

NESLI MARC records: an experiment in creating MARC records for e-journals.

Ross MacIntyre

Abstract

This article concerns an experiment in producing USMARC records for e-journals within the UK's National Site Licence Initiative (NESLI). This involved working with an aggregator, Swets Blackwell, to produce sample records that were made available for comment. Some other related developments are also mentioned.

Keywords NESLI, Electronic Journals, E-journals, Cataloguing, MARC

Ross MacIntyre, M.Sc.
Senior Project Manager,
MIMAS,
Manchester Computing,
University of Manchester,
Oxford Road,
Manchester, M13 9PL,
UK.
Tel: +44(0)161-275-7181
Email; ross.macintyre@man.ac.uk

Acknowledgement: Robert Bley, Swets Blackwell, consolidated the feedback responses included herein.

NESLI MARC records: an experiment in creating MARC records for e-journals.

Introduction

NESLI¹ stands for National Electronic Site Licence Initiative and was established in 1998 by the UK Higher Education Funding Councils' Joint Information Systems Committee (JISC). The overall aim of NESLI is to facilitate and promote the use of e-journals in UK Higher Education Institutions (HEIs). NESLI might be described as being a (voluntary) virtual consortia of UK HEIs and specifically the only one approved by JISC to negotiate for full-text e-journal content on its behalf.

Following a formal tendering process, a consortium of Swets & Zeitlinger and University of Manchester was appointed to act as the Managing Agent (MA) for the initiative. The MA reports to a Steering Committee composed mainly of senior UK academic librarians. The MA performs the day-to-day service operations: publisher negotiation, communicating offers, taking orders, service delivery (where orders have been placed with Swets Blackwell), support and a small amount of service development in related areas - and it is this later aspect that is the focus of this paper. It must be noted that the funding for any development activity reduced to zero in the third year of the initiative. A summary of the initiative by Woodward² was included in this journal recently.

MIMAS stands for Manchester InforMation and Associated Services. It is a section within Manchester Computing at the University of Manchester, home to one of the UK's three National Data Centres. MIMAS hosts data and applications principally for the UK Higher Education community, though users also include the Research Councils, Further Education and other European academic institutions. MIMAS are one-half of the MA, the other being Swets Blackwell.

The technical developments were principally undertaken by MIMAS, with the support of Swets development staff.

Technical Developments

A part of the technical development of the service, a number of research areas were proposed. An indicative summary of results in each area follows:

1) Implementing UK education's standard authentication mechanism (ATHENS) for access via *SwetsnetNavigator* (SN).

Result: The mechanism implemented is best explained as consisting of three elements: authentication, conversion and connection. The user enters their ATHENS username and password on the login screen and these are authenticated against the ATHENS system for the resource type of 'NESLI'. The call to ATHENS is made by the server hosted at Manchester. Once authenticated, the ATHENS identifier and resource type is converted to SN account data, again on the Manchester server. This data is then passed to the SN server, based in the Netherlands, and a user session is initiated. The user is now interacting directly with the SN NESLI delivery system.

2) Negotiating with publishers to allow 'trusted' access. There seemed little need to perform a third authentication check, typically IP-based, once both ATHENS and the Swets subscription record had okayed the access, as described above.

Result: Comparatively little progress has been made. However, three publishers were willing to take forward and discuss at a technical level. Interestingly, one publisher still wanted the IP address of the access passed as a parameter, as this was used for compiling usage statistics.

3) Creation of journal title 'link lists' for all journals on offer, i.e. the URLs of the material in SN.

Result: Alphabetic lists of the journals available from each publisher making a NESLI offer are freely available on the web site. They include links to each journal's 'home' page on the publisher's web site and also to the journal's issue list page within SN. They are provided in a choice of formats.

4) Preparation of aggregated usage statistics for the NESLI Steering Group.

Result: The main SN log file contains data relating to the company's global user base. Once a month Swets Blackwell in the UK create an extract of usage data relating solely to NESLI customers and pass to MIMAS who produce the spreadsheets requested by the NESLI Steering Committee. This is described in more detail in MacIntyre³.

5) Off-site authentication for IP-authenticated linking to SN.

Result: The direct links to SN are IP-authenticated. This can mean that a user browsing their library's OPAC from home, say, can be refused access to licensed material. A mechanism has been implemented that prompts for an ATHENS username and password when a NESLI title is requested, but refused, as a result of an IP check..

6) Creation of MARC records for each NESLI subscriber.

Result: The project started using existing data files from Swets, as this was the most pragmatic approach. MIMAS undertook an initial mapping to USMARC format and code was developed based on this. It was planned to also offer UKMARC assuming there was a demand and that the data proved acceptable. Feedback was sought from USMARC user sites. (This feedback is summarised in the detail section that follows.) Though there were many comments about the precise contents of each field, the most significant additional data element requested was 'local data' detailing the individual site's subscription. This would have required an extract from the subscription database on a per customer basis by Swets. This was done for a sample site, but, unfortunately, no funding remained to take the development further.

Further Detail - MARC Record

The file from Swets had one record for each title with the following fields:

- 1) Title
- 2) Paper ISSN
- 3) Electronic ISSN
- 4) Publisher
- 5) Country of Publication
- 6) Language
- 7) Subject 1
- 8) Subject 2
- 9) Frequency (issues/year)
- 10) Earliest Year in Swets

There were also 5 internal (to Swets) fields.

From the above is generated a USMARC file with the following fields:

001 Control number. Use 'NESLI-' followed by the ISSN.

005 Date and time of creation of record (YYYYMMDDhhmmss.s)

008 Bytes:

- 0-5 Creation Date of record (YYMMDD)
- 6 Publication status = 'c' (Current)
- 7-10 Start year of journal. Use field (10).
- 11-14 End date of journal = 9999, which means continuing
- 15-17 3 letter code for country of publication. Use field (5).
- 18 Frequency - one letter code. Derived from (9).
- 19 Regularity = 'r' if regular, 'x' if not. Derived from (9).
- 20 ISDS center = blank
- 21 Type of serial = 'p' (Periodical)
- 22 Form of original item = blank
- 23 Form of item = blank
- 24 Nature of entire work = blank (Unspecified)
- 25-27 Nature of contents = blanks (Unspecified)
- 28 Government publication = 'u' (Unknown, could say blank => isn't)
- 29 Conference publication = '0' (isn't)
- 30 Title page availability = 'u' Unknown
- 31 Index availability = 'u' Unknown
- 32 Cumulative Index availability = 'u' Unknown
- 33 Script (i.e. alphabet) = blank (Unspecified, could say 'a' or 'b' => Roman)
- 34 successive/latest entry = '1' (latest)
- 35-37 3 letter code for Language. Use field (6).
- 38 Modify status = space (not modified).
- 39 Cataloguing source = 'd' (Other)
- 022 \$a ISSN (usually there are 2 '022' fields). Use fields (2) & (3).
- 040 \$a "NESLI"
- \$b "eng"
- \$c "NESLI"
- 041 \$a Language turned into code. Use field (6).
- 245 \$a Title up to " - " string. Use field (1).
- \$b Rest of title. Use field (1).
- \$h "PDF"
- 260 \$a Publisher Address (Only country is available)

\$b Publisher Name. Use field (4).
 310 \$a Frequency (translated to English text). Use field (9).
 650 \$a Subject (in separate fields). Use fields (7) & (8).
 856 \$a "swets2.nesli.ac.uk"
 \$u href to journal in Swets (calculated from ISSN)

Figure 1 - USMARC File Description

The following 5 additional (sub)fields had been requested when an initial trawl for opinions was made, but could not be derived from the existing data supplied:

082 \$a Dewey Decimal Number
 210 \$a Abbreviated title
 260 \$c First date of publication (print and electronic)
 710 \$a Any relevant/appropriate Corporate Body mentioned elsewhere in the record (e.g. Learned Society)
 780 \$a Previous title of journal (if any)

Figure 2 - Additional USMARC Fields Requested

For example a record from Swets that contains:

- 1) Acta Physiologica Scandinavica - Internet
- 2) 0001-6772
- 3) 1365-201X
- 4) Blackwell Science Ltd
- 5) ENGLAND
- 6) English
- 7) Anatomy and Physiology
- 8)
- 9) 12
- 10) 1998

becomes the USMARC record:

```

001 'NESLI-0001-6772'
005 '19991018145732.0'
008 '991018c19989999enkmr p u0uuu 1eng d'
022 '0'
    $a '0001-6772'
022 '0'
    $a '1365-201X'
040 ', '
    $a 'NESLI'
    $b 'eng'
    $c 'NESLI'
041 '0'
    $a 'eng'
245 '00'
    $a 'Acta Physiologica Scandinavica'
    $b 'Internet'
  
```

```

    $h    '[computer file]'
260    '0'
    $a    'ENGLAND'
    $b    'Blackwell Science Ltd'
310    ' '
    $a    'Monthly'
650    '4'
    $a    'Anatomy and Physiology'
856    '40'
    $a    'swets2.nesli.ac.uk'
    $u    'http://swets2.nesli.ac.uk/link/access_db?issn=0001-6772'

```

Figure 3 - Example USMARC File

Feedback Received

Some initial views and suggestions were sought at the time of the data mapping to USMARC. The following is a consolidated list of the feedback received from the sites following the production of some sample data.

Pricing

One response suggested a price could be charged, though it should be in same region as charged by OCLC.

USMARC Standard Conformance

Field 001

One respondent said that they would want to change to 001 field to be ISSN.

Field 245 Caused some controversy.

One respondent said there must be punctuation between the title and the format description, i.e. "Acta Zoologica Internet" should be "Acta Zoologica – Internet version".

Another respondent said that they would prefer "Acta zoologica – electronic serial"; as the title should all be in lower case except for the first word and proper nouns.

Another said "Internet" should not be included in the title, this should be a \$h[computer file]. Also pointing out that this was soon to change to \$h[electronic resource], to conform to ISBD. So the correct format would be, e.g. \$aAfrican journal of ecology\$h[computer file]

Field 260

Comments:

- "Should be 260.##, not 260.0# , also should not have ENGLAND in \$a"
- "Get rid of "Ltd." after each publisher name"
- "Should have country publisher is located in, not country of origin of journal"
- "Both indicators should be blank"
- "The correct format is, e.g. \$aOxford (England) :\$bBlackwell Science n.b. USMARC requires punctuations."

Field 650 – Three respondents said this should have true LC subject headings, e.g. \$aPhysiology\$vPeriodicals and that they couldn't load records with this field as it is. Another said that this was "OK as is".

Field 846 - One respondent wanted subfield a removed.

Field 856

One respondent would prefer to have all possible URLs included here if there are options e.g. direct to publisher's site - not just access via NESLI/SN. "It is easier to delete URLs we do not want than to add new ones."

There was general agreement that it is a good idea to have the URL. One respondent wanted to add a |z field, explaining "accessible via XXXX service".

"Would be useful to have a subfield to say which ISSN is paper, and which electronic, but this is not in the MARC standard".

One respondent said that they would not want this field to be filled if they only have subscriptions at Table of Contents (ToC) level. They only want it if they have subscriptions at full text level.

One respondent would also prefer subfields 'd' & 'f' to be included in the 856 field, containing file path and file name respectively, i.e.

'\$aswets2.nesli.ac.uk\$dlink\$faccess_db?issn=0001-6772

\$uhttp://swets2.nesli.ac.uk/link/access_db?issn=0001-6772' - using the url in the text example.

Additional USMARC Fields Wanted

Additional fields that had already been requested, see Figure 2, but which were not available.

In addition, the following were requested after the test data had been circulated:

Field 258 (Computer file characteristics) which should contain 'Computer journal' in subfield 'a', taken from ISBD (ER) (International Standard Bibliographic Description (Electronic Resources)), although not yet adopted by AACR2;

Field 542 (Mode of use note) with preferably 'Mode of access: World Wide Web' in subfield 'a';

Field 537 (Program note) detailing any system requirements, if applicable.

Field 538 Addition of a "System details note" field would be useful \$aInternet access or similar wording.

Field 785 (succeeding title) when applicable.

Two customers asked for "first issue date" field. The Managing Agent could supply this on a per-journal basis, but not easily on a per-customer subscription basis. It was not clear if this "local" data would be best added locally in any event.

One customer asked for a "previous title of journal" field. This could be supplied.

One customer said that inclusion of the "dates of coverage" would be very useful although they realised that this is somewhat difficult when some publishers are adding older volumes retrospectively and libraries' ability to access back/current volumes could change dependant on their subscription status.

They gave the following example from OCLC which might be useful:

362 1 \$aCoverage as of June 7, 1999: Vol. 4, no.1 (1997)-

Note that the first indicator must be 1 (Unformatted note).

Other Matters

Ideally sites would want to be able to pick up complete files from time to time to keep holdings up to date.

One respondent was concerned about how the URLs in the 856 field would work for off-campus users. They stated they would try and add local cache authentication or some other local scripted link to ensure access for off-site users.

Overall

The sites were asked to say that having seen this initial example, would they be interested in receiving MARC records as part of a NESLI deal.

Respondents saying “yes” – they would want MARC records from the Managing Agent.

# of respondents	Their comments
3	Interested if their suggested changes (see Additional USMARC Fields Wanted above) can be incorporated.
1	If records can be delivered by FTP, or by use of a Z39.50 client for downloading
1	No comment
1	A high priority, but will only take it up if the price is right, and the format is USMARC
3	Will use if format is USMARC
1	So important that they will not subscribe to any NESLI deals without it
3	Will use if UKMARC is used (1 of these “depending on cost”)

Table 1 - "Yes" Respondents

Respondents saying “no” – they would not want records from the Managing Agent

1	Because too much editing would be required to get it into their library system
2	Because they already have access to free MARC records from the union database provided by their library system vendor.

Table 2 - "No" Respondents

Conclusion

As was noted above, some of the additional data could have been provided, but this unfortunately remained an experiment.

Many libraries are now cataloguing their e-journals separately and wrestling with implementation issues, such as which ISSN to quote (and note that UKMARC does not allow Field 22 = ISSN to repeat). Although there may be differences of opinion

regarding adherence to MARC standards and local preferences, it was clear that these records would be valuable to the institutions.

There was nothing particularly ground-breaking here, certainly there are many suppliers providing cataloguing records (it may be interesting to see a list compiled if one does not exist already, with cost included, naturally). It is notable that the provision of electronic records for cataloguing is now being taken into account when purchasing decisions are made (admittedly a long way behind price).

Understandably, institutions want to move closer towards an ideal of "auto-cataloguing", i.e. being able to easily add/delete e-journal records to the catalogue on subscription/cancellation, to ease some of the administrative pain that currently is associated with managing electronic journal access provision.

Notes

1. The NESLI web site: <http://www.nesli.ac.uk>
2. Hazel Woodward, "NESLI--Gathering Momentum", *The Serials Librarian* 41, no. 1 (2000).
3. Ross MacIntyre, "Electronic journal usage data within NESLI", *Serials* 13, 3 (2000):161.