to the institute/funder repository

Submissions   Peer Review Management   APC Billing and Collection   Publication

Editorial Screening   Post Peer Review Screening   Production Management

Roadmap for ROR

**Initial ROR integration
Set for Q1 2021**

Here the real rockers come on stage!

What other wild/exotic/useful user cases for ROR - Research Organization Registry at a publisher can you think of?

https://forms.gle/EXDGJ1dSvdfmNxJt5

The Open Science Layer Cake

Informed decisions

Understanding

Open analytical tools and applications

Open abstracts and bibliographic metadata

Open publications

Open publication infrastructures

Open and FAIR data with persistent unique identifiers

IfOA Launch Webinar, Oct 2020   © 2020 David Shotton,   Creative Commons Attribution-ShareAlike 4.0 International License

Hindawi

My research as an Erasmus Intern at Hindawi
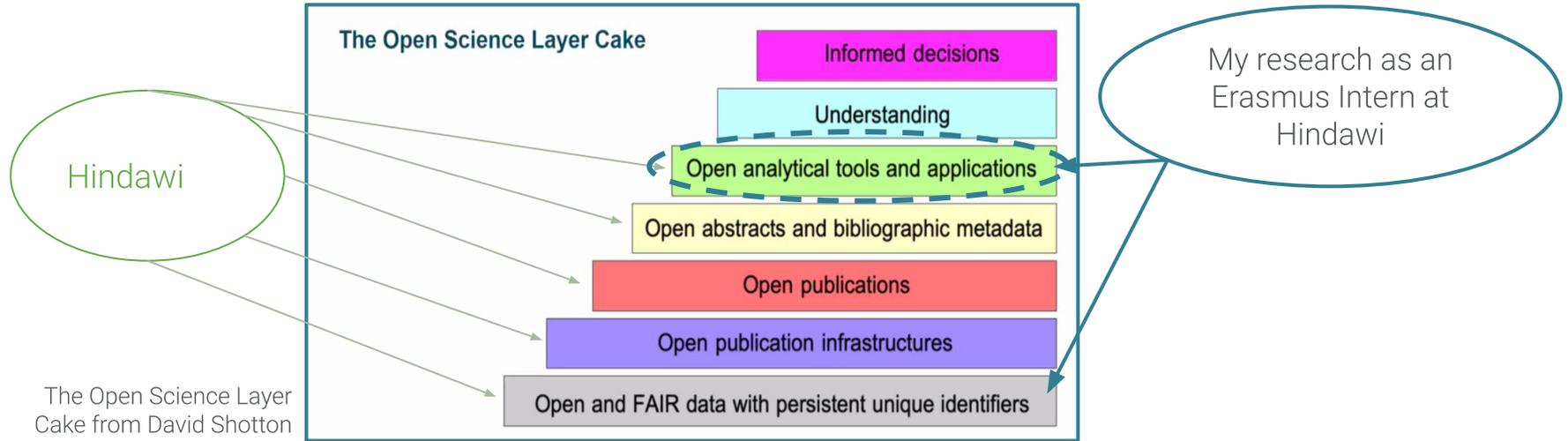
The Open Science Layer Cake from David Shotton

Existing monthly journal reports

journalreports.hindawi.com

Hindawi Journal Reports

View a summary of statistics for Hindawi journals.

Search journals...

	2018	2019	2020
Abstract and Applied Analysis	2018	2019	2020
Active and Passive Electronic Components	2018	2019	2020
Advances in Agriculture	2018	2019	2020
Advances in Astronomy	2018	2019	2020

Submissions
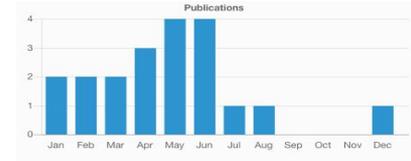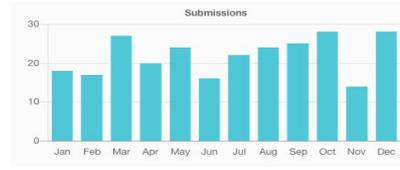
Publications

Most Viewed Articles

Article	Views
A degree theory for compact perturbations of monotone type operators and application to nonlinear parabolic problem	10,039
Some characterizations of the Cobb-Douglas and CES production functions in microeconomics	5,630
Analysis of Water Hammer with Different Closing Valve Laws on Transient Flow of Hydrogen Natural Gas Mixture	3,335
FOURIER SERIES OF THE PERIODIC BERNOULLI AND EULER FUNCTIONS	1,462

Most Viewed Special Issue Articles

Article	Views
Numerical methods for solving Fredholm integral equations of second kind	3,884
Picard Successive Approximation Method for Solving Differential Equations Arising in Fractal Heat Transfer with Local Fractional Derivative	3,241
A note on fractional order derivatives and Table of fractional derivative of some specials functions	2,509
A Generalized Solow-Swan Model	1,622
Solving Nonlinear Partial Differential Equations by the sp-	

41 Total authors

20 Total published articles

263 Total submitted articles

28 Total accepted articles

94 DAYS — Average time from submission to acceptance

33 DAYS — Average time from acceptance to publication

13

Why the dashboard?

The dashboard **business goals:**
1. Encourage readership and convert readers into authors
2. Attract high quality editors/reviewers
3. Represent the brand and its quality
4. Promote Open Access, Open Science and Open Infrastructure

The dashboard is a **data visualization tool** that displays metrics and key data points to monitor specific process result in Hindawi, in particular how:

● Articles undergo a **transparent editorial process** that respects standards and integrity principles.

● Articles are **read, cited** and contribute to the increasing of knowledge in the research field and have a **huge impact** on society.

● Articles are **findable, accessible, interoperable and reusable**.
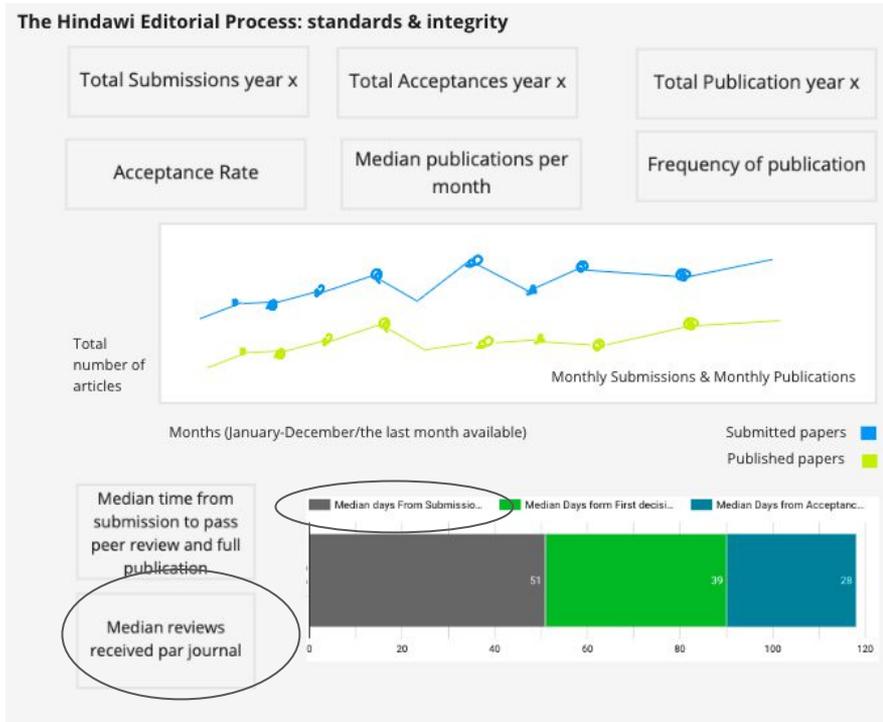
The dashboard - Introduction

The information that will be contained in the dashboard reports have to be evaluated accurately in order to not overwhelm the users.

The reports may be filtered by journal name and year to allow an easy and intuitive navigation for the users, respecting also what is the current structure of our reports.

Our users must be able to immediately **identify the journal** and recognise some **static and essential information**.
For these reasons, we will display the journal name, the journal homepage URL, the publisher name, a validated ISSN (unique persistent identifier), the discipline area of the research articles and the APC price.

The dashboard - Editorial process

The Hindawi Editorial Process: standards & integrity

| Total Submissions year x | Total Acceptances year x | Total Publication year x |

| Acceptance Rate | Median publications per month | Frequency of publication |

Total number of articles

Monthly Submissions & Monthly Publications

Months (January-December/the last month available)

Submitted papers
Published papers

Median time from submission to pass peer review and full publication

Median reviews received par journal

Median days From Submissio...   Median Days form First decisi...   Median Days from Acceptanc...

51   39   28

0   20   40   60   80   100   120

We should provide a transparent description of our **journals editorial policies** and decision making processes.

Basic statistics must be published annually, covering in particular the number of submissions, the number of reviews received, the acceptance rate and the average time between submission and first decision.
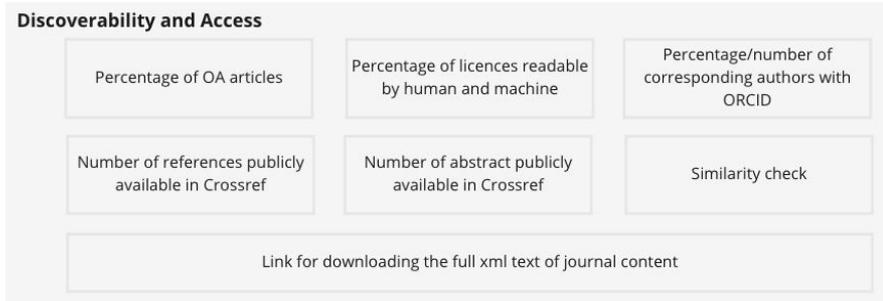
16

The dashboard - View and Reuse
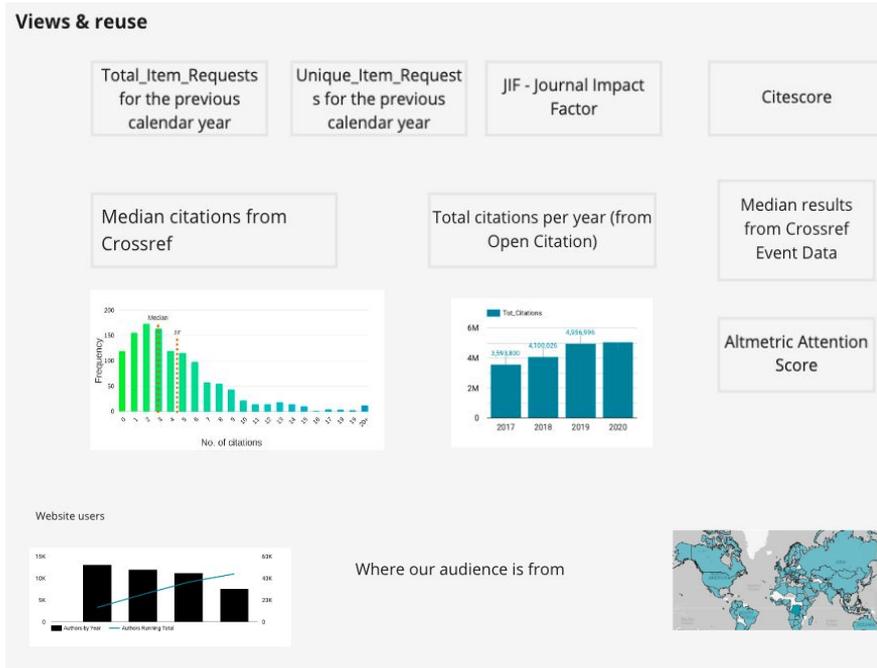
Studying a different approach to the research output evaluation, Hindawi must use its reports to show also the impact of its articles on the society.

There is an increasing need to change the approach through which scientific research is evaluated and **Hindawi can push innovation**, becoming the standard-bearer of the metrics for the evaluation of the research. This will raise awareness of the broader picture that JIF conceal.

In particular, I extracted some information on the impact of a publication in the society through services like **Crossref Event Data** and the number of citations through the years thanks to **Open Citation** APIs.

The dashboard - Discoverability and Access

Discoverability and Access

| Percentage of OA articles	Percentage of licences readable by human and machine	Percentage/number of corresponding authors with ORCID
Number of references publicly available in Crossref	Number of abstract publicly available in Crossref	Similarity check
Link for downloading the full xml text of journal content		

The inspiration came from **Crossref Participation Reports** and I mainly used the *Metadata Retrieval Crossref, REST APIs* to extract these information.

We can use the dashboard to rise the awareness of users about the importance of open infrastructure for discovery and reuse.

This is a value-add service by publishers for Open Science that makes **research more open, global, collaborative, creative and closer to society**.

The dashboard - promoting Equity, Inclusion and Diversity - a global Open Science Network
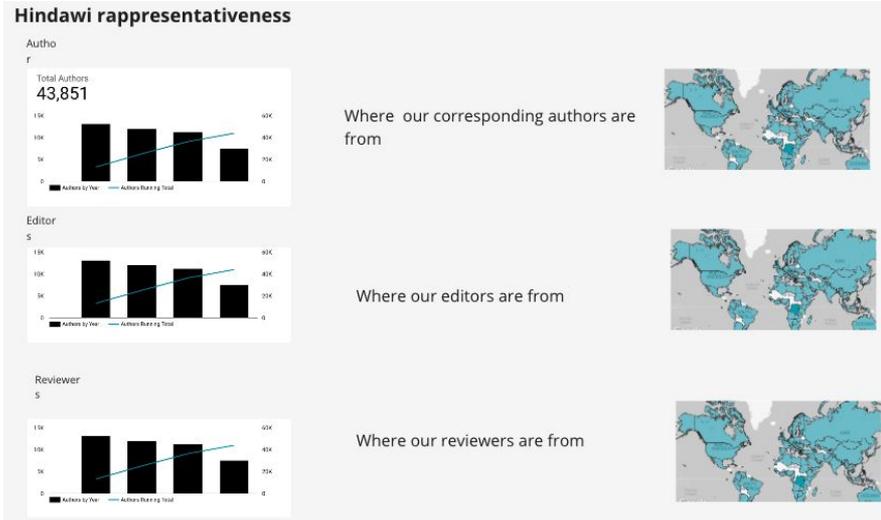
Hindawi collect papers from different authors from many different countries and it is important to show how **globally representative** Hindawi articles are.

As well, we may show where our reviewers and editors are from, letting the users of the dashboard know how much we encourage and research **collaborations from all over the world**.

We need a global Open Science network and, Hindawi as a publisher, embraces global diversity.

The spotlight's on you!
We want to hear your song

What information and/or metrics do you think will (or should) be of most interest to authors? Why?

How can we make researchers more aware of the importance of open science and open infrastructure?

Where else should we showcase open data and responsible metrics from?

https://forms.gle/EXDGJ1dSvdfmNxJt5

The following slides were added at the end of #Pidapalooza21 and summarise the feedback we received on the google form to the questions and in the chat during the presentation

The form

We shared a google form with our four questions with the audience, allowing them to answer the questionnaire during the presentation and for all the duration of the PIDapalooza 2021.

Our audience

We received twelve responses from:

First question

What other wild/exotic/useful user cases for ROR - Research Organization Registry at a publisher can you think of?

1. Topic clusters - orgs associated with keywords - articles on this topic were published here
2. Way to QC (quality control) and correct/standardise affiliations; link the institutions more easily with authors and funding and research publications; measure our global reach for attracting diverse authorship; review our peer review process to look for inequalities and potential bias
3. ORCID (for authors/contributors); Grant DOIs (Crossref) to acknowledge funder awards directly (not just the funder organization but the specific grant(s)
4. There are many - we are holding workshops to collect this info around usage data and they intersect a lot https://educopia.org/data_trust/
5. Like ORCID permits disambiguation. My institution found another one in the US with the same name
6. For soliciting potential guest editors for special issues
7. Open Source Intelligence for use by academic/scholarly researchers; law enforcement (LEO); the Intelligence community (IC) and citizen journalists & scientists
8. List of connections among ROR based on co-authors
9. Would prefer to focus on the badly-needed but not at all wild use case of correctly associated all authors of a publication with their institutional affiliation(s) in a machine-readable way that's open and easily consumed by other systems

Second question

How can we make researchers more aware of the importance of open science and open infrastructure?

1. Show them the benefits for them and their colleagues and make it easy for them and ensure they feel it will not compromise them or their research potential by sharing their work and allowing others to build on it
2. Hit them with a big stick that has "pids are good" stamped on it
3. Work with funders - researchers follow the money
4. Speak of this. Concrete use cases
5. Make participation in it free, or linked to incentives like fee waivers, etc.
6. Guerrilla marketing and product placement in popular and niche entertainment and infotainment; chalk art in front of research libraries and related institutions; PR -- seed stories in public radio (e.g. in USA NPR; etc) and find sponsors in the political sphere who will cite its benefits
7. Make it obligatory

Third question

What information and/or metrics do you think will (or should) be of most interest to authors? Why?

1. So many. I'm happy to share what we're learning from our author workshops
2. Self-citations; top-six most co-cited collaborators (this is available in commercial tools, of course); # of open data sets provided at author and paper/publication level (eg an indication of sharing enthusiasm)
3. Reach
4. Downloads and Views of the paper/data/code. The structure of viewers/downloaders, too.

Fourth question

Where else should we showcase open data and responsible metrics from?

1. Help funders understand how open their funded researchers have been.
2. Twitter
3. The corporatesphere

« Regarding providing author-level information to funders, my take is that is not the role of publishers (as it is inherently limited to the output with that publisher), but publisher can play an important role in contributing to such information as open (meta)data» Bianca

«The use cases we're developing via the OA eBook Usage (https://hcommons.org/deposits/objects/hc:24148/datastreams/CONTENT/content) data trust intersect quite a bit. » Christina

«Design thinking workshops sound like a great approach to create dashboards» Kristian

«Showing information from Open Citations and Crossref Event Data, rather than using more closed alternatives - will send a strong signal not just to researchers, but to the scholarly communication community at large - creating awareness and (hopefully) support.» Bianca

«Can you help funders with metrics to assess how "open" their funded researchers are?» Jamie

«This conversation is right in line with what we at Make Data Count (https://makedatacount.org/) are trying to address - developing proper indicators for research data in a way that makes sense for researchers & research stakeholders. Required that upstream we all support these PIDs, data citations, etc so that we have usage/citation data to expose and evaluate.» Daniella