IFLA WLIC 2019 Athens

Submitted on: 26.09.2019

Title of the Satellite Meeting: Data intelligence in libraries: the actual and artificial perspectives
Date: 22-23 August 2019
Location: Deutsche Nationalbibliothek, Frankfurt, Germany

Leaving Comfort Behind: a National Union Catalogue Transition to Linked Data

Bodil Wennerlund
Information Systems Department, National Library of Sweden (Kungliga biblioteket), Stockholm, Sweden. bodil.wennerlund@kb.se

Anna Berggren
Information Systems Department, National Library of Sweden (Kungliga biblioteket), Stockholm, Sweden. anna.berggren@kb.se

Abstract:

The last few years the National library of Sweden has been involved in a major infrastructure project. In June 2018, a new version of Libris, the Union Catalogue, was deployed. Libris (XL) is a custom-built open source system based on Bibframe 2.0 and linked open data. With linked data at its core, it creates new ways of opening up data previously locked away in library-specific formats and structures. Thus the information in the library catalogue can be more easily understood by the rest of the web. The catalogue data can be linked to information both inside and outside the library domain. The content of the Swedish Union Catalogue is a joint effort of cataloguers working in more than 500 member libraries. Data is collaboratively edited within the central system. The transition to new Libris was complex and involved migrating 10 million bibliographic records from MARC to BIBFRAME. The participating libraries have their own local library systems and are dependent on MARC data, why a transition back to MARC from BIBFRAME is necessary before the library systems can digest the incoming data. Because of that all of our import and export routines are still based on MARC. The cataloguing client is the main surface where we expose the new forms and discuss improvements. Cataloguers working in Libris have taken on a major challenge. The new cataloguing editor is format driven and introduces a completely new way of representing the information in Libris. Adopting our new means of description requires experience and we needed to get the new system in place and improve structures within it. The advantages of linked data has not been immediately obvious. Huge efforts has been made by the cataloguers working in Libris as well as by staff at the National library to improve the system. Extra support has been offered, a lot of expertise has been invested in the preparation of

technical documentation, help texts, instruction movies, experts has participated in developing new features and loads of user feedback has been given to the development team.

Keywords: National union catalogue, Sweden, BIBFRAME, Linked data, Software development

---

In June 2018, the National Library of Sweden (KB) launched a new generation of LIBRIS, the Swedish union catalogue. The new LIBRIS, named XL,[1] is a custom-built open source system based on BIBFRAME 2.0 and linked open data. The system is developed at KB by a team of developers, designers and specialists on data modelling and metadata. With linked data at its core, Libris XL creates new ways of opening up data previously locked away in library-specific formats and structures. The information in the library catalogue can be more easily understood by the rest of the web, and link to information both inside and outside the library domain.

This paper focuses mainly on the transition to LIBRIS XL, particularly on the launch and implementation from 2017 to date.

1.1 Libris – the union catalogue of Sweden

Initiated in the 1970s, Libris is the joint effort of more than 500 Swedish university and special libraries, the National library and, from 2011, public libraries. Data is collaboratively edited within the central system and then exported to local library systems. This way, a publication is catalogued only once and then made available to all Libris libraries, who only need to register their holdings. The National Library (KB below) is responsible for the development, maintenance and support of Libris. Libris currently stores around 11 million resources, including the Swedish national bibliography and authority file. Also, the member libraries have added ca 35 million holdings related to the titles. It is worth noting that primary cataloguing is not the most common cataloguing work flow as most records are created in Libris via imports from book vendors and other partners. Cataloguers enrich basic records already stored in Libris, alternatively download records from other sources such as the British National Bibliography and modify them. Also, the number of e-resources has increased substantially in recent years in relation to books and other physical resources. For example, around 100,000 bibliographic records describing e-resources are imported each week.

Libris also offers a public OPAC[2] and an inter-library lending system.

Over its 47 year lifespan Libris has run on different platforms – 2002 to 2018 on Voyager® from Ex Libris. The internal format of Voyager® was MARC21 which triggered a conversion of the Libris data from legacy format LIBRISMARC to MARC21 in connection with the migration. Librarians used the Voyager® desktop cataloguing tool to edit Libris, with an interface largely in MARC21. Many Libris cataloguers are very experienced with

---

[1] LIBRIS was first launched in 1972 and celebrated its 40th (XL) anniversary in 2012 when prestudy activities around the new system were in progress.

[2] The Libris OPAC http://libris.kb.se/ is the public search interface of Libris, including information on which libraries hold the title. The OPAC is populated with data from LIBRIS XL. Its user interface has not been redesigned within the LIBRIS XL project.

MARC21 and know exactly how a Libris record will land in their local system.

This was all about to be challenged by LIBRIS XL.

1.2 Libris at a crossroads

Around 2008, KB began to seriously consider the future of Libris. Ideas related to linked data, Web 2.0 and the semantic web began to develop in the wider world. The Libris content had become increasingly diverse and the MARC model was not suitable for describing new types of data, e.g. e-resources. The existing platform based on MARC21 was difficult to develop and maintain, and proposals for new services around Libris were put on hold. Historic layers also meant that old and new records had different structures. Performance issues were becoming problematic due to the architecture of the system. In effect, the Libris data was locked into a system past its prime.

Meanwhile, the MARC format was being questioned within the library community. In its 2008 report *On the record* the Library of Congress stated that MARC is "based on forty-year old techniques for data management and is out of step with programming styles of today."

As a consequence, KBs architects and developers began experimentally exploring ideas for a new system. A few high-level objectives were defined early on:

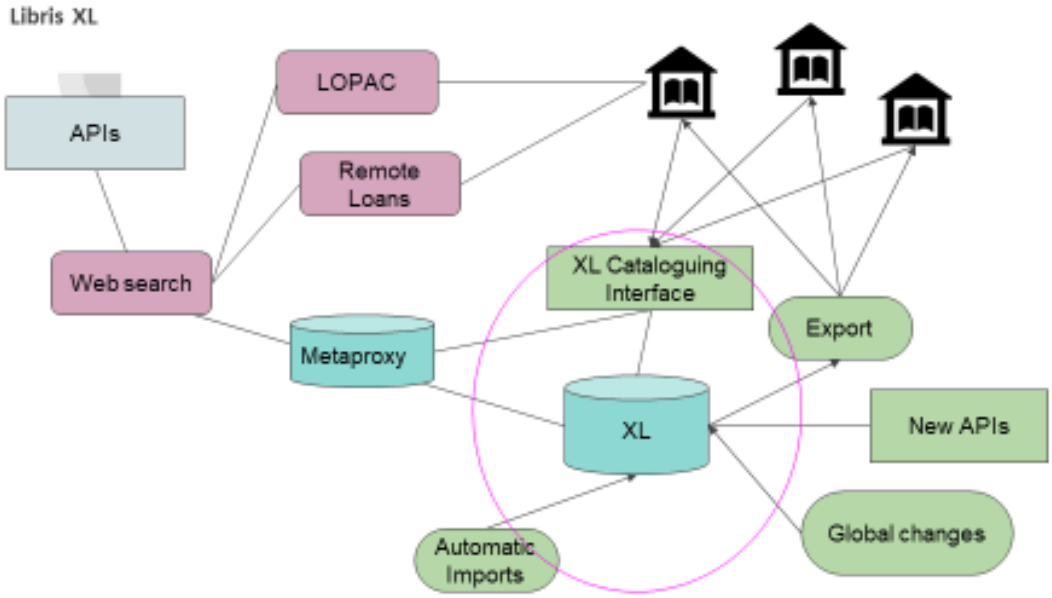
- A new, structured data model to efficiently describe new and diverse materials, e.g. electronic publications, music, or films.
- The new system should be built on linked data
    - Linked data will open up library material to users, applications etc. inside and outside the library community.
    - Linked data can be used to incorporate data from other systems and contexts into LIBRIS.
- The new system must be able to handle the rapidly growing data volumes.
- The new system must be an open system –via open APIs[3] and publicly available code (open source).

To realise LIBRIS XL as a Linked data system, KB turned its attention to the BIBFRAME initiative aiming at replacing MARC. In 2012, the Library of Congress released a draft version and the design process of LIBRIS XL began. However, it was not until the 2016 release of BIBFRAME version 2.0 that the architecture of the new system fell into place and a version of BIBFRAME 2.0 adapted to LIBRIS XL, with a Swedish vocabulary was modelled.

KB now decided it was time to take action and launch an implementation project.

The new system and project was named LIBRIS XL.

[3] Application programming interfaces (API) offer other systems ways of communicating with Libris and using its data.



Architecture of LIBRIS

1.3 The implementation project

A radical plan was conceived to abandon MARC21 in favour of Linked data with JSON-LD (RDF) as a storage format in the new LIBRIS XL core.[4] However, since large parts of the library community are likely to use MARC for years to come, KB still has a responsibility for making data available to libraries in a form they can use. Also, many of the surrounding library and discovery systems are complex with modules working on MARC data. These modules may not be changed in the near future. In fact, at the moment no local library system in production in Sweden receives linked data or records in JSON-LD from LIBRIS XL.

As a consequence, the new system must be able to import and export data in MARC21. The LIBRIS XL project conceived a transformation mapping intended to provide local systems with LIBRIS XL data **similar** to the old system. However, it was clear from the start that the MARC21 transformed from LIBRIS XL would not be identical to the old MARC21 records. For example: BIBFRAME does not use punctuation at the end of elements which means that is impossible in some cases to add punctuations in the transformation.

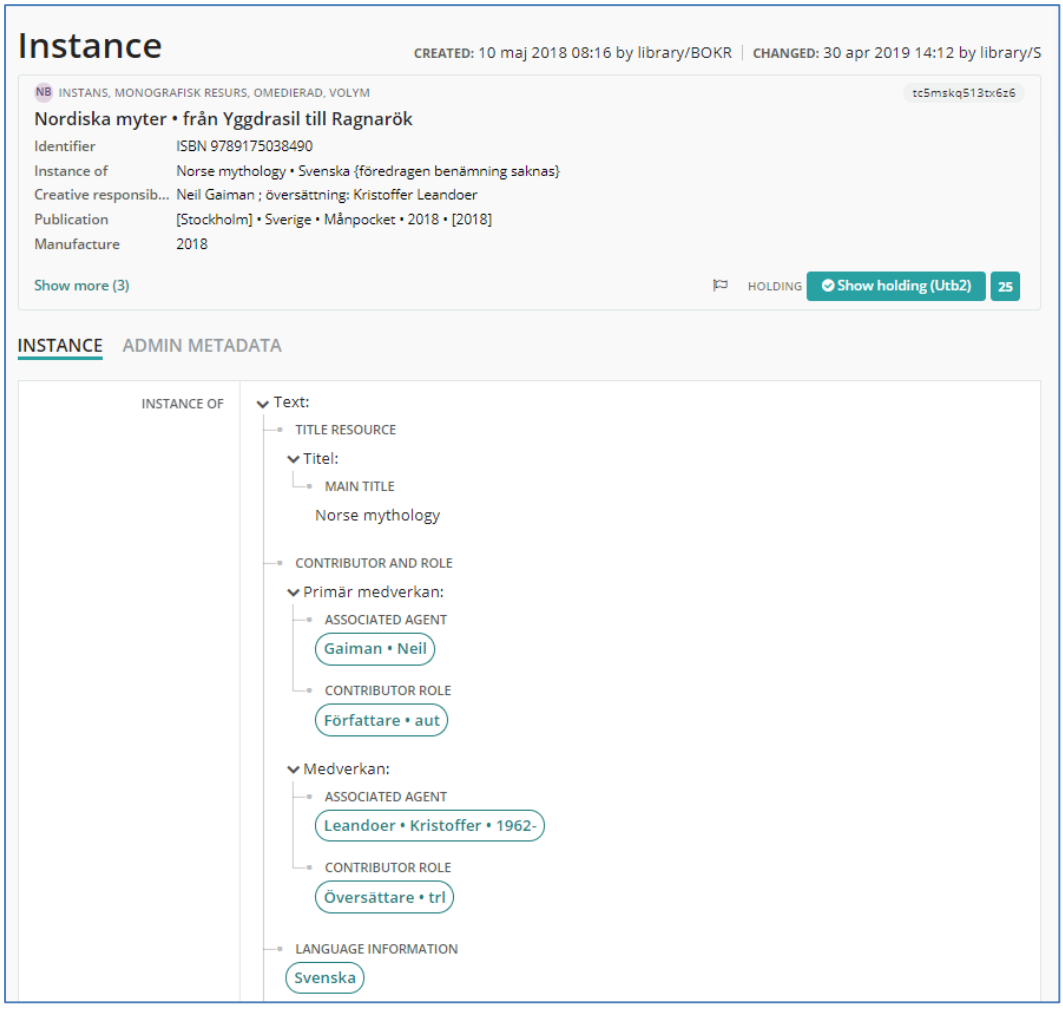
Lessons learned from the transformation mapping are

- Local library systems interpret the MARC21 data very differently which means that the project in some cases has had to implement library specific export transformation mapping from LIBRIS XL, and
- Mapping activities (analysis, development and testing) were more labour intensive than estimated and required expert resources on metadata, local library systems and MARC21 and BIBFRAME.

[4] JSON-LD (JavaScript Object Notation for Linked Data) is a method of encoding Linked Data. JSON-LD is mapped to RDF, the Resource Description Framework.

The migration and transformation of MARC21 data to LIBRIS XL were a big part of the implementation project. KBs metadata experts provided guidelines for the migration, one was to leave certain metadata behind (but not discard it), and that quality enhancing activities should be on improving functionality, not to make it look nicer.

1.4 What about the cataloguers?

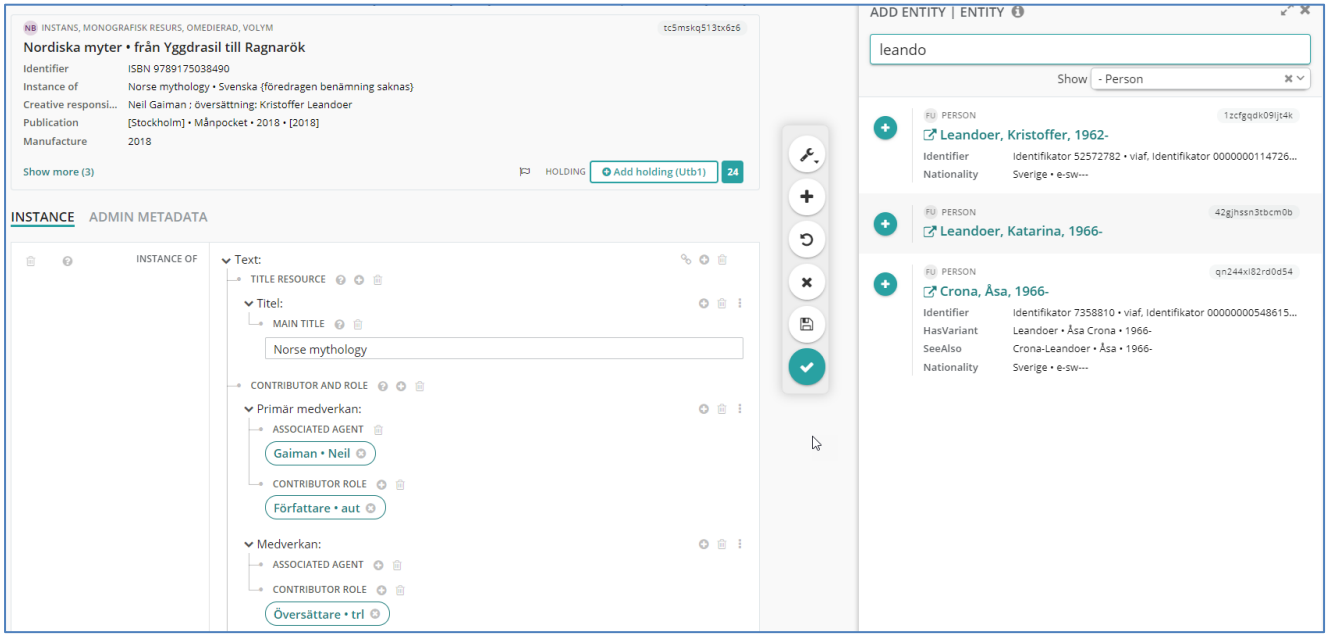


A monograph *Nordiska myter* in LIBRIS XL. Links are indicated as pills.

A new web-based cataloguing tool was designed by KB, with a model-driven (JSON-LD) interface providing a fairly unfiltered view of the LIBRIS data. As a consequence, bibliographic information, holdings and authorities are presented in a completely new way to the LIBRIS cataloguers who are familiar with MARC21. The cataloguing interface is open and available for browsing at libris.kb.se/katalogisering, but only Libris users can edit in the tool. The cataloguing interface is not intended as a discovery or circulation tool and the existing Libris OPAC is still live at libris.kb.se – populated with data synchronized from LIBRIS XL.

Cataloguers working in Libris have met with a major challenge. The new tool is a radically different experience from the previous MARC21 tool. Since it is format driven it introduces a completely new way of representing the information in Libris. Adopting the new means of description requires experience and KB needed to get the new system in place and improve structures within it. The BIBFRAME data takes up more space than the previous, very compact, MARC21 records. Cataloguing in LIBRIS XL requires linking to authorities,

subject headings etc. This in its turn means scrolling and clicking on links, as opposed to typing.

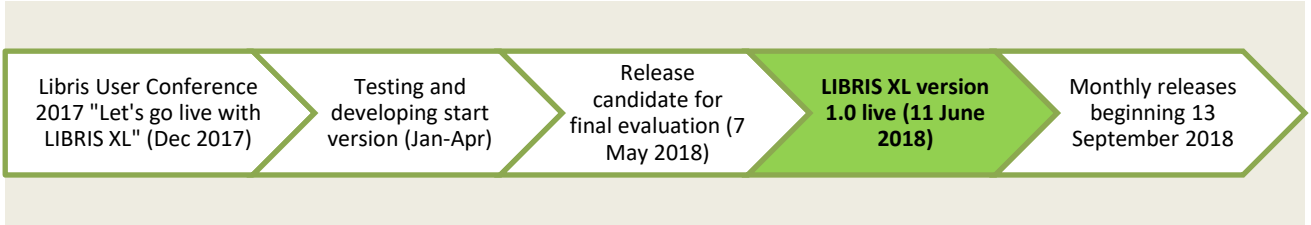


Linking to an authority in https://libris.kb.se/katalogisering

1.5 The big leap - launching LIBRIS XL

In 2017, the project governing structure, consisting of KB's internal steering group and the Libris development council, decided it was time to launch LIBRIS XL. A beta version of the cataloguing tool was made available for those who wanted a preview. However, a lively discussion within the LIBRIS network led to expressed concerns that the new system was not yet functional or robust enough to go live. Many stakeholders would also have preferred to keep the old system in production parallel to LIBRIS XL, which would enable a more cautious shift, and more time to get used to the new format. However, parallel operation of two Libris systems was not considered realistic by KB due to, primarily, the massive task of maintaining data flows to and from two systems.

A big concern was whether the new editing tool would be functional for librarians to keep up their processes and not fall behind cataloguing incoming resources. After evaluating the beta version many libraries had serious doubts – the tool appeared to have inadequacies and lacked some functionalities. KB responded by defining the new tool as a **start version** and indicated that new features would be released very soon after the initial launch. For example: although the start version only had cataloguing templates for basic monographs, serials and holdings additional templates would be added in upcoming versions. The spring of 2018 was spent adding important cataloguing features and refining XLs BIBFRAME format, mappings and integration flows.



Timeline of LIBRIS XL launch

After lengthy discussions on the stability of the system and the best, or least problematic, point in time for the system switch, it was decided that June 2018 would suit the libraries. The switch meant one week's downtime where neither old nor new LIBRIS would be open (users could still search the OPAC though).

A release candidate (the start version) was pushed to the Libris community for testing and evaluation in May 2018. Activities such as test sessions, workshops and product demos were organized by KB and the Libris libraries. Many libraries actively prepared for the big switch, setting up their own work groups to deal with potential issues caused by the new system.

4 June 2018 all cataloguing in LIBRIS was suspended and the Voyager® platform shut down for edit. The LIBRIS XL team then began the migration and transformation of more than 10 million records from MARC21 to BIBFRAME XL. Functionalities were deployed, including the new cataloguing tool and backend functionalities for imports and exports to the new system. LIBRIS XL was indexed for searching with the search engine *Elasticsearch*. The first editing in the new cataloguing tool was made by KBs National Bibliography Unit on 8 June 2018 to verify that the new system was operational. At the time of launch LIBRIS XL had cataloguing templates, mapping and transformation to MARC21, including MARC21 views of the records, exports to local library systems as well as documentation and user help.

When LIBRIS XL went live 11 June, the information in the catalogue was opened up to the rest of the web and KB became the first national library in the world to switch to cataloguing in BIBFRAME.

1.6 Communication and support

Professional communication, support and user assistance proved a key success factor early on in the LIBRIS XL implementation. In 2017, KB assigned a team to deal with these issues as part of the implementation. Activities included authoring help texts and guidelines for cataloguing in LIBRIS XL. Existing social media channels were used throughout to reach out and collaborate with libraries. The Libris blog was essential and used for weekly communications around LIBRIS XL, e.g. information on upcoming releases.[5]

An extended support team was set up, meeting daily to manage and assign incoming issues for action. The Libris support site was used especially for questions, suggestions and discussions. Since the launch of LIBRIS XL more than 1,000 discussions have been initiated on the Libris support site.[6]

It was decided initially not to produce traditional manuals for LIBRIS XL. Guidelines, help and user assistance would either be included in the new system, or published on social media channels. One example are YouTube instruction videos focussing on specific topics.[7]

1.7 Preparing for chaos

After years of discussions on the future of LIBRIS, the member libraries, including KB, found themselves facing the big shift. The mood was apprehensive, a mix of anticipation and awe. Would LIBRIS XL really be launched as planned? Cataloguers feared the inevitable

[5] The Libris blog is from June 2019 part of the new KB website https://www.kb.se/samverkan-och-utveckling/libris.html New content is added regularly.

[6] Supportforum för nya Libris https://kundo.se/org/librisxl/

[7] YouTube playlist *Instruktionsfilmer för Libris katalogisering* https://youtu.be/9MzVriPQUA8

drop in productivity and potential problems with quality and MARC mappings from the new system. There were known remaining issues waiting to be solved in the data flows from LIBRIS XL to local systems. However, many libraries had at this point decided to initiate activities to facilitate the LIBRIS XL implementation and make it less stressful.

The National bibliography cataloguing team at KB, although optimistic about the new system, decided to prepare for potential disaster by assigning a task force to test, evaluate risks and adjust the cataloguing processes to the new situation. As a result, productivity targets were temporarily dropped. By the launch of LIBRIS XL, they were ready for chaos. As anticipated, the team's productivity during the summer of 2018 was indeed only a fraction of their normal average. The usual focus on cataloguing excellence was abandoned in favour of exploring and learning. In parallel, the cataloguers created help texts, guidelines and instruction videos in cooperation with the LIBRIS XL development team and other experts. Problems and suggestions were coordinated and channelled in writing to the LIBRIS XL project team. In the end disaster and chaos did not occur for this team of experienced cataloguers and towards the end of 2018, they were back almost to normal productivity.

The Libris member libraries have of course met different challenges depending on size, processes, local systems, subject focus etcetera. Some larger libraries, like KB, assigned special teams to ease the implementation of LIBRIS XL, for example by organizing workshops and supporting learning and transfer of know-how. Some libraries needed to focus on further testing and evaluation of how data from the new system worked in their local systems including mapping issues from Bibframe to MARC. Some libraries lacked templates adapted to their type of resources and had to find workarounds, or decide whether to pause cataloguing of certain resources.

The impact of LIBRIS XL from the perspective of reduced cataloguing throughput varied between libraries. It depends on a number of factors, for example the types of material the library holds – the more diverse, the longer the period of low productivity.

1.8 Post launch – developing while in production

At the time of launch in June 2018, expectations on the KB development team to release additional functionalities without delay was great. Many libraries felt they had been pressured to accept the new system and now wanted guarantees that their most wanted features would be released soon. KB received lists of requirements, especially from the large member libraries.

In fact, KBs development team was already busy working on new features, templates and bug fixes. Beginning September 2018, new versions of LIBRIS XL have been released monthly. To make this possible, the team follows a strict agile (Scrum) process with biweekly sprint planning and daily scrum meetings.[8] Since all user input channelled via KB:s social media, customer support and e-mails is screened by the support team, only genuine candidates for development will be added to the project backlog. This has taken care of duplicate requests and issues not related to the LIBRIS XL system, as well as allowed the development team to stay focused. Issues for potential development are analysed, documented and categorised continuously based on feedback from team members, product owner and other stakeholders. The agile development process creates a method for incorporating new ideas and changing

[8] Scrum is a widely used agile framework for software development. https://en.wikipedia.org/wiki/Scrum_(software_development)

requirements as well as reconsidering prioritisations based on input from the member libraries.

Analysis and development of new features has in many cases been done jointly by designers, developers and experts/librarians in the Libris network. For example, templates for cataloguing special materials such as sheet music, maps or computer games have been developed and verified in close cooperation with experts from Swedish libraries, often in workshops.

A roadmap is available on https://www.kb.se/samverkan-och-utveckling/libris.html to indicate future development of LIBRIS XL on a high level.

1.9 LIBRIS XL one year on – where are we now?

The advantages of linked data have not been immediately obvious. Huge efforts have been made by cataloguers working in Libris as well as by staff at the National library to improve the system. Extra support has been offered to users, a lot of expertise has been invested in the preparation of technical documentation, help texts and instruction videos. Experts have participated in developing new features, and plenty of user feedback has been given to the support and development team.

The transition to LIBRIS XL was complex and involved migrating over 10 million bibliographic records from MARC. The participating libraries have their own local library systems and are dependent on MARC data, why a transition back to MARC from BIBFRAME is necessary before library systems can digest the incoming data. Because of that all of our import and export routines are still based on MARC. The cataloguing tool is the main interface to expose new structures and discuss improvements.

Some of the lessons learned were quite unexpected:

- The new system forced re-evaluating the processes: Why did we do this? What's the use of this piece of data? As a consequence, some requirements in the LIBRIS XL backlog are no longer valid, or could be implemented a different way.
- Developing the new system has inspired many interesting new collaborations in the Libris network. The potential to open up and expand collaborations around future development is great.
- The effect of new systems on people: "It's been tough, but very interesting how a new system has affected our organization." One example is more equality between junior and senior cataloguers.
- In the future, it is reasonable to assume that fewer cataloguers will need to do more work. It follows that manual work should be limited to the intellectual parts of cataloguing and that more automation needs to be implemented. What can be eliminated?

Future activities include modelling of LIBRIS XL so that the Works level of BIBFRAME can be introduced, bringing benefits such as improved data quality, search and linking opportunities. Ideas for opening up the library data to the rest of the world are emerging and libraries are considering their own processes.

At this time there is definitely a feeling that Libris belongs to the libraries: Libris is our system. There is also a sense of satisfaction in the development team that, even though the start version lacked functionalities, the challenge was overcome.

In summary, LIBRIS XL is not done, but live. Over time, the LIBRIS XL data will be reused more and more by stakeholders outside the library community.

References

Kungliga biblioteket. Libris information start page. 2019. https://www.kb.se/samverkan-och-utveckling/libris.html (downloaded 2019-07-15)

Kungliga biblioteket. 2018. *Startversion för nya Libris*. [blog] https://www.kb.se/samverkan-och-utveckling/nytt-fran-kb/nyheter-samverkan-och-utveckling/2018-05-07-startversion-for-nya-libris.html

Kungliga biblioteket. 6 July 2018. *KB becomes the first national library to fully transition to linked data. Pressmeddelande.* https://www.mynewsdesk.com/se/kungliga_biblioteket/pressreleases/kb-becomes-the-first-national-library-to-fully-transition-to-linked-data-2573975

Kungliga biblioteket. *Nationalbibliografin*. 2019. https://www.kb.se/om-oss/det-har-gor-vi/nationalbibliografin.html

Thomas, Barbro (red.) (2012). *Att bryta ny marc: LIBRIS 40 : en jubileumsskrift.* Stockholm: Kungliga biblioteket

The Library of Congress Working Group on the Future of Bibliographic Control. 2008. *On the record* https://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf

The Library of Congress. April 21 2016. *Overview of the BIBFRAME 2.0 model.* https://www.loc.gov/bibframe/docs/bibframe2-model.html