RÁKÓCZI THEMATIC DIGITAL ARCHIVE AT THE NATIONAL SZÉCHÉNYI LIBRARY

Németh Márton; nemeth.marton@oszk.hu; (National Széchényi Library)

In this article a new project is being presented from the National Széchényi Library about the Rákóczi Archive. In a project framework in 2019 by developing a demo app we could demonstrate the integration of digitized and digital-born contents. It is useful to find a historical event for this purpose that can represent our common past in the Carpathian Basin and can be represented by a multilingual digital interface. This Rákóczi project offering a good opportunity to present archived web content in the same context with digitized materials from other digital library collections. The target group is broad, it can be also useful for researchers, students, and librarians.

INITIAL GOALS

Libraries are building archives from digitised and digital-born resources. Web content is a digital-born resource. Archiving of this kind of document type has become a permanent service in some national libraries in Europe. These web archives usually are separate collections. In this article we would like to offer an

Figure 1 Integration of major digital content resources as setting pieces of a whole puzzle

example about the integration ways of these resources. (Figure 1).

The web is a complex medium, harvesting, long term preservation and retrieval of web materials is a serious challenge from an IT perspective. Web documents saved by National Széchényi Library web archive are being harvested and retrieved by various software products.

By building-up a thematic collection, the strong connection through the representation of digitized and digital-born resources from various collections such as Hungarian Electronic Library, Electronic Periodical Archive and Database, Digital Archive of Images, Web Archive can be presented.

In a project framework in 2019 by developing a demo app we could demonstrate the integration of digitized and digital-born contents. It is useful to find a historical event for this purpose that can represent our common past in the Carpathian Basin and can be represented by a multilingual digital interface.

The Rákóczi Uprising and Freedom War (1703 – 1711) is a part of our common historical past from the former Kingdom of Hungary (an integral part on that time of the Habsburg Empire) and heavily effects the territory of the current Slovakia. Francis Rákóczi II. the leader of the freedom war was finally buried in Košice (Kassa) in the 19th century. If we want to describe the remembrance of this historical period through a digital content service it involves a need of representation of a digital interface in at least three languages: German, Hungarian and Slovak. Moreover, historical connections are clear to some Polish territories, to the power policy of the French Kingdom. An English interface is being needed for the general international audience.

RÁKÓCZI ARCHIVE

The Hungarian parliament proclaimed 2019 a memorial year of Francis Rákóczi II as he were proclaimed to the Prince of Transylvania in 1704. That is why a digital archive (http://rakoczi2019.webarchivum.oszk.hu/) has made in 2019-2020 on the Francis II. Rákóczi Memorial Year at the National Széchényi Library (NSZL), by the framework of the Public Collection Digitization Strategy (KDS). (figure 2) A major project of this strategy implied

Figure 2 General information about the Rákóczi Archive in the web portal

Figure 3 List of results of a metadata search with green padlocks (open access) and yellow padlocks (restricted access, only on a dedicated library network)

the involvement public libraries to web archiving tasks in a project framework called KD-K. We wanted to establish a collaboration framework with public libraries in order to get more digital materials that can be found in the local collections.

The primary goal to present the technology of web archiving and presenting the integration of the web archive with other digital collections through a demo application. The content is focusing on the webpages and websites related to the memorial year, to the Prince and to his family. Furthermore, it contains born digital or digitized books from the Hungarian Electronic Library, articles from the Electronic Periodical Archives, photos, illustration and other image-based documents from the Digital Archive of Images. This collection offers an extensive set of resources about the Hungarian social and military history from the 17th and 18th centuries, for education and research.

Browse and search functions are available for metadata of each collection. (figure 3) In the hit lists different colour codes shows the origin of items by collections. Padlock icons shows the availability status of a document. Red or yellow padlocks appears on those archived webpages where the website owner has not or not yet granted authorization for public access. However, for these kinds of not publicly available documents the metadata, a small screenshot image and links to relevant items on the Internet Archive are being offered. Full text search function is only available for the publicly available items in the web archive.

The collection of the archived materials ended in February this year, last year we digitized books in the Hungarian Electronic Library, in February 2021 we searched for 45 websites for additional websites, thus the collection ended. The archive will be available in the future as a non-updated permanent content service of the National Széchényi Library.

WEBSITE

In the last quarter of 2019, we created an independent website for the Rákóczi Archive, which also served as a prototype for the new website of the NSZL web archive. Previously, we only published the news of the project in a temporary, manually edited HTML file, and the public demo collection, bibliography, wiki, annual workshop page and recommendation form were available for anyone to suggest website related to Hungary and Hungarian culture.

The new website is WordPress-based, and the interface of the Rákóczi collection was also created with this content management system, for which we also developed various search and browser functions.

The materials of the archive can be listed by thematic categories, document types and collections. You can search all metadata at once or by combining several fields, and the result lists can be sorted by four criteria. Several kind of statistics are also available about the content types of the collection (figure 4). For legal and technical reasons, the full-text search engine only covers websites and only those that are publicly available. This variety is partly justified by the fact that Rákóczi and the impact of the war of independence were of European significance, and although most of the documents in the archive are in Hungarian, it may be useful to draw the attention of foreign researchers to them. The other reason is that we believe that libraries in other countries are interested in such a service to demonstrate how to turn a content service out of archived websites and other digital or digitized documents.

COOPERATION

Apart with the creation of the Rákóczi website, another main element of the project was focusing on collaboration with other institutions. We were looking for librarians who could participate on the KDS-K project. By the request to appoint a contact person with whom we can discuss the tasks that can be performed in the framework of the cooperation. Fortunately, most institutions are not completely unfamiliar with this area, as some of their staff members have already participated in the 404 Not Found – Who Preserves the Internet? annual web archiving workshops, or one of the courses organized by the Library Institute entitled Archiving the Internet as a Public Collection Task. We reviewed the work that could be done by the end of May in the form of video chats with contacts and contributing colleagues in late February and early March 2020. We primarily asked for help in selecting sites worth archiving, as local colleagues are more familiar with important online sources of information in the region, as well as institutions with their own website, blog, Facebook or Instagram page, public figures, artists or even individuals who publish interesting content on the Internet for a wider audience.

In a longer term, we would like to involve colleagues working in county and city libraries in the licensing process, as well as in quality control and metadata, at least on an experimental basis, so that Hungarian librarians also have practical experience in this field. In addition, on request, we are ready to hold external lectures and presentations or even courses. We hope that the professional contacts that are now being established will continue after the deadline of the KDS-K project, and that we will be able to continue working
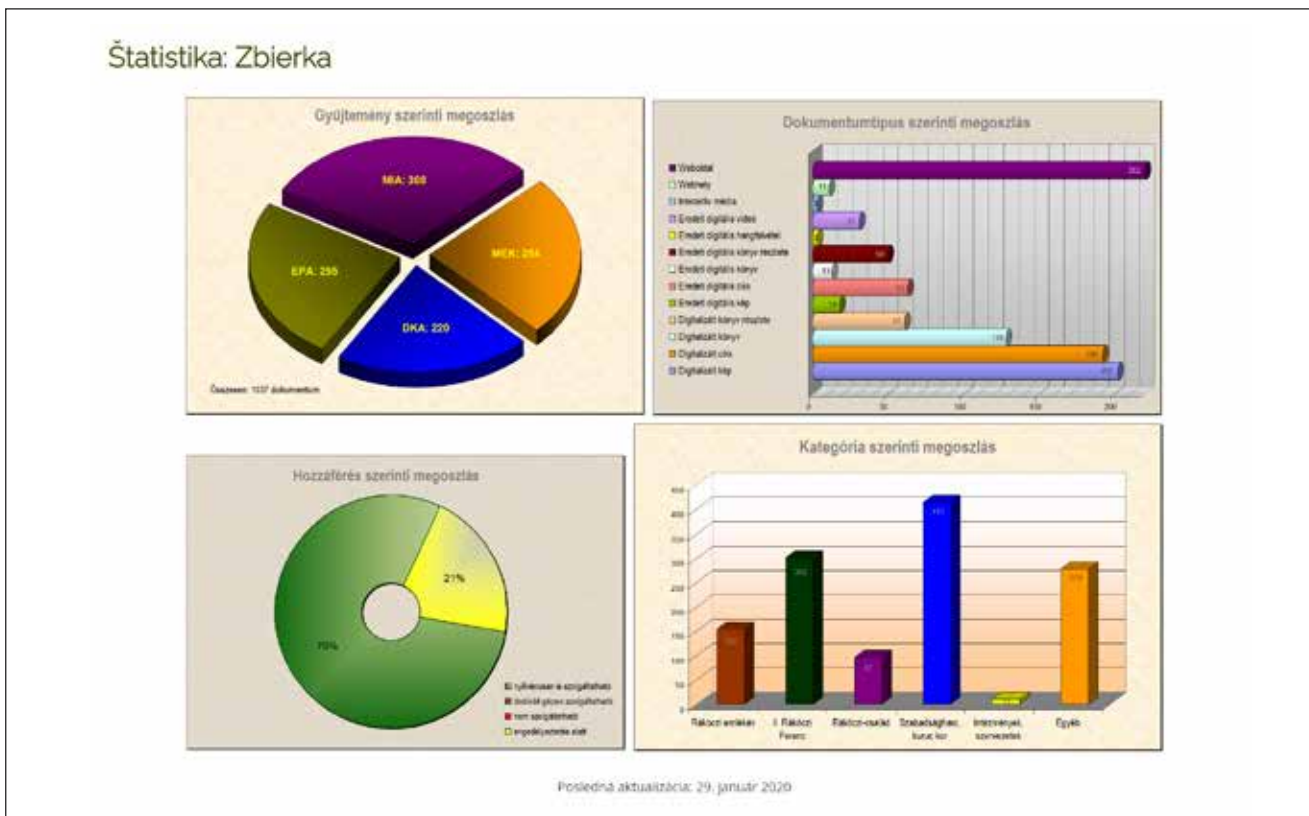
Figure 4 Resources statistics. The first figure shows segments by major collection types, the second figure refers to statistics by document types, the third figure illustrates the percentage of open access availability of content, the fourth one takes an overview about major subject categories of content.

with these libraries in some way. Archiving the national web space is a task of such size and complexity that carried out in cooperation with several institutions. In the first period, a co-operation can be imagined in such a way that the archiving and the conclusion of contracts are carried out by the NSZL, and the Hungarian public collections help in the other work phases. Among the archived web contents, libraries can integrate the ones that are important to them into their own digital services, either by linking them from the public archive or by using the NSZL points planned in the OKP project, i.e. the terminals placed in the libraries, for the non-public collection gives users access.

CONCLUSION

This Rákóczi project offered a good opportunity to present archived web content in the same context with digitized materials from other digital library collections. The target group is broad, it can be also useful for researchers, students and librarians. The layout concept of the interface of this collection is also rather new, reflecting to the topic, to the user's needs, and to the appearance of each collection segment by content types. By getting the experience from the build-up of the website of this special archive, we could also make a new complex portal for the permanent web-archiving service. Another major advantage that we could take contact with other libraries starting to create a broad collaboration framework in web archiving field.

References

Drótos, László L. & Visky Ákos László (2020). Rákóczi-archívum. Könyv, Könyvtár, Könyvtáros 29(3), 35-48. http://ki2.oszk.hu/3k/2020/05/rakoczi-archivum/ https://epa.oszk.hu/01300/01367/00326/pdf/EPA01367_3K_2020_03_035-048.pdf

Németh, Márton (2020). Az OSZK Webarchívum új honlapjának felépítése és szolgáltatásai, Könyv, Könyvtár, Könyvtáros 29(6), 16-26. http://ki2.oszk.hu/3k/2020/08/az-oszk-webarchivum-uj-honlapjanak-felepitese-es-szolgaltatasai/ https://epa.oszk.hu/01300/01367/00329/pdf/EPA01367_3K_2020_06_016-026.pdf

Visky Ákos, László (2020). Gyorsmérleg az OSZK Webarchívum és a KDS-K pályázat nyerteseinek együttműködéséről, Könyv, Könyvtár, Könyvtáros 29(7-8), 24-38. http://ki2.oszk.hu/3k/2020/10/gyorsmerleg-az-oszk-webarchivum-es-a-kds-k-palyazat-nyerteseinek-egyuttmukodeserol1/ http://epa.oszk.hu/01300/01367/00330/pdf/EPA01367_3K_2020_07-08_024-038.pdf