



---

# Federal Demonstration Partnership Meeting

## January, 2012

Laure Haak, PhD  
Chief Science Officer  
Custom Analytics and Engineered Systems  
13 January 2012



THOMSON REUTERS

# RESEARCH REPORTING WORKFLOW

---

- Identify people, concepts, and organizations
- Determine connections
- Track research activities and outputs



---

# DEFINING THE PROBLEM

# NAME AMBIGUITY

How do I get credit for my work?

Peterson, J

FIND



1839 articles

Peterson, J

FIND



• Janet Peterson  
55 articles  
• Janet Peterson  
37 articles

Disambiguation software can help find out who wrote what. But an author's work is often broken up into multiple groups, for instance, because they changed jobs. And papers published under misspelled or former names aren't found.

1965496543

FIND

→ 96 articles

A number identification system could help find all the papers written by an author, independent of changes in fields, affiliations, or name. It could also help keep contact information up-to-date.

## JANET PETERSON

Alternative Name:  
JANET HOLDEN



E-mail Address:  
JPETERSON@UFA.EDU

Current Affiliation:  
UNIVERSITY OF FLORIDA AT ARCADIA



THOMSON REUTERS

CREDIT: J. NEWFIELD/SCIENCE

## RESEARCHER

---

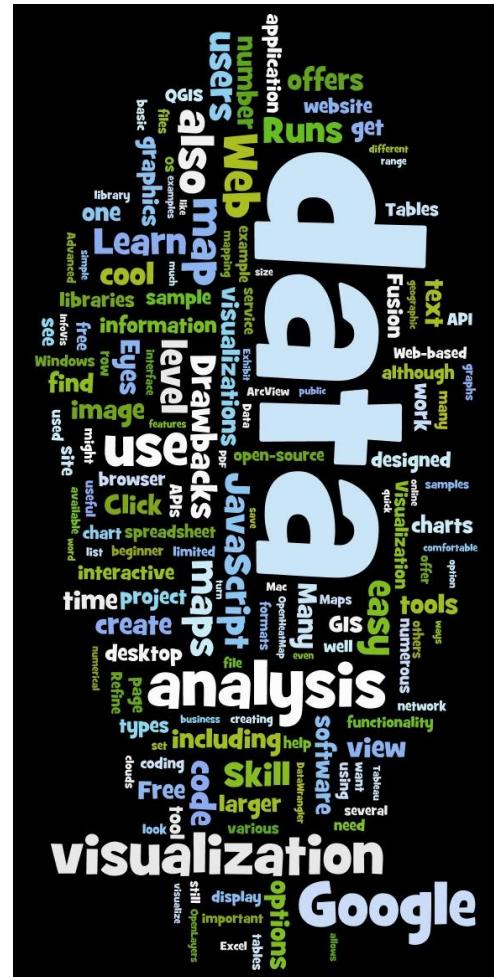
*If it provides a more efficient way to manage my profile data and institutional and funder reporting requirements, I'll use it.*



- I need to submit a paper
- I need to list my scholarly output in a grant application biosketch
- I want to know what other papers were published by these authors
- I want to find potential collaborators

# REPORTING CHALLENGES

Funding, personnel, and output data sources were not designed for the purpose of tracking or reporting



# RESEARCH INSTITUTION

---

*If it provides a more efficient way to manage expertise and impact data and supports seamless institutional reporting, we'll use it.*

- We need to understand our institutional research strengths and impact
- We want to know how we compare to other institutions
- We need to track faculty publications for our institutional repository and/or profiling system.
- We want to know with whom our faculty are collaborating



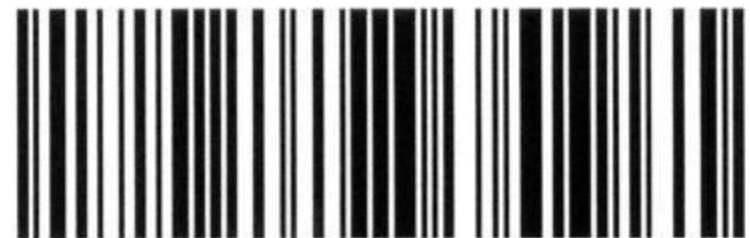
# FACING THE CHALLENGE

---

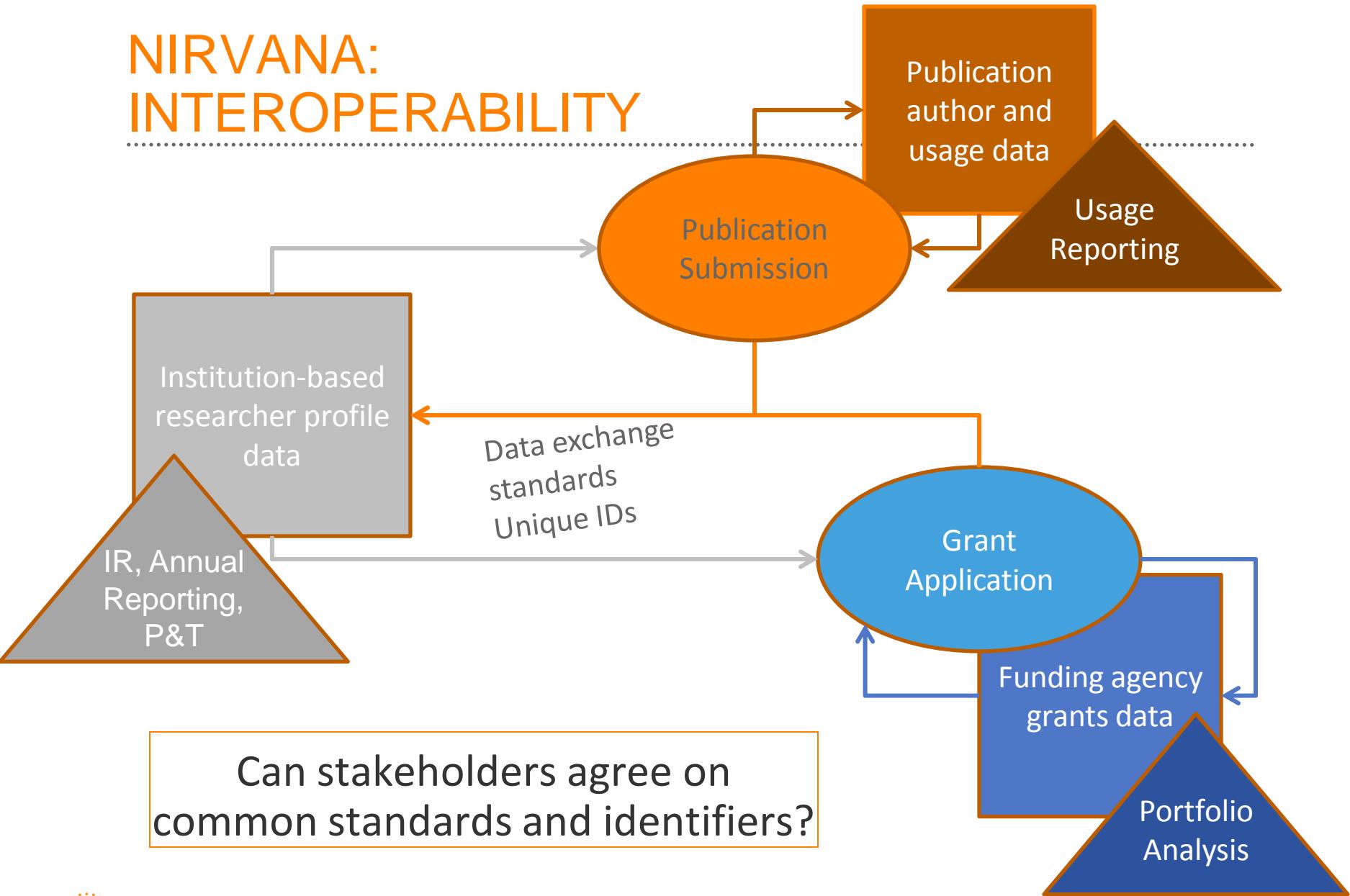
Adopt definitional standards within and across data sources and types

Apply unique ID to person records across data sources and types

Enlist all sectors in data collection and standards effort



# NIRVANA: INTEROPERABILITY



---

# STANDARDS AND UNIQUE IDENTIFIERS



## ORCID

---

ORCID is a permanent  
registry of unique contributors

ORCID provides a  
disambiguation service

ORCID is the crux of a  
disambiguated record of  
research activity

Example ID: <http://orcid.org/0137-1963-7688-2319>





## WHY ORCID?

---

**TRUSTWORTHY:** Inclusive and transparently governed not-for-profit

**GLOBAL:** Not limited by discipline, institution, geography

**OPEN:** Data and source code available under recognized open licenses

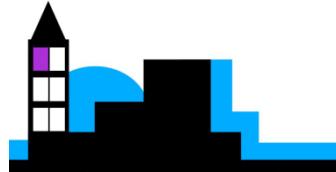
**EFFICIENT:** Third-party seeding of profiles

**INTEGRATED:** Part of institutional, publisher, and funding agency infrastructure



# ORCID IN CRITICAL WORKFLOWS

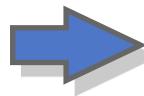
Researcher



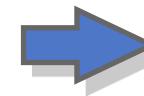
Joins faculty  
Joins student body



- Track output of researchers
- Locate collaborators



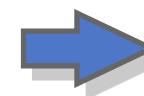
Applies for grant



- Streamline application process
- Support research assessment



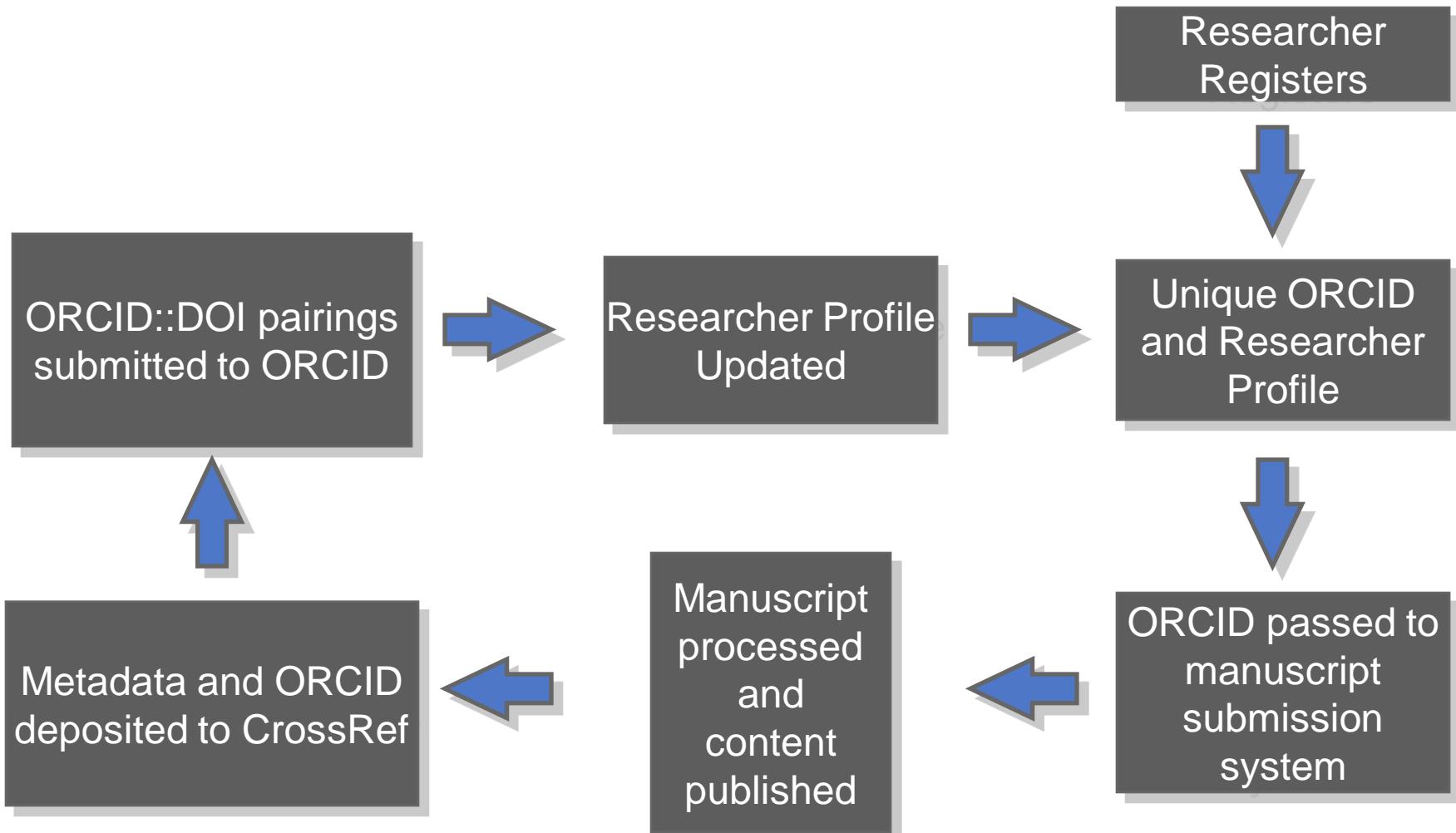
Submits  
manuscript



- Streamline data input
- Create author links



## AUTHOR→ORCID→PUBLISHER WORKFLOW





# ORCID PROFILE

Researchers create and maintain an ORCID ID and profile free of charge, with fine-grained control of privacy settings

Personal Information    My Affiliations    Password    Options    [Return to My Researcher Profile](#)

Update your information. Optionally, use the "public" setting to indicate which information to display on your Profile page. Clear a public check box to make that information private. Note: Click the "Submit Changes" button on each page. \* required fields.

|   |  |   |
|---|--|---|
| <input checked="" type="checkbox"/> Public (Master) | First / Given Name: <input type="text" value="Martin"/> *  | Public (Master): Clear the "Public (Master)" check box to make your information private. Only you, when logged in, will be able to view your information.   |
|   | Last / Family Name: <input type="text" value="Fenner"/> *  | Name: Your name will be searched and displayed on your profile page.  |
|   | Middle Initials: <input type="text" value="H"/>  | ORCID ID: This is your unique identification.   |
| <input checked="" type="checkbox"/> Public          | Other Names Used by You: <input type="text"/> Add <input type="button" value="Remove"/>                        | Other Names: If you've published under different names or if you are known by more than one name, enter Other Names. These will be searched if made public. |
| <input checked="" type="checkbox"/> Public          | URL: <a href="http://www.orcidsandbox.org/rid/AAA-1034-2010">http://www.orcidsandbox.org/rid/AAA-1034-2010</a> | URL: The URL can be used to link directly to your profile page.   |





# RESEARCHER IDENTIFICATION STANDARD

ORCID will interact with other researcher identification systems



THOMSON REUTERS  
ResearcherID



Author Resolver

SCOPUS

dblp.uni-trier.de  
Computer Science  
Bibliography



ACM DL DIGITAL  
LIBRARY

VIVO

Microsoft  
Research

RePEc

INSPIRE  
Fermilab SLAC



THOMSON REUTERS



# ORCID BRINGS STAKEHOLDERS TOGETHER

- ❖ ORCID members include publishers, research institutions, funding organizations, and data aggregators.
- ❖ ORCID governance is majority non-profit by design
- ❖ Software developed by ORCID will be released under an Open Source Software license.

wellcome trust



nature publishing group npg

INSPIRE  
Fermilab SLAC

THOMSON REUTERS



crossref



MHH  
Medizinische Hochschule  
Hannover

arXiv.org



OCLC  
The world's libraries.  
Connected.

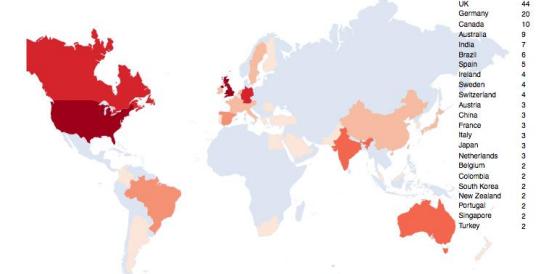
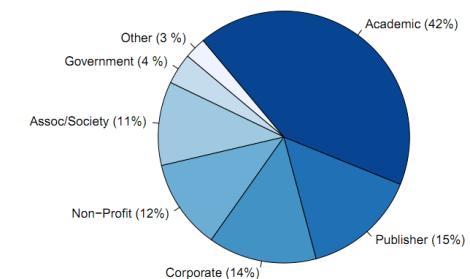
MIT Libraries

ACM  
Association for Computing Machinery

Advancing Computing as a Science & Profession

WILEY  
KNOWLEDGE FOR GENERATIONS

NII  
日本学術情報研究所  
National Institute of Informatics





# TIMELINE 2010

Feb    March    April    May    June    July    Aug    Sept    Oct    Nov    Dec

Build  
Sandbox

Alpha Prototyping

ORCID Members  
Demo

Organization  
Established

Wellcome/  
MIT Survey

Principles/Scope  
Defined

Alpha Testing

Profile Exchange Research & Development





## ORCID SYSTEM

---

In August, 2011, ORCID and Thomson Reuters reached an agreement that enables ORCID to start building the ORCID service based on Researcher ID code, giving ORCID the critical technology to create its system.





# TIMELINE 2011-12

---

Q1 2011

Q2 2011

Q3 2011

Q4 2011

Q1 2012

Q2 2012

Q3 2012

Q4 2012

Build Phase 1

API available  
11/11

VIVO Technology Research

Mellon Market  
Research

Sponsorship  
Drive 1

Obtain Loans and  
Sponsorship Drive 2

Staff  
Hired

Start Collecting Fees?

Profile Exchange Research & Development





## ORCID PHASE 1 SYSTEM

---

- ❖ Researcher self-claim, delegated management, and institutional seeding of data
- ❖ Fine-grained control of privacy settings (public/private/protected)
- ❖ Data exchange into grant/manuscript submission systems
- ❖ ORCID identifier resolution and metadata search via GUI and REST API
- ❖ API mockups posted at <https://github.com/ORCID>



# PHASE 1 QUERY TYPES

| Name   | Key  | Returned                                   | Description   |
|--------|--|--|---|
| Bio    | ORCID  | Profile metadata                           | Given an ORCID, give me name and affiliation data.  |
| Works  | ORCID  | List of work metadata                      | Given an ORCID, tell me what works they have contributed to.  |
| Full   | ORCID  | Profile metadata, work metadata and ORCIDs | Given an ORCID, tell me what works they have contributed to, name and affiliation data.                 |
| Work   | Work identifiers (e.g. DOIs)                 | ORCIDs & associated metadata               | Given a work, tell me who contributed to it.  |
| Search | ORCID, Work identifiers, or profile metadata | ORCIDs & associated metadata               | Given whatever metadata I have, give me a ranked list of potential parties identified by that metadata. |



## ORCID PHASE 2

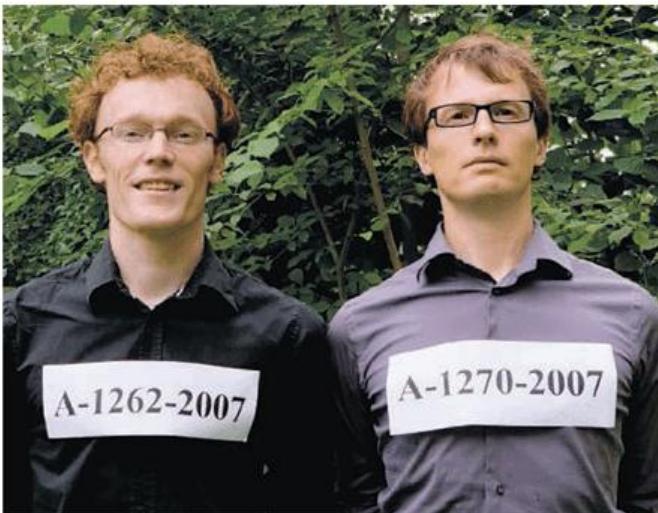
---

- ❖ Non-university, third-party deposit of data
- ❖ Automatic de-duplication of records
- ❖ Ability to collect and protect sensitive data
- ❖ Ability to claim non-publication contributions
- ❖ Ability for institutions and/or publishers to verify researcher claims
- ❖ Ability to capture contribution roles (for example, second author, editor, reviewer, etc.).

# QUESTIONS?

## Are You Ready to Become a Number?

Science 27 March 2009:  
vol. 323 no. 5922 1662-1664



### Identity solution.

Jochen Cals (*left*) and Daniel Kotz argue that a global system of unique identifiers, such as their ResearcherIDs, will help scientists better search research literature and work with colleagues.

CREDIT: APPIE DERKS/MAASTRICHT UNIVERSITY MEDICAL CENTRE

## Laure Haak

Chief Science Officer

Discovery Logic/Thomson Reuters

[laurel.haak@thomsonreuters.com](mailto:laurel.haak@thomsonreuters.com)



| Name           | Institution(s)  | Country / Territory | Researcher ID |
|----------------|-----------------|---------------------|---------------|
| Haak<br>Laurel | Discovery Logic | United States       | C-4986-2008   |



THOMSON REUTERS

FDP 13 January 2012

---

# APPENDIX





## Scope of Research Data Standards

- Master entity profiles with data elements, semantic definitions, structural specifications and codes tables
- Business documents as 'point-in-time' snapshots of profile data
- Exchange packages grouping related business documents

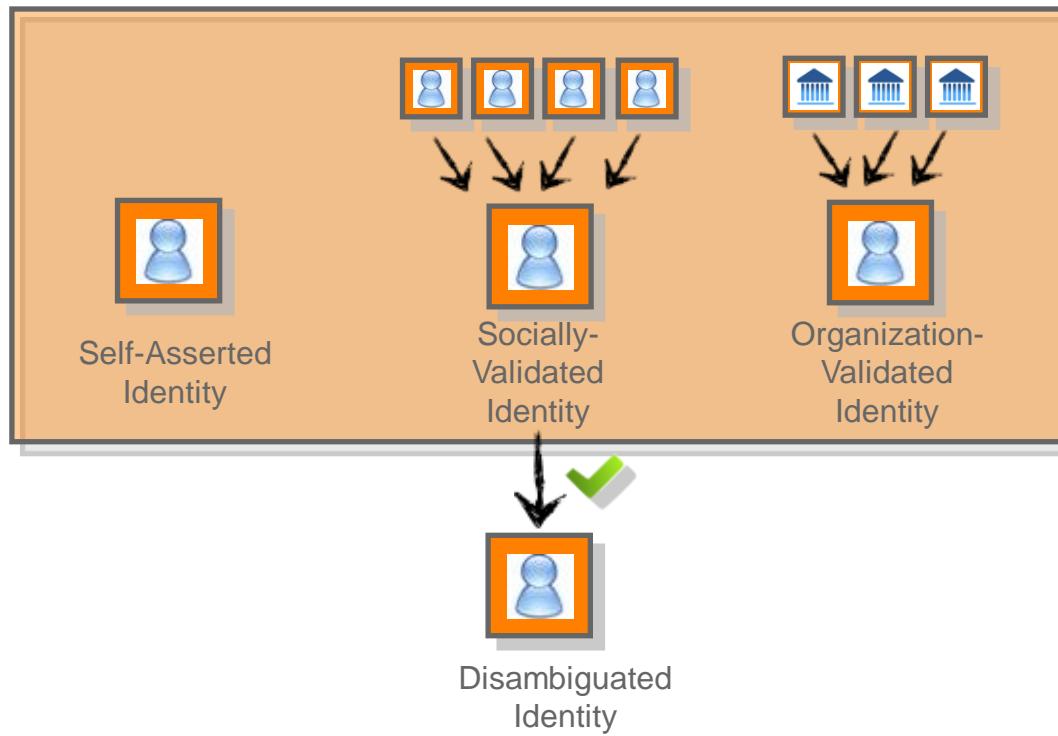
- Research Personnel Data
- Research Activity Data
- Research Organization Data
- Research Impact Data
- Research Classification Taxonomy

<http://dictionary.casrai.org>

- ORCID for Personnel
- DOI for Research Outputs
- Exploring ID options for Activities and Organizations

# COMMUNITY-VALIDATED PROFILES

***Self-asserted + socially-validated + organizationally-asserted attribution  
= more credible assertion***



## MICRO ATTRIBUTION

---

WorkClaim: *for a particular work (i.e., a DOI), tell me not only what contributors (ORCIDs and associated metadata) are responsible for it, but also what each contributor actually contributed (e.g., “wrote paper”, “designed study”, “collected data”, “performed experiments”, “contributed reagents”, “performed statistical analysis”, “acquired funding”, etc.)*



## ORCID FUNDERS TO DATE

---

VIVO – to model VIVO-ORCID profile exchange

Mellon – to conduct business feasibility study and planning

NSF – to examine the value of ORCID in advancing the science of science policy, via NSF-ORCID-UChicago/Harvard profile exchange

Wellcome Trust – market research, and forthcoming workshop on micro-attribution



## OTHER SPONSORS

---

APS  
Wiley  
Elsevier  
NPG  
Hindawi  
Mendeley  
Bloomsbury Qatar Foundation Journals  
American Society of Civil Engineers  
Copyright Clearance Center  
Digital Science  
PNAS  
California Digital Library / Innovation Alliances and Services  
Royal Society Publishing  
Emerald Group Publishing  
American Institute of Physics  
Springer  
CSIRO Publishing  
MIT  
Portland Press Limited  
HFSPO  
American Astronomical Society  
Royal Society of Chemistry  
Aries Systems

American Chemical Society  
SAGE Publications  
Oxford University Press  
California Institute of Technology  
APA  
IOP Publishing  
ACM  
Thomson Reuters  
Wellcome Trust  
PLoS  
Taylor and Francis  
National Institute of Informatics  
University of Vienna - Vienna University Library  
Ringgold  
Annual Reviews  
AAAS  
Hogrefe Group  
Cornell University Library  
Harvard University  
Cambridge Crystallographic Data Centre  
eJournal Press  
SSRN  
Microsoft Research  
Atypon  
JSTOR





## DECadal Membership Targets

Exhibit 2: Projected ORCID Members, By Segment

| Market Segment                     | Participating Organizations |
|------------------------------------|-----------------------------|
| For-profit Publishers              | 60                          |
| Nonprofit publishers               | 40                          |
| Governmental & Private Funders     | 10                          |
| Universities & Research Institutes | 325                         |
| <b>Total</b>                       | <b>435</b>                  |





## MEMBERSHIP TIERS

- Membership is at the institutional – not library – level.
- Tiers tied to level of research intensity, a function highest degree granted and number of faculty

Exhibit 3: Fee Ranges, By Segment

| Market Segment                     | Annual Fee/Capitalization Fee |                   |
|------------------------------------|-------------------------------|-------------------|
|                                    | Low Tier                      | High Tier         |
| For-profit Publishers              | \$500/\$550                   | \$44,100/\$48,510 |
| Nonprofit publishers               | \$400/\$440                   | \$35,280/\$38,810 |
| Governmental & Private Funders     | \$12,500/\$13,250             | \$58,000/\$63,800 |
| Universities & Research Institutes | \$600/\$660                   | \$9,400/\$10,340  |