

INFORMATION STANDARDS QUARTERLY

FALL 2013 | VOL 25 | ISSUE 3 | ISSN 1041-0031

# ISQ

TOPIC

## DATA CURATION

---

DATA CURATION ISSUES IN THE  
CHEMICAL SCIENCES

---

DATA CURATION IN THE OPENAIRE  
SCHOLARLY COMMUNICATION  
INFRASTRUCTURE

---

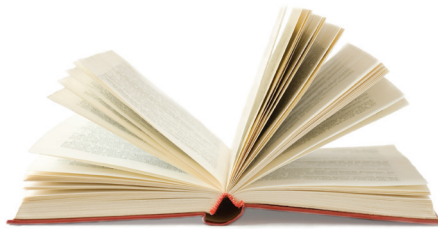
PRESERVING THE GREY  
LITERATURE EXPLOSION:  
PDF/A AND THE DIGITAL ARCHIVE

---

ENSURING THE LONG TERM IMPACT OF  
EARTH SCIENCE DATA THROUGH DATA  
CURATION AND PRESERVATION



Charles, Prince of Wales |



Charles Prince, Author |



Prince Charles, R&B Artist |

The ISNI standard (ISO 27729) was created to disambiguate names and act as a link for the data about an identity. So, for example, information about Charles, Prince of Wales is distinct from information about Charles Prince, an author, or Prince Charles, an R&B artist, and the information attached to, say, "Tolstoy" is also attached to "Tolstoi."

## ISNI Assignments Top 6.5 Million

More than 6.5 million assignments of the ISO-certified naming standard International Standard Name Identifier, known as ISNI, are ensuring the right information gets connected with the right name. Published in early 2012, the standard applies a unique 16-digit code to public identities, providing a single identifier that can be leveraged across many applications, helping to sync alternate or disparate spellings of the same name, and eliminating confusion when names are alike. It's been quickly accepted by information organizations, including Wikipedia.

The ISNI standard (ISO 27729) was created to disambiguate names and act as a link for the data about an identity. So, for example, information about Charles, Prince of Wales is distinct from information about Charles Prince, an author, or Prince Charles, an R&B artist, and the information attached to, say, "Tolstoy" is also attached to "Tolstoi." ISNIs are especially important for organizations administering rights, simplifying identification and administration of royalties.

The ISNI International Agency, a worldwide group of organizations that serve researchers, rights management organizations, authors, musicians, and other public contributors, manages the implementation of the standard and appoints registration agencies. Organizations and individuals can apply for an ISNI for any public identity—real or fictional—through these registration agencies, such as Bowker and Ringgold. ■

ISO 27729:2012 standard: [www.iso.org/iso/home/store/catalogue\\_tc/catalogue\\_detail.htm?csnumber=44292](http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=44292)

ISNI International Agency: [www.isni.org](http://www.isni.org)

ISNI Registration Agencies: [www.isni.org/content/isni-registration-agencies](http://www.isni.org/content/isni-registration-agencies)

Full press release: [www.proquest.com/en-US/aboutus/pressroom/13/20130827.shtml](http://www.proquest.com/en-US/aboutus/pressroom/13/20130827.shtml)



## NFAIS Issues Recommended Practice on Discovery Services

The relative newness of Discovery Service technologies—defined as the provision of a single search box to access a central index of pre-indexed metadata and/or full-text—has generated questions and concerns from both content owners and libraries. NFAIS issued in August 2013 a set of guidelines to “help maintain an equitable balance of the interests of all participants in the information distribution process” associated with discovery services.

“Acceptance of these Recommended Practices will help facilitate full disclosure and transparency for the ultimate benefit of information seekers so that they will be able to know the parameters of the information that they are searching and to which they have access, and to become aware of the complex relationships among the participants in any Discovery Service arrangement.”

Following a brief background on what discovery services are and a discussion of the concerns and issues regarding their use, the document provides recommendations on the rights and obligations of five different participants in a Discovery Service arrangement: the Content Owner, the Platform that hosts the content, the Discovery Service, the Subscriber to the service, and the user of the service. Eighteen different services are shown in a matrix against the five participants with a designation of Right (R) or Obligation (O) or both for each. The services cover a diversity of issues related to content identification and retrieval, usage reports, service components, privacy, and training. Each of the 18 services is then described in a more detailed narrative.

Appendices provide a bibliography about discovery services and a glossary of terms. ■

🔗 NFAIS Discovery Services Recommended Practice: [info.nfaais.org/info/Recommended\\_Practices\\_Final\\_Aug\\_2013.pdf](http://info.nfaais.org/info/Recommended_Practices_Final_Aug_2013.pdf)

## W3C Launches Web and Mobile Interest Group

The World Wide Web Consortium (W3C) launched in August 2013 a new Web and Mobile Interest Group that is chartered to accelerate the development of Web technology so that it becomes a compelling platform for mobile applications and the obvious choice for cross platform development. The forum is intended to include organizations that commission such products and services, designers, developers, equipment manufacturers, tool and platform vendors, browser vendors, operators, and other relevant participants in the value chain that creates and operates such products and services. Participants will focus on a wide range of sectors including retail, advertising, technology, network operators, content creation, and content distribution.

Initial planned deliverables include a new version of the Core Mobile Web Platform 2012 report, a gap analysis of the differences between the Web as a platform on mobile and other popular platforms, a broad look at all the Web technologies under development that are particularly relevant to mobile devices, and additional reports on use cases and scenarios for context-relevant user experiences, multi-device and cross-device user experiences on the Web, and usability and efficiency considerations. ■

🔗 Announcement: [www.w3.org/News/2013#entry-9917](http://www.w3.org/News/2013#entry-9917)

Web and Mobile Interest Group:  
[www.w3.org/Mobile/IG/](http://www.w3.org/Mobile/IG/)

## Find the Registration Agency for any DOI

CrossRef is one of the best-known Registration Agencies (RAs) for Digital Object Identifiers (DOI) due to their focus on scholarly documents and reference linking. But there are close to 10 agencies authorized by the International DOI Foundation to assign DOIs, each of which may provide different services or address different needs. For example, DataCite assigns DOIs to datasets. The metadata associated with DOIs is generally stored and maintained by the RA that assigned the DOI and until now it was difficult, if not impossible, to determine which agency is associated with a particular DOI.

CrossRef recently released a new Application Programming Interface (API) tool to allow anyone to look up the Registration Agency (RA) for a given DOI or a group of DOIs. To find the registration agency, simply add "http://doi.crossref.org/doiRA/" before any DOI in your browser address bar or program.

For example, when searching for the doi: 10.1037/0003-066X.59.1.29, the API returns:

"RA": "CrossRef"

While a search for the doi: 10.1594/PANGAEA.57055 in the API returns:

"RA": "Data Cite"

Multiple DOIs can be searched simultaneously by separating them with commas. ■

🔗 List of DOI Registration Agencies: [www.doi.org/registration\\_agencies.html](http://www.doi.org/registration_agencies.html)

CrossRef: [www.crossref.org](http://www.crossref.org)



## Thomson Reuters Whitepaper Tackles Issues of Uniform Access to Global Scholarly Data

Thomson Reuters convened an Industry Forum in April 2013 for STM publishing leaders to discuss trends, developments, and strategies related to the management and use of research data. The first output of that forum is the whitepaper *Unlocking the Value of Research Data*. The paper outlines many of the challenges that are being faced in managing research data, highlights organizations working in this arena (e.g., figshare, the Research Data Alliance (RDA), the International Council of Science (ICSU) Data Publication Working Group, and Thomson Reuters with its Data Citation Index), and recommends creating a consortium of publishers and associations to develop standards that will guide new modes of communicating and sharing data.

Future topics that the Forum intends to address include open access, content enrichment, post publication peer review, and research workflows. ■

🔗 Announcement: [thomsonreuters.com/press-releases/082013/global-scholarly-data](http://thomsonreuters.com/press-releases/082013/global-scholarly-data)

Unlocking the Value of Research Data white paper: [collaborativeresearch.thomsonreuters.com/researchanalytics.thomsonreuters.com/m/pdfs/1003903-1.pdf](http://collaborativeresearch.thomsonreuters.com/researchanalytics.thomsonreuters.com/m/pdfs/1003903-1.pdf)

## ORCID Project to Develop Use Cases and Best Practices

ORCID, the non-profit organization that provides an open registry of unique identifiers for researchers, has received funding from the Alfred P. Sloan Foundation to standardize and streamline the ORCID identifier integration process, collect and document use cases, develop open source code samples, and provide case studies of working integrations. Use cases and integration best practices will be disseminated through an Outreach meeting and CodeFest, to be held in Chicago in May 2014.

While the Sloan-funded award focuses on organizations based in the United States, it complements an initiative at

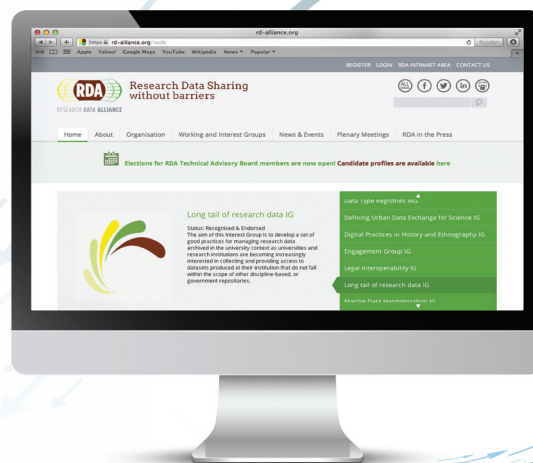
developing possible ORCID identifier implementation use cases being carried out by Jisc in the United Kingdom.

ORCID launched its Registry in October 2012. Since then over 160,000 researchers have registered for an identifier, and several systems including manuscript submissions and grant applications have begun to embed the identifiers. ■

🔗 Announcement: [orcid.org/blog/2013/06/13/orcid-awarded-grant-alfred-p-sloan-foundation-support-university-and-professional](http://orcid.org/blog/2013/06/13/orcid-awarded-grant-alfred-p-sloan-foundation-support-university-and-professional)

ORCID website: [orcid.org/](http://orcid.org/)

The Research Data Alliance (RDA) goal is to accelerate international data-driven innovation and discovery by facilitating research, data sharing and exchange, use and reuse, standards harmonization for specific communities and across scientific disciplines.



## Research Data Alliance Formed to Facilitate Research Data Sharing and Exchange

The Research Data Alliance (RDA) is a newly formed virtual organization funded by the European Commission and the US and Australian governments in their respective regions. Its goal is to accelerate international data-driven innovation and discovery by facilitating research, data sharing and exchange, use and reuse, standards harmonization for specific communities and across scientific disciplines. These goals are achieved through the development and adoption of infrastructures, policies, practices, and standards with the support of a structure that includes a council, plenary, secretariat, non-governmental structures (NGSs), and working groups.

Organizations may join the RDA. Individuals may participate through Discussion Forums/Organic Groups, Interest Groups, Working Groups, and the plenary bi-annual meetings. A Launch Event and First Plenary were held on March 18-20, 2013 in Gothenburg, Sweden and a 2nd Plenary was held September 16-18 in Washington, DC. 714 people have formally agreed to the *RDA Principles* and officially signed up as members of RDA.

Six WGs have been formally recognized:

- 1 Data foundation and terminology
- 2 Data type registries
- 3 Metadata standards directory
- 4 Persistent identifier information types
- 5 Practical policy
- 6 Standardisation of data categories and codes

Two additional WG case statements are in review (Community capability model and Data citation) and several more are out for community comment.

Thirteen Interest Groups (IGs) are formally recognized and active: Agricultural data interoperability, Brokering, Certification of digital repositories, Digital practices in history and ethnography, Engagement group, Legal interoperability, Long tail of research data, Metadata, Preservation e-infrastructure, Publishing data, Research data provenance, Structural biology, and Toxicogenomics interoperability. ■

🔗 Research Data Alliance: [rd-alliance.org/](http://rd-alliance.org/)

RDA Principles: <https://www.rd-alliance.org/get-involved.html>

| NW | doi: 10.3789/isqv25no3.2013.06