

**Submission of Copyright Clearance Center in response to Request for Information: NASA Public Access Plan for Increasing Access to the Results of NASA-Supported Research** issued May 18, 2023 by the National Aeronautics and Space Administration; Agency/Docket Number Notice: 23-051, Document Citation 88 FR 31827.

CCC welcomes the opportunity to submit this response to Question 4 of the RFI and welcomes NASA's interest in improving use of PIDs and metadata to improve the experiences of institutions and researchers.

### **Background on CCC.**

CCC is a not-for-profit organization founded in 1977 at the suggestion of Congress to facilitate collective copyright licensing for the text sector. Presently, among other lines of business, CCC provides licenses to content from over 10,000 rightsholders for whom we serve as an agent. We provide these licenses to more than 35,000 business organizations (Business Users) around the world. CCC is a supplier of knowledge management software called RightFind®, which is used by a subset of these Business Users to manage and access content. We also provide (1) other software services, (2) library staffing, (3) content enrichment, data and metadata services, and (4) content delivery. On October 19, 2021, [U.S. Secretary of Commerce Gina Raimondo announced](#) that we were awarded a Market Development Cooperator Program grant, administered by the Commerce Department's International Trade Administration, to support our work with standards development organizations.

Our fastest growing business is managing the agreement- and fee-administration process on behalf of publishers who collect fees or otherwise track usage from authors, institutions, consortia, government and other funding bodies for immediate open access to scientific articles (OA). We do this primarily through our RightsLink® for Scientific Communications software platform (RLSC). RLSC is by far the market leader in managing open access agreements and payments, doing so for many top publishers.

### **PIDs and Metadata.**

Through both our knowledge management work with Business Users and our work on behalf of publishers, CCC experiences firsthand the promise of persistent identifiers (PIDs) when applied early, consistently and persistently. We are also painfully aware of the problems related to the entropy that results from lack of early, consistent, and persistent application thereof.

A healthy research and publishing ecosystem requires PIDs and robust, rich, quality metadata to make connections among people, organizations, places, and digital objects. For example, in RLSC alone, we depend on dozens of author, institution, and manuscript metadata elements to apply the appropriate business logic and workflows necessary to automate and scale OA on the path toward open science.

Even within a seemingly unified sector such as scientific publishing, it is sometimes necessary to accommodate multiple PIDs serving the same purpose, such as organizational identifiers. While in some ways accommodating multiple PIDs increases work and decreases interoperability, PIDs have different scope, attributes, and audiences. Some users prefer PIDs with ISO certification, while others prefer PIDs with established business models to ensure sustainability and maintenance, while others focus on ability to use without cost to access PIDs. When one PID has been selected for use by a stakeholder as part of master data management, being forced to accommodate a different PID can have significant costs and introduce unnecessary friction. Accordingly at CCC, we accommodate a variety of organizational IDs in RLSC and have long preferred the features of Ringgold for our primary use.<sup>1</sup>

### **Mapping metadata management across the research lifecycle.**

In late 2022, CCC and Media Growth Strategies undertook a thorough examination of metadata management across the research lifecycle. This review builds on an existing body of work to uncover multiple system complexities and breakages, which – separately and together – create missed opportunities for the communities for whom OA and open science models are designed to serve.

CCC has made this information publicly available in interactive infographic form at <https://www.copyright.com/stateofmetadata/>. Drawn directly from research interviews, the infographic depicts the significant economic impact that a fragmented metadata supply chain is having today on researchers, institutions, funders, and publishers. Researchers in particular shoulder a significant administrative burden that ultimately disrupts and delays the process of scientific discovery.

The infographic is a living document which will be updated and modified based on ongoing community feedback.

As the scholarly communications community continues its shift to OA and open science, stakeholders require a robust network of interoperable systems for making critical and necessary improvements, and much progress is underway. In that environment, a dedication to data stewardship across each stakeholder group, and the service providers supporting them, will lead to greater data sharing; reliable, trustworthy metrics on research impact; and a responsive, equitable rewards system. NASA can lead the way.

### **Question 4**

**Input on considerations to increase findability and transparency of research. NASA seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs (such**

---

<sup>1</sup> CCC adopted Ringgold as its preferred organizational PID approximately 8 years ago. CCC acquired Ringgold in 2022 so that we could ensure its continued viability given its broad adoption in scholarly communications and importance to ourselves and our clients.

**as ORCID) and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.**

First, we recommend that NASA review the [State of Scholarly Metadata: 2023](#) report referenced above which, among other things, summarizes where metadata breakages occur throughout the research lifecycle and how they impact various stakeholder groups.

As a research funder, NASA is well positioned to help research and lead by example by requiring PIDs at all appropriate points. As can be seen in the report above, grant application is one of the first organized parts of the lifecycle where PIDs can be effectively mandated. Once mandated and used, PIDs can flow throughout the lifecycle to improve use of research. We urge NASA to review the infographic, sign up for updates, and provide feedback should NASA believe there are amendments and changes needed.

**In addition, we have three specific recommendations with respect to mandated use of PIDs.**

- 1. NASA should mandate that grant applications include organizations IDs for the institutions(s) affiliated with each researcher listed on the grant application, and Funder Registry IDs for the distinct funders of the grant of applicable. The requirement should insist that grant applications include at least one of the following organizational identifiers used in the scholarly publishing ecosystem and NASA should make metadata fields available for all four:**

A. Ringgold- a proprietary global organization identifier system owned by CCC with over 600,000 unique records and rich hierarchical metadata used today by (1) most large and mid-sized commercial and non-commercial publishers, (2) a number of leading standards development organizations, and (3) a range of critical infrastructure providers in the publishing ecosystem. For publishers, Ringgold often is part of a master data management strategy. Ringgold is also used by some funders, academic institutions, and consortia. Ringgold maps one-to-one with ISNI and the Funder Registry.

B. ISNI- ISO standard name identifier system with 1,697,000 unique organizational records of which a minimum of 500,000 are relevant to the research sector. ISNI is free to use and has been adopted by many national libraries. It lacks the hierarchical metadata of Ringgold but enjoys the rigor and authority of ISO accreditation. The relevant organization records in ISNI map one-to-one with Ringgold.

C. ROR- Research Organization Registry (ROR) is a global, community-led registry of open persistent identifiers for research organizations. ROR is free to use and has been adopted by some publishers, institutions, and overseas funders. It contains 104,000 unique identifiers and some hierarchical metadata. It can map to ISNI and the Funder Registry, but not on a one-to-one basis.

D. Funder Registry (formerly known as FundRef) –Funder Registry is an open registry of grant-giving organization names and identifiers, with 32,000 unique identifiers for funders. It is donated by Elsevier to CrossRef and is updated approximately every 4-6 weeks. The Funder Registry ID can be used for author affiliations where the funder and affiliation are one and the same.

2. **NASA should mandate that grant applicants include one or both of the following individual identifiers for all researchers in grant applications, and NASA should make metadata fields available for both.**
  - a. ORCID- ORCID, which stands for Open Researcher and Contributor ID, is a global, not-for-profit organization sustained by fees from member organizations. ORCID is the most broadly adopted identifier system for individuals in scientific publishing.
  - b. ISNI- While not as well adopted as ORCID in research and science, ISNI has been broadly adopted in adjacent and non-adjacent fields.
3. **NASA should mandate that appropriate PIDs be used at each stage reporting, while remaining flexible as to which PIDs it mandates, and should reevaluate its mandated PIDs on an ongoing basis.** New PIDs such as RAiD (Research Activity Identifier) and DataCite (DOI-based system for research outputs) are being developed regularly and can help connect people, places and research. Appropriate PIDs should be mandated at each stage of the workflow, while recognizing that the needs of researchers and the availability of PIDs change over time.

**As a final recommendation,** we suggest that NASA register grants for DOIs. This will help enable connectivity of PIDs and the discoverability of the grants, maximizing return to US taxpayers.

Respectfully submitted for Copyright Clearance Center by,

*Roy S Kaufman*