



Publishing digital resources

Marinela Covaci

► To cite this version:

Marinela Covaci. Publishing digital resources. ELPUB 2020 24rd edition of the International Conference on Electronic Publishing, Apr 2020, Doha, Qatar. <10.4000/proceedings.elpub.2020.12>. <hal-02544313>

HAL Id: hal-02544313

<https://hal.science/hal-02544313v1>

Submitted on 16 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Publishing digital resources

Marinela Covaci

The Digital Library

- 1 Within the roles of the national libraries resides providing full access to voluminous national documentary heritage, regardless of the support type of the documents, according to the Law 334/2002 for libraries, the National Library of Romania coordinates the establishment of the National Digital Library. Therefore, the institution drives the creation of the digital collections by applying the digitization strategy, management, integration and by promoting them at the national and international levels. The reference documents in this regard are Digitization Guide (2009), MINERVA (2003) and Europeana Data Model (EDM, n.d.).
- 2 Despite the shortage of skilled personnel, especially in IT, the National Library is permanently preoccupied about the development of modern services that facilitate access to the documentary heritage in a user-friendly mode. The National Library offers online access to information and builds the Digital Library on three platforms presented below.

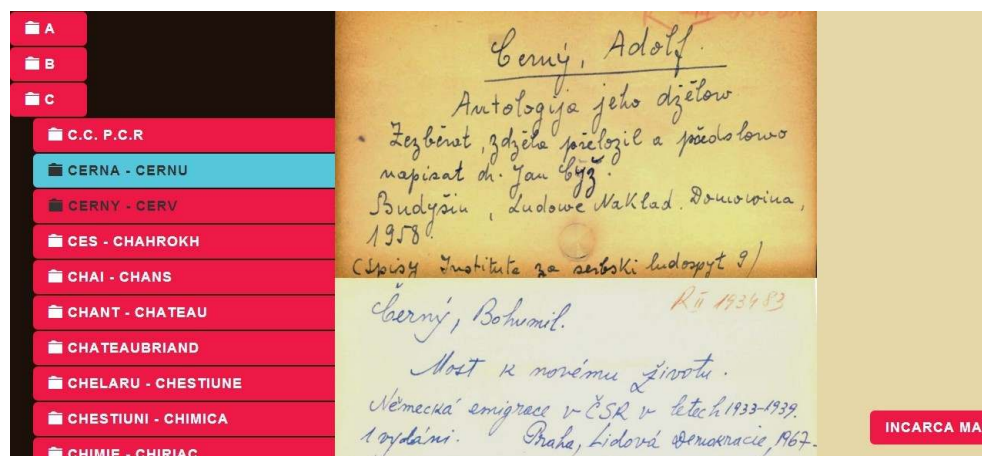
The Virtual Traditional Catalogue of books published before 1993

- 3 The catalogue is accessible through the *Book-a-Book* platform, available via <http://bookathon.bibnat.ro/>, after the authentication check. It has two functions: one of consulting the Traditional Catalogue only, second implying a voluntary contribution to the creation of metadata. The Virtual Traditional Catalogue is the virtual image of the traditional catalogue, which is organized alphabetically by the name of the author in the appropriate folders – the “*physical drawers*”. Currently, there are 1,600,000 cards structured in 1,600 folders. Therefore a folder contains on average 1,000 scanned cards. Although there are no access points (title, author, etc.) in this catalogue, the platform

can provide browse and search functions (proximity search by creating an index consisting in fact of the traditional catalogue drawers' alphabetical tags).

- 4 The goal of the first project was to import the XML files containing metadata into the bibliographic database, but most of them are handwritten. The lack of financial support has made impossible the fulfilment of the project started in 2008, about the scanning of the traditional library cards. Only 300.000 cards were completely digitized, as they were scanned, OCR processed and had the metadata extracted in XML files linked to JPG files. But most of the cards are handwritten, and the OCR process was impossible.

Figure 1. Screen capture: virtual image of the drawers



(Image rights: CC BY-SA)

- 5 The second project, called *Book-a-Book* started in 2016 and was aimed to create a platform capable of providing both search and retrieval of the bibliographic information from the virtual catalogue of books published before 1993; it was also meant to offer the possibility of indexing the books and of creating a database based on MySQL with multiple facilities. Later, the XML files associated with the cards would be imported into the library's system. The platform was developed through a partnership with the *Tech- Lounge Association* of the students and professors from the Faculty of Automatic Control and Computer Science of the Politehnica University of Bucharest (UPB). *Book- a-Book* is a collaborative project. Anyone can voluntarily participate to the development of the catalogue by creating associated metadata for the scanned catalogue cards.

Figure 2. Screen capture: creating metadata based on a digitized catalogue card

BOOKa Clasament

Â â Ă ă Î î Ș ș Ț ț

Collins, Colette 190589
Always and forever: [Novel]/Colette Collins. — Eugene, Oregon: Harvest House Publishers, 1985.
 192 p. — (Rhapsody Romances; 19)
 Extras din catalogul colecției la afișarea
 cărții.

Cotă

Nume Autor 1

Prezume Autor 1

Adaugă autor

Titlu

(Image rights: CC BY-SA)

- 6 The volunteers can learn to fill correct bibliographic information by attending tutorials associated with each field. The goal is to obtain a brief description with 16 bibliographic fields. If the volunteer receives a hardly visible card or a card written in a non-Latin alphabet, there is the possibility to report it and to receive a new card image. These files are archived in a separate folder and processed afterwards by a librarian. Because the volunteers are not experts in the bibliographic records, the system has a check-document routine before saving the record in the database, which requires a matching of the data in three fields, such as Location, Title and Authors, provided by at least two volunteers.

Figure 3. Tutorial on location information

Cotă

Cota este poziționată în partea dreapta-sus a cartelelor. Este alcătuită din 2 părți:

- un număr roman de la I până la V (I, II, III, IV, V), o combinație de majuscule sau ambele
- un număr format din maxim 6 cifre arabe (0..9)

Diaconescu, Ștefan

COVALIU, DORIN și DIACONESCU, ȘTEFAN
 Reglementări privind organizarea și funcționarea
 cooperăției meșteșugărești. / Redactat de Covaliu Dorin și Diaconescu Ștefan/. — București, 1973.
 199 p. 205 x 147 mm. (Uniunea Centrală
 Vezi fișa urm.
 334.68(498)(094):34 /Nr.68/74/

II 157263

(Image rights: CC BY-SA)

- 7 There was a crowd-cataloguing event organized, on 14th May 2015, called *Bookathon*, aimed to form a volunteers network to be involved in creating metadata for the rest of the cards. More than 100 people participated in this event, representing 4,5% of the people who subscribed. The final results consisted of 624 valid records and 1,573 records pending validation, representing 0,05% of the total number of cards.
- 8 The *Book-a-Book* platform is an excellent tool to be used for getting acquainted and practising cataloguing description. Subsequently, more people have joined it, as

volunteers, undergraduate and graduate students from the Faculty of Letters of the University of Bucharest. Today, the platform has 908 users, and 11681 valid records in the database. Until the final moment, when the database will have all data loaded, it is a long way.

The Online Catalogue of books published after 1993

- 9 The online catalogue¹ is the web component of the integrated library system Aleph. It contains bibliographic information about the documents published after 1993, which are part of the library collections: books, periodicals, multimedia documents. There are two digital collections, with restricted access for copyright reasons. These include:
 - **PhD Theses**, which provides access to electronic content thereof (by digitizing paper or archiving the electronic content of the CD attached by the authors).
 - **The Digital library for blind people** (created within the project *the Sound of pages*), through which disabled users have access to digitized works, on-demand.
- 10 The digital content of the doctoral theses is available in the library's network only. Since 2013, the digitization of the doctoral dissertations is done on demand. Today, the PhD theses sub-collection contains 3500 digital objects, which can be consulted in this way.

Figure 4. Screen captured full view on the online catalogue

Vizualizare completă a înregistrării

Alegeți formatul: [Format standard](#) [Fișă de catalog](#) [Citare](#) [Nume etichete MARC](#) [Etichete MARC](#)

Înregistrarea 1 din 2809

Toate exemplarele	Toate exemplarele
Localizare/exemplar	Depozit legal - BNR
Cota	V 95839
Autor	Zvâncă, Alina-Gabriela
Titlul	Strategii de resurse umane pentru scăderea și depășirea rezistenței la schimbare în organizațiile românești
Tipul documentului	[Text tipărit]
Localitate	Iași
Editură	Universitatea Tehnică "Gheorghe Asachi"
An	2011
Descriere fizică	250 f. : graf., schem., tab. în parte color ; 30 cm. + 1 CD-ROM
Note responsabil.	Antetitul: Universitatea Tehnică "Gheorghe Asachi" din Iași. Școala Doctorală a Facultății de Textile-Pielărie și Management Industrial
Bibliografie	Conține bibliografie
Note disertații	Teză de doctorat. Cond. de doctorat: prof. univ. dr. ing. ec. Costache Rusu
CZU	65.012.4.658.3+061(043)(0.034)
Alte clasificări	65
Autor secundar	Rusu, Costache (cond. șt.)
Extern	tezedoctorat.bibnat.ro/V/95839.pdf

(Image rights: CC BY-SA)

- 11 Initiated by The National Library and funded by the *Orange Foundation* through the “*world through colour and sound*” Project, the *Sound of pages* aims to create and develop the most extensive digital library in Romania for the blind. The collection will be grown exponentially by continually adding digitized documents on demand and making them accessible.

Figure 5. Bibliographic data: querying the online catalogue

Înregistrările selectate: Vizualizarea selecției Salvare/eMail Creare subset Adăugare în eRaft Salvare pe server						
Intregul set: Selectează tot Deselectează Relevanță (Rank) Rafinare Filtrare Cerere SDI						
<p>te pentru W-title= povești S-a selectat mai mult decât un set; Sortat după: 01---D02---A</p> <p>de sortare: • Autor, titlu • Autor, an • Titlu, an(A) • Titlu, an(D) • An, autor(A) • An, autor(D) • An, titlu (D) • An, titlu (A) • Subiect, titlu</p> <p>privind formatul: • Titlu - 951 • Autor - 952 • Scurt - 953 • Afisare prescurtată - tabel</p> <p>rări 1 - 1 din 1 (afișate și sortate maxim 10000 înregistrări)</p>						
<div> <input type="text"/> Salt la text <input type="text"/> Salt la # </div> <div> <input type="button" value="Pagina anterioară"/> <input type="button" value="Pagina următoare"/> </div>						
Autor	Titlu	An	Editură	Localizare/Exemplare	Cota	Media
rcia Márquez, Gabriel	A trăi pentru a-ți povesti viața	2010	Editura RAO	Sediul central (1/0) Sediul central (dublet) (1/0)	II 554870	pdf

(Image rights: CC BY-SA)

- 12 Choosing pdf file from Media column, if the users have installed special software on their PC, than they can *hear* the text.

The Digital platform

- 13 The National Library contributes to the European heritage by publishing digital collections, a project that started in 2009 using the Digitool² system, and also through content aggregation for the Europeana portal which commenced in 2014. The system has the protocols OAI-PMH³ and Z39.50⁴ which made possible the metadata harvesting. The digital collections was created by digitizing the documents from the public domain and are organized according to the topics or to the library events such as exhibitions, for instance. The selection criteria applied to prioritise digitization are a combination of:

- legal (copyright and broadcasting) rights;
- degree of interest and addressability;
- documentary value;
- representativeness for a particular area, a certain time-period or a specific geographic region, both nationally and internationally;
- long-term preservation of the original documents.

- 14 The workflow comprises the following steps:

- scanning documents;
- archiving digital objects;
- cataloguing documents in Unimarc bibliographic format, in the integrated library system;
- ingesting metadata in MARC21 bibliographic format in the digital library software;
- preparing topical collections (creating sets of records extracted by a unique identifier and check data quality);
- publishing on the own library's platform, National Digital Library;
- harvesting metadata in Dublin Core⁵ format which is needed for the Europeana portal.

- 15 So far, there have been about 25 TB of digitized content. Currently, it includes the following digital collections, obtained by scanning documents from public domain:

Historical documents from 15th – 19th centuries

- 16 There are 2416 resources created from family archives – Cantacuzino-Pășcanu, Sturza-Bîrlădeanu, Beldiman Magdalena, Stephen Bonachi, Gregory Buicliu, Elena D. masts, Gh. Ghibănescu, N. Docan and refer to villages, monasteries and country estates in Wallachia and Moldova. The majority of them are sale or purchase documents, court

documents, donations, tax, property rights reinforcements etc. The funds mentioned here are Kogalniceanu, Saint-Georges, Brătianu.

Old and bibliophile Romanian books (738 resources), relevant to the history of the Romanian books

- 17 The collection includes 738 resources, relevant to the history of the Romanian books. There are the first Romanian printed book (Slavonic Liturgy, Târgoviște, 1508) and The Bible (Bucharest, 1688).

Figure 6. Sample of an old book from the NLR's digital collection



(Image rights: CC BY-SA)

Old foreign books and bibliophile books

- 18 The collection includes 1119 resources capturing historical and artistic prints representative for 16th – 20th centuries. It also includes works of classical authors from antiquity to the present time, published in the most important printing centres in Europe. One example is from the legacy of the famous humanist Erasmus of Rotterdam – with his most popular work, *The Praise of madness* – published in 1515 or 1518 in Basel.
- 19 Incunabula include 149 works representative for the 15th-century European cultural landscape, comprising chronicles, historical and religious texts. The oldest Incunabula preserved in the collection is *Apophthegmata* by Plutarch, in the Latin translation of Franciscus Philelphus, printed in 1571 in Venice.

Old Romanian periodicals

- 20 This collection includes 668 resources from journals published between 1862 and 1942.

Photographs

- 21 The collection includes 2482 resources, and it is the result of several years of research, identification, cataloguing and scanning of over 400 images and two topical albums taken by photographer Carol Pop Szathmari (1812–1887). It is composed of photographs taken in Szathmari's workshops in Bucharest and photographs taken during his travels around the country, as well as photos of the War of Independence 1877 – 1878, grouped in two albums. In the portrait gallery, we can find political personalities of public life.

Figure 7. Historical photograph the NLR's digital collection



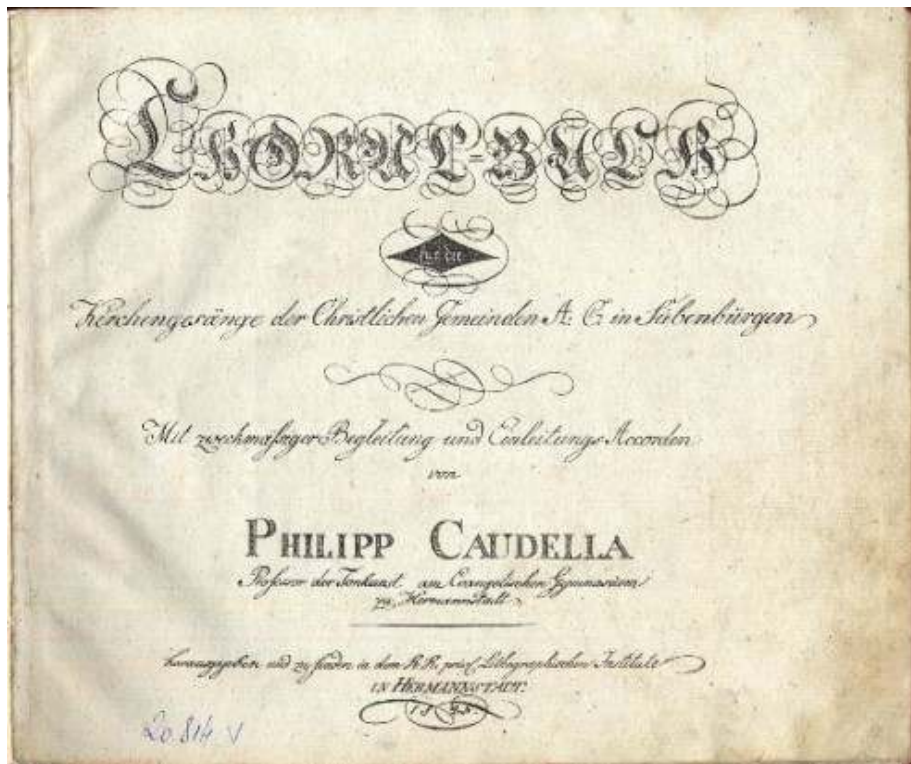
Screen capture: photo from the NLR's collections
(Image rights: CC BY-SA)

Musical scores

- 22 The collection has 1191 resources from the XIX-XX centuries. The value of these scores lies in the importance of artistic or historical works, their age, the author, the places where they were printed. The oldest printed scores are the *Chorus-Buck* by Philip Caudella and *Irmologhion or Musical Katavasieriu*, both from 1823. But we will also find old Romanian scores from the XIXth century, which have become treasures through

their historical value. Among others, we can find the *Union March* by Alexander Flechtenmacher dating from 1850 and Ciprian Porumbescu's compositions.

Figure 8. Title page of music sheets of Philip Caudella's work



(Image rights: CC BY-SA)

Manuscripts

- 23 The collection includes 706 resources in a variety of language across two categories of manuscripts: old documents and relatively recent modern works. Among the first category, we notice the several volumes of Carol Nicolas, entitled *Panis emendicatus animae* written in Latin with drawings in pen and pencil. These are 18th century religious manuscripts. Closer to us, are the manuscripts written in the Greek language, medical books written in the 19th century.
- 24 Visually spectacular are the Romanian religious manuscripts of the 19th century such as *Archdiocese Liturgy*, copied by Veniamin Catulescu Cernican, *The Great Canon of Andrew of Crete* and *Coligate of religious songs from the Orthodox religion* copied by Picu Pătruț Procopie from Săliște Sibiu.
- 25 One of the most beautiful and valuable contributions of the National Library of Romania to the cultural heritage in the virtual environment is *Codex Aureus*, the most famous mediaeval manuscript illuminated from the library collection and also from Romania. It is a fragment of Latin gospel on parchment from 810. (MS II.1), made at the command of Carol the Great (768-814). The manuscript is rightly famous both by the exceptional quality of the illustration and by the fact, no less famous, that it was

written entirely with gold ink. In the Hungarian and Romanian literature, the fragment from Alba Iulia is also known as *Codex aureus*.

Graphics

- 26 This collection contains 268 documents from the 19th–20th centuries in the time of writing in various sub-collections such as:
- 27 **Ex-Libris** with 163 resources, includes both original prints and printed reproductions, signed by Romanian, German, Czech, Hungarian, Spanish, Portuguese, Japanese and Chinese contemporary artists.
- 28 **Japanese woodblock prints**, with 105 resources. The most beautiful and well-preserved album of the Japanese woodblock prints collection is *One Hundred Aspects of the Moon*, published between 1885-1892, created by Tsukioka Yoshitoshi, one of the most famous Japanese artists of the 19th century. The album is a unique and valuable source for understanding the Japanese traditional culture, full of mythological figures, where the artist uses the symbolism of the various phases of the Moon, to reflect on the human condition.
- 29 In 2015, TEL⁶ and Europeana disseminated on the social networks, during the Open Collections event, Open Access event and Halloween week, four of The National Library of Romania's collections. Some examples include *Ex-Libris*, *Musical Scores*, *Japanese Woodblock prints* and *Incunabula*.
- 30 This poster was also disseminated via twitter (bit.ly/1T36MhW). Published in August 2015, it has enjoyed incredible success reaching an audience of 69,920 hits.

Figure 9. NLR Japanese Woodblock Prints: Image on Europeana Blog



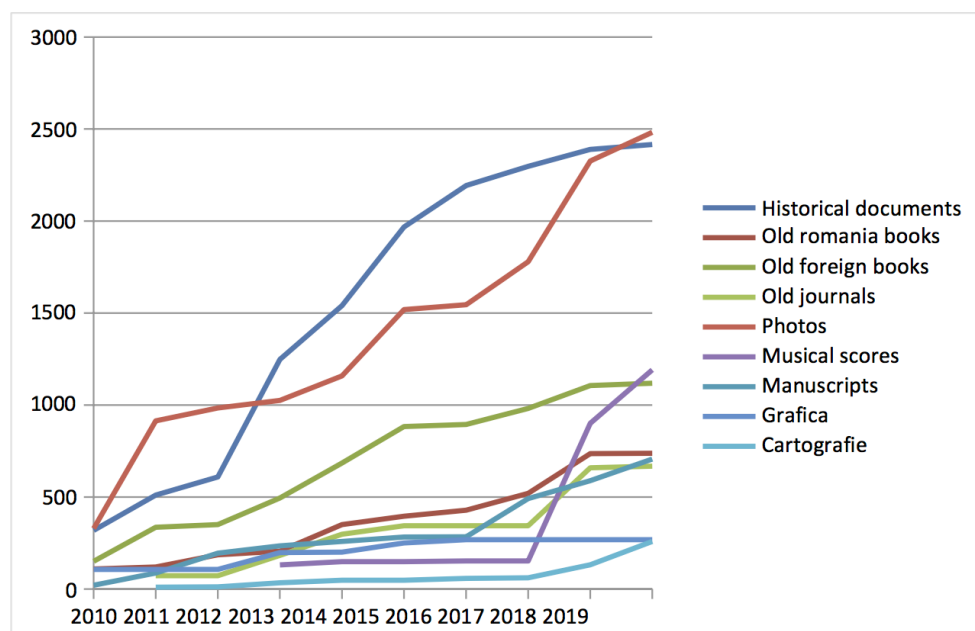
(Image rights: CC BY-SA)

Evolution of the Digital Library

- 31 The digital library platform was born in 2010 when the digital collections created in-house were published, containing documents belonging to the patrimony of the National Library of Romania. Currently, nearly 11000 digital objects are on display. Unfortunately, from the beginning until now, the office had only two employees, hence

the small number of digital documents published on the digital platform, compared to other national libraries.

Figure 10. Growth chart on the NLR digital collections



(Image rights: CC BY-SA)

- 32 In the last two years, there was an increase in the collections of photographs and historical archives, since Romania celebrated its centenary in 2018.

Digital Libraries worldwide

- 33 At the global level, there are many big digital libraries similar to those presented in previous chapters, built generally by libraries with various partners, having general or specific collections.
- 34 A famous digital library is World Digital Library, available at www.wdl.org. The World Digital Library is a cooperative project of the Library of Congress, the United Nations Educational Scientific, Cultural Organization (UNESCO), and partners such as libraries, archives, educational and cultural institutions from around the world. The project brings together on a single website 19,147 resources (rare and unique documents: books, journals, manuscripts, maps, prints and photographs, films and sound recordings) from 193 countries. National Library of Romania contributed with 15 precious resources among which we mentioned *Latin Bible with interpretations of Hebrew Names in Alphabetical Order*.
- 35 From the point of view of preserving the European cultural heritage, we should mention the Manuscriptorium portal, available at <http://www.manuscriptorium.com>. It is operated by National Library of the Czech Republic. The National Library of Romania is one of the data providers for it, contributing 871 digital resources.
- 36 From the point of view of the digital library for the blind, a good example is available in China, where the visually impaired can listen to electronic books, music or on-line

lectures for free. Opened in Beijing, at National Library, the facility was jointly set up by the Information Center of the China Disabled Persons' Federation, the National Library and China Braille Publishing House.

- 37 From the point of view of the accessibility online of the library traditional catalogues, as well as crowd-cataloguing, some examples of successful projects are the *Scanned Card Catalogue (pre-1980)* from State Library of NSW, Australia, available at <https://www2.sl.nsw.gov.au/eresources/scc/?imageid=0063672> and the Europeana-related projects, based on communities of volunteers, *Transcribathon*⁷. A crowd-sourced transcription of digital materials from the First World War had been completed by Europeana 1914-1918. The volunteers' network adds the benefit of creating a vast and fully digital record of personal documents. Romania participated at the first marathon of transcription of unpublished documents; the National Library was the pioneer in this domain and the first organizer of such an event in Romania.

Conclusions

- 38 The digital resources introduced in this paper offer opportunities for research, learning and generating and dissemination knowledge, thus contributing to the development of a creative economy. As an example, one of the digital objects presented in the *Japanese woodblock prints collection* was used in an award-winning entry in the GIF IT UP 2017 competition organized by Europeana, DPLA, Trove and DigitalNZ. In the Public Choice section, among the winners, was Hannah Langford Berman – Vancouver, Canada who created an animated picture based on one digital image (see Fig. 11) from the collection mentioned above⁸. It is one of the excellent examples demonstrating how the accessibility to the NLR collection online is creating opportunities for innovative uses of the objects.

Figure 11. A Japanese woodblock image used in a GIF IT UP 2017 entry



(Image rights: CC BY-SA)

- 39 Through these projects, the National Library of Romania has offered innovative opportunities for people as to get familiarized with its platforms and to become active participants in the voluntary activities run within the Library. We believe that this is a way to form a community of volunteers which together with our library experts will enhance the accessibility of the library holdings for a large category of users interested in research and e-learning.

BIBLIOGRAPHY

Borună, Adriana-Elena. "Cărți accesibile persoanelor cu deficiențe de vedere, în cea mai mare bibliotecă digitală dedicată acestora". *Biblioteca: revistă de bibliologie și știința informării*, anul LXVIII, serie nouă – anul XXVII, nr. 11, 2016, p. 350.

Covaci, Marinela. "Colecțiile digitale ale Bibliotecii Naționale a României expuse în Europeana". *Colecțiile de patrimoniu în era digitală: lucrările celui de-al doilea workshop*, București, Editura Academiei Române, 2017, pp. 23-31.

Covaci, Marinela. "De la Catalogueul tradițional la Bookathon: Proiectul Book-a-Book". Presentation at Romanian Librarians Association ABR, 27 mai 2016, Bucharest, Biblioteca Națională a României, 2016, abr.org.ro/wp-content/uploads/2015/05/p.-book-a-book.pdf.

Digitization Guide – thematic pillar "Libraries" (in Romanian: Ghid de digitizare – pilonul tematic "biblioteci" – versiunea 01/30.10.2009), 2009, bibnat.ro/dyn-doc/Ghid%20de%20digitizare_Pilonul%20tematic_Biblioteci_versiunea01_25_11_2009.pdf.

EDM. *Europeana Data Model Documentation*, (n.d.), pro.europeana.eu/page/edm-documentation.

Gruia, Luminița. "Catalogueul tradițional virtual al Bibliotecii Naționale: acces online la informație și crowd-catalogueing prin platforma Book-a-Book". *Biblioteca: revistă de bibliologie și știința informării*, anul LXVIII, serie nouă – anul XXVII, nr. 12, 2016, pp. 380-382.

MINERVA. *Technical Recommendations for digital content creation programs in the cultural field*, 2003.

Website of the National Library of China, (n.d.), www.china.org.cn/china/national/2008-10/15/content_16614900.htm.

Website National Library of Romania, (n.d.), www.bibnat.ro.

Website Transcribathon, (n.d.), transcribathon.com.

Website World Digital Library, (n.d.), www.wdl.org.

NOTES

1. Accessible on <http://alephnew.bibnat.ro:8991/F>.
2. Digital asset management and preservation solution.
3. Protocol for Metadata Harvesting (OAI-PMH) is a protocol developed for harvesting (or collecting) metadata descriptions of records in an archive, available via <https://www.openarchives.org/pmh/>.
4. Z39.50 is an international standard communications protocol for searching and retrieving information from a database over a TCP/IP computer network, available via <https://en.wikipedia.org/wiki/Z39.50>.
5. The Dublin Core Metadata Element Set used in resource description, available via <http://dublincore.org/documents/dces/>.
6. The European Library.
7. <https://transcribathon.eu/>.
8. <https://blog.europeana.eu/2017/11/winners-of-gif-it-up-2017/>.

ABSTRACTS

The National Library of Romania shares digital resources, providing access to knowledge and cultural heritage, on three online platforms. These are:

- a) The Traditional Virtual Catalogue for the books published before 1993. It is the virtual image of the traditional catalogue which is organized alphabetically by author name in the appropriate folders physical drawers. The platform and database was designed by the students from the Technical University Bucuresti. Accessible at <http://bookathon.bibnat.ro>;
- b) The Online Catalogue on Aleph system, which offers bibliographic information about documents in the library collections since 1993. There are two logical databases, The Doctoral Theses Referential,

which provides access to online content thereof (by digitizing paper or archiving electronic content of CD attached by author) and digital library for the blind (project Sound of pages) that disabled users have access to requested digitized works. Accessible <http://alephnew.bibnat.ro:8991/F>; c) The Digital platform on Digitool system. It started in 2009. It is comprised of digital collections created by digitizing special collections of documents in the National Library of Romania, organized by themes or after events. Accessible at <http://digitool.bibnat.ro/R>. The databases have grown steadily and have become researching and learning tools for some categories of users. The systems provide open resources, learning and navigating tutorials and contribute to a creative society.

INDEX

Keywords: digital library, metadata, open access, online services, cultural heritage

AUTHOR

MARINELA COVACI

National Library of Romania

marinela_covaci@yahoo.com