

Expert Advisory Group on Institution Specific Data Meeting

Minutes

Wednesday 12 December 2007
Libraries Australia Meeting Room
National Library of Australia

Attendees: Sue Clarke (Monash University); Stephen Dunne (TAFE South Australia); Paul Pretor (University of Ballarat); Wan Wong (National Library of Australia); Rob Walls (Database Services, Libraries Australia) (Chair), Bernal Rajapatirana (Database Services, Libraries Australia)

Russell Varney (State Library of Queensland) via teleconference

Minutes: Lenore Kennedy (Database Services, Libraries Australia)

1. Welcome and introductions

The meeting opened at 10.30 a.m.

Rob welcomed the members to the first meeting of the new Expert Advisory Group.

2. Review the terms of reference of the Expert Advisory Group

Rob opened the discussion by noting that the Libraries Australia Advisory Committee (LAAC) had already endorsed a paper prepared by the National Library of Australia (NLA) requesting that Libraries Australia accept certain of their ISD tags so that the Libraries Australia database could be used as a surrogate public catalogue for the NLA or other libraries. In this context, while it was within the remit of the EAG to make recommendations about the desirability or otherwise of adding ISD to Libraries Australia, the EAG should work with the assumption that ISD will be added and therefore should also make recommendations regarding what institution specific data elements should be supported and how ISD would be accommodated within the system.

Sue suggested the group should take some time to discuss the business model for LA acceptance of ISD and the benefits such a step might or might not bring to users other than NLA. Sue was concerned that the allocation of resources to the implementation of ISD would divert resources from other enhancements that could benefit a larger number of users.

The EAG's draft report will be sent to LA users for comment, the feedback would be discussed by the EAG either at a meeting or by teleconference and then the final report would be sent to the LAAC. This process would provide an opportunity to gauge the level of interest in ISD amongst LA users.

Some reasons why the proposal had been endorsed were the interest of NLA (LA's largest user) and the benefits of using LA to syndicate Australian library data to other services including Google, Yahoo and WorldCat. If implemented the proposal could provide better support for the ISD that is already in Libraries Australia (e.g. URLs) and better support for the exchange of holdings data. Paul asked whether WorldCat Local could be used instead of Libraries Australia to support the public catalogue function required by the NLA.

3. Should Libraries Australia support additional ISD data elements?

Discussion turned to the type of ISD that could be shared in the database, from NLA's local subject headings and series statements to ISD URLs that are currently accommodated in the holdings area in the database architecture but output as shareable tags in MARC records. This led on to a discussion of the M21 Holdings format and whether implementation of this should be the first, or only, concession to hosting ISD in the database. LA will need to continue to support the current method of sending and receiving holdings information in the bibliographic 984/850 tag for the foreseeable future even if the M21 Holdings format was adopted.

Paul made the observation that there was a difference between simple ISD that could be accommodated within the current bibliographic format and more complex ISD that would require the implementation of M21 Holdings and other system enhancements. It was up to individual libraries to decide what sort of data they might want to contribute and up to the group to formulate guidelines on the type of data that belongs in each format.

Rob said that bibliographic metadata was increasingly being created by non-library organisations and by library users, for example annotation data, user reviews, tag clouds, etc. He suggested that investigation of schema to exchange this data could form part of the EAG discussion if it was thought necessary. The point was **noted**.

The group considered a report from the Victorian Libraries Australia User Group meeting that had been provided by Paul. Paul explained that the meeting had shown there was a tension between the opportunity provided by the uploading of ISD and the investment of library resources required to create the data in a standard format.

The discussion around the Victorian User Group report centred on the critical mass required for a site to have an impact on the Web. It is increasingly likely that the initial point of access for the users of any given library to that library's resources will be a web search engine or other web site rather than the library's own web catalogue. It is therefore advantageous if data such as ISD URLs are included in records supplied to Libraries Australia, and syndicated by them to other web services such as WorldCat and Google, because it can lead a library's own users back to that library's own resources.

4. Discussion of data elements that could be supported

Rob said that implementation of ISD would require a major enhancement of the Search Service and significant reconfiguration of the Cataloguing Service.

Wan explained the ISD fields that NLA would like implemented. Currently Australia-centric subjects on Music Australia records use NLA's own music thesaurus and local subdivision practice and are coded as 690 and 691; these tags are converted to non-specific 653s on import to LA and so lose their subdivision structure. Geographic subject headings used for Australian maps data use a local thesaurus and local period subdivision. NLA has suggested use of regular MARC subject tags with Indicator2=4 for the exchange of this data with LA. NLA creates local series headings, for example for formed collections; these are currently tagged 830 and 930, with the 930 tags not being sent to LA.

Paul expressed concern about the standing of LA with the MARC Standards office if we wanted to implement a non-standard approach to ISD. Rob replied that LA could set an exchange standard within Australia without seeking the imprimatur of the international library community; however, the EAG should consider whether user requirements could be met if LA was to fully exploit the ISD provisions that are already present in the M21 format. Bernal commented that the expanded use of the \$5 subfield proposed in the background paper (i.e. use of subfield \$5 for most M21 bibliographic fields) would only need approval if the data were to be exchanged internationally. Bernal suggested that ISD should not be supported for the tags required by the Libraries Australia minimum record standard.

The group expressed support for basing the LA ISD model on the M21 format. The group **noted** a point for further investigation that was raised by Bernal: that there is no problem in using the existing \$5 subfields within the LA network, but syndicated data would require international NUC symbols in the \$5 in order to identify the data owner when records are viewed outside Australia.

The discussion now focused on the principles at Section 6 of the Background Paper: using the M21 format would "ensure that the chosen model provides sufficient specificity in tagging and functionality ..."

To ensure that "the chosen model supports identification of the contributing organisation" the chosen model would need to use the \$5 convention in order to facilitate the tailoring of displays and the export of data to the appropriate institutions. The contributing organisation would not necessarily have to supply the \$5 subfield - it might be possible to add this through the Record Import Service (RIS) rules.

In reply to a question from Stephen, Rob explained that \$5 data might or might not be visible to catalogue users depending on the implementation. It would be possible to mask the \$5 data from some views in both the web and the client. Stephen suggested that the option to view \$5 data or not would be required by both the end user and the data provider - does the user want to see it versus does the library want it to be displayed. Rob replied that this functionality could be linked to user profiles. The minimum development would be to link the display of ISD to library authentication.

The "the impact of ISD on the Authority File" was then discussed. The group was asked whether LA should support some ISD in authority records, or perhaps support the contribution of completely institution specific authorities. Bernal was keen to support as a minimum the storage of local system numbers in authority records to facilitate their exchange between LA and contributing libraries. Bernal also described

how the entries in People Australia would contain data from many sources which was, in effect, ISD, such as links to digital objects within particular collections.

Paul made the observation that ISD authorities are likely to be uncontrolled and uncontrollable. He emphasised that libraries faced an issue in terms of resources that they could afford to commit to maintaining their authority records and the stability of any ISD contributed to the NBD. Wan explained that the NLA's ISD authorities belong to a controlled in-house thesaurus that is maintained. Rob suggested that having institution specific authorities would add to the complexity of the initial implementation. Sue commented that the contribution of ISD authorities would not be a priority for Monash. The concept was **noted** as a possible development for Stage 2.

Rob also asked that it be **noted** that any library contributing ISD (including holdings data and URLs of electronic resources) is encouraged to keep it up-to-date for the benefit of Libraries Australia users.

5. Discussion of exchange formats for ISD

Rob explained to the group that the exchange format for ISD would not necessarily have to be the same as the format used in the online cataloguing interface; he used the current manipulation of ISD URLs as an example of this.

Sue argued for implementing only the standard M21 support for ISD - \$5 subfield in the allowed fields and 6XX ind2=4. Rob wondered if sticking with standard M21 might mean LA could miss out on valuable data in fields where \$5 is not supported.

Paul agreed that the list of M21 tags was limited and suggested that the M21 Holdings format could supply the shortfall. Paul suggested the group look at groupings of ISD fields to see where they might fit in to the standard:

- topical subject headings (e.g. NLA's Music Australia headings)
- geographic subject headings (e.g. NLA's Australian map headings)
- series headings (e.g. Pandora collection records, individual records in a digitised photo collection)
- collection level records (e.g. the main record for a digitised photo collection)

Wan explained that the NLA ISD subject fields contain ISD period subdivisions which are currently dropped when the 690/691 tags are uploaded to LA and converted to 653s. NLA proposes to use the Ind2=4 convention for the in-house music subject headings, not the map headings.

Bemal pointed out the current issue with using 69X tags for ISD - the "9" can only replace 1 character (600, 610, 650 all become 690) and there is no provision for indicating the institution that contributed the data. There is also no way of limiting who gets the data in their downloaded records, who can edit the data (or remove it) and who can view it. Paul made note of the very real difference between use of \$2(thesaurus code) and \$5(contributing institution). Rob reiterated the issue of the gap between what can be accommodated in standard M21 and what is required by the NLA, i.e. the fact that \$5 is not currently supported in the 6XX fields. He mentioned the possibility of asking the MARC Standards Office to extend the use of \$5.

The ease or difficulty of configuring individual library management systems to export data in specific fields or formats should be considered.

After further discussion it was agreed that the EAG draft report would suggest the model of using \$5 in all M21 fields that currently support its use, plus other fields required to meet the requirements of Libraries Australia users.

6. Discussion of MARC21 holdings format

Paul noted that the most useful M21 holdings format data elements are:

- 852 - Location - detailed information with many subfields
- 856 - Electronic Location and Access – including URLs
- 866 - Textual Holdings – Basic Bibliographic Unit - few subfields, the most likely repository for the summary holdings data currently held in field 850
- 868 - Textual Holdings – Indexes
- 876 - Item information – Basic Bibliographic Unit
- 878 - Item Information – Indexes

Bemal acknowledged that even if LA implements the full M21 Holdings format, member libraries won't necessarily use all the available fields. LA would have to provide input guidelines.

Rob assured Paul there was no hidden agenda to withdraw support for tag 984 while it was still being used by local systems. Paul suggested that the group consider expanding the number of subfields defined for tag 850. Bemal agreed that there could be scope for that in a staging process between the simple 850 and the M21 holdings. Bemal added that the mapping of 850 to the M21 holdings format was bound to be difficult and the best solution would be for libraries to map the data in their local systems and then send it to LA for loading.

Paul was concerned that the addition of multiple holdings records with ISD might push the records beyond system length limits. The potential of this happening was acknowledged although there is currently no known limit to record length in LA other than the absolute MARC constraint of 99999 characters in both bibliographic and holdings formats. The issue was **noted**, although Bemal added that the expected number of institutions that would take up the option to add ISD and the number of fields involved would probably not be of a magnitude to create a problem.

Rob added that if data were created on a local system then that local system would be able to cope with it. Exchange of M21 holdings could be made reasonably transparent for RIS users although the online system would require some changes. The current system was supplied with the functionality to support multiple holdings for each customer but this was not implemented on migration. Removing the current restraint raises questions regarding design changes to the interface, e.g. how to display holdings records in brief view and then be able to see a full view.

It was **noted** that the proposal from the EAG would be to implement the full M21 holdings format with extensive guidelines to the Australian library community on how to use it. Bemal commented that minimum standards for holdings records would

have to be developed. Requirements for a minimum holdings record could be included in the draft report for discussion.

There was then a brief discussion on the elements of system functionality that would be required to support the exchange and/or exploitation of this data: what would be essential to include in Stage 1 and what could be left until later. Sue commented on the difficulty of knowing who would use the facility. Paul asked for an indication of the number of libraries that use ISD URLs. These ISD URLs are usually not shareable data; they rely on IP authentication at a specific institution. To provide data against which to compare the uptake of ISD URLs Rob said that LA currently has 80 RIS users, 300 users of the Cataloguing Client, and 500 users of the Web Interface. Approximately 800 users are adding holdings.

To wrap up the discussion the EAG returned to the Victorian User Group meeting notes to check that the meeting had covered all the concerns raised. Rob reported he was waiting on a response from Pica regarding the technical aspects of client functionality with regards to M21 holdings format. Bernal gave the group an overview of the three levels of data that are supported in the native LA Cataloguing system (CBS/Pica format) records - level 0 tags contain shareable bibliographic data, level 1 (not currently implemented) contains ISD in tags that reflect the shareable data tags, level 2 consists of location (holdings) data. Bernal will seek advice from OCLC as to whether Level 1 tags can be viewed by users other than the contributing library.

Standardisation by way of a recommended schema for contribution of ISD will take care of points 5 and 7. LA will have to produce a prescriptive "Beginner's kit" for guidance. Bernal pointed out that all fields can be indexed for LA Search but in the cataloguing Client certain data can only be searched by the contributing institution. Expansion of access to this data could possibly be raised with OCLC EMEA as an enhancement. Rob said that he will include suggested Z39.50 Holdings attribute index mappings into the group report.

It was assumed that contributed ISD would be overlaid, as is currently the case with holdings. Wan added that NLA expects updated records to overlay existing ones. RIS reception profiles and RES output files would have to be set up for each RIS/RES user depending on what ISD they were sending/receiving. Rob also said that some development work might need to be done for the Z39.50 server - if libraries had the opportunity to tailor their views in the Cataloguing Client and in Search they would expect the same functionality via Z39.50 searching.

It was **noted** that when LA disseminates a draft schema they should also ask libraries if they have data that they would want to contribute as ISD *a*) in bibliographic tags e.g. using \$5 subfield and/or *b*) in M21 holdings records.

Wan commented that any development would need to keep in mind the requirements of LA Document Delivery. Rob replied that retention of a summary holdings display format would probably be required and that LADD could use this, so no change would be apparent.

7. Devise a workplan and set schedule for teleconferences

The products of the day's meeting would be:

- Draft minutes for circulation to the group
- A paper for dissemination based on the day's discussion
- A teleconference to discuss the draft paper.

The meeting closed at 4.00 pm.