

IFLA Namespaces Task Group

IFLA Namespaces - Requirements and Options

Mar. 2010

(Reference links updated Feb. 2011)

Background

Following a preliminary meeting with the IFLA Web master and representatives from several stakeholder groups (Anders Cato for the Cataloguing Section, Patrice Landry for the Classification & Indexing Section, Gordon Dunsire for ISBD XML, Barbara Tillett for FRBR, FRAD, MulDiCat, and ICP), the Classification and Indexing Section of the Division of Bibliographic Control discussed the topic of IFLA standards in the new IFLA professional structure during its Standing Committee meetings at the 75th IFLA General Conference and Council in Milan, Italy, August 2009. The discussion is recorded in items 12 and 18 of the minutes of those meetings¹. As a result, the Section initiated a working group to recommend a structure to coordinate, maintain, and promote standards from a technical point of view, to be submitted to the IFLA Professional Committee before its meeting in April 2010. The Section wished to include members from all relevant Sections (Bibliography, Classification and Indexing, Cataloguing, Information Technology, and Knowledge Management) dealing with bibliographic standards, on a voluntary basis.

The topic was also discussed during Standing Committee meetings of the Cataloguing Section (item 16 in the minutes²) and the Bibliography Section (item 3.1 in the minutes³); both Sections generally supported the initiative.

The resulting IFLA Namespaces Task Group contains representatives of the Bibliography, Cataloguing, Classification and Indexing, and Information Technology Sections, and the FRBR Review Group, ISBD Review Group, and ISBD/XML Task Group. Details of members are given in an appendix to this report. The Task Group has used the discussion paper "Identification of potential requirements for managing IFLA namespaces" as a basis for this report. The paper is Appendix IV of the minutes of the Cataloguing Section Standing Committee minutes already cited.

The focus of that paper, and of this report, is on the management of IFLA bibliographic standards in relation to the Semantic Web. This reflects current activity within the IFLA community, such as

- the FRBR Review Group's project on a namespace for FRBR entities/relationships in RDF,
- the Cataloguing and Classification & Indexing Sections MulDiCat initiative to post a SKOS version of the Multilingual Dictionary of Cataloguing Terms and Concepts on the IFLA Web site,
- the collaborative work of ICADS with both UNIMARC and MARC formats, and
- the work of the ISBD/XML Task Group,

¹ Available at: http://www.ifla.org/files/classification-and-indexing/minutes/august-2009_rev.doc

² Available at: http://www.ifla.org/files/cataloguing/reports/meeting_2009.pdf

³ Available at:

http://www.ifla.org/files/bibliography/Documents/Bibliography_Section_Meeting_minutes2009_Milan.pdf

as well as in other communities with strong links to IFLA standards, such as RDA: Resource description and access, and the CIDOC Conceptual Reference Model. There is general consensus within the relevant IFLA groups that IFLA standards have a significant potential role in the development and utility of the Semantic Web.

The term "namespace" is used here with a broader scope than its usual technical definition, and covers all aspects of compatibility with the Semantic Web.

Requirements, recommendations, and options

1. *Technical Group*. There is a need for a group which can help coordinate technical issues associated with all relevant standards, and act as a single point of liaison between IFLA standards groups as well as with third-parties, especially where the third-party is interacting with more than one standard. An example is IFLA's interaction with the Vocabulary Mapping Framework (VMF) project, which has so far directly involved FRBR and FRAD along with external standards such as RDA, which themselves have some interaction with IFLA standards. VMF has also indirectly involved ISBD and UNIMARC. The project has not yet required any formal authorisation of its mappings to IFLA standards, but this is an identified future development that would be more effectively carried out if there was no necessity for each IFLA standards group to nominate a representative. The group could also coordinate and implement the development of common approaches to technical issues, some of which are identified below.

Recommendation 1: Create a technical group to coordinate IFLA standards and develop common approaches to their maintenance, promotion, and interaction with third-parties. The technical group could be a sub-group of a more general IFLA Standards Group/Program, to which it would report.

2. *Dedicated IFLA Web Space*. There is a need for an IFLA Web space for publishing RDF/XML (or other format) files which declare a namespace, along with associated documents including reports, diagrams, guidance and mappings to other standards and namespaces. Documentation for IFLA bibliographic standards is currently scattered throughout the IFLA website, reflecting the various groups responsible for those standards but with little indication of the many alignments and other relationships between standards. This conceals the effort that groups have undertaken to work together within IFLA, and hampers promotion and uptake of IFLA standards by other communities.

Recommendation 2: Develop a standards "portal" on the IFLA website. This is a specific area of the website, and might comprise links to documents stored elsewhere on the website or storage of documents that are linked to from elsewhere on the website or a mix of both. The portal could also accommodate inter-group documentation and be the basis for future interactive development of standards using social networking (Web2.0) technologies. A single focus point will not only improve promotion, access, and uptake of standards, it will also help IFLA groups to identify maintenance issues such as review schedules, gaps, and synchronisation with related standards. The portal can use the URL domain iflastandards.info, which has already been registered and is intended for use as a base domain for namespaces (iflastandards.info/ns/). Portal content should be maintained by nominated representatives of each IFLA standards group, and in liaison with the IFLA webmaster and coordinated through the general IFLA Standards Group/Program.

3. *Dereferencing Services*. There is a need for dereferencing services that accept a URI identifying an element of an IFLA standard, for example an entity, attribute, or relationship, and return detail about the element, such as a label, definition or scope note. Dereferencing services can operate at different levels of granularity. A basic service might return the RDF/XML associated with a URI, leaving it up to the client to process the metadata to extract the required part. A more advanced service might carry out an extraction specified by the client, for example returning the RDF/XML for a specific language or "property" such as the scope note. An advanced service might also offer a choice of formats for the returned metadata, for example JSON. There are well-established techniques for achieving this, involving an http redirection of a URL containing the URI to HTML and RDF documents containing the detail. The availability of a dereferencing service will encourage adoption of IFLA standards by developers of end-user applications. However, dereferencing services can be misused by poorly-designed web crawlers mistaking URIs for URLs, which may result in overloading the server hosting the dereferencing services, and this needs to be taken into account when implementing such services

Recommendation 3: Provide basic dereferencing using HTML and RDF documents representing IFLA standards. Set-up and maintenance of dereferencing services should be carried out by the IFLA webmaster in liaison with IFLA standards groups and coordinated through the more general IFLA Standards Group/Program. Adoption of common guidelines in creating HTML and RDF documents will significantly reduce the need for close liaison and allow maintenance to be carried out as a routine process. The dereferencing service should be aligned with the base domains used for IFLA URIs. Taking into account the portal requirement, this suggests that the portal itself and the associated dereferencing service should be mounted on a server separate from that/those used for the rest of the IFLA website.

Recommendation 4: Consider developing more advanced dereferencing in due course.

Options: Dereferencing services could be provided by a third-party in close liaison with the IFLA webmaster.

4. *Third-party Namespace Maintenance Services*. There is a need for an infrastructure for maintaining namespaces as a result of review/revision of standards or the provision of translations. While review processes are generally scheduled and controlled by IFLA standards groups, the provision of translations tends to be more ad hoc, so a flexible infrastructure is required. Common approaches are to use either a relational database system or RDF "triple store" to store namespace metadata. This allows elements to be amended individually, and provides a flexible basis for producing output in the form of RDF and HTML versions of the namespace declaration files and other output as required. An important requirement for IFLA is that the interface is easy to use by non-technical staff so that content maintenance does not depend on technical knowledge of RDF. However, this would require significant resourcing from IFLA and would take some time to develop. The namespace for FRBRer has been created (and could be maintained) using the NSDL Metadata Registry⁴. This service is freely available and has many required and desirable features, such as web access, authentication, automatic generation of URIs, provision for translations, change tracking, and instant output of RDF documents. It lacks some features such as simple HTML output, although they could possibly developed under contract from IFLA. Alternatively, IFLA could develop the software

⁴ Available at: <http://metadatarregistry.org/>

itself, as the software is open-source (and uses a relational database). There may be other similar services available and further investigation is necessary.

Recommendation 5: Further investigate the provision of namespace maintenance services by third-parties. This need not delay the creation of new namespaces, as the NSDL registry is available and workarounds for missing features could be developed.

Options: An infrastructure could be provided by IFLA, if resources were available. This would require expertise in database development, metadata analysis, web publishing of database content, interface design, and formatting for database output.

5. *Version Control.* There is a need for an infrastructure for version control and change alerts. Version control is extremely important for promoting standards and fostering trust. Developers adopting IFLA standards and namespaces will have clients who could be other developers, end-users, or service providers who will need to know what version is being used, and what later versions might be available. Version control can be passive or active: it is passive when users are not specifically alerted to an update and are expected to check for themselves; it is active when users are alerted when an update occurs and told what specific change has been made. Passive version control needs to be applied to source documents, individual elements of standards, and output documents, in the form of change histories, etc. Active version control is becoming an expectation of developers, who will benefit from RSS feeds, blog posts, etc. alerting them to changes as they occur. Active version control is closely connected to namespace maintenance procedures, and is best carried out using the same infrastructure (see Recommendation 5). However, passive version control can be readily carried out in-house.

Recommendation 6: IFLA standards groups should be responsible for passive version control, in-house, as monitored and coordinated through the general IFLA Standards Group/Program.

Recommendation 7: Develop general guidelines for in-house control to assure a common approach within IFLA. These could be based on guidelines produced by W3C and other relevant agencies.

6. *IFLA Infrastructure.* There is a need for an infrastructure for ensuring that IFLA standards contribute to the open linked-data environment of the Semantic Web. This is essential for the promotion and uptake of standards. The greatest utility of linked-data is likely to take the form of instance data, that is, metadata content from catalogue records with structures based on IFLA standards. The IFLA structural standards (metadata models such as FRBR, parts of ISBD, UNIMARC, etc.) will be an important component of the process of parsing existing catalogue records into linked-data, so the IFLA namespaces need to be available and accessible to linked-data developers. This can be achieved via a dereferencing service and by outputting IFLA namespaces in a linked-data format (Semantic Web triples).

Recommendation 8: Implement Recommendations 1 and 3 for a technical group and dereferencing services.

7. *Future Software Service Requirements.* Inevitably there will be advanced, future requirements including the provision of Applications Program Interfaces (APIs), widgets and other software to help implement the use of standards by other communities (for example, XSLTs for parsing instance data such as catalogue records or adding punctuation to metadata displays), application profiles

specifying required or repeatable elements, SPARQL end-points, and linked-data triple stores. The provision of such software services would require significant resources and additional technologies. Some of these requirements can be accommodated immediately, if necessary, in the form of documents in the portal, but a longer term coordinated approach is likely to be beneficial.

Recommendation 9: Plan for future requirements when providing more immediate requirements.

General options: All recommendations except the formation of a Technical Group can be met by contracting third-party suppliers in place of any in-house development.

Appendix: Composition of the IFLA Namespaces Task Group

Emmanuelle Bermès	emmanuelle.bermes@bnf.fr	Information Technology
Leda Bultrini	leda.bultrini@arpalazio.it	
Anders Cato	Anders.Cato@kb.se	Cataloguing
Charlene Chou	cc179@columbia.edu	Bibliography
Gordon Dunsire	g.dunsire@strath.ac.uk	(Chair) Classification & Indexing
Massimo Gentili-Tedeschi	massimo.gentilitedeschi@beniculturali.it	
Lynne Haworth	lynne.howarth@utoronto.ca	
Sally McCallum	smcc@loc.gov	Knowledge Management
Pat Riva	Patricia.Riva@banq.qc.ca	FRBR Review Group
Elena Escolano Rodríguez	elena.escolano@bne.es	ISBD Review Group
Barbara Tillett	btil@loc.gov	Classification & Indexing
Mirna Willer	mwiller@unizd.hr	ISBD/XML Task Group