

Institutional identifiers and the Journal Supply Chain Efficiency Improvement Pilot

Based on a briefing session given at the 30th UKSG Conference, Warwick, April 2007

Institutional identifiers are nothing new. The problem in the e-content world is that there is no single schema that covers all the identifiers needed. Previous identifiers have been for physical location purposes which are no longer relevant for new groupings. Publishers now need to know exactly who their customers are and the customers need to be able to identify themselves and their 'licensing unit' to the publishers and other members of the journal supply chain. A Journal Supply Chain Pilot including the major participants in the supply chain started in 2006 and is continuing to look at the identifiers, metadata and transactions that take place in the supply chain.



HELEN HENDERSON
Ringgold Ltd

Introduction

Institutional identifiers have been discussed in the library information industry for more than 20 years. Pat Harris, formerly of the National Information Standards Organization (NISO), remembers when this was first raised at a MARC standards meeting in the 1970s. It is not that institutional identifiers do not exist, there is a plethora of them, but each has its own small constituency for specific business applications. Most identifiers, such as the standard address number (SAN) or global location number (GLN), are designed to deliver physical objects to known locations. Similarly, MARC and NUC codes expedite inter-library lending. The International Standard Identifier for Libraries and Related Organizations (ISIL) would also expedite physical deliveries. Other financial or business identifiers would include the D-U-N-S number, tax identifiers or authorization codes. If you want to see an example of what a typical academic organization has to quote in terms of their identifiers, go to the Medical College of Georgia website and you will see:

Institutional Identifiers:

2. What is our DUNS number?
3. What is MCGRI's Entity Identification Number, IRS Employer Identification Number (Tax ID) and Organization Type?
4. What is MCG's Entity Identification Number, IRS Employer identification Number (Tax ID) and Organization Type?
5. What is MCG's Institutional Profile Number?
6. What is MCG's Congressional District?
7. What is our Animal Assurance Number?
8. What is our Institutional Human Subjects Assurance of Compliance number?
9. What is our institution's Radioactive License number?
10. What address should grant and contract checks be mailed to?

Each of these identifiers has a specific purpose, but there is no one schema which covers all types of institutions which use information resources, and with whom publishers have to deal on a daily basis. Delivery of content was not much of an issue or problem in the past because in the print world mailing labels pretty much guaranteed delivery (although sometimes materials arrived in poor shape!). In the electronic world, with complex licensing negotiations, e-content delivery is a different matter. An identifier is needed for every 'licensing unit', which you can consider as being the equivalent of the 'delivery address'. This identifier, therefore, would potentially be used for

licensing, marketing, customer analysis, authorization and authentication, depending on the individual needs of the participant in the supply chain of journal materials. The most important use for such an identifier for the publisher (or content provider) and the user is authorization: who is authorized to access the content?

Grouping customers

In this day of purchasing e-content by academic consortia, knowing who is included in any grouping is important for both pricing and access. Some consortia are opt-in, some are opt-out. Publishers and librarians now agree that transparency and the sharing of information about entitlements is essential for customers to get access to the content they have paid for. This applies equally to large corporations and governmental units who wish to provide access for certain divisions, which need to be separately identified; not an easy task when there is a single IP address, or address range, for the global organization. In particular these corporations are also constantly merging, acquiring, changing names, changing domains and, therefore, are a real problem for publishers to track. Government departments now license on behalf of their agencies and, as everyone knows, they too are given to reorganizing. For example, the UK National Health Service has changed its structure three times in the past three years, contributing to a huge amount of work to stabilize and ensure access rights.

Publisher benefits

In 2002 Oxford University Press decided they needed to better identify and classify their customers, timing this decision to coincide with a move to a new fulfilment system. Information Power – now part of Ringgold – was selected as the consulting firm to carry out this work and create a master database of institutions which subscribed to Oxford journals. As other publishers picked up on this example, it was clear that a master database of organizations which subscribed to academic journals would be useful not only for grouping, but also for publishers' marketing and customer support.

The type of work publishers can perform with the Ringgold 'Identify' database includes studies of the overlap between the subscribers of primary publishers and aggregators. Perhaps most publishers are really interested in looking at market penetration levels and doing gap analysis to identify prospects. Supporting such work in terms of 'best practices', Ringgold is compiling standardized reports for those publishers who subscribe to its web service. Because industry standards are being followed for usage reporting, it is possible to merge reports incorporating institutional data from all channels, including Ringgold's Identify.

There is a need to be able to authoritatively identify the institutions in different reports, perhaps generating new licensing and pricing opportunities. It also allows a publisher to have an overall institutional view; publishers typically have different levels of relationships with their customers, even holding information about authors, editors and reviewers in different databases. The ability to pull all this information together with the subscription information gives the 'big picture' view of all institutional relationships.

With the increase in open access journals, knowing if an author is part of a subscribing institution is important for pricing page charges. With title changes, subscription agent changes, and even institutional reorganization, subscription lapses are counted as new, unidentified subscriptions, which causes an interruption in the supply, even with gracing periods.

Customer benefits

The benefits of identifying institutions are not limited to publishers; the potential benefits for their customers are huge. If subscribers can identify themselves easily and more accurately to their providers, it will ensure continued supply of individual subscriptions by eliminating incorrect reporting of a subscription renewal as first a cancellation and then as a 'new order'.

The existing electronic resource management (ERM) services (e.g.: Ex Libris, Serials Solutions) perform important services, particularly for libraries, by their consolidation of lists of serials for libraries, incorporating both e-content and print subscriptions and populating their link resolvers.

As publishers and agents gain experience with improved institutional identification, one could expect them to be providing more accurate information to their customers, directly as well as to ERM services. Libraries and other institutional subscribers can drive this effort forward by asking agents and publishers to deploy local numbers in their fulfilment service activities.

Greater use of content for which one has already paid benefits everyone. The ability to provide authoritative lists of subscriptions by publisher will be mutually important when negotiating a 'big deal', e.g., when considering changing the supply chain for subscriptions to journals/journal bundles.

Institutional registry

There has been discussion about the need for a central registry for libraries in which to lodge information about their institutions to encompass supply to all of its information providers where today such information must be provided singly to every supplier. Such a new paradigm would allow existing 'many- to-many' relationships to become 'one-to-one' relationships. OCLC is developing the WorldCAT registry for libraries and we can hope that it will be used for this purpose, at least in the public and academic library arena.

Authentication is changing and IP addressing will continue for some time, but managed federations like Athens and Shibboleth are being implemented and it is even more important that institutions can be identified within such a federation.

Journal Supply Chain Pilot

These initiatives are all happening in a piecemeal way, so a group of organizations got together in late 2005 and decided to see if it was possible to take the journal supply chain, map it and look at the transactions that take place and how they could be simplified and improved. The imperative for this was the increasing complexity of work flows in the electronic era and the increasing number of participants. Instead of the old print supply chain being managed by agents, there are now distributors, hosting services, authentication services, fulfilment systems, ERM vendors and

aggregators to be taken into account. Publishers have more direct contact with customers and there is an increase in large bulk negotiations. Institutions are combining into many types of grouping, not just library consortia, and may be members of many different purchasing groups.

This group formalized itself into the Journal Supply Chain Efficiency Improvement Pilot (JSCEI) with the initial participants being: the British Library; HighWire Press; HighWire Publishers; Ringgold; Swets and UK Libraries. There were nine work packages and these covered auditing UK subscriber information for the participants, mapping to Ringgold's Identify database, mapping transactions between participants, and evaluating the effort needed to standardize transactions. The benefits are being evaluated and other participants are joining in different initiatives. Six work packages are completed and an Interim Report has been produced.

One of the main issues that came up was the level of granularity that should be used for an institution. Different participants identify different levels within an institution, from individual, through departments, countries and campuses. For the purpose of managing printed content, it was important to identify the physical location for delivery. For electronic content, it is the virtual institution identified by authentication and authorization levels that is important for identification.

Another issue was how the many systems used to exchange information would handle an institutional identifier. At present there will be different identifiers used by agents, publishers and hosting services, and to standardize the identifier and the exchange format is crucial. The International Committee for EDI in Serials (ICEDIS) currently works on message formats for information exchange and is actively involved in the Pilot. A new standard XML message for activation transactions is being developed. NISO's new Business Information Topic Committee is looking to set up a working group on institutional identifiers.

Finally, there is an issue about an appropriate business model to support any standards. In the case of an identifier there is a database to be maintained, which is an expensive process without government or other subsidy. This means that a revenue source has to be identified. Up to now the publishers have provided this revenue for value-added services they receive from Ringgold, but

one of the questions going forward is stability and sustainability. This is one of the issues that the Journal Supply Chain Pilot group is still discussing and the next steps include the expansion of the test database outside the UK, and the inclusion of other participants.

All of us, as participants in supply chain work in one way or another, are also looking at analysing the metadata for describing institutions, to see how it can be split into the different elements needed for different activities, and the levels of security needed for the different elements. The

mission of Ringgold is to make the market for journal supply more efficient, thereby increasing revenues for suppliers of all types and improving access to end-users.

Article © Helen Henderson

■ Helen Henderson
Managing Director
Ringgold Ltd
E-mail: helen@ringgold.com

To view the original copy of this article, published in *Serials*, the journal of the UKSG, click here:

<http://serials.uksg.org/openurl.asp?genre=article&issn=0953-0460&volume=20&issue=3&spage=180>

For a link to the table of contents for the issue of *Serials* in which this article first appeared, click here:

<http://serials.uksg.org/openurl.asp?genre=issue&issn=0953-0460&volume=20&issue=3>