

Curating Data from/to Publishers in Latvia: paradigm shift in metadata reuse exploiting new Portal for Publishers

Dace Ūdre

Bibliography Institute, National Library of Latvia, Riga, Latvia
dace.udre@lnb.lv

Evija Krūmiņa

Bibliography Institute, National Library of Latvia, Riga, Latvia
evija.krumina@lnb.lv

Elza Ungure

Bibliography Institute, National Library of Latvia, Riga, Latvia
elza.ungure@lnb.lv

Ansis Garda

Bibliography Institute, National Library of Latvia, Riga, Latvia
ansis.garda@lnb.lv

Anita Goldberga

Bibliography Institute, National Library of Latvia, Riga, Latvia
anita.goldberga@lnb.lv



Copyright © 2019 by Dace Ūdre, Evija Krūmiņa, Elza Ungure, Ansis Garda, Anita Goldberga. This work is made available under the terms of the Creative Commons Attribution 4.0 International License: <http://creativecommons.org/licenses/by/4.0>

Abstract:

This paper describes how introduction of a new tool – the National Library of Latvia (NLL) Portal for Publishers (Portal) – has affected data management that involves publishers and the NLL who have been using it to perform day-to-day tasks since 2018. Data submitted and created in the Portal later are reused for the purpose of the national bibliography and union catalogue, national digital library, legal deposit copy distribution, gathering and distributing publishing statistics, doing research, management and remuneration for reprographic reproduction etc. The introduction of the Portal has strengthened the cooperation between publishers and the NLL as well as provided the technological basis for implementation and development of automated metadata creation and reuse, and other processes that save a lot of redundant manual work.

Keywords: Portal for Publishers; reuse of metadata; online services for publishers and researchers; national bibliography; publishing statistics

Introduction

Introduction of the Portal has not only improved online services for publishers and facilitated a dialogue with them, but also has fostered reuse of data whilst compiling official publishing statistics of the Republic of Latvia and running services such as the national bibliography and union catalogue and the national digital library. Based on the data from the Portal, management and remuneration for photocopying is also performed protecting the rights of creators and publishers in cooperation with the national reprographic reproduction rights organization LATREPRO¹.

Within the Portal, publishers can apply for international identifiers (ISBN, ISMN and ISSN) and barcodes generated from them. They can also request to receive a Universal Decimal Classification (UDC) index and an author notation. Legal deposit of online publications and data on publishing output as well are submitted to the Portal by publishers.

Development of the Portal's functionality

Before the Portal was introduced in 2018, in order to receive ISBNs, ISMNs, ISSNs, and UDC indexes and author notations, publishers had to fill out a separate application form for each service providing baseline information on publication. Applications could be filled out on NLL website or in a document template, which then had to be e-mailed to NLL staff. Many publishers chose to receive NLL services in person instead. In addition to that, an IT system developed by a partner company facilitated issuing ISBNs to larger publishers. All this considered, NLL services were not user-friendly, especially in cases when, for example, publishers needed to receive an ISBN, a UDC index, and an author notation.

In order to provide modern services and improve the work process of the NLL including data reuse potential, the NLL made a decision to develop a new IT system. The goal was to create an online space where publishers could not only request and receive international standard numbers and barcodes generated from them, as well as UDC indexes and author notations but also submit legal deposit copies of online publications and reports on publishing activities for statistical purposes. For the NLL staff it was also important to oversee and control the process of legal deposit copy acquisition, to register received legal deposit copies, and to automatically generate a deed of delivery and acceptance for other legal deposit libraries. Functionalities that allowed editing metadata supplied by publishers and to easily compile statistical data on publishing in Latvia, and to export data from the Portal to other IT systems. These expectations were fulfilled when the NLL staff and publishers from Latvia began using the Portal in 2018 at full capacity.

During development stages of the Portal, exchange of experience and best practices was sought from colleagues from the National Library of Estonia since similar e-services portal for publishers was introduced by them back in 2016. Although the functionality imagined for the Portal by the NLL differed from that created and used in Estonia, this dialogue helped to understand the needs of publishers and the NLL staff better. In general it allowed avoiding significant mistakes and encouraged taking time whilst seeking the best technological and design solutions for more complex elements of daily tasks of publishers and the NLL staff, e.g., delivery of files containing e-serials within the Portal and other aspects that are related to large volumes of objects since, as Raivo Ruusalepp (2017: 4) has pointed out based on empirical

¹ <https://latrepro.lv/>

evidence, they are likely to require different “tools to present the material in more manageable forms”.

Whilst the Portal was in development stage, awareness of publishers’ plans to introduce the Books in Print system in Latvia resulted in an agreement between publishers and the NLL, which provided that both systems should be interoperable. Therefore, ability to migrate data in ONIX format was built in the Portal. However, since the idea of Books in Print system is not introduced yet, this functionality of the Portal has not been fully explored.

Publishers and the Portal

As publisher registers to the Portal, a publisher’s profile is created and a designated responsible user profile is linked to it. If necessary, responsible users can register new users and attach certain roles and rights to them, e.g., submission of requests for ISBNs, ISMNs, ISSNs, UDC indexes and author notations, payment for services, delivering legal deposit copies in mediated and unmediated form, and reporting on publishing activities. Typically, a single user is linked to a publisher’s profile. However, larger publishing houses, where responsibilities are divided or shared among several employees, appreciate the opportunity to register several users under single publisher’s profile.

When requesting an ISBN, ISMN, ISSN or a UDC index and an author notation, first, a publisher has to select from a menu the type of publication (e.g. a book, a magazine, a calendar, sheet music etc.) they plan to publish and also choose the format/-s (e.g. printed in hard or soft cover, digital in CD or online etc.). These are the first data publisher submits to the Portal. See Figure 1 for national data creation and reuse exploiting NLL Portal for Publishers. Depending on selections, corresponding available services are then presented to publisher. For example, if publisher plans to publish a book, publisher’s user is supposed to select “Book” as the type of publication. Based on that they then have an option to apply for an ISBN and/or a UDC index and an author notation. Thus, a chance of erroneous use of international standard numbers is lowered.

In the next section of request application, the user provides metadata and other information² by filling out required and – if so desired – optional fields, which are generated based on chosen type and format of publication and requested services and hence will differ from publication to publication. They then submit the application and it is registered within the system. The NLL then makes sure that the application is filled out correctly and, if it is, confirms it. If a fee applies to any of the services selected (i.e. ISBN, ISSN barcode or additional ISSN barcode) and publisher is not an ISBN subscriber, who pays annual fee only once a year, designated responsible user and user who submitted the application receive an automatically generated application confirmation e-mail. Then they can automatically generate an invoice within the Portal (an indefinite number of items and all services can be included in a single invoice) and pay for the services using services provided by their financial institution. When payment is received or if a publisher has not selected a service to which a fee applies, responsible user and user who submitted the application receive an automatically generated e-mail about service having been delivered.

² Type and subtype of publication; Creator (author, editor or compiler); Title of publication; Language; Provisional issue date; Annotation; Number of authors of publication; Subject of publication; Periodicity; Image of title page (for books), cover (for magazine) or first page (for newspaper); URL of digital publication.

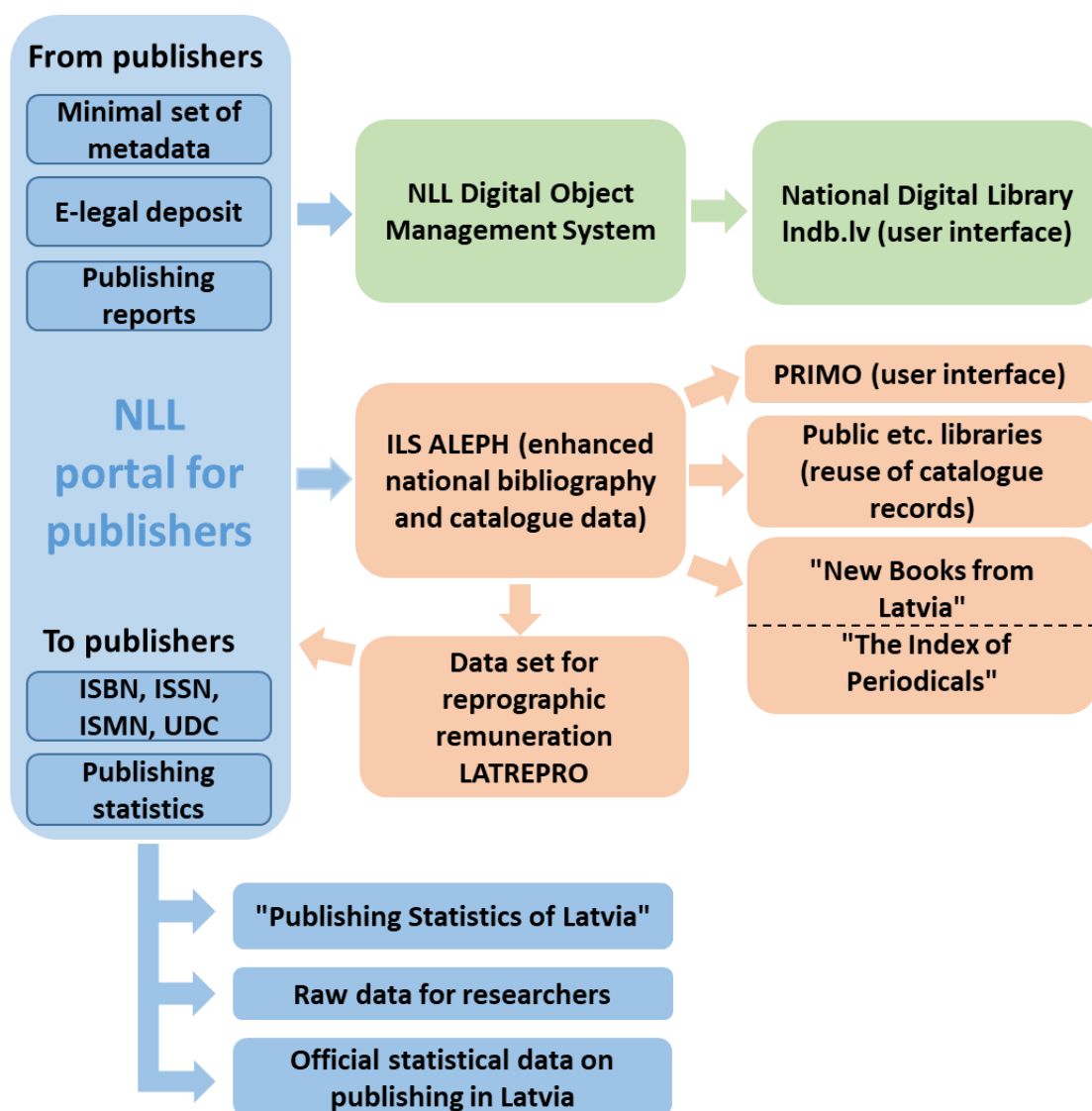


Figure 1. National data creation and reuse exploiting NLL Portal for Publishers.

When the NLL changes status of an application or when payment is received and invoice is marked as paid, a publication entry with a status “In production” is automatically generated. Since more than one publication format can be selected in application form (e.g. a printed publication in hard and soft cover, an online publication in PDF, and ePUB), any required number of publication entries can be generated from a single application. Apart from “In production”, four other statuses can be assigned to publication entries: “Published”, “Will not be published”, “Received”, and “Processed”. Each status enables a specific functionality guaranteeing that all processing steps are performed successively. In case a publisher applies for one service at first (for instance, an ISBN) but decides to request another service (a UDC index and an author notation, for instance) after application has been confirmed, it is possible for them to add another service to existing application allowing them to save time.

For various reasons during the past couple of decades, the number of self-publishers has increased globally as noted, for example, by a journalist Porter Anderson (2013) in an article in online business magazine *Publishing Perspectives*. Latvia has been no exception to this trend (NLL 2019; see also Shorthose 2004 on theoretical discussion and Ho et al. 2011 for a more applied study of this phenomenon). In addition, more and more organizations and institutions,

whose principal activity is not publishing, are releasing some publications. Given that, informing publishers about their obligation to submit legal deposit copies is becoming more challenging and, thus, also more important. In order to improve the ability and efficiency of legal deposit submission, publishers are required to state the estimated time of publishing when applying for service/-s. When the stated estimated time of publication release comes, publisher receives an automatically generated e-mail reminder to submit legal deposit copies of that particular publication. If estimates have not been accurate and publication has not yet been released, publisher can change the estimated time manually.

Portal, national bibliography and union catalogue

When the NLL staff receives a legal deposit copy, they find the publication entry in the Portal and, if necessary, edit the data submitted by publisher. If there is no corresponding publication entry (use of international standard numbers is not compulsory in Latvia), staff member creates a new entry. After entry is edited, data are exported to the union catalogue of the libraries of national importance. I.e., the data are created according to MARC 21 format and “Resource Description and Access” (RDA) rules and exported into the integrated library system (ILS) ALEPH. Then scanned book covers and table of content is added for books, as well as subjects, authority data, holdings information etc.

In Latvia, academic libraries use the union catalogue of the libraries of national importance and public, school, and special libraries use ILS ALISE, which also supports data import from the ILS ALEPH. Hence, it is possible for all libraries to use the data created by the NLL. In addition, the data submitted by publishers to the Portal form the foundation of the national bibliography. Before, two bibliographic records were created for each national publication: one – in the union catalogue of the libraries of national importance, the other – in the National bibliography database. However, since 2018 when the RDA cataloguing standard was applied to cataloguing practice in Latvia, only one bibliographic record is being created for each publication by adding to the baseline data imported from the Portal into the union catalogue. Twice per month the NLL releases an online publication “New Books from Latvia”³, which is composed of the data initially extracted from the Portal. The publication is used by libraries as they plan the acquisition of publications and also by other interested parties. Every year, the NLL also publishes “The Index of Periodicals”⁴ composed of extracted data on publications of previous year.

Portal and digital library

E-books are also represented in the union catalogue of the libraries of national importance and in the publication “New Books from Latvia”. In order for data on e-publications to be extracted, first, e-publications have to be delivered by publishers through the Portal. Then their entries have to be reviewed, edited if necessary, and exported to the Digital Object Management System DOM (metadata mapped according to the Dublin Core Metadata Element Set). After a successful export, DOM ID number is automatically added to the publication entry in the Portal. Then, data are exported from the Portal into the ILS ALEPH. Through the Portal, an ALEPH ID number is added to the publication entry in DOM, which is necessary for the data interoperability between the two systems. For e-serial delivery, the Portal has a special built-in

³ <https://lnb.lv/en/publishers/bulletin-new-books-latvia>

⁴ <https://lnb.lv/en/publishers/index-national-bibliography-index-periodicals>

functionality. Entries of e-serial issues are generated based on frequency and time of publishing and they are linked to main e-serial entry.

Publishing statistics

Aside from tasks regarding the national bibliography, the NLL is also responsible for gathering and dissemination of the official publishing statistics of Latvia. According to the Cabinet of Ministers regulation No. 281 (Ministru kabineta noteikumi Nr. 281, Likumi.lv 2017), publishers are bound to submit a report on their publishing activities within a fixed period twice per year. Reports must include data on all publications issued within the previous six months and these are, at the moment, the only source of data on print runs. Data on print runs are also the kind of data, which is more difficult to get, and it is also harder to evaluate their accuracy. It is so because due to historical context publishers have become more cautious in their practice. At first, a publication is typically released in a small print run and then, if deemed necessary, additional copies are made, which can happen a few times in case a publication is a success. Ensuring accuracy and reliability of statistics has become more challenging, because the NLL can only learn about additional copies if publishers willingly give data on them. Technological solutions, however, have led to at least some improvements with regards to this situation.

Previously publishers had to submit their reports in printed form by filling out a document template twice in a year. This order of affairs was cumbersome to both – publishers and the NLL staff as it required them to manually collect, process, review, and edit a massive amount of data in a relatively short period of time. Quality control and ensuring that timetable is abided by were proving to be the most challenging aspects of statistics collection process. Therefore, when the Portal was developed, solutions to these issues and other possible ways to ease and improve the work process for publishers and staff were sought. As a result, the workflows are now automated, reports are automatically generated using the data submitted to the Portal, and data can be reused.

Shortly before the period of report submission begins, publishers receive automatically generated reminder e-mails. When they log into the Portal, a form is available to them containing a list of all publications on which the data on print runs are required and all publishers have to do is fill in the blanks and generate a report document. Publishers also have an option to provide data on additional copies of any publication made within the reference period by giving the minimum amount of data – a title and size of additional print run. The raw data gathered in the Portal can be extracted and exported into a spreadsheet data set, which is a valuable resource for publishers, researchers, academics, professional organizations, and other parties interested in and studying publishing in Latvia. In addition to the data set, the NLL annually compiles an online publication “Publishing Statistics of Latvia”⁵, which includes general information on the number of publications and their print runs in aggregate form. The NLL also annually provides the Central Statistical Bureau of Latvia with data covering the main indicators of publishing industry as part of the Official Statistics Programme⁶, the terms of which are laid out in the Cabinet of Ministers regulation No. 763 (Ministru kabineta noteikumi Nr. 763, Likumi.lv 2018).

⁵ <https://lnb.lv/en/publishers/publishing-statistics>

⁶ <https://csb.gov.lv/en/documents/official-statistics-system/programme>

LATREPRO

Initial metadata from the NLL Portal for Publishers are used not only as foundation, from which to build full level national bibliography entries in ILS ALEPH. These are also a relevant element towards an improved system of management and remuneration for photocopying to observe compliance with legal norms regarding intellectual property rights and copyrights of creators and publishers in cooperation with the national reprographic reproduction rights organization LATREPRO. LATREPRO is a body authorized by Ministry of Culture of the Republic of Latvia with the aim to ensure that authors and publishers receive rightful remuneration for reprographic reproduction of works in accordance with the Copyright Law (Autortiesību likums, Likumi.lv 2000, Section 35). In this sense, fair remuneration is composed of a fee that is supposed to be paid by every owner and user of a copying machine. Collected funds are then distributed among those authors and publishers whose printed publications have been delivered to the NLL as legal deposit copies in accordance with the Legal Deposit Law. In 2018, based on data provided by the NLL, LATREPRO remunerated authors and publishers for reproduction of 6380 objects of property.

Conclusion

The introduction of the Portal has strengthened the cooperation between publishers and the NLL as well as provided the technological basis for implementation and development of automated metadata creation and reuse, and other processes that save a lot of redundant manual work. It is also the first time in the history of the NLL that the annual official publishing statistics of the Republic of Latvia covering the main indicators of publishing industry is compiled exploiting the initial data created in the Portal by publishers themselves. The wide range of online services has eased the daily endeavours for publishers in Latvia and the NLL staff, as well as introduced new opportunities for data mining and processing.

References

Anderson, P. (2013) Self-Publishing and the Industry: Implications and Impact. Publishing Perspectives, September 24, 2013. <https://publishingperspectives.com/2013/09/self-publishing-and-the-industry-implications-and-impact/>

Ho, H., Wang, L. and Cheng, H. (2011) Authors, Publishers, and Readers in Publishing Supply Chain: The Contingency Model of Digital Contents Production, Distribution, and Consumption. Systems Engineering Procedia 2 (2011): 398-405. <https://www.sciencedirect.com/science/article/pii/S2211381911001421>

Likumi.lv (2000) Autortiesību likums, Latvijas Vēstnesis, No. 148/150 (2059/2061). <https://likumi.lv/doc.php?id=5138>

Likumi.lv (2017) Ministru kabineta noteikumi Nr. 281 “Noteikumi par izdevējdarbības pārskata oficiālās statistikas veidlapas paraugu un veidlapas aizpildīšanas un iesniegšanas kārtību”, Latvijas Vēstnesis, No. 107 (5934). <https://likumi.lv/ta/id/291156>

Likumi.lv (2018) Ministru kabineta noteikumi Nr. 763 “Noteikumi par Oficiālās statistikas programmu 2019.–2021. gadam”, Latvijas Vēstnesis, No. 246 (6332). <https://likumi.lv/ta/id/303613-noteikumi-par-oficialas-statistikas-programmu-2019-2021-gadam>

NLL (2019) [Indicative data on number of self-publishers extracted from annual statistical dataset on publishing in Latvia, 1995-2018], unpublished raw data accessible by selected members of the NLL staff.

Ruusalepp, R. (2017) *Preserving digital legal deposit – new challenges and opportunities*. Paper presented at: [IFLA WLIC 2017 – Wrocław, Poland – Libraries. Solidarity. Society.](#) in Session 210 – Preservation and Conservation (PAC) Strategic Programme. <http://library.ifla.org/1677/1/210-ruusalepp-en.pdf>

Shorthose, J. (2004) A more critical view of the creative industries: Production, consumption and resistance, Capital and Class No. 84., Special Issue “The Creative Industries: Production, Consumption and Resistance”.