Virtual Challenge to a Library: Displaying Digitised 19th-Century Books Online

Ekaterina Dikova
Bulgaria

Introduction
The current European Union driven situation in virtual heritage has reached the level of striving for more attractive online displays: semantically enriched digital exhibitions, immersive and interactive experience communicating knowledge in an entertaining, exciting and creative way by means of toolkits, 3D modelling, graphics and computer animation. The most characteristic library items – the books (including medieval codices and earliest prints) still need to be read line by line and page by page. Yet there is a considerable number of people (not just scholars) addicted to studying them. This means that the best way to present books online is as digital objects that are easily searchable and accessible, convenient to refer to, browse, read and investigate – line by line and page by page. This is proven also by the success of self-publishing projects like Gutenberg and the Internet Archive. Consequently, I am presenting in this paper a multilingual template reflecting as fully as possible the respective physical items, some 19th-century books, including covers, preliminary and last pages, illustrations etc. as well as my personal idea of a virtual exhibition to come. The languages involved – English, French, German, Russian, Romanian, Greek, and Bulgarian – are the languages in which the books in the respective physical collection were published, and so the need arose to have those terms translated.

The presentation of 19th-century books online is part of the National Academic Library an Information System Foundation (NALIS) initiative to digitise library items kept at the libraries taking part in the NALIS consortium. The main goal of this project, financially supported by the America for Bulgaria Foundation, has been to create, maintain and en-
While waiting for the times when a real digital library will be virtually possible, we can still only play with digitised cultural content, looking for the most attractive ways to present it, selecting fragments from books that would be most interesting for a larger audience or that we would like to popularise.

Digitisation of NALIS libraries’ items began at the end of 2011 and initially involved were the Library of the Bulgarian Academy of Sciences (CL-BAS) and the Sofia University Library (SUL) with their collections of rare and valuable books and archival materials. SUL enriched its digital content with copies of theses held at the Sofia University, with the new issues of the Annual of the Sofia University and a few 19th-century manuscripts. The Library of the New Bulgarian University (NBU) joined the initiative in 2012, when, also
NALIS began constructing a massive collection of valuable photographs (in relation with the Europeana Photography Project). Quite recently, in July 2014, also the Plovdiv Public Library took part with a small in size but extremely valuable collection of medieval Slavonic manuscript fragments. The total number of digital objects is about 20,000 (including the Bulgarian National Library’s digital collections created for another project, but successfully integrated by the NALIS foundation into the union catalogue. Therefore, the main aim of this initiative has been to make all the digital copies mentioned easily searchable and fully accessible through the Union Catalogue of the academic libraries in Bulgaria – the NALIS Union Catalogue.

The strategy for digitising the holdings of the NALIS founding libraries appeared spontaneously, but was legitimised at the NALIS Board of Directors Meeting (Minutes 23/01.08.2011 and 28/01.02.1012). The first step was to plan the coordination between the Central Library of the Bulgarian Academy of Sciences (CL-BAS) and the Sofia University Library (SUL), so that the first institution embarks on digitisation of mainly the 19th-century books, and the latter – of the 19th-century periodicals. The National Library joined at a later stage and with ready digital collections in which part of the independently prepared material from SUL and CL-BAS turned out to be duplicated.

Initially, CL-BAS and SUL began scanning their most valuable items – the Bulgarian early prints. Later NBU set as its goal the digitisation of materials most used by their students. Thus, the selection of items for digitising has been based on two main requirements: value or frequency of use, on the one hand, and coordination in order to avoid duplication of digitised objects on the other. A decision has been made to digitise the entire volumes, including the later covers, flyleaves, and pastedowns. And, of the various ways of online visualisation of items with complicated structure, the parallel presentation of content and image has been chosen i.e. each line of the table of contents is a link to the respective image.

**Digital object – image and metadata: digital collection and physical collection**

As a non-digitally born object is a combination of a copy of a given physical item and metadata for this copy, several types of data should be distinguished:
- data about the item⁴ – general bibliographic and authority information, usually in MArchine-Readable Cataloging (MARC) format, but the simpler Metadata Object Description Schema (MODS) and Dublin Core (DC) are also acceptable,
- data about the contents of the item – according to the MARC scheme it may be inserted in subfield 2 „Scope and content” of 520 field „Summary, etc.” note,
- the content itself (i.e. structural metadata, especially important when digitising materials with a very complex structure such as the periodicals, books or manuscript codices (in contrast to photographs or other one- or two-page items), reflected online by means of the Metadata Encoding and Transmission Standard (METS)⁵.

Without the last type of data, the images of the pages or, respectively, the text files building up the digital material, cannot be used fully by the patron. The kernel of the METS files – and their only compulsory element⁶ – is precisely the structuring information by means of which each line of the table of contents turns into a link to a particular image (a digital copy of a page).

No matter whether the user decides to choose a) the option for the structural (or “logical”) view, when the contents are given parallel to the images, or b) the page turner, every page will be always linked to its metadata. While the bibliographic description of the item in MARC 21 is necessary to locate it in the catalogue (and has been integrated in the catalogue so that this could happen), the METS file, attached to the digital counterpart of the item, is necessary to ensure convenient examining and page-turning. METS files have been created by librarians with the help of an especially designed application that generates them⁷.
Both the digital objects and the digital collections may or may not fully match the content of their physical counterpart. Digitisation is an opportunity to easily reorganise any of the files and the information about them. The missing pages in a physical item, for instance, may well be restored with the relevant digitised pages of the copy of the same edition, kept in another institution. In a similar way, it is possible to amend (by transposing and adding) the content of a digital collection. The CL-BAS' digital collection of books is comprised of digital copies of items physically located in four collections, which is clear from the first element in their call-number.

The template - structure, rendering and editing
The main emphasis in the part related to presenting books online is the creation of a template which to reflect as fully as possible the common in the structure of the 19th-century books online of the physical items themselves. A Multilingual Template was prepared in relation to the NALIS initiative in digitisation, used for the creation of METS files for each of those digitised books and published on the Library Standards page on the NALIS website. The main obstacle when constructing this template was the lack of a single multilingual library dictionary, or at least a similar dictionary for each language, where all the necessary terms are present with accurate definitions.

The terms are not new to librarians, but mainly through practice and not as concepts. This is partially due to the fact that some of the terms are related rather to printing (and book design) as well as book binding (and restoration) of volumes than to library studies. This is why for the initial creation of the structure of this template (in 2012) some of the decisions were difficult and when there were no terms for a notion, descriptive phrases were used instead.

The template prepared by NALIS is applicable not just for the 19th-century books, but also for a wider range of editions, since it reflects the specifics of any volume. It can be used as a basis when creating METS files for digitisation not just of non-periodicals, but also of manuscript codices and periodicals. Certainly, then some of the elements of the template should be removed, others - to be added or slightly paraphrased. In the case of the periodicals, for instance, there should be just one more line per each year.
First digital exhibition planned

It is crucial to take this first step towards a digital exhibition of books – to build the basis for their online presentation, make them searchable for the end user and multilingually operable for the librarian, convenient to use and refer to, and robust in the face of new standards and future types of information media. Needless to say, all drawbacks – of the template in our case – should be removed at an early stage. We need to synchronise our watches before finding the best way to arrange digital books in virtual exhibitions. Ideally, the library of the future is going to be presented as a virtual tour allowing the user to pick up items he or she is interested in and read through them. There is a long way to go from the current state of so much digitised content of cultural importance, scattered here and there, often poorly presented, hardly searchable at all. Yet the online presentation of books through the NALIS Union Catalogue is heading the same way.

While waiting for the times when a real digital library will be virtually possible, we can still only play with digitised cultural content, looking for the most attractive ways to present it, selecting fragments from books that would be most interesting for a larger audience or that we would like to popularise. As far as the digital collections created in relation to the NALIS project are concerned, one should be aware of the fact that the majority of the complex-structure items – and, currently, the most valuable in terms of both history and language – are exactly the 19th-century books.

Another important circumstance is that the 19th century was the period of national liberation after 5 centuries under the Ottoman Turks, and that this included, among other things, strong educational activity. Thus, a vast proportion of all the books of the time are educational and not only for children. There are many grammar books and primers, histories, geographies, mathematics, physics, and chemistry textbooks, but also catechisms, books in cosmography, zoology, botany, propriety; on law, economics, social study, and politics. That is why it seems to me that the most valuable subject of a digital exhibition – at least at this current stage – is the illustrations...
such an exhibition would be much narrower.) One of the major characteristics of those books, as far as their decoration is concerned, is that they are definitely imitative of other nations’ books from the Age of Enlightenment and yet there are numerous local peculiarities that would be of interest for any art historian. How an image or a concept has developed, where it originated, what influenced it — such an exhibition may well help in assembling the puzzle.

The manner of arranging such a digital exhibition — of illustrations found in 19th-century educational books — should not be uniform. Moreover, digitality is ideally suited for it. Multifaceted presentation, i.e. various options to be chosen by the user according to his or her needs, should exhibit the objects not only thematically and chronologically, but also by origin and other specifics. In addition, the...
content should be searchable by keywords (year, artist if known, author of the book, topic, what is depicted, etc.). As the illustrations on the title page and back covers would also be included (there are hardly ever any frontispieces in the collection in question), there should be classification also by their type. Whatever the arrangement, any opportunity should be used to exhibit treasures hidden in old books – whether images, scripts, decorations or pieces of language – which would also attract the attention of the audience to the whole books and collections these books are kept in.


2. METS reflects the structure of the object without decoding its text, for which ALTO is usually used. Online presentation, search and use are. The combination of these two standards is used in the largest projects in digitisation of periodicals.

3. For details about the structure of the METS files see e.g. Cantara 2005 (cf. endnote 1).

4. By Evgeni Dimitrov, a NALIS system administrator.