MESUR: implications of usage-based evaluations of scholarly status for open repositories.

Johan Bollen, Herbert Van de Sompel
jbollen@lanl.gov , herbertv@lanl.gov ,
Digital Library Research & Prototyping Team
Research Library
Los Alamos National Laboratory

The MESUR project at the Digital Library Research and Prototyping team of the Los Alamos National Laboratory's was funded by the Andrew W. Mellon Foundation to investigate the possibilities of usage-based scholarly evaluation metrics. MESUR's investigation is conducted on the basis of a large-scale reference data set that combines usage, citation and bibliographic data obtained from a wide variety of publishers, aggregators and institutions (Figure 1). At this point the MESUR database contains information on more than 1 billion usage events, and all associated bibliographic and citation data. Preliminary investigations have yielded a comprehensive survey and comparison of citation- and usage-based metrics of scholarly status leading to detailed information on how they can be used to indicate different facets of scholarly status. The MESUR project's final 2008 deliverable is a set of guidelines and specifications for the application of usage-based scholarly evaluation.
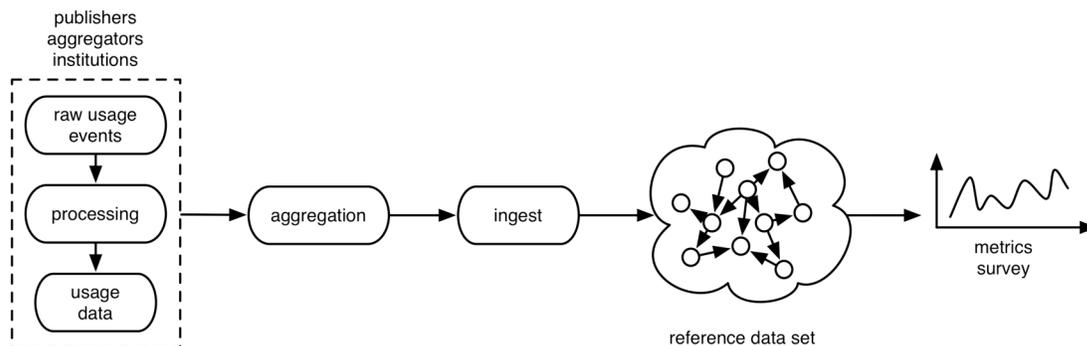

Figure 1: Overview of MESUR project work plan.

Scholarly evaluation is at this point mostly conducted on the basis of metrics based on citation statistics, e.g. article citation counts, the Impact Factor and the h-index. Although citation data, possibly due to its origin in the published, peer-reviewed literature, has acquired an aura of validity and reliability, a number of shortcomings have been repeatedly identified that affect open repositories. Publication delays can cause citation data to lag scholarly developments by significant periods of time leading to an under appreciation of services that shorten the scholarly communication cycle. Most importantly, citation data frequently ignores the growing body of grey literature or non

textual scholarly objects that exist outside the realm of scholarly journals and can thus not assess the growing impact of scholars that publish their research outside the limited set of journals tracked by Thomson Scientific or any particular commercial publisher collections.

It has become apparent that usage data, in particular when aggregated across a multitude of scholarly information sources, is a promising complement to traditional citation data. This is particularly true for open repositories. Usage data can be recorded for a wide variety of resources and user communities. It is subject to only minor publication delays as it can be recorded as soon as the resource in question becomes available rather than when the next generation of citing articles is published. In addition, compared to citation data, usage data generally contains extensive contextual data on user motivations and intentions. Property rights can generally be exerted by the institutional entity that records the usage data. Usage data can thus open up the domain of scholarly evaluation that so far has largely bypassed non-traditional modes of publication.

In spite of the many possible advantages of usage data, serious issues need to be resolved for it to find the same level of acceptance and trust as citation data has achieved over the past 30 years. The MESUR project attempts to address these issues. First, how do the communities and services for which usage data has been recorded, i.e. its sample, affects consequent impact evaluations and how can sampling problems be addressed by wide-scale aggregation? How can such aggregation be effectively instrumented and conducted? Second, given that adequate and reliable samples can be achieved that include non-traditional publication models and services, which metrics best express the various modalities of their scholarly impact, e.g. status, trust, popularity, educational impact, etc (Fig. 2)? Third, how can usage data best be combined with other sources of information to arrive at more reliable models of evaluation?

This presentation will discuss the main architecture, objectives and preliminary results of the MESUR project. Attention will be paid to MESUR's approach to usage data acquisition, aggregation and ingestion. We will discuss the various problems that were encountered by MESUR to achieve a representative sample of the scholarly community and the solutions that were adopted. We will show preliminary results that outline how detailed surveys of citation-, usage and hybrid-based metrics of scholarly impact can be conducted to assess the various modalities of scholarly impact. In particular, attention will be paid to the implications of this novel approach to scholarly evaluation for the open repository movement and future trends that this community will be subject to.
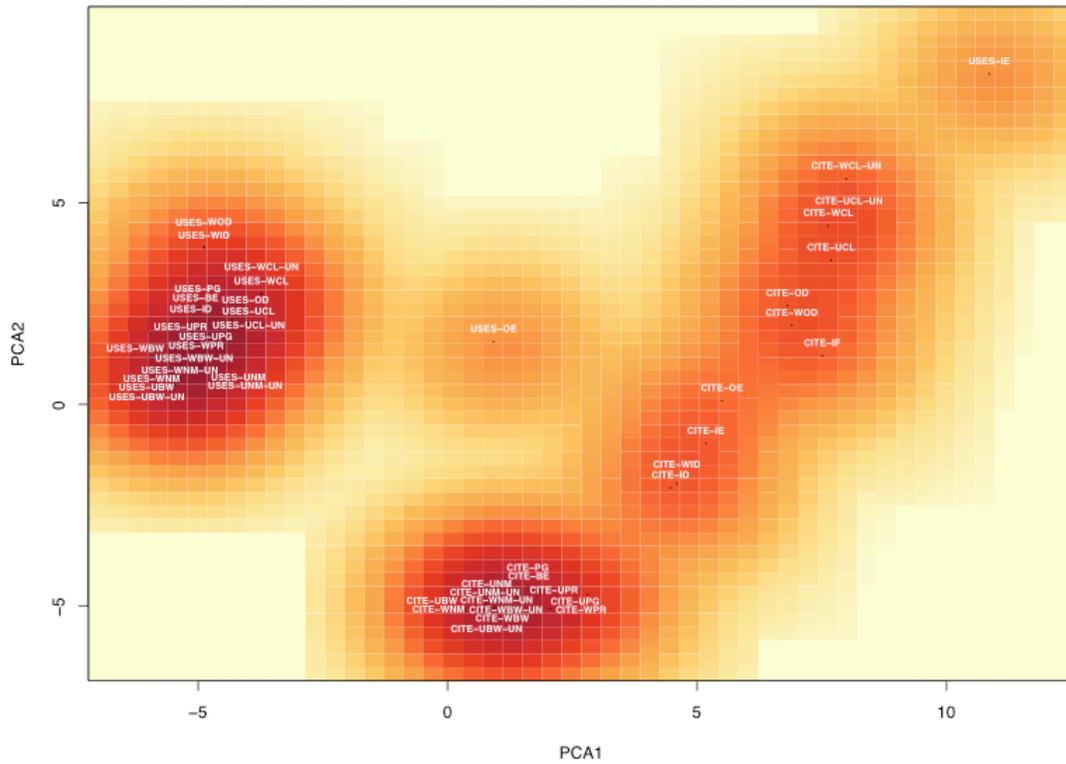
Figure 2: Survey of similarities and dissimilarities of various usage- and citation-based metrics of impact.