

For a system based on unique researcher identifiers, all parties must trust that the identifier system is reliable, authentic, and immutable. Accordingly, the development of systems that are unable to generate a significant level of trust is “likely to limit and fragment any effort to coordinate, federate, or integrate differing identity solutions in the research space. Therefore interoperability of any developed system with the wider web must be a prime consideration” (Neylon 2010b). One possible method for establishing trust and ensuring proper attribution is encouraging the use of an identification system like ORCID (see Exemplar), and establishing publishing practices that make the “creation and capture [of unique identifiers] an integral part of the editorial process” (Trasande and Hannay 2010).

3.2. Technical Issues

The development of a persistent, trusted, ubiquitous, and interoperable centralized repository for housing the unique researcher identifiers may provide a “trusted broker” for promoting knowledge access and attribution (Trasande and Hannay 2010; Neylon 2010b). Currently, a number of identification tools exist or are under development, including ORCID, Vivo, Lattes Platform, Public Library of Science (PLoS), and PubMedCentral. For example, ORCID proposes to create a “central registry of unique identifiers for individual researchers and an open and transparent linking mechanism between ORCID and other current author ID schemes” (<http://www.orcid.org/>). Lattes, a fully developed researcher database that allows for verification (Aragão 2010; Lane 2010) has now been adopted in 17 countries in Latin America, Europe, and Africa [Aragão].

3.3. Role of Funding Agencies

Funding agencies worldwide can play a critical role in encouraging knowledge access and the implementation of an identification system to facilitate attribution. Agencies are uniquely positioned to require data and code sharing in publicly funded work, and they support the infrastructure and tools for data and code sharing. Participants felt that funding agencies should embrace the creation of identification systems and ensure their adoption by requiring registration as a prerequisite to applying for agency funding (Trasande and Hannay 2010). Participants also thought that agencies could support a research library coalition that would provide an international open-standard data set for bibliometric information for all published work worldwide (Conlon 2010).

4. Conclusions and Next Steps

At the conclusion of the workshop, participants agreed on a set of attributes of the “ideal” attribution landscape 5 years into the future [Greer]. It would include a framework of openness and international standards for data and knowledge; reliable and unique identifiers for each researcher, organization, publication, and the

relationship to each other; a link between all publications and their appropriate data; continuous investment for data preservation and access; and formal and informal training of students, researchers, and personnel at funding agencies.

In white papers submitted before the workshop, and during presentations and in discussions at the workshop itself, participants identified a set of actions that would achieve this vision:

- (1) Establish a system of persistent identifiers for both researchers and their outputs. The following specific suggestions were made about the characteristics of such a system:
 - Create taxonomies of scientific data—Enable the cataloguing, tagging, and parsing of data sets for automated recall.
 - Create incentives—Encourage and offer incentives to researchers to routinely use the standardized identification schemes to annotate their data, a process that will be aided by the further development of software tools. Provide researchers with incentives to encourage them to make data sets available to the wider research community through the development and use of attribution systems. Help ensure that data sets are linked to subsequent publications and other research outputs, further aiding attribution and the reproducibility of research. Publish data and code to facilitate assessment and certification of quality and allow data sets to become part of the citable “scientific record.”
 - Create independent standards—Establish federally funded platforms for data and code sharing that are independent of institutions and individual researchers, and use standards of unique identification for citation and version control.
 - Create a legal framework—Develop an Open Research License (ORL) to resolve conflicts between reproducibility and copyright law.
 - Create a registration mechanism—Encourage the development, implementation, and use of standardized identification systems to facilitate attribution by requiring system registration as a prerequisite to applying for agency funding.
- (2) Develop national and international pilot projects that compare different technical solutions for developing and maintaining open data platforms, fostering the replication of scientific research, and ensuring attribution for the intellectual contributions of researchers.

- (3) Foster formal and informal training to ensure that open data and knowledge systems are maintained.

Workshop participants agreed that engaging in these efforts will provide opportunities to work across counterpart funding agencies to encourage international cooperation and the dissemination of knowledge and data.