

Book Industry Communication

7 Ridgmount Street · London WC1E 7AE
Telephone 020 7255 0516 · mail@bic.org.uk



Minutes of the meeting of the BIC Product Metadata Group held on Wednesday 9 November 2011 at The British Library, London NW1

Present: Evelyn Nicholas, Bowker (Chair)
Graham Bell, EDItEUR
Sarah Crossley, Wiley
Peter Kilborn, BIC
Albert Peck, Wiley
Julian Sowa, Nielsen
Glenn Summers, Wiley
Gabrielle Wallington, Waterstone's
Samantha Watson, Nielsen
Keith Walters, Bibliographic Data Services
Terry Willan, Capita Software Services

1. Minutes of the last meeting and matters arising

The minutes of the last meeting on Wednesday 18 May were approved.

Identification of print on demand titles: Julian Sowa and Graham Bell clarified that Nielsen, the UK ISBN Agency and the International ISBN Agency were not working on best practice guidance for POD products. The group was asked whether best practice guidelines were needed. The main issue was around reviving out of print titles and making them available using POD. Using the original ISBN contravened the basic principle that a book once made OP could not subsequently be reported as in print. The ISBN standard itself was silent on this issue. It was felt that publisher practice would vary according to individual business needs: some would favour using the existing ISBN in order to preserve life sales history; others would prefer to treat the title as a new edition and assign a new ISBN, especially as distinctive terms might apply to a POD title (for instance, firm sale). It was suggested that further specific examples should be identified to see whether there was any de facto consensus on the issue. Gabrielle Wallington proposed that some issues around POD practice would be further explored in the price and availability group.

2. ONIX

ONIX for Books: Graham Bell reported that codelist issues 14 and 15 had been published since the last meeting. National groups had been invited to submit requests for issue 16 by the beginning of December, with publication following during January. New features of the codelists over the last three issues had been the expression of EU Toy Safety warnings and of accessibility features for print-impaired

readers, as well as additions that enabled use of ONIX in Japan, China and the Arab world. It was intended to maintain the pattern of new codelist releases every three months.

The main specification documents for ONIX 3 had been combined into a single enhanced specification, considerably simplifying access and understanding of the specification. This is available as a PDF or as an HTML file (or as a Word file if required for translation purposes). International best practice guidelines for ONIX 3 had also been published; and these, together with the main specification, would form the basis of ONIX documentation in the future.

There had been good progress on implementation in Japan, with over 200 publishers and 70 recipients (retailers and logistics organisations) now using ONIX 2.1 revision 04 in just nine months. There had also been significant progress in China; and a new Arab world national group, based in Egypt, had been formed. There were also early signs of interest in India.

Matters arising from the International Steering Committee meeting: At the meeting held at the Frankfurt Book Fair, outline agreement had been reached on what will become ONIX for Books 3.0 revision 1. This will incorporate a number of fully-compatible additions to the schema that mirror changes that were made for ONIX 2.1.4 for Japan, and will make it possible for Japan and China to use ONIX 3. The opportunity will also be taken to simplify the handling of reissues – which would involve deprecating the existing reissue composite - and make an improvement in the way titles are handled. A draft would be distributed to national groups towards the end of the year, with publication expected at the end of January.

Howard Willows had expressed concern on behalf of the UK group about the slow take up of ONIX 3 – still only two live feeds going to Nielsen. Bowker reported a growing interest in the US, but it was clear that many publishers did not see a good reason to migrate, certainly for describing their printed products. This had been a major reason for an agenda item to discuss proposals for a 'sunset date' for ONIX 2.1. An explanatory document had been circulated to national groups to explain the options and what the implications of each were.

UK ONIX guidelines: Peter Kilborn reported that the existing UK guidelines for ONIX 3 were in the process of being updated to move UK practice closer to the EDItEUR international recommendations. It was expected that these would be published by the end of November.

Other ONIX developments: Graham Bell reported that the ONIX for ISTC schema had been updated; and a new release of ONIX for ISBN Registration had been published to bring it into alignment with ONIX 3.

There had been substantial progress on defining the next priorities for development of ONIX for Serials, to streamline the new orders and the price quoting mechanisms with potential new ONIX for Serials family messages for both, and to continue development work on structured addresses and organisational identifiers, and on price and payment reconciliation.

3. Metadata for digital content

ISBN for digital products: The best practice guidelines which had been under development by the Book Industry Study Group were close to publication. It was understood that BISG's advice would not differ materially from that which had been given to BIC members in its existing guidelines.

Julian Sowa reported that a new user manual was in the process of being drafted by the International ISBN Agency. This would bring together guidelines which had been put out since the last issue of the manual, but would not extend them further.

Metadata Futures Group: The group was in temporary abeyance, having reached the end of the first phase of its work. A report on this was available on the BIC web site¹. A second phase agenda was under discussion and would be made available shortly: it would focus on identification of fragments and chapters and the future systems needs for publishers.

Metadata for digital preservation and legal deposit: A report had been submitted by Jan Ashton in her absence and is appended to these minutes at Annex 1.

4. BIC Standard Subject Categories

International scheme update: The new multilingual scheme – IBIC – had now been announced following a meeting in Frankfurt with representatives of the book industries of Spain, Italy, Germany, Portugal, Sweden, Norway, the Netherlands and the Arab world. It was anticipated that the main subject categories would be internationally standard, but the qualifier codes would be greatly extended to accommodate geographical, historical, educational and religious national variations. Publication was planned for next October. Although it might be necessary to publish a revised version to accommodate minor changes affecting UK usage, it was not anticipated that the book trade in the UK would be seriously affected by the internationalisation process.

BIC and BISAC: The Book Industry Study Group, though well aware of the IBIC development, was not planning to participate. There was, however, a clear wish to understand more about the way in which search and discovery happens and how subject schemes contribute to it. Consequently, BIC had agreed in principle to take part in a research project with BISG which it was hoped would clarify the best way ahead for subject taxonomies. This might lead to the development of a completely new subject scheme alongside – or in due course replacing - both BISAC and IBIC. For the immediate future, however, there was no prospect of BISAC being abandoned, as it was believed by BISG and its members to serve the current needs, at least of bricks and mortar stores in North America.

UK Standard Library Categories: Peter Kilborn reported that the e4libraries subject headings, which had been in draft form since 2008, were now in the process of being

¹ <http://www.bic.org.uk/77/Metadata-Futures-Group/>

republished under this new name. A steering committee had been set up which would oversee development of the scheme and address change requests.

5. Reports on other current metadata activity

BIC price and availability working group: Gabrielle Wallington reported that a best practice guide for recipients of P & A data had been published in the summer². The work of the group was now concentrated on guidelines for the assignment and interpretation of existing availability code lists. The group was scheduled to meet again in January when a draft document would be available.

BIC Discount Group Codes: The P & A group had also developed guidelines for the assignment and use of these codes³. There had not previously been any proper documentation of how the scheme was to be used in practice; and this had led to cases where the necessary uniqueness of assigned codes was being jeopardised.

Product Data Excellence Awards: There were now 30 publishers accredited under the scheme, three additions – BRF, Green Books and Kuperard - having been made. In addition Penguin had been promoted to the Excellence Plus level, joining Little, Brown and Orion.

ISTC: Julian Sowa reported that the International Agency web site had been updated and now incorporated a search mechanism to enable users to see which ISTCs had been assigned. The number of registration authorities remained at ten, but three more were in discussions. So far 13,500 codes had been assigned, thought to represent around 100,000 manifestations.

The Nielsen agency was in discussion with Springer about assignment of ISTCs to all its titles; and both Nielsen and Bowker reported continuing interest from a number of small to medium sized publishers. There was, however, no evidence of assigned codes being used yet in external applications.

The Book Industry Study Group was planning to begin a project on ISTC when its identification of e-books project was completed. BIC's Metadata Futures Group had identified ISTC as being a valuable option for publishers wishing to co-locate versions of titles within their own catalogue and also the component elements within titles, though this type of usage was not strictly within the scope of the standard.

Discussions were continuing about the possibility of expanding the scope of the standard to non-textual content. This did, however, raise issues for the ISO standardisation process, as there was a risk of conflict with standards relating to other media; and this was likely to slow down progress in this regard.

Bibliographic standards group: Terry Willan reported that the group had met in June. The current focus was on how MARC could respond to the new RDA (Resource

² <http://www.bic.org.uk/files/pdfs/110721recipients%20best%20practice%20final.pdf>

³ <http://www.bic.org.uk/13/Discount-Group-Codes/>

Description and Access) cataloguing rules. The Library of Congress had decided not to implement RDA before the beginning of 2013 and then only if significant progress had been made towards establishing a replacement for MARC. It was not yet clear what form that might take, though there was growing acceptance that MARC was inadequate for modern-day needs. A new approach based on linked data and RDF was most likely, but it was hard to see how this would work within the existing library supply chain.

British Library report: A report had been submitted by Jan Ashton in her absence and is appended to these minutes at Annex 2.

6. ISO identifier standards

Three outstanding standards – International Standard Name Identifier (ISNI), Digital Object Identifier (DOI) and International Standard Collections Identifier (ISCI) – had all gained approval in the ISO balloting process. All were now held up from publication by a dispute over the contracts ISO was proposing for the various registration authorities which were claimed to be unworkable and unacceptable. This issue, until resolved, was likely to delay the progress of any new identifier standard or revisions thereof.

7. Date of the next meeting

The next meeting would take place on the afternoon of Thursday 17 May, once again at the British Library beginning at 2.00 pm.

Annex 1

Metadata for Digital Preservation - British Library Legal Deposit Report

Consultation by DCMS took place in late 2010/early 2011. Key issues for publishers: security, uncertain cost of depositing, governance. Included in scope: offline publications (e.g. CD-ROM), web harvesting of UK domain, e-publications whether freely available or behind pay walls, deposit of e-publications by mutual agreement.

Potential Risks and Challenges

External	British Library
<ul style="list-style-type: none">• Delays due to Parliamentary timetable or another consultation?• Publisher resistance?• Differences between Legal Deposit Libraries?• Funding constraints on other Legal Deposit Libraries?	<ul style="list-style-type: none">• Funding constraints, and unknown budget settlement for 2013-2015?• Techniques and costs for harvesting behind pay walls?• Investing in delivery options/facilities, but how many publishers will agree to deliver?• Addressing publishers' security concerns?• Challenge of managing expectations?

Key principals that have been agreed by the Legal Deposit Libraries are:

- The Legal Deposit Libraries shall share deposited content;
- There will be agreement on common approaches / policies, and on content priorities;
- Costs of implementing non-print legal deposit shall be shared by the LDLs;
- The shared technical infrastructure shall be based on the Digital Library System;
- National Library of Wales and National Library of Scotland already have DLS nodes;
- Access shall be available at Bodleian Library, Oxford and Cambridge University Library;
- Trinity College Dublin shall have access, subject to confirmation in the regulations.

Deliverables

Legal Deposit Programme	Legal Deposit Libraries Jointly
<ul style="list-style-type: none">• Publisher deposit facilities• Domain crawl pilot• Document harvesting pilot• E-books ingest pilot• Strategic ingest platform• Legal deposit discovery and access• Legal deposit user support• Legal deposit security• Communications	<ul style="list-style-type: none">• Governance with publishers• Governance between Legal Deposit Libraries – new Memorandum of Understanding• Joint Non-Print Deposit Policy:• Evidence gathering for review 2016-17

Current Activities

- Supporting DCMS preparations (definitions, policies);
- Stakeholder communications: officials, MPs, media, staff, users;
- Setting up governance arrangements;
- New Memorandum Of Understanding between all six Legal Deposit Libraries;
- Launching the BL's implementation programme.

Annex 2

British Library Report - BIC Product Metadata Group Meeting November 2011

Digital Content and Ingest Programme – Non Legal Deposit materials

This programme aims to increase the range and volume of non legal-deposit digital content available to our users. We aim to increase digitisation, content partnerships and collaborative projects e.g.

- The Find My Past genealogy project where external partners will be digitising India Office materials and Electoral Registers. This has involved mapping metadata in our local EAD-like IAMS format into MODS for the digital archive. We will provide access to the digitised images via our Primo search service.
- The World War 1 project. This project is carried out under the auspices of Europeana, with a target date of 2014. The bulk of the digitisation of the Library's materials and provision of the related metadata will take place during 2012-13.

Linked Open Data

The Library of Congress' announcement about '*A Bibliographic Framework for the Digital Age*' notes that its approach will be focused '*... on the Web environment, Linked Data principles and mechanisms, and the Resource Description Framework (RDF) as a basic data model.*

At the British Library we have been experimenting with making samples of our data available as linked data, wrapping BNB records in RDF and linking to external linked data sources. Our records are converted from MARC to XML to enable enriching them with a series of links derived from a mixture of traditional and new generation bibliographic resources. These include:

- The Library of Congress SKOS version of its subject headings
- The OCLC linked data representation of the Dewey Decimal Classification
- The RDF Book Mashup

RDA Development

Delegates of the ISSN Network and ISBD Review Group attended the Joint Steering Committee Meeting 1-4 November for a series of discussions on harmonisation. The Library is in the process of configuring its Aleph implementation environment to support new MARC fields. This will be rolled out to the production system once tested.

ONIX Preservation Holdings

The Library has reviewed and commented on the latest version of the standard and our technical colleagues have output a sample data file. The sample file has been shared with the working group members, including PEPRS. Tim Devenport (EDItEUR) and Fred Guy (EDINA) will provide feedback on the file. At a voice conference on 21st October, it was agreed that the standard would be made available as a draft 0.3 for comment.

ONIX Code lists in SKOS

Metadata Standards team have received a set of the ONIX code lists in SKOS and we will be commenting and providing feedback on their potential use within a linked data environment.

JISC Names Project

The project team are working to develop the name authority prototype into a form that will be useful for repository services, using a variety of data sources. The immediate focus will be on contributions from institutional repositories and individual researchers. There have been discussions on collaboration with The National Archives and with the research councils. For an overview please see the article <http://www.niso.org/publications/isq/2011/v23no3/danskin>.