ISQ

INFORMATION STANDARDS QUARTERLY

TOPIC: LINKED DATA FOR LIBRARIES, ARCHIVES, AND MUSEUMS
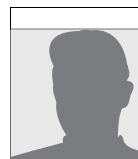
LINKED DATA VOCABULARY MANAGEMENT

EUROPEANA MOVING TO LINKED OPEN DATA

LINKING LIVES END-USER INTERFACE

JOINING THE LINKED DATA CLOUD

OCLC'S LINKED DATA INITIATIVE

LC'S BIBLIOGRAPHIC FRAMEWORK INITIATIVE

NISO
How the information world CONNECTS

CR [CONFERENCE REPORT]

Thomas Elliott

Sebastian Heath

John Muccigrosso

THOMAS ELLIOTT, SEBASTIAN HEATH, AND JOHN MUCCIGROSSO

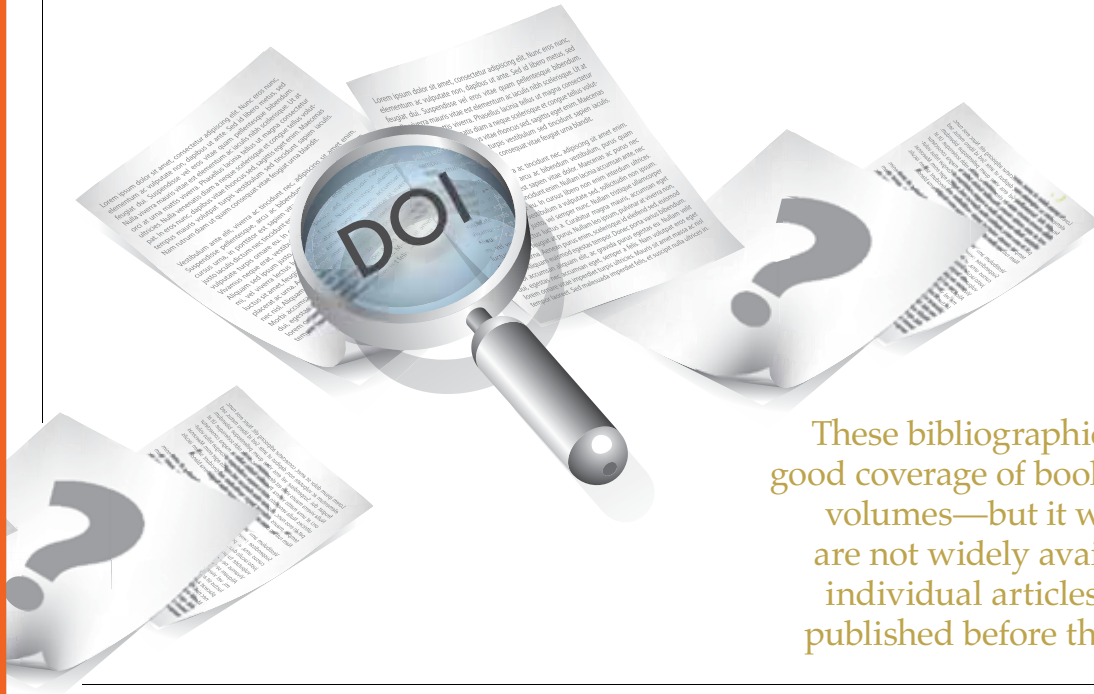# Report on the Linked Ancient World Data Institute

From May 31 to June 2nd, 2012, the Institute for the Study of the Ancient World at New York University hosted the *Linked Ancient World Data Institute* (LAWDI), an internationally attended workshop funded by the National Endowment for Humanities' Office of Digital Humanities (Grant number: HT5004811). This three-day event mixed longer presentations by invited speakers with presentations by twenty applicants who had submitted statements of interest on why their work would benefit from intensive interaction with colleagues also pursuing digital publication of scholarly resources on the public internet. This was the first of two LAWDI sessions, with the second to be held at Drew University from May 30 to June 1, 2013.

LAWDI's intellectual scope is the ancient Greek and Roman Mediterranean and the Ancient Near East. Taken as both geographic and chronological markers, these terms encompass modern academic disciplines that have long histories of creating digital resources, many of which are already accessible via HTML-based websites. And while it is too optimistic to say that these disciplines have always maximized the potential of interdisciplinary work, there is a continuity of cultural development and a degree of contact that gives many commonalities to the study of the early civilizations of Mesopotamia, the subsequent eras of Greek and Roman cultural prominence, and the ongoing reworking of ancient precedent by later Byzantine and Syriac societies. Accordingly, one premise of LAWDI is that publication of well-structured and reusable digital resources will benefit all scholars, as well as the interested public who are working in these areas.

In addition to being academically inclusive, LAWDI also took a very open approach to the concept of Linked Open Data. Most of the attendees at LAWDI were experts in the content and methods of their respective disciplines, rather than in the technical details of web architecture. Additionally, many present were also included museum professionals, librarians, and archivists who curate so-called ancient world data, such as archaeological fieldwork archives and bibliographic resources. Again, there was no expectation that participants came with experience in implementing

These bibliographic resources provide good coverage of books—that is physical volumes—but it was noted that there are not widely available identifiers for individual articles, particularly those published before the adoption of DOIs.

Linked Open Data so that the organizers recognized that a three-day workshop was not enough time to develop complete technical proficiency.

Accordingly, LAWDI began with a focus on two aspects of current best practices.

1. Early sessions stressed the importance of establishing stable URIs that allow fine-grained access to scholarly resources. Examples of current work included the URIs that Pleiades is establishing for ancient geographical entities and URIs for numismatic concepts established by Nomisma.org.

2. Presenters stressed that progress going forward depends on high-quality, automatically parsable data being available when those URIs are de-referenced. Of course, RDF can play a role here, and attendees were introduced to basic concepts as "triples" and "things, not strings." But there was also discussion of RDFa, JSON, KML, and Atom as reasonable formats that allow machine-based reuse of ancient world data. As an example of such reuse, more than one presentation discussed the Pelagios Project, which is aggregating references to Pleiades URIs via the Open Annotation Collaboration RDF vocabulary—and is currently one of the best examples of the potential for Linked Open Data to enable new forms of discovery of scholarly resources.

In particular, the overlap between geographic named entities and discovery of ancient textual sources that refer to those entities is being pushed forward by the participation of the Perseus Digital Library and Google Ancient Places in the Pelagius consortium.

While participant presentations were spread throughout the three-day program, it quickly became apparent that many of the applicants came to LAWDI with very basic questions, all of which were very welcome. In general, those of us working on the digitization of the Ancient World recognize the importance of reusing existing vocabularies. But we also recognize that it is very easy to push the limits of what generic vocabularies such as the Dublin Core allow us to communicate. For example, does dcterms:creator refer to the webpage being de-referenced at a URI or to the ancient artist who created the object being described by the document found there? Such a question is recognizable as falling under the rubric of "HTTP Issue 14"—now open as "Issue57: Mechanisms for obtaining information about the meaning of a given URI"—and the invited speakers strove to highlight such issues and to illustrate both the "fragment identifier" and "303 redirect" mechanisms for solving them. It is not clear that the expertise to implement such solutions is widely available in the academic computing environments that were typically available to LAWDI participants.

Discussion during the workshop also highlighted the availability of bibliographic linked data as a pressing need for scholarly initiatives. The Library of Congress' id.loc.gov server, OCLC's VIAF service, and the resources OCLC makes available via WorldCat were all highlighted. These resources provide good coverage of books—that is physical volumes— but it was noted that there are not widely available identifiers for individual articles, particularly those published before the adoption of DOIs. While JSTOR does provide retrospective
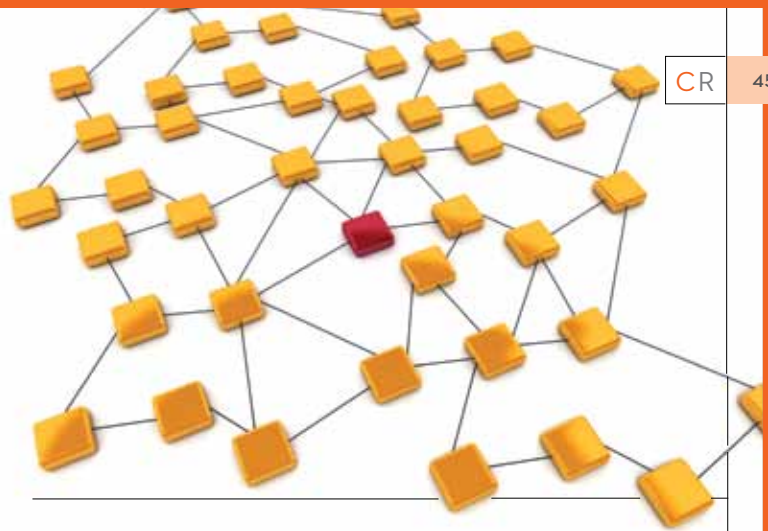
identities, it is far from comprehensive and is not open to the general public. The fact that humanities scholarship refers to work that can be decades and even centuries old was a theme of the bibliographic discussion at LAWDI.

One premise of LAWDI was that face-to-face interaction would lead to greater interlinking and reuse of digital resources in the future. LAWDI presenters frequently made the point that pointing to a stable digital resource is a form of endorsement that encourages yet more reuse of that same resource. This in turn can lead to interoperability of data as common identifiers form a basis for linking together disparate work. Essentially all LAWDI participants were eager to show resources that provide stable URIs or to ask for advice on what is currently available. But both the participants in and organizers of LAWDI recognize the need to take active steps to grow the number of high-quality digital resources. That will require ongoing outreach as well as clear examples of how Linked Open Data benefits both creators and users. As we plan for the 2013 session, it will be important to pay attention to tools that make it easier to take the first steps into publishing Linked Open Data. These tools may include cloud-based services such as GitHub.com, Nodester.com, and database hosting services such as Mongohq.com or Mongolabs.com. Likewise, technical developments such as RDFa and JSON-LD should increasingly take on their intended role as lower-cost entry points for Linked Data-based projects.

Basic information about LAWDI is available at the event's website, which is hosted by *The Digital Classicist*, a decentralized community that supports digital initiatives within Classical Studies. The page *LAWDI 2012 Websites* shows the very wide range of projects that participants are working on, some of which already implement Linked Open Data principles. The *LAWDI 2012 Documents Presentations* webpage has links to the extensive Twitter feed of the event, to a selection of the presenters' slides, and to blog posts that both preceded and followed LAWDI's three days of intensive conversation. These documents do much to capture the spirit of the event and to suggest ways that Linked Data will continue to transform research within the study of the Ancient World.

THOMAS ELLIOTT (tom.elliott@nyu.edu) is Associate Director, Digital Programs and Senior Research Scholar, and **SEBASTIAN HEATH** (sebastian.heath@nyu.edu) is a Research Assistant Professor, both with the Institute for the Study of the Ancient World at New York University. **JOHN MUCCIGROSSO** (jmuccigr@gmail.com) is Associate Dean & Professor, Department of Classics, Drew University.

> LAWDI presenters frequently made the point that pointing to a stable digital resource is a form of endorsement that encourages yet more reuse of that same resource. This in turn can lead to interoperability of data as common identifiers form a basis for linking together disparate work.

Google Ancient Places
googleancientplaces.wordpress.com/

Library of Congress Linked Data Service
id.loc.gov/

Linked Ancient World Data Institute
wiki.digitalclassicist.org/Linked_Ancient_World_Data_Institute

Open Annotation Collaboration
www.openannotation.org/

Pelagios
pelagios-project.blogspot.com/

Perseus Digital Library
www.perseus.tufts.edu/

Pleiades
pleiades.stoa.org

Virtual International Authority File (VIAF)
viaf.org/

WorldCat
www.worldcat.org

RELEVANT LINKS