International Federation of Library Associations and Institutions

Bibliography Section

December 2004

NEWSLETTER

Editorials

From the Chair of the Bibliography Section

The Bibliography Section has a major focus to assist national bibliographic agencies in preparing their countries' national bibliography in the digital environment.

As chair of the Section, and to aid the Section in addressing this focus, I established a working group during IFLA's World Library and Information Congress held in Berlin in 2003. The Working group on Guidelines for (Electronic) National bibliographies, chaired by Maja Žumer had its first meeting during the IFLA World Congress in Buenos Aires in 2004. The main goal of that first meeting was to establish the terms of reference for the working group and lay the groundwork for the more intensive meeting to be held in Prague in fall 2004. The working group initially began its deliberations with a broad mandate: to provide general guidance for national bibliographic agencies on all aspects of national bibliography production. With this background, the working group convened for three days in

Prague to refine further its goals, objectives, and work products. I am pleased to offer the minutes of the working group's very productive Prague meeting as the Section's December Newsletter. I trust readers will find these preliminary notes instructive until final documents are produced.

Bohdana Stoklasová
Chair, Bibliography Section
National Library of the Czech Republic

From the Chair of the Working Group on Guidelines for (Electronic) National Bibliographies

During the Berlin IFLA conference in 2003, Bibliography Section organised a program on electronic national bibliographies. Papers presented gave an overview of the situation, but also focused on future trends. The conclusion was that guidelines for electronic national bibliographies are needed. Many national bibliographic agencies have already started publishing national bibliographies in electronic form (online, CD-ROM or both, several even stopped publishing the printed version). They would need guidelines for the future development. But these guidelines would be even more important for those who are in the planning process.

An informal working group was established in 2003, but the real work started during the Buenos Aires conference, when the working group was formally established. While the general mission of the working group is to develop guidelines for electronic national bibliographies, several intermediate goals were proposed:

- selection principles for electronic resources
- specification of data model and access points
- functionality of the electronic national bibliography

It was agreed that an analysis of users of national bibliographies and their contexts of use is the essential first step. The minutes of the meeting in Prague reflect the work accomplished so far. In addition to the work on electronic national bibliographies the working group is planning to develop general guidelines for national bibliographic agencies in order to provide a general framework, useful also for those who are now being established.

Maja Žumer, Chair, Working Group on (Electronic) National Bibliographies
University of Ljubljana, Slovenia

Minutes of meeting of IFLA Working Group on (Electronic) National Bibliographies, 18th-20th November 2004, Prague

Working group membership and attendance

Present:

Whitney Coe (Retired)

Alan Danskin (British Library*)*

Tuula Haapamaki (Helsinki University Library, National Library of Finland)

Agneta Holmenmark (Royal Library/National Library of Sweden)

Unni Knutsen (Oslo University College)

Christian Lupovici (Bibliothčque Nationale de France)

Bohdana Stoklasová (National Library of the Czech Republic)

Regina Varniene (National Library of Lithuania)

Beacher Wiggins (Library of Congress)

Maja Žumer, **Chair** (National & University Library Slovenia)

Apologies: Randi Diget Hansen (Danish Bibliographic Centre)

Users and Uses of National Bibliographies

The paper, with the above title, prepared by Žumer was the starting point for the discussion. The latest version of the paper, **Appendix A**, was the product of the discussion.

Librarians

More prominence was given to the requirements of librarians with responsibility for collection development. It was determined that the information needs of this role for searching were more complex. In addition to the data elements defined, it was recognised that language and country of publication were also required. For acquisitions librarians, "ordering information" was broken down into those pieces of information required to order material, including specifically: publisher, distributor, addresses, and publication status. The value of a link to books in print products was identified. Cataloguer needs were well defined in the draft. It was confirmed that any controlled access point should be a valid search strategy. Although cataloguers are searching for known items, they may need to establish the context of these items within a subject area and therefore subject and classification searches would be useful additions. It was agreed that reference librarians act on behalf of "End users" and should be classified in the same way. There was some discussion of the use of national bibliography by preservation librarians to determine trends in publishing for future planning. This falls under the use class "Bibliometrics and other statistical purposes."

Endusers

Who are the endusers of national bibliographies? The discussion revealed considerable variation between different national models. In the UK, for example, the national bibliography is a commercial product and is not widely used by the general public. In contrast, in certain Scandinavian countries it underpins the national union catalogue and is very widely used. The expectation of users is that they should be able to obtain access to resources. Direct access to resources has not hitherto been a function of national bibliography, but user expectations must be taken into account in the electronic environment. It was agreed that, in FRBR (Functional Requirements of Bibliographic Records) terms, the national bibliography describes entities at the manifestation level, it should also enable users to obtain items.

Booktrade

This group was defined as including publishers and booksellers. In this context, "publisher" is meant in the most general sense. It is one of several terms that will require definition in the working group's final document. It was suggested that the behaviour of booksellers has much in common with collection development librarians and reference librarians. Knutsen reported that she has been running courses for the booktrade, teaching booksellers to use national bibliography and union catalogue. The reuse of records by the booktrade (e.g., for books in print products)

was also noted for discussion under uses. The importance of cataloguing-in-publication information as advertising for publishers and potential requirements for booksellers was stressed. The potential significance of national bibliography for the booktrade as a tool for market analysis and evaluation of competition was discussed.

Uses of national bibliographies
Certain uses of national bibliography were identified that were not specific to any user group, but that had specific data or structural requirements.

Bibliometrics and other statistical purposes
Several official statistical uses of national bibliography were reported. For example: in Norway, the national bibliography has been used on a trial basis to register authors' rights. The national bibliography also forms the basis of the official books statistics. In Sweden, the NBA registers publishers' rights to specific serial titles. For these purposes comprehensive, consistent, normalised data are essential. Target audience was suggested as an additional search requirement.

Reuse of records
This was defined as bulk transfer of data between databases. In addition to libraries and scholarly resources the importance for commercial agencies was recognised. The key issue was that records should be authoritative and timely.

Rights management
Title was added to the search criteria.

Computer software
This was accepted as presented in Zumer's paper.

Conclusions
The WG agreed on the following premise:
- WG is developing *Guidelines* for creating an electronic NB
- From the e-NB *Guidelines*, the WG will extract guidelines for creating a basic NB
- From this work, the WG will develop selecting principles that will cover e-resources

The following possible mission statement was discussed:
"To record reliably the national output that documents the national heritage of a country."
1. The data model of e-NB must be derived from user requirements. User requirements will also determine functionality.
2. User expectations are different in the "Google-Age;" e-NB must reflect these expectations.
3. e-NB does not need to do everything and record every data element that every user may require. It ought, however, to be possible to link to other products and services that can deliver specific requirements.
4. National interests take precedence over international interests--this reflects political and economic realities.
5. The continuing production of national bibliographies cannot be assumed--the case has to be made.

Issues
1. What is the mission of the national bibliography? The Copenhagen principles are rather vague. The ambition of the UNESCO 1979 principles was limited by available technology.
2. What does the national bibliography deliver that other products and services cannot/do not?
3. There is a lack of hard data in respect to certain users and uses of national bibliography. There is good understanding of how they are used in and by libraries. The reliability of the data declines for other groups. Users and uses vary according to

the coverage and the business models of individual national bibliographies.
4. There is overlap with other products and services, such as books in print, national union catalogues and institutional catalogues.
5. Are manuscripts in scope? Yes – in some countries.
6. Does the national bibliography record the book as it was when first published or should it be updated, e.g. to reflect price changes.
7. Must the national bibliographic agency that creates the NB hold the item? Can an NB list a website that its NBA has not harvested? At a minimum, is access to a website listed in an NB a requisite?

Selection criteria for electronic resources
Constraints
The following constraints were identified, which helped to focus the discussion:
- Electronic resources cannot be treated generically. Electronic resources covers a spectrum of resources ranging from hand held media, such as CD-ROMs to complex integrating online resources.
- It was agreed that online resources represent the principal challenge and should therefore be the main focus of discussion.
- There are four levels of metadata/cataloguing:
1. Basic: no information is added to the harvested documents.
2. Enhanced: metadata are automatically generated.
3. Enriched: (simple) metadata are assigned manually.
4. Authoritative: full cataloguing, implementation of authority control.

The key distinction between levels 3 and 4 is the provision of controlled vocabulary for access points in level 4.

Relevant experience
Each institution reported on its experience and issues.

BNF
Legislative framework is in place for digital document on media and a project for a web archive is under consideration. The BNF has been harvesting snapshots of the "French" Web since 2000 as experiments preparing the future regulation. The harvester follows the links and collects all linked sites and documents. In this case, the bibliographic unit is the website. Content of the sites is analysed to evaluate relevance. The software reports when it encounters deep web or access restrictions. These problems have to be resolved by manual intervention and agreement with publishers for deposit. The whole "French" web is intended to be harvested annually; in addition more frequent snapshots can be taken depending on criteria recorded in the metadata created for each site.

BNF is leading an international consortium (IIPC) for the Internet preservation. In this framework, BNF with the British Library will develop a "smart" crawler. This will employ language analysis. Initial research suggests that automated procedures may be more consistent than human intervention.

BNF is seriously considering whether to exclude web resources from full level cataloguing. The volume of web data is such that it cannot be catalogued and support for web archiving is already taking resources away from print. There is a perception that print resources are in any case more stable than web and whereas print resources cannot be found unless they have been catalogued there is always at least a minimal level of access to web resources.

National/University Library of Finland
Legislation is pending. Similar approach to BNF. Full text indexing. Documents defined as "important" are fully catalogued for national bibliography. The data model includes two databases: Web Archive and National Bibliography.

Denmark (Danish Bibliographic Centre and The Royal Library)
Legislative framework is in place.A legal deposit law from 1998 exists where publishers/producers are obliged to notify „ static" internet publications, and they are all fully catalogued for the National Bibliography. In 2005 a new law will be passed which allows harvesting.

National Library of the Czech Republic
Czech Republic does not have a legal deposit law for web resources. Voluntary scheme has been operating for several years. Subject categories are distributed across different institutions to enrich the web resources coverage by the Subject Gateways model.

National Library of Lithuania
Following Finnish model. Publishers have been trained to assign Dublin core to resources. Records are included in national bibliography.

Royal Library Stockholm
Legislation is pending. Sweden does not yet have legal deposit for web resources, only for e-resouces such as CD-ROM, diskettes, etc. Sweden is "allowed" to do harvesting and has agreements with several publishers. The law is expected by 2006. All records are included in the national bibliography. Certain publishers contribute XML records.

Library of Congress
No national bibliography. Recent research project has analysed cataloguing requirements for electronic resources. Conclusions include the recommendation that more of the available staff resources should be expended on subject analysis, rather than description. Description is less critical for web resources because the resource can be linked to the record.

British Library
Legislative framework in place, but enabling legislation is required before web resources are subject to Legal Deposit. Voluntary scheme has been operating for several years. BL is mainly cataloguing hand held media, but has some experience of cataloguing web resources for collection. Consideration is being given to treating web resources as collections because it is often impossible to describe all the components of a website.

Cataloguing levels
Metadata levels were related to the users and uses

User/Use	1	2	3	4	Comments
Librarians: Collection Development			X	X	4 for specialist uses only
Librarians: Cataloguers				X	
Endusers	X	X	X	X	
Booktrade			X		
Bibliometrics		X	X	X	
Reuse of Records			X	X	
Rights Management				X	Name authority control essential

What type/level of material should be collected? The principle that emerged from discussion was that everything should be harvested, but filtered by increasingly stringent selection criteria. The objective is to refine the resources requiring manual interpretation as effectively as possible. The metadata available to this filtering process vary according to record level.

The issue then is to determine what criteria are used to determine which level 2 records should be extracted for level 3 (manual) enhancement and level 4 (manual) enrichment. Various types of criteria were considered.

The recommendations in *Guidelines for the national bibliographic agency and the national bibliography / prepared by the IFLA International Office for UBC. Paris : Unesco, 1979* assigned significance to breadth of distribution or number of copies produced. There was discussion concerning whether this could have a web analogue. It was suggested that as the concept of "copies" is not appropriate in the web environment, links to the site might be considered evidence of widespread distribution, therefore a large number of links would represent many copies. The Paris recommendations also make judgements regarding the length (i.e. number of pages) and the content of a publication. Metadata for identifying "quality" content were discussed: these included presence of identifiers, such as ISBN or ISSN; creator information may be tagged in the record or it may be inferred from domain names. Type/Format/Genre were identified as further criteria that may facilitate automated selection in line with national selection policies. One example discussed was "games." These are (currently) explicitly excluded from BNB, but are included in other national bibliographies. Another possibility might be "no cookies." More information on formats is required. The richness of metadata already associated with a resource may be an indication that high value is already attributed to the site. More subtle refinement will probably depend on linguistic analysis being sufficiently subtle to identify the significance of a document in cultural, heritage or social terms.

Conclusions
- Level 3 records appear to satisfy many requirements.
- Co-operation with publishers (e.g., exploiting the ONIX format) is essential.
- Granularity is dependent on sophisticated analysis tools.
- NBA is responsible for setting the boundaries of harvesting.
- National bibliographies include records at different levels.
- Automated selection processes are essential to filter harvested resources for:
  - Inclusion in national bibliography;
  - Addition of enriched or authoritative metadata.
  - Discussion of automated selection processes yielded the following points:
    - At a minimum, automated means should capture any existing metadata that can be reused.
    - Automated means equal web crawling/harvesting? Frequency of crawls affected by type, e.g., serial/integrating resource crawled frequently—monthly or quarterly; monographic sites crawled less frequently—annually?
    - Automated means can be done by focusing on and setting boundaries related to country domain; by public vs. private domain; by language; by frequency of the crawl; by frequency in changes of the website; by number of links to and from a website.
    - After automated means have been exhausted, the following manual criteria were discussed: number

of links to and from a website; official registration
(e.g., types of standard numbers); extent of metadata;
importance (as determined by publisher, research value,
cultural relevance, business relevance, reflection of
evolving technology, contemporaneous relevance; quality
(as determined by content; demand; specialness of the
website); monographic vs. serial/integrating resource.

Issues

- Is IFLA minimal element set adequate for minimum level 3
  definition?
- How can NBAs share authority control data with publishers?
- The extension of legal deposit to electronic resources is
  blurring the distinction between the library and the archive
  – or, at least it offers new possibilities to libraries to collect
  ephemera

Survey etc.

The purpose of the survey is to gather more comprehensive
information on selection criteria. It was agreed that the survey
should be limited to Europe on pracital grounds. This would
give an appropriate "regional" focus for the Oslo World Library
and Information Congress. The survey could be re-run for the
Far East in advance of the Seoul World Congress in 2006. The
following topics were identified for inclusion on the survey:

The full survey is available at:
http://www.ifla.org/VII/s12/pubs/s12-Survey-Europe-e-
resources-Programs.pdf

- Legal Deposit
  o Legislative framework—legal requirement? voluntary?
  o What materials included (format/content)
  o Agency with responsibility for the materials
- Selection Criteria: what is collected; manner collected
  o URLs of published criteria
- Types of material included in national bibliography, reflecting
current situation and future plans.
  o Text
  o Sound
  o Images
- Cataloguing
  o Levels
  o Material type
  o Reuse of data
  o Identifiers

  o Metadata
  o Enhancements
- Access to documents
  o Archiving/preservation of resources
  o International property rights issues/restrictions to
    access
  o Guaranteed?
  o Persistent?
  o Technical framework
- Current research
- Future plans
- Co-operation
  o Extent of responsibility of NBA
  o Other agencies involved
For all of the areas above,
  - Any collaboration?
  - Any experiments/research pilots being
    conducted or planned?
  - Provide URLs for any documentation that can
    be shared
- The survey will be publicised through the *Newsletter*, by
  e-mail, personal contact and as a news item on the website.
  The survey will be sent to NBA directors.

Section's Program in Oslo

Survey results	Beacher Wiggins
Keynote presentation	John Byrum
SMART Crawler and BNF experience	Christian Lupovici
Interim report of WG on e-NB	Maja Žumer

Future activities of working group

Deliverables/priorities
1. Selection criteria for e-NB.
2. Guidelines for e-NB
3. Beginner's guide to National Bibliography

The *Guidelines* for e-NB are the principal deliverable for the
group. The selection criteria are a component of that broader
deliverable and should be possible to wrap them up with a further
meeting in Oslo. It was agreed the WG will meet again next year
after IFLA, to be arranged in Oslo.

Appendix A
Users and uses of national bibliographies (NB)

*Prepared by Maja Žumer (with the help of working-group
members)*

Introduction

**General statement of the purpose and scope of this
document**…

Definitions:
Publication (published resource)
Publisher
End-user
Author
Harvesting
Access

Archiving
Preservation

Users (and 'uses'; see 7) of NB:

1. Librarians
- **Collection development and acquisitions librarians**: to
  analyse available publications, to select and order publications,
  including awareness of future publications (e.g. using CIP
  records)
  Searching:
  - known items: search by title, author, publ. number (ISBN,
  ISSN, ISMN…)
  - on a topic: search by subject headings, class. number or
    keyword
  - date of publication

- language
- publication type/genre/format

Data needed:
- full bibliographic record for collection development
- information needed to order publications
- price
- availability
- terms/condition (including rights, technical requirements and status of publication)

- **cataloguers:** for copy cataloguing or cataloguing support

  Searching:
  - known items: by title, author, publ. no.
  - authority controlled fields

  Data needed:
  - full bibliographic record
  - full authority record

- **reference librarians:** on behalf of end-users (see **end-users**)

- **administration of legal deposit**: to control, to plan preservation processes

- **preservation librarians:** to assist with preservation needs

2. **Endusers** (including groups, corporate bodies etc, which use NB as an information source)**:** to obtain information on authors and their publications, on topics…, identify publications or authors

   Searching: all access points, particularly:
   - author name (any form)
   - title words
   - language
   - publisher
   - publication year
   - subject headings or keywords
   - publication type/genre/format

   Data needed:
   - to identify, select and enable to obtain publications (see FRBR)

   User expecations:
   - direct access to electronic resources (taking into account rights, authentication and authorisation etc.) or obtaining location information for traditional publications

3. **Booktrade (publishers and booksellers) including non-commercial sector and government and official publishers:** to analyse the market and competition, to order publications and awareness of future publications (similar to acquisitions librarians), possibly to refer customers to libraries for out-of-print publications (similar to reference librarians).

   Searching:
   - on a topic: search by subject headings, class. number or keyword

- date of publicationpublisher
- publication type/genre/format

Data needed: basic publication data

4. **Statistical and analytical purposes (metrics)**
- **funding bodies:** to analyse publications
  Searching:
  - language of publication
  - language of original
  - country of publication
  - classification code
  - date of publication
  - audience
  Data needed: basic publication data (citation format)

- **official statistics:** to analyse publications
  Searching:
  - language of publication
  - language of original
  - country of publication
  - classification code
  - publication type/genre/format
  - audience
  Data needed: basic publication data (citation format)

5. **Re-use of records:** downloading or export of records into other databases and other computer applications (CERL, Index Translationum, library catalogues (retrospective conversion!),…)
   Searching: criteria of selection
   Export: full record in required format

6. **Rights management (collecting agencies):** to track publications and identify authors
   Searching:
   - title
   - author name (any form!) (primary, secondary)
   - relator code
   Display: basic publication data (citation format) + relator code or authority records

7. **Computer software** (federated searching, distributed searching, harvesting etc.):
   Searching:
   - standard access points (Bath profile)
   Export:
   - standard formats (XML), different metadata schemes
   Specific requirement: compliance with standards such as Z39.50, SRU/SRW, OAI ….

In addition to those common users and uses there may be others with specific information needs such as printers trying to identify publishers of certain formats or types …

IFLA's Three Pillars: Society, Members and Profession

At its meeting this week, the Governing Board decided to endorse a new > model for IFLA's operations, the three pillars, which recognises that IFLA's core functions relate to the societal contexts in which libraries and information services operate, IFLA's membership and professional matters. These three pillars are supported by the infrastructure offered by IFLA HQ, IFLANET and the Federation's governance structures.

The Society Pillar focuses on the role and impact of libraries and information services in society and the contextual issues that condition and constrain the environment in which they operate

across the world. Those issues are addressed currently through FAIFE, CLM, Blue Shield, and our advocacy in the World Summit on the Information Society (WSIS) and other arenas. Next year, of course, we will need to place particular emphasis on the World Summit on the Information Society in which IFLA has invested so much and to such good effect. We must capitalise on that investment and those results so that we take full advantage of the second phase leading up to the November 2005 summit meeting in Tunis and to the achieve the longer term goals.

The Profession Pillar focuses on the issues covered by the long established Core Activities - ALP, ICABS, PAC, UNIMARC - and the Sections and Divisions. They lie at the core of our professional practice and help libraries and information services to fulfil their purposes and to shape responses to the needs of clients in a rapidly
changing global environment.

The Members Pillar is of course central to IFLA. It includes the services we offer to members, management of their membership of IFLA, conferences and publications. We must work together to make IFLA more vibrant and attractive and beneficial for members throughout the world.

Of course, all three pillars and the underlying infrastructure are interdependent and not mutually exclusive. They offer a way of understanding and presenting IFLA holistically to the library and information sector and to governments and the wider community. The pillars should also provide the framework for IFLA's management and their focus should be reflected in program based budgeting.

Working parties of the Governing Board are focussing on the Society Pillar, the improvement of IFLANET and the new approach to budget management with actions to be identified early in 2005. These priorities recognise the crucial significance of WSIS, WIPO and other current advocacy initiatives, the vital importance of IFLANET as a medium for engagement and communication within the Federation and beyond and the need for more contemporary approaches to management and finances. The other pillars and areas of activity will be tackled subsequently.

As announced, Mr Ramachandran has resigned from the position of Secretary General. Mr Sjoerd Koopman, the Coordinator of Professional Activities has been appointed acting Secretary General while a new Secretary General is being sought. Sjoerd will continue the process of change and support the reframing of IFLA's programs over the coming months. Any issues relating to IFLA activities should be referred to Mr Koopman at ifla@ifla.org.

At this time, as the process of change in IFLA accelerates, it is important that we should all work to strengthen our Federation. We need to build the resources to undertake fully the priorities which IFLA members have signalled especially a strong advocacy program, flourishing professional activities, valued membership benefits and effective infrastructure. All members need to join with IFLA staff and elected leaders to help secure the required financial strength which will allow us to achieve these goals. All need to promote IFLA and its priorities - and particularly, at this time, the World Summit on the Information Society.

And, finally, if we are to achieve these aims we need the best possible governance so we encourage the nomination of committed colleagues from around the world to the positions on Standing Committees, to the Governing Board and for President elect by the closing date of 2 February 2005. They should be nominated not as sectoral representatives or representatives or particular countries or regions but as strong candidates who represent the best of our profession, reflect its diversity and will be able to make significant contributions to its development through a strong and effective IFLA.

We both wish you all the very best for 2005:
Kay Raseroka
President

Alex Byrne
President elect

The Hague, 10 December 2004

For more information on the Bibliography Section please consult:

http://www.ifla.org/VII/s12/index.htm

or the chair

Dr. Bohdana Stoklasová
Director of Library Collections and Services
National Library of the Czech Republic
Klementinum 190
110 00 Prague 1
Czech Republic
Tel: +420 221663293
Fax.: +420 221663301
e-mail: bohdana.stoklasova@nkp.cz