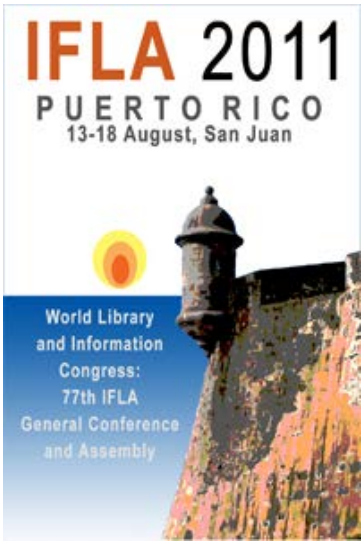


Convergence and Interoperability: a Linked Data perspective

(Convergence et interopérabilité : l'apport du Web de Données)

Emmanuelle Bermes
Bibliothèque nationale de France (BnF)
Paris, France

Meeting: **149 —*Bridging domains, communities and systems* — Classification and Indexing Section**

Abstract:

The web provides a seamless environment where the user can navigate through resources regardless of their provenance: search engines, public and governemental institutions, commercial websites, social networks, etc. There is a strong contrast between this reality of a seamless web, and the cultural heritage approach of data dissemination. Metadata modelling based on domain-model requirements has led to incompatible standards that make it really difficult to share data between libraries, museums and archives. Cultural heritage institutions maintain catalogues that exist like silos, isolated from one another, and isolated from the wider ecosystem of the web.

In the cultural heritage domain, providing relevant services for Web users requires to seek convergence, and to bridge the gap between libraries, museums, archives and other cultural institutions. Despite the efforts that have been conducted during the past several years,

data interoperability is still an open issue. Federated search is permitted by protocols like Z39.50 and Web services & APIs, but the kind of service that can be provided by this technology doesn't seem to fit the needs of Web users. Other interoperability strategies, like

exchanging simple Dublin Core records through OAI-PMH repositories, have shown their limits : by seeking the smallest common denominator between disparate data, they create a low-quality data environment with limited perspectives.

The Linked Data provides a perspective for a different kind of interoperability, based on Web architecture principles. Linked Data interoperability is designed to support heterogeneous description models, which is necessary to handle the very different data from libraries, museums and archives. The Linked Data

cloud is built following a bottom-up approach, allowing each institution visibility and ownership on their own data.

Yet, in order to create this new interoperability framework for linked data, we need to build strong links between datasets. These links rely on the description of "real things" : persons, objects, concepts, places, things that have been described as "authority data" or "subject data" by libraries. The new standards for future catalogues, FRBR, FRAD, RDA, tend to dramatically increase the importance of authority data in the library data landscape. Beyond libraries, these data will provide the hub where cross-domain links can be anchored.

There are still a number of open issues, regarding the creation of a semantic network of data that would be able to connect information from libraries, museums and archives, in a seamless way. Some of these issues are related with the need to be able to convert existing metadata into semantic web enabled data. Some other issues are specific to the Linked Data environment : how to create alignments between datasets, what kind of properties should be used to create those links, etc.

This paper aims at exposing the challenge of interoperability and convergence, and the perspectives offered by Linked Data to address this issue. It will be illustrated by implementation examples such as VIAF, Rameau & LCSH, Europeana, etc. We want to propose a way to renew our perspective on interoperability, based on a user-oriented approach, and an emphasis on the need for cross-domain links for subject data. Our goal is to demonstrate that services built on Linked Data will be more efficient and relevant to the end user than services built on traditional library interoperability frameworks.

The added value of the Web environment, the motivation of its universal adoption and its reign as the major information media nowadays, lies in its interoperability and globality. The Web relies on a set of standards, allowing it to be completely independent of hardware and software environments. In this regard, the Web is the greatest interoperable environment.

Users are now familiar with hypertext navigation and use of search engines, therefore libraries and other cultural organizations should embrace a user-oriented approach when designing their online services. They should no longer focus on how to serve the user once he's inside the library, but rather adopt a global approach of the user's needs. What is he trying to do ? Learning ? Searching ? Enjoying himself ? Is he trying to answer a question for his everyday life, about his health, his house, his foods ? All these questions can be part of the user's experience of the library, but none should be considered as a domain where the library can be the unique selling point. The library must be in the flow of Web usage, and not expect users to come purposefully to the library.

In this context, convergence between libraries, museums and archives is an even more important issue, as we can't expect the user to understand organizational barriers between institutions, and to accept them. A user who plans to visit a museum should find books about Picasso as well as digitizations of his paintings ; a genealogist seeking his great-grand-parents should be able to browse seamlessly resources from libraries and from archives.

Even further, we could wish that our user wouldn't have to look for these resources, but that they would be pushed towards him during his standard Web navigation, through results in his favorite search engine, links in Wikipedia pages, references from his favorite websites.

All this seems quite obvious when speaking about websites : why isn't it so when speaking about data in our catalogues ? Searching and browsing proves efficient when we look as documents, like Web pages or PDF files. In order to expose library data, we need to apply these principles to the data itself : that's the point of the semantic Web and Linked Data.

Let's take an example. Nowadays, every library has her website, with links to other pages or websites, making it a part of the global Web. However, the library's data, like the catalogue records, are usually part of the “deep Web” : they are stored in databases, and the only way to access them is through a query form, thus making it impossible for search engines and Web crawlers to access the data. If the user wants to access this data, he'll have to display the page and query the database through the online form. If the resources of interest to him are scattered in several databases, he'll have to repeat this operation for each one.

Of course, the issue has been known for years, and libraries have already elaborated different technologies in order to address it.

1. Interoperability and cultural data : a complex situation

In libraries, most interoperability models have been built based on the fact that the objects they manage are multiple (there are several holdings for the same book, held by several libraries). The main challenge here was to avoid redundant cataloguing effort, which is achieved by transferring records from one catalogue to another. In order to do so, libraries needed to share the same data format, or they had to build complex crosswalks from one format to another (Marc21 to Unimarc, for instance.)

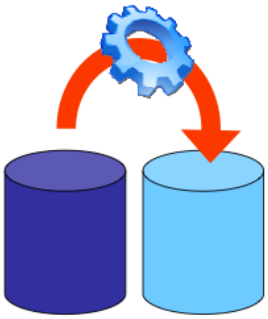
The Z 39.50 protocol was developed in the 1980's, then followed by a next generation relying on Web standards (SRU/SRW). These protocols allow to query several databases synchronously. However, their use is still limited to the library domain, and as it hasn't been implemented by others, it doesn't help to bridge the gap with museums and archives. Moreover, synchronous querying raises technical issues such as query detail levels, duplication of results, response times, etc.

Z 39.50 is still very useful in a professional context, for record exchanges between libraries. This practice creates a type of interoperability we will call “map and duplicate interoperability” : if the resources are stored in different formats, a complex mapping allows to transform them with a minimum data loss (but no loss is impossible when converting the data) and store them in a single database, this database being used for providing the user with query services.

[*Addition from the author – not in the french version*

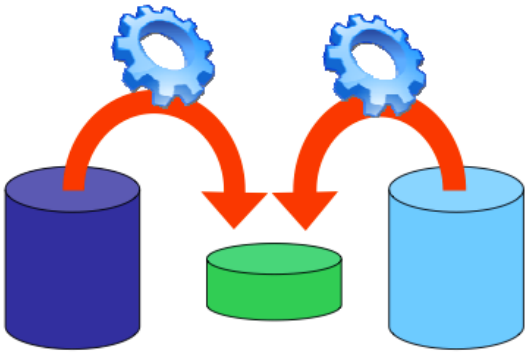
Of course, this kind of synchronous protocols can be used to create a portal for the end user and query the databases in real time. But in that case, all the process is managed by the application – the portal – that handles the service. If the portal becomes obsolete – or too expensive ! - and you need to change technology, you'll have to start the work from scratch because the intelligence of the interoperability relies in the application, not in the data. This model could be called “application-based interoperability”. Besides, it still raises the issues listed above : query detail levels, duplication of results, response times, etc.]

Illustration 1: Map and duplicate interoperability



The OAI-PMH protocol, created in the 1990's, proposes a different way of solving the issue. As it emerged from the open access movement, the need for a convergence between research data (in particular preprints stored in open archives) and library data or other cultural data was taken into account from the beginning. In order to guarantee interoperability between data from different domain, the protocol requires the use of the Dublin Core standard, following the simple XML expression of this vocabulary. The data is harvested and ingested in a new database, which is used to provide the user with access to the data. This is interoperability built on the smallest common denominator.

Illustration 2: smallest common denominator interoperability



Again, there are drawbacks to this approach. The data has to be simplified in order to fit in the common format, which leads either to data losses, or to concatenation of various data in a generic field, which will be more difficult to process. The Dublin Core itself is not to be blamed here, it's the way the OAI-PMH uses it that leads to that kind of consequences. For example, regarding library data, OAI Dublin Core generally looses all the links from bibliographic data to authority data [*Note from the author : these links actually exist in most french catalogues, they don't rely only on headings to reference authorities*]. Hence all the information is presented flat and without links or granularity levels.

Moreover, when it comes to aggregating data not only from libraries, but also from museums and archives, the differences between their models have to be taken into account.

The library data model relies on two basic concepts : bibliographic records for describing documents, and authority records for describing entities (persons, places, organizations, concepts, etc.) shared by several documents.

The archives data model emphasizes the concepts of hierarchy and context. The EAD format, based on the ISAD-G description model, allows to represent finding aids as a hierarchy of components, thus respecting the organic nature of archival “fonds”. Notions such as author, or title, are less important in an archival context than in the library domain.

Finally, museum data is characterized by the fact that it mainly describes unique objects. The objects are usually described, not only in regard of their context of creation like in libraries and archives, but also with respect to the different events that may have affected their lifecycle as museum objects, starting from their creation, but also encompassing their preservation, restoration, changes in ownership, and presentation in exhibitions. The event concept becomes central in the museum model : it's the case in the CIDOC's CRM model.

These differences in data model within the cultural heritage community make convergence even more challenging. When trying to merge these data, the specificities of the different domains are often lost, and common services are reduced to their minimal expression. Moreover, these interoperability styles fail to take into account the Web model as we described it earlier : they still suppose that the user will come to the portal, and require that he already knows what he's looking for before he can start searching it...

2. Linked Data and interoperability based on links

The Web of data proposes a style of interoperability which doesn't rely on synchronous query of separate databases, nor on reducing databases into a common format, but on the creation of a global information space, using links to browse seamlessly between resources.

The Web of data, or Linked Data, is an extension of the actual Web allowing to create a global information space, beyond Web pages, for data. Best practices for the Linked Data have been established by Sir Tim Berners Lee and adopted by the SWEO (Semantic Web Education and Outreach) group from W3C and they include four major rules :

- use URIs as names for things : every resource about which you want to assert something needs to be named with a Web identifier, a URI;
- use HTTP URIs so that the URIs can be dereferenced to retrieve data: URIs should be actionable so that looking up those names leads to the actual data about the resource;
- when dereferencing a URI, provide useful information using Web standards such as RDF, SPARQL
- Create links between datasets : you need to link your data to other existing data, so that it can be found by browsing those datasets, and build an ecosystem based on links.

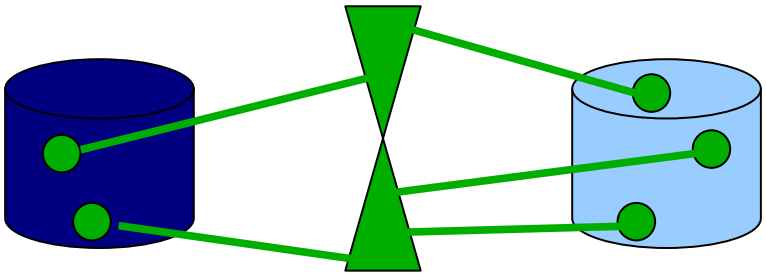
The overall objective is to create a global information space where the data is described following a common model : RDF, and linked by active links which can be used by machines as well as humans. As an inherent feature of the RDF model, the links are qualified, which means that it is possible to know the exact nature of the relationship between two resources : similarity, aboutness, other. Using this approach, it is possible to create links between resources that are described following different models, as long as the common grammar is RDF.

Two interoperability models derive from this new way of handling data : hub-and-spoke interoperability, and follow-your-nose interoperability.

Vocabularies will play en important role in the Linked Data, especially when it comes to linking data from several different domains. On the Web, the user can browse information without any knowledge of the underlying technical structure, and the browsing experience is seamless even when linking from one website to another. Similarly, with Linked Data, it should be possible to browse datasets, and link from one dataset to another, even if they are stored in different places and in different formats.

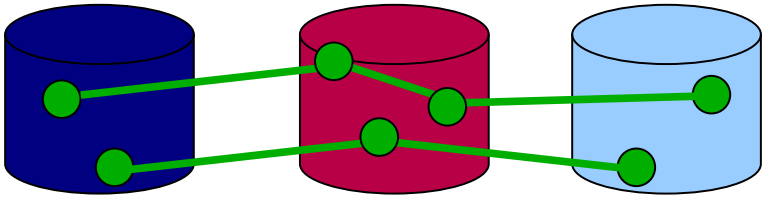
Vocabularies are thus associated with the hub-and-spoke interoperability model : they act like a hub, or a spine, for anchoring data expressed according to different data formats. In the Linked Data, this hub makes it possible to navigate from one dataset to the other by following links, or URIs, even if the data that is anchored to the hub is diverse.

Illustration 3: Interoperability based on links : "hub and spoke" model



Moreover, in the Linked Data, any dataset providing resources that can be reused by others can become such a hub. There isn't a centralized, unique hub, but rather several hubs connected together. Browsing through these hubs and the data that they link together allows to discover new information in an intuitive way, « following your nose » as the inventors of Linked Data themselves stated.

Illustration 4: Interoperability based on links : "follow your nose" model



3. Vocabularies and the Web of data

Vocabularies hence form the basis of interoperability based on links. In its final report, the Library Linked Data incubator group of the W3C (LLD XG) defined 2 different types of vocabularies : metadata elements sets, and value vocabularies.

A metadata element set defines classes and attributes used to describe entities of interest. In the linked data terminology, such element sets are generally made concrete through (RDF) schemas or (OWL) ontologies, the term “RDF vocabulary” being often used as an umbrella for these. Usually a metadata element set does not describe bibliographic entities, rather it provides elements to be used by others to describe such entities. Dublin Core metadata terms are a good example of such a vocabulary : they provide a set of classes and properties to describe bibliographic entities (examples of classes : Agent, Document... examples of properties : Creator, Format...) In some way, metadata element sets contribute to the creation of links in the Linked Data space, by allowing to share types of things and relationships between them. The Vocabulary Mapping Framework project (http://cdlr.strath.ac.uk/VMF/) was an interesting application of this principle : different metadata element sets were aligned with a common matrix in a hub-and-spoke fashion, and the matrix helped manage crosswalks between them.

Value vocabularies are sets of terms organized according to a knowledge information system or KOS. A value vocabulary defines resources (instances of topics, art styles, authors...) that are used as values of elements in metadata records. A value vocabulary thus represents a “controlled list” of allowed values for an element. Examples include the LCSH (Library of Congress Subject Headings), or the code list for languages ISO 639-2, both published as RDF datasets on the website http://id.loc.gov maintained by the Library of Congress.

The Linked Data principles imply that if 2 datasets, described in different formats, share the same value vocabulary, it will be possible to browse from one dataset to the other through the links to this vocabulary. The only feature that is required in order to do so is the use of HTTP URIs.

In libraries, this navigation principle already exists between bibliographic and authority records, when they are actually linked (and not only relying on name strings to maintain the consistence between them). The evolution of cataloguing models and rules towards FRBR, FRAD, FRSAD and RDA also implies to identify more information that can be shared between records through the use of links, and also within a single bibliographic entity. In these new models, a growing number of entities can be considered as nodes for links, shared between different records : works, expressions, persons, corporate bodies, families, subjects... These entities are also the ones that are easier to share with other communities outside the library domain, thus contributing to expose library data on the Web.

In archives and museums, the data is focused on describing unique objects. Maybe for this reason these communities haven't had an early interest in authority data. However, a similar concept is developed today within archives with the EAC – encoded archival description. In museums, value vocabularies for describing graphical objects include thesauri and classifications (e.g. thesauri from the Getty for subjects, places, artists... and the Iconclass classification system).

New data models focus on developing links between entities, in order to facilitate the discovery of new resources by following those links. Entities such as persons, events, places, concepts are key in this regard, and they correspond to library authority data. They can be shared in a cross-domain way. Classifications including Iconclass, Dewey, UDC... act as a specific type of vocabularies, and as they use coded values based on numbers, they provide a framework for multilingual services [Dunsire, 2010].

Shared use of common vocabularies creates hub-and-spoke interoperability without any additional developments. For instance, let's consider the bibliographic data from the National library of France. This data contains a reference to a high level classification scheme described using the Dewey classification. Converting this reference to URIs provided by the http://dewey.info is very easy, as the URIs are

constructed on the numbers (ex. for french fiction literature : http://dewey.info/class/843/). This single link makes the BnF data part of the Linked Data, and links are possible with other datasets connected with dewey.info [Wenz, 2010].

Other datasets that are not designed for being used as value vocabularies are so popular that they end playing a similar role. If someone publishing a dataset decides to reuse existing URIs rather than mint new ones, he naturally reaches the follow-your-nose interoperability by allowing to browse seamlessly from his dataset to the one he reuses. Dbpedia, the RDF extraction of Wikipedia created by researchers from Frei Universiteit Berlin and Leipzig University in Germany, is such a hub for the Linked Data : being an encyclopedic resource, Dbpedia is often chosen by data owners as the first anchoring point for their data. If a library decides to use Dbpedia URIs instead of local URIs, it will be possible to navigate from this library's data to Wikipedia, and also to the data from any other institution, library, museum, archive, that would have made the same choice.

Finally, the same kind of crosswalks can be created through the alignment of different vocabularies. A french example : the “Thésaurus W” published by french national archives as Linked Data (http://www.archivesdefrance.culture.gouv.fr/thesaurus/). This thesaurus provides links to RAMEAU, the french subject headings maintained by the BnF. Hence it would be possible to link an archival resource and a library book through these two thesauri and the links they share.

4. Examples

Here are a few examples building applications on these principles in order to facilitate data convergence.

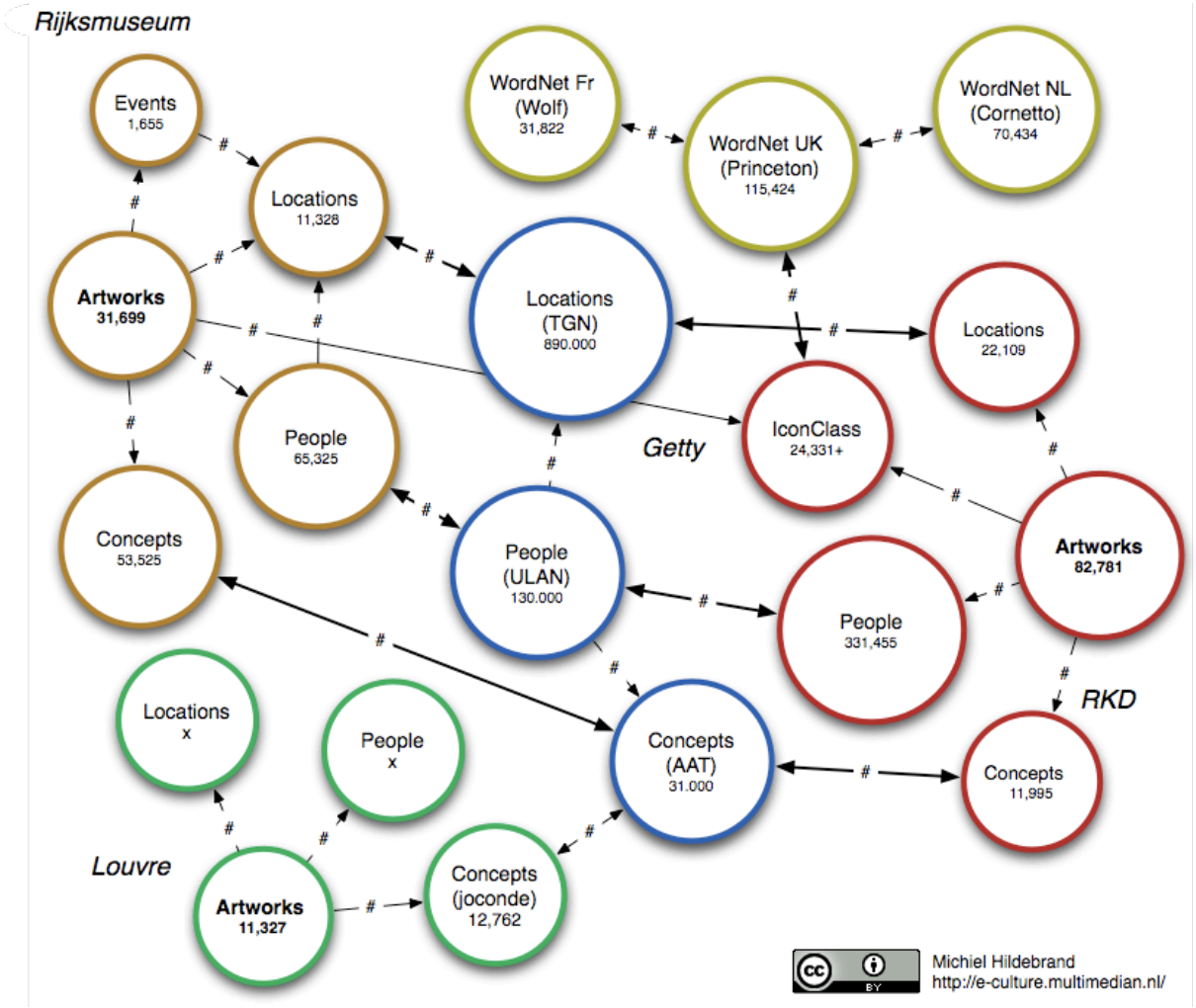
Europeana

The Europeana project (http://www.europeana.eu) has been created to federate data from libraries, archives, museums and audiovisual archives. The first Europeana portal prototype was built upon the ESE standard (Europeana Standard Elements), an extension of Dublin Core elements, to provide a first attempt at interoperability in a smallest common denominator style. Additions to the simple DC include mainly provenance and rights information, as well as links to primary digital objects on stakeholder institutions' websites.

In complement to that first experiment, Europeana developped the EDM model (Europeana Data Model), a high level ontology aiming at applying Semantic Web principles to cultural heritage data, and at creating a semantic layer of information to link the different datasets aggregated by the digital library [Doerr, 2010]. EDM allows to aggregate document-centric descriptions as well as event-centric descriptions, thus fitting one of the main convergence issues. The semantic layer, i.e. the set of interlinked vocabularies describing persons, places, concepts, etc. is responsible for providing links between resources.

A prototype demonstrating these principles can be seen in the Europeana Labs (http://eculture.cs.vu.nl/europeana/session/search). This prototype aggregates data from the Rijksmuseum in Amsterdam and the Musée du Louvre, from the Joconde database (french Ministry of Culture) and from the Rijksbureau voor Kunsthistorische Documentatie (Netherlands Institute for Art History) in The Hague. Thesauri are provided for places (The Getty Thesaurus of Geographic Names), persons (ULAN –

The Union List of Artists Names), concepts (WordNet et AAT – Art and Architecture Thesaurus). Links are also created from the use of the Iconclass classification system.



The Pompidou Center

This example demonstrate that the interoperability strategy based on links can also prove useful in an institutional context. Since 2007, the Pompidou Center, major modern art museum in Paris, developed a new digital strategy aiming at providing a global platform for online digital content, the Virtual Pompidou Center. This platform provides access through a unique access point to the whole digital production of the organization and associated institutions (Bpi, Ircam) : digitized works of art, documents about art and art history, videos and podcasts, library books records, etc. Art content (works from the museum, recording of musical performances) is linked with event-based information (exhibitions, performances, conferences) and with other relevant resources (posters, photographs, books, archives, etc.) thus allowing to browse the website and discover all these resources in a serendipitous manner.

One of the main challenges of the project lies in the need for creating a global and common information space, with links, from data extracted from several databases which all have their own structure (EAD for archives, MODS and DC for libraries, local systems for museum contents and audiovisual material). In order to handle this, an RDF ontology was created, and binds around a few main concepts (work, document, person, place, collection, event and resource) all the data from these different databases.

The user interface of the Virtual Pompidou Center demonstrates the added value of these technologies to the end user.

5. Issues

However, there are still a number of issues regarding the creation of a global information space for libraries, museums and archives.

The first barrier lies in the mapping of original data to RDF. Adopting the RDF model leads to questioning existing models, partly because some information may be difficult to express using triples, partly because the triple model opens new possibilities and challenges the former models. For libraries, for instance, the Semantic Web model based on links naturally leads to compatibility with the FRBR model (and other FR** models). However, current catalogues are sitting on huge masses of existing data that doesn't fit this model, and in particular, a lot of links are missing in order to straightforwardly convert library data to RDF [Koster, 2011]. Similar issues will be raised regarding the hierarchical EAD model, which has to evolve towards a graph model.

Another issue is the attribution of URIs. In RDF, each resource that needs to be described, and each relationship binding resources together, must have a URI, a persistent and actionable Web identifier. So that raises the issue of identifier maintenance.

With respect to metadata elements sets, organizations in charge of publishing standards (such as IFLA regarding the FR** standards) should convert their standards to Linked Data and provide URIs for each elements, so that they can be reused in other contexts. If these organizations don't commit to do it themselves, it may be done by others in a non-authoritative way, as had begun to happen with FRBR. IFLA has now started this work within the cataloguing section (FRBR review Group, ISBD/XML group), under coordination of the Namespaces task group.

Regarding the datasets themselves, URIs are also needed for documents, for persons, but also for concepts that are declared in value vocabularies. These value vocabularies must not only provide name strings, as it is often the case, but must provide each resource with a URI in order to manage it in an interoperable way and allow the creation of links.

A set of usual issues are associated with this task : how to make URIs persistent, should URIs contain human-readable strings... [Bermes, 2009]

The modeling of value vocabularies in RDF is also a challenge in some respect [Vatant, 2010]. The work of the LLD XG shows that library authority data have a dual nature : they are both lists of “names” or labels, with recommendations on how to name specific kind of entities, and knowledge systems providing information on real-world entities like persons. Complex models as the one elaborated by VIAF are

necessary in order to render this situation : the information on the person and the information on the label are treated as distinct elements.

This complexity is not a pure modeling issue. In the Linked Data, consistency is required from the alignment point of view. As stated before, in order to be part of the Linked Data, datasets need to provide links to other existing datasets. When one wants to publish data, he has 3 main preoccupations :

- identify common assets between his data and existing data,
- identify the precise nature of the relationships between his data and the data he wants to create links with,
- develop the best way to create those links, manually or automatically.

The representation of links or relationships between datasets, in particular value vocabularies, is sensitive in terms of logic [Bergman, 2010]. It is not sufficient to declare a strict equivalence between resources (for instance, Victor Hugo in the BnF dataset and Victor Hugo in Dbpedia) : the owl:sameAs property, fit for this kind of relationship, implies that all information that is true for the first URI will be true for the other too. In this context, providing information such as the creation date of the authority record for Victor Hugo in the BnF catalogue would generate an inconsistency, because this date is not applicable to the Dbpedia resource.

For this reason, some metadata element sets provide looser equivalence relationships, such as skos:exactMatch and skos:closeMatch for managing automatically generated alignments, which are not always 100% trustworthy. There are generic similarity relationships such as umbel:isLike or rdf:seeAlso. Finally, datasets can also be linked with one another through properties like dc:creator or dc:subject, according to the follow-your-nose interoperability model.

How to generate the alignments when we're talking about huge masses of data such as the ones held by library catalogues? Ideally, they should be created automatically, but often, when it comes to bridging between different topics or domains, string matching is not enough. At the Pompidou Center, for instance, we observed that string matching on first and last name proves efficient for artists names. But it is impossible to align events based on their titles, because of the disparate information provided by the various sources (“la subversion des images”, “exposition la subversion des images”, “exposition au Centre Pompidou : la subversion des images”, “subversion des images 2010” etc.) In that case, automated processes generate a lot of duplicates, and even worse, a lot of false positives (“conférence dans le cadre de l'exposition la subversion des images”). Even when speaking only about names of persons, huge datasets like VIAF are confronted to a number of homonyms. Likewise, the best way to ensure smooth alignment of similar entities from various datasets is the availability of unique keys : identifiers that may not be URIs, but are standardized enough to be reused in different contexts. ISO identifiers such as ISSN, ISBN, ISNI will thus continue to play a major role in the Linked Data, and it is vital for local catalogues to make sure that they are available locally, as a preliminary stage to future alignments. Other strategies are built on links : in VIAF, matching algorithms take into account bibliographic data linked with each authority data in order to increase the level of confidence of automated alignments.

6. Conclusion

Linked Data is a matter of institutions publishing their data, but even more, of creating links between datasets in order to build interoperability in a hub-and-spoke and follow-your-nose fashion. Thesauri, authority data, knowledge organization systems of all kinds are more than ever needed to serve as hubs

between datasets described according to different data formats. This new interoperability style, built on Web principles, will allow to offer our users new services, more adapted to serendipitous discovery.

7. Références

[Bergman, 2010] Mike Bergman, « Bridging the Gaps: Adaptive Approaches to Data Interoperability. » Keynote presentation for DC 2010, Pittsburgh, Octobre 2010.
<http://dublincore.org/workshops/dc2010/DC-2010_20101022_Bergman_keynote.pdf>

[Bermes, 2009] Emmanuelle Bermes, « Linking Open Data : a case for releasing library data on the Semantic Web. » Emerging trends in technology, IFLA satellite conference, Florence, August 2009.
<http://www.ifla2009satelliteflorence.it/meeting3/program/assets/EmmanuelleBermes.pdf>

[Doerr, 2010] Martin Doerr, Stefan Gradmann, Steffen Hennicke, et. al. « The Europeana Data Model. » IFLA 76th congress, Göteborg, August 2010.
<http://www.ifla.org/files/hq/papers/ifla76/149-doerr-en.pdf>

[Dunsire, 2010] Gordon Dunsire, « Not just numbers on shelves: using the DDC for information retrieval ». EDUG meeting, Alexandria, April 2010.
<http://www.slainte.org.uk/edug/meetings.htm>

[Koster, 2011] Lukas Koster, « Missing links : the challenge of generating linked data from legacy databases. »Common Place, March 28th 2011.
<http://commonplace.net/2011/03/missing-links/>

[LLD XG, 2011] Report from the Library Linked Data W3C incubator group (in progress).
<http://www.w3.org/2005/Incubator/lld/wiki/Main_Page>

[Vatant, 2010] Bernard Vatant, « Porting library vocabularies to the Semantic Web, and back. A win-win round trip ». IFLA 76th congress, Göteborg, August 2010.
<http://www.ifla.org/files/hq/papers/ifla76/149-vatant-en.pdf>

[Wenz, 2010] Romain Wenz, « data.bnf.fr: describing library resources through an information hub. » Semantic Web In Bibliotheken 2010, Cologne, October 2010

About the Author
Emmanuelle BERMES, head of Prospective & data services at the National library of France (Bibliothèque nationale de France, BnF), has been working at BnF since 2003, first in digital libraries, then in metadata. These past two years, she's been involved in a semantic Web project, data.bnf.fr. Emmanuelle Bermes is a member of IFLA IT section where she's contributing to the creation of a Semantic Web special interest group (SIG). She's also a co-chair of the Library Linked Data W3C incubator group.