

HUNGARIAN NATIONAL LIBRARY PLATFORM

The Freedom and Responsibility of Choice



Modular and modern collaboration platform for the Hungarian library community

In the past decades the wealthier part of humanity that has access to the internet, has occupied the large and continuously broadening space of the world wide web in terms of communication, publishing, cooperation, work and entrainment alike. The majority of the users have transferred from the scope of passive reception to that of active production and publication; consumers have become contributors, procreators of new contents and active attendants of processes including multiple participants. The physical limits earlier set by space and time have vanished on communication channels through which information can be shared and accessed, and a global connection has been set up between the most diverse fields.

A significant part of culture is being now created in the online space, the larger part of written cultural heritage now exists exclusively in this form that is (or is considered to be) technologically more volatile and vulnerable. The predominant part of the manuscripts treasured by the national library is in digital format; the collected web contents and webpages today add up to a decisive proportion of cultural treasures. The national libraries of other countries have also expanded their services to different extents in order to collect, process and service contents born electronically. The international practice has also provided a colourful overview showing to what extent external representatives are being involved into enriching, or verifying as such, the data stored in library systems. It can be generally confirmed that in developed and modern societies the relationship between cultural institutions and other representatives (publishing houses, printing presses, copyright societies, commercial database providers, scholars, researchers, civilian experts etc.) has been expanded significantly; the majority of the collected material and the legal deposits are being received in digital formats; the users use the majority of the services electronically. The library catalogues are no longer

the descriptions of repositories published on the web, but the catalogue itself has become part of the cyberspace, its components are the parts of the web, they exist exclusively there, 'they are made of web' – they can be generated, searched and maintained as elements of the online space.

During the development of information technology, the robust software solutions making the online communication and supply possible were made many years ago. The use, connectedness and expansion of these can result in such a library system that could advance the services of libraries in a dynamically scalable manner to service-oriented levels users have already got used to by other fields and would satisfy the users' raised expectations. The most comprehensive service is necessary in the most individual manner.

In 2016, the National Széchényi Library received authorisation (government decree nr. 1605/2016. [XI. 8.]) to establish the HNLP in collaboration with the KIFÜ (Governmental Information Technology Development Agency). This governmental intention overlapped with those aspirations for renewal included in the letter of intent of the most significant Hungarian libraries, according to which: 'the Hungarian libraries find it necessary to establish a new-generation, unified library platform based on innovative technology that could guarantee effective collaboration...'. (Letter of Intent, 7 December 2016)

During the EU tender discussions, the representatives of the libraries who took part in the system's planning phase soon realized that the current library systems available would not meet the requirements, cannot be freely extended as expected and are so expensive that they cannot be managed in the cultural sector on the long run. However, the FOLIO (Future of Libraries is Open) set exactly the aims mentioned above, and has already executed some of them. The most important founders of the FOLIO are the Open Library Foundation, a community of library professionals, EBSCO, one of the world's leading content provider companies, and the Index Data, which has decades-long experience in the technological development of library systems. The NSZL has taken part of this from the beginning, and is member of

the product council that is the most important professional decision-making body in the developmental directives of the platform. Thanks to the EBSCO's significant financial support, such an international community took shape under the name FOLIO, where by the collaboration of librarians and IT professionals a solution was established that proved to be a 21st century one in every aspect. It has become clear through the process of the tender that the FOLIO platform is the only available choice, therefore it was decided that the Hungarian National Library Platform had to be based on it. Unfortunately, due to some complications in the tender process, the actual work could only be launched in April 2019. The first library to 'join' with all of its library functions will be the KSZK (Library and Information Science Library) in the summer of 2020 according to the plans, if – by replacing the ILS (Integrated Library System) AMICUS – the HNLP will overtake all the existing ILS functions and will complete these with new, modern and comprehensive services. The notion of platform also reflects these additional options: the Library Platform will fulfil a wider variety of services in a more modern manner, than the earlier ILSs. According to the contract, this comprehensive collaboration platform will be ready by the end on 2020 with all of its functions, in which the services of the KSZK and the NSZL together with national services will be available to the users, both to libraries and readers. In line with the plans, the platform will be able to meet all the requirements of any kind of libraries who wish to join; in case the parameters are not sufficient, new modules and minor functional units can be connected to the system.

Compared to the currently operating integrated library systems, this platform has several new elements. Even though partially modern solutions are being used by many current software suites now, there is still no such platform that, similarly to the HNLP, would consistently fulfil the following revolutionary innovations in a native manner, i.e. from the beginning of its birth and comprehensively by applying to all the elements of the system as opposed to the earlier Integrated Library Systems:

- **flexible metadata formats:** customised data format can be generated through parameters. As a result of the appropriate parameters, the HNLP can manage multiple data formats simultaneously: the legacy of MARC (Machine Readable Code), the BIBFRAME data exchange format currently taking shape, the slightly less complex structure of Dublin Core just like any other

formats specific of museums or archives. The system's inner data model is regardless of the specifications of the given field, it can be modified and extended as necessary. This freely rich data set can be adjusted to the unique formats of specialized systems, making the standardized data exchange possible. The stored data can become excessively richer at a given data exchange format, as in this case the 'additional information' can be linked in form of a remark or note to the HNLP's external systems fulfilling a narrower scope of possibilities;

- **entity data model:** in order to make particular data elements and their connections identifiable, and make complex, implicit, semantic searches possible, the implementation of so-called namespaces is necessary. In the namespaces mutually managed by institutions in collaboration, we keep record of the entities via clearly identifiable unique identifications of people (authors, contributors), geographical units, works etc.;
- **overall integration of analogue and digital contents:** the HNLP does not differentiate between the media of a given unit, it can uniformly manage any type of it;
- **freely determinable workflows:** the joining institutions and partners will be able to freely generate by parameters the most suitable workflows for their needs. The components of the system are the smallest functional units possible which can be addressed from any given point of the process and they branch off according to the given situation. The processes taking place can be subject to conditions (depending on whether these conditions are met or not, the workflow will branch off into multiple directions), triggered by events (the occurrence of some kind of event causes the subsequent event) or/and depending on authorizations (the ones taking part in the system can execute different operations according to their authorization);
- **computer-aided processing:** by certain well-defined algorithms computer processing can effectively support manual workflows, either by automatized recognition (text mining, image identification etc.), or by automatically fulfilling embedded subtasks. In the light of the results achieved, the user is able to provide further clarifying information for the computer's algorithm, therefore the machine will be able to learn continuously;
- **freely customisable online user interfaces:** the libraries can individually generate user interfaces of any number and type through parameters, on which

the range of the data, their visualization and design can be freely determined. By this it will be possible to come up with home pages for subcollections or completely unique design of the library's information homepage, while in the background, from large mutual metadata storage and object collection, the system will be able to select the serviced content according to its authorizations both by dynamic or static selection or by the combination of these two approaches;

- **versions/qualification levels:** the system allows for the storage of multiple versions of the same data. Based on the indication of the source by the publisher of the data it can differentiate between multiple versions of the same data down to the smallest data unit – the user will clearly know the source of the given data all the time. Within the collaboration platform the generated data will be classified into the appropriate qualification levels based on professionalism and credibility of the collaborating members' classification. The user can choose whichever level of the data set raises their interest, namely whether to see everything or only the data from certain sources or exclusively those the credibility of which reaches a certain qualification level;
- **degrees of connection:** the system allows for the libraries to move in all of their existing functions and by this to replace its current ILS. It is possible for the HNLP to only fulfil certain selected functions, such as maintaining and publishing only a fraction of the digital collections. Even less constrained connections are possible, when the connecting library synchronizes its metadata alone in the HNLP, and 'redirects' to the collection stored in its local cloud. The degree of connection is freely scalable;
- **free choice of modules:** the HNLP implements the functions of all types of libraries in a comprehensive manner, namely it has at disposal all kinds of basic modules that we have managed to identify through surveys. As it has a modular built, the integration of modules developed by others into this system is also made possible, moreover, modules of the same type can operate simultaneously (for example, the operation of multiple cataloguing modules at the same time can be implemented, as well);
- **scalability and load sharing:** the software is made of robust components of almost entirely open-source codes. It can operate in hybrid cloud: the library's own cloud, the cloud service operated by the KIFÜ or solutions offered by professional providers (for example

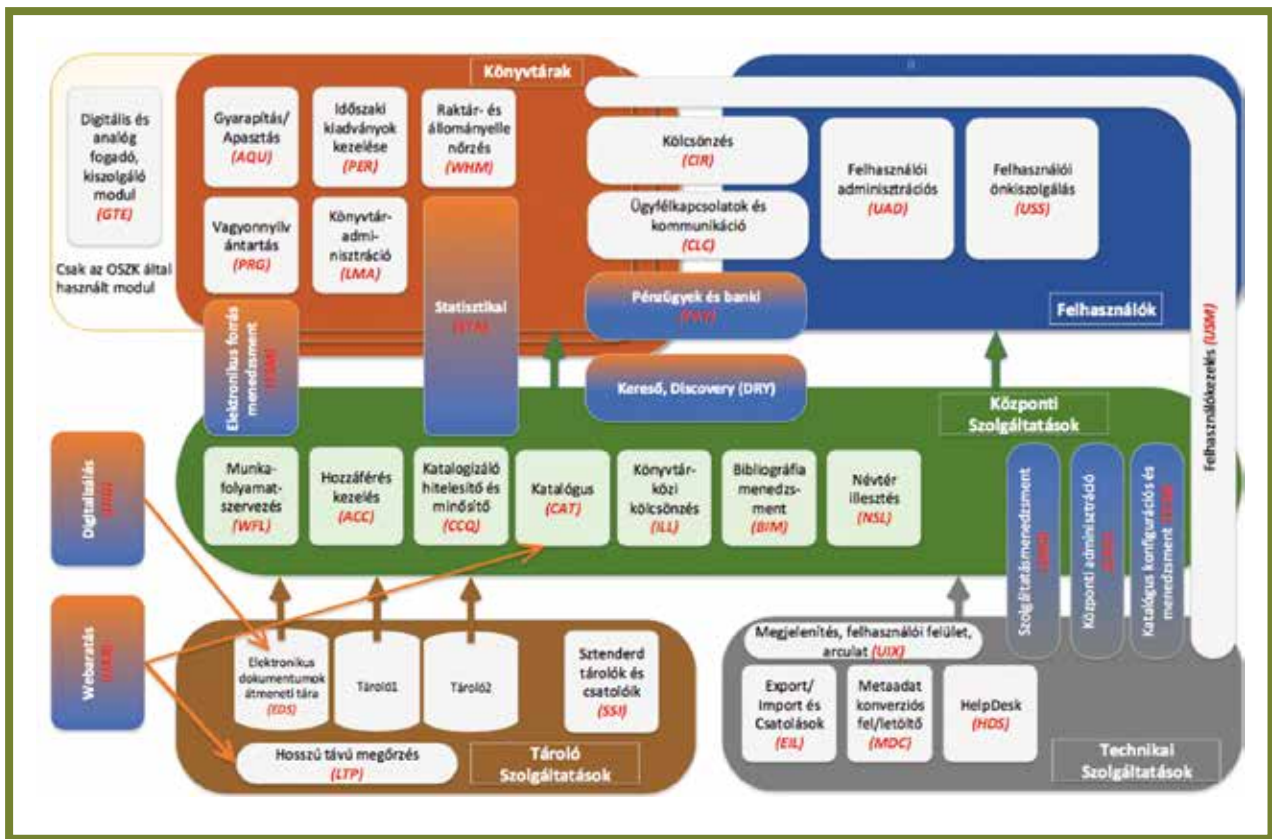
Amazon AWS, Microsoft Azure, Google One) can all be mounted into the operation, even simultaneously. In line with the increasing load the system is able to clone and launch components automatically, in order to satisfy the increased demand by concurrent operations. The infrastructure for the HNLP system is provided by the KIFÜ in a georedundant manner, i.e. it is mirrored through systems in geographically secure, remote locations.

The system utilizes the ultimate solutions of information technology, making even the most specifically individual configurations possible. By this a comprehensive platform can be generated which includes the traditional library functions, focuses on the modern service for the users, it is service-oriented and integrates the result of the workflows of digitizing, taking care of the long-term preservation of the most important contents at the same time.

The greatest challenge that occurs during the implementation of the HNLP in 2020 is: the manner we can expediently and effectively migrate the current data set (metadata and digital objects), and how the libraries will be able to move into this platform in a way to understand the opportunities given by these new data models, profit from the advantages of this new approach and not repeat those bad practices that might exist currently.

The data that was generated in the libraries in the past decades are of very different quality when considering their extent of completion or accuracy. Regarding the migration of the data into the system, it is fundamental to determine which data will be accepted by the system, what sort of data reparation we will be able to accomplish while uploading and into what qualification level we will classify certain data. In case of the entity-based data model it is practical to generate entities out of given data elements and apply unequivocal identifiers on the linked data. (We try to use the cataloguing standard RDA – Resource Description and Access – already at migration.)

Taking the division of labour into consideration, the new system might revolutionise the collaboration among institutions. By the consequent development of the MNN (Hungarian National Namespace) a given entity has to be generated only once, and all the processors will be able to connect to this (moreover, making the unambiguous connections possible, excluding any duplicates). The authors and publishers can be involved in the description of the works already on the interface of the ISBN Office, and the data provided by them will be ready for use when



Functions of the HNLP

generating the catalogue entry. It will also be enough to generate a catalogue entry once and it becomes available and referable at all times for each member. The metadata and digital objects will be reusable as well when developing collections, exhibitions and also for setting up thematic web pages.

During the past few years, starting from 2016, the developments have focused on the HNLP. In order to take advantage of its opportunities as much as possible, the consortium of the NSZL and KIFÜ have generated many other components within the National Library System project. This is how the whole hardware infrastructure has been renewed, including the equipment for the readers and staff members of the NSZL, and three server parks and storage units were established: in the NSZL's building in the Buda Castle and in the KIFÜ data centers located in Budapest and Debrecen. This establishment guarantees high availability and georedundancy during operation. A digitizing centre including one of Europe's most modern set of tools was established by the NSZL in order to provide digital access to analogue collections, and it developed a software called DTK (Digitization Support Framework System) which streamlines the digitizing processes (units originating from the most diverse sources can be digitized

through a variety of workflows, in the most colourful forms of collaboration – avoiding any redundancies and rationalizing the division of labour). Like HNLP, the DTK facilitates the national collaboration between the institutions, therefore it might be suitable to support the members of KDS (Digitizing Strategy for Public Collections). The formation of the software of the Hungarian National Namespace has been a significant development, the data of which will be maintained by multiple institutions in a form of a consortium.

The developments of the FOLIO community are taking place in many countries. The leaders of its implementation are the Swedish, German and American libraries. As far as functionalities are concerned, the Hungarian development is more comprehensive (for example the integration of the ISBN Office or connecting web archiving) and technically more robust than the implementations of other countries. However, the Hungarian deadlines are tighter than in other countries. Therefore, it is obvious that in the case of the development of some modules, a closer integration will be possible for the solutions developed simultaneously which are not possible now due to the limited time available.

At the moment, within the HNLP project the software development, the data migration, the training of the

colleagues, the testing and the course material production are running simultaneously. As we have mentioned, by the end of 2020 the HNLP software suite will be ready and the migration of the data of the national services will be completed – including among others the national common catalogue (MOKKA) and the data regarding the copies of the interlibrary document supply (National Document Supply System), by collecting (harvesting) them automatically on time, and requesting them in form of files from the databases of the partner libraries. The Library and Information Science Library and the NSZL will both be integrated into it with all of their functionalities.

This indeed new generation and without exaggeration the world's first such system hides great potentials. The key

to success is the activity of the librarians open to modern solutions. If we are able to learn about all the opportunities, as well as reshape and renew the everyday processes, by the end of 2020 the libraries will use effectively a software that indeed measures up to the 21st century and opportunities not seen before will open up for readers and researchers when it comes to learning about the treasures of Hungarian culture.

Additional information online: <http://okp.oszk.hu>

Miklós Lendvay
lendvay.miklos.at.oszk.hu