Persistent Identifier Interest Group – PID-IG

Date: 2014/03/11
TAB Conditional Approval: March 2014.

Rationale: An effective environment for research and innovation would have researchers willingly and easily sharing research findings, and the community able to access findings and identify contributors for potential collaborations. At present, researchers do not have much incentive to share results, except in the form of written journal articles. And even then, tracing contributions to individuals has been hampered by lack of coherent standards for naming people and objects. To effectively address this issue will require that the research community coordinate efforts to adopt and manage interoperability of existing data exchange standards. At the same time, scientific data centers and infrastructures are put under significant pressure to manage an ever increasing amount of research data. The pervasive use of persistent identifiers is seen as a foundational element in anchoring and referencing such data in an interoperable way. Applications for persistent identification begin to cover a broader range of the data life cycle, which calls for a coordinating effort.

A primary goal of the interest group would be to engage the community to adopt common standards for computer-computer (API) and person-computer (UI) interfaces to support referencing and citation of research products, linkage of those products to the persons who created them and supporting efficient data management throughout the life cycle of research data. Such interfaces, if implemented, would decrease barriers to sharing, create incentives to sharing (e.g., acknowledgement through citation) and help to establish a common data fabric.

Goals and Activities: The purpose of the group is largely pragmatic, with two goals: (i) to meet on a regular basis to coordinate PID initiatives, leverage existing efforts, and reduce duplication of effort, and (ii) to determine when it is necessary to form a Working Group to discuss specific issues in more depth, determine prototypes for community dissemination, and standards recommendations. Some possible working group topics include determining plans for harmonizing organizational identifier structures, for supporting PID linkages to narratives, for validation of linkages between PIDs, providing PIDs for workflows and operations and harmonizing resolver usage. The group would meet quarterly, twice by teleconference and twice at the RDA plenaries, would be responsible for developing a PID session at each plenary.

Linkages with other RDA Groups: There are several RDA groups with intersecting or overlapping goals in the area of persistent identifiers. The PID-IG will coordinate with these groups to ensure adequate representation of findings and recommendations between efforts, and to inquire about joint presentations at RDA plenaries. In particular, the PID-IG will solicit the following groups for input and representatives.
>   Data Foundation and Terminology (DFT) Working Group
>   PID Information Types (PIT) Working Group (via Tobias Weigel)
>   Metadata Standards Directory Working Group (via EuroCRIS and CASRAI)

Membership: The group will include organizations that provide PID services to the international research community and organizations that implement or otherwise use PIDs. Tobias Weigel, DKRZ and PID Information Types RDA WG co-chair; Laure Haak, Executive Director, ORCID; and Adrian Burton,

Director, Australian National Data Service (ANDS) have agreed to serve as co-chairs of the group. Membership is open to the research community. Parties that have stated interest so far include:

- David Baker, Executive Director, CASRAI
- Paolo Bouquet, President, OKKAM
- Jan Brase, DataCite and International Council of Scientific and Technical Information
- Laura Dawson, Product Manager, Identifier Services, Bowker/ISNI
- John Kunze, California Digital Library
- Larry Lannom, Director of Information Services and Vice President, CNRI
- Maurizio Lunghi, APARSEN
- Maryanne Martone, Force 11
- Reagan Moore, UNC Chapel Hill / RENCI
- Andreas Rauber, TU Vienna
- Ulrich Schwardmann, EPIC/GWDG
- Ralph Shoffner, CEO, RingGold
- Ed Simons, President, euroCRIS

Working Groups. The PID IG will foster working groups to address specific topics, and in particular to address and define emerging PID use cases in the domain of data, and whether the research community would benefit from a global open identifiers for persons, data objects, etc. Based on a survey of participants and other interested parties, carried out in February/March 2014, the overall most pressing concern is the need for interoperability across different PID systems. This also includes interoperability of metadata to be exchanged horizontally and vertically. A second important concern is to develop policies for using identifiers, tailored to different use cases and application areas. Where applicable the PID-IG will collaborate with existing RDA working groups, or it will establish new working groups based on priority areas as identified at RDA plenaries.

Communications: The PID-IG will use RDA-supplied group email and document repository tools.

- Group Email: rda-pid-ig@rda-groups.org
- Group Document Repository: https://www.rd-alliance.org/filedepot/folder/132?fid=359

Related publications
- Interagency Data Stewardship/Identifiers Wiki page.
  http://wiki.esipfed.org/index.php/Interagency_Data_Stewardship/Identifiers
- Duerr, et al. (2011) On the utility of identification schemes for digital earth science data: An assessment and recommendations. Earth Science Informatics 4(3):139-160.
  http://link.springer.com/article/10.1007%2Fs12145-011-0083-6
- DFT Working Group models and data structures: https://www.rd-alliance.org/filedepot/folder/132?fid=99
- A Data Oriented (and Beyond) Network Architecture (2007).
  http://www.cs.berkeley.edu/~istoica/papers/2007/dona-sigcomm07.pdf
- Force 11: Joint Declaration of Data Citation Principles (2014). http://www.force11.org/datacitation