LEGAL DEPOSIT MANAGEMENT INFORMATION SYSTEM

Katalin Bubnó[1], Zsuzsanna Jeszenszky[1], Georgina Molnár[1], Gellért Nagy[2], Viktor László Takács[3]

[1]University and National Library University of Debrecen (HUNGARY)
[2]Eötvös Loránd University (HUNGARY)
[3]Institute of Applied Informatics and Logistics University of Debrecen (HUNGARY)

Abstract

Legal deposit is one of the most significant collections of the Hungarian national cultural heritage and a very important part of the collection of the University of Debrecen University and National Library (DEENK). The first packages of the legal deposit collection arrived 70 years ago, in 1952. Anyone interested in science has the opportunity to use this library collection for research. This collection is invaluable. Nevertheless, due to low interest, little attention is given to the fact, how much work the cultural institutions, and within that to DEENK, have to do with it. There has always been a need for librarians to accurately document their work. That is why during the library processing work, the book carts have long been accompanied by the so-called ”cart sheet”. The paper-based cart sheet is the same age as the legal deposit collection in DEENK. Nevertheless, earlier there had been no need to process and extract information from the sheets about the parts or the whole of the workflow. In 2020 the administration of the current processes were digitized and in 2021 a manager information system was set up to bridge this gap.

The Library uses an integrated library system called Corvina. However, there are more work phases around the legal deposit that are not documented in the Corvina system, so a modernized version of the cart sheet accompanying the book carts should be kept. In this paper we introduce our system and also the final dashboards as provided for the managers of the Library. A hybrid modeling methodology was applied. On one hand, we planned from the top by mapping and formalizing management questions, but on the other hand we used the bottom-up approach by building a data cube through an Extract - Transform - Load (ETL) process from the online cart sheet filled by the colleagues in an everyday routine. Intermediate steps are also presented: requirement analysis, source analysis, integration phase with the transformation processes. With the help of an entity-relationship model the working of the system is also demonstrated, as well as the dashboards of the final system.].

Keywords: Legal deposit, University of Debrecen University and National Library, Data warehouse hybrid modeling methodology, Management information system, Reporting system.

1 INTRODUCTION

1.1 Legal deposits in Hungary

It is the responsibility of every printing house and publisher that produces a document in Hungary to provide legal deposit copies, and this includes not only Hungarian, but also foreign language documents. There is only a quantitative limitation on the number of copies produced: they are obliged to provide 6 copies if at least 50 copies are published. Under 50 copies only 2 of them are enough. Printed publications must be provided by the printing house, but it is the responsibility of the publishers to ensure

this. In all other cases, the publisher is also obliged to provide the copies, e.g. if publications were “produced abroad and published in Hungary” or published in an electronic form. Based on the above, DEENK is forced to accumulate a large number of documents due to the regulation, because it operates as a library with a national collection.

According to the current decree, the distribution of legal deposit copies is as follows: two out of six copies remain in the National Széchényi Library (OSZK). One of them is intended to preserve the national cultural heritage (archived copy), the other one is to serve its function in library services. The University and National Library of the University of Debrecen (DEENK), as a library with a national collection, also performs a service function, for which purpose it receives only 1 copy. Copies received have to be assigned with accession numbers and provided for readers to borrow or for interlibrary loan or for in-library use. The Hungarian Central Statistical Office Library (KSH) receives 1 set of complete copies on the basis of the “Act on Official Statistics”. The Library of the Hungarian Parliament (OGYK) and the National Library of the Foreign Literature and Music Collection (OIK) receive documents belonging to their collection, while one series of copies is shared by the libraries of the University of Pécs and that of the University of Szeged. (Government Decree 717/2020. (XII. 30.))

Other libraries with the right to make selections about the remaining documents and books have the opportunity to view them and to sort them in accordance with their needs at the times specified by the OSZK.

The six legal deposit copies are considered to be outstandingly high at the international level. In countries larger than Hungary, publishers are obliged to provide two or four copies. In some of them publishers and libraries decide on preservation by mutual agreements. (Bázing, 2013) There are countries where the number of legal deposit copies is high, also in relation to the territory and population of the country. Two of these are Sweden and Denmark. In the first one 7 copies are currently required to be provided, and in the second one 2 copies of printed documents, although these two are indirectly given to two specific libraries. (Majoros, 2012)

1.2 Legal deposit in the DEENK

Collecting the legal deposit copies has always caused difficulties for the library of the University of Debrecen. Not only because of the amount of material, but also the interpretation of the regulation has caused some difficulties for library managers in many cases. The decree currently in force (Government Decree 717/2020 (XII. 30.)) does not specify the purpose of collecting the legal deposit. The definition of the "real" purpose can be found in a previous regulation:

4. § The purpose of the legal deposit service is for press products

a) collection and preservation by the national library,

b) national bibliographic register and state statistical census,

c) making it available in the library system. (Decree No. 17/1986 (VIII. 20.) MM)

The Library of the University of Debrecen has previously asked whether it is necessary to store all publications physically. Thanks to the National Document Supply System, researchers living in any part of the country have the opportunity to access any documents of scientific value. However, the storage of publications without such value is unnecessary. The number of books and documents published has increased significantly since the 1970s, and even more especially after the end of communism. This is a huge difficulty for the Library as the storage capacity is finite. Further difficulties were reported in 1991, problems that have been persisting to this day. In many cases, books for preservation do not arrive properly due to library services, e.g. borrowing. (It is cheaper to pay the publishers a penalty than to deliver the publication.) It becomes difficult to control the service, or the documents do not get to the libraries properly (incomplete, bad copies). (Gomba, 1994) Tracking and claiming tasks give extra work to the libraries receiving the legal deposit.

The DEENK (then DEK) became a national library under the direction of Máté Kovács. He held the position of director of the University Library in Debrecen between 1949-1956. His aim was to make the Library a major factor in the country. In accordance with the government decree issued in December 1951, in addition to the OSZK, the University Library of Debrecen also became a complete collection place. This made it the second most important library in the country, presenting new tasks and challenges for the library. Due to the increased responsibilities, new classes and groups were set up. In 1952, the cataloging and reference and circulation services separated. The former was even expanded with a scientific bibliography group. At that time, the so called "Cart Companion Sheet" (or cart sheet)

was developed, the purpose of which was to measure and record the lead-time of books. (Reports) All phases of work were marked on the cart sheets from equipping the documents to storing them. Although in the 1990s the predecessor institution of the Library switched to a computer-based system, the cart companion sheet remained and accompanied the documents during the processing journey. However, the data stored on the sheets were never aggregated and the information that could be extracted from them was never displayed/transmitted to the managers/decision makers. Thus, we only had impressions and conjectures about the work of the colleagues working in the process, their daily activities, and the lead-time and other indicators of the work processes.

In 2001, a year after the integration of the universities, the libraries were also integrated and today's DEENK, one library with 6 units, was established. From the 2000s onwards, the thematic distribution of legal deposit copies began in the Library of National and General Collections (now Arts and Sciences Library, or BTEK) among other member libraries in order to ensure that the documents were as close as possible to their potential readers, while relieving the burden on the Main Library.

Until 2014, the DEENK's libraries operated as independent units, with their own management, their own acquisition, cataloging, as well as reference and circulation services. The real integration took place in 2015, with the creation of departments organized according to the different workflows, independent of geographical location, and the posts of associate directors to coordinate other projects.

In 2016, the Quality Management Council was established, with several subgroups out of 25 employees of the Library. Most of them were not managers but colleagues. Their task was to prepare the material for the Quality Award and the Qualified Library title application as well as to reconsider the strategy of the Library and, accordingly, its operation. Several working groups were organized and a business process-based review of DEENK's earlier operations was initiated. The Process Management Group played a key role in this, and because of their work, a process inventory of 125 processes was completed. In the process inventory, each process received a textual and a standard BPMN (Business Process Management and Notation) description, as well as a process form, in which the process parameters were defined (e.g. input-output events, critical points, indicators, target values, labour recently, the DEENK has undergone further structural transformations. Although the libraries on the various campuses were transformed into one organization as early as 2001, the services, tasks, and working methods used could not be standardized everywhere. The services were standardized first, as it was important for the Library that UD citizens and external readers using the Library should have access to the same services on all campuses under the same conditions. After that the standardization of internal working methods and work processes could begin.

Today's organogram of DEENK shows the reorganization of the Library according to the process inventory. According to this, the organization has three divisions: Customer Relations (Front), Collection Building (Back), and Support Section.

[kép]

Defining process indicators and target values is an important part of process management. However, it is not enough. The fulfillment of the assigned goals must be regularly monitored.

2 METHODOLOGY

To describe the processes, we use the BPMN (Business Process Model and Notation), which has brought a new way of thinking to the Library. BPMN is “a standard business process model” in which business processes are presented graphically. The marking system is conventional, which can be interpreted by anyone, to describe complex processes, thus creating a “harmony” between the actors who take part in the workflow (performing, planning, leading). BPMN represents a series of processes and events, which is why it is considered event-driven. The tools used are wide-ranging, easy to learn, and one does not have to know all the tools available in order to use them. The figures themselves contain a standardized XML markup language that allows one to choose the software. The Library uses the Yaoqiang editor for BPMN 2.0, which is built up of 3 levels. At the first level, the known signs of the process descriptions, at the second level, in addition to expanding the possibilities of event control, the standard transaction control, and at the third level, the existing elements are expanded with additional elements based on common properties. The system design was preceded by the design of a business process that helped normalize the completed system. (Bátfai-Lakatos-Molnár-Takácsné, 2016)

The DEENK now consists of 9 units. On seven campuses there are 120 colleagues working in the Library to carry out both university roles and national responsibilities. The accession and cataloging of

legal deposit copies are mainly the work of the departments of the Content, Collections and Discovery Section (Collection Development Department, Metadata Department, Collection Management Department), supplemented by the Subject Specialist Group belonging to the Education and Research Support Department. In smaller units, the colleagues of the Reference Department sort the documents for the open shelf area and do the administration of the process. So we can say that in the workflow, …% of the staff is involved in some way in their everyday work.

The DEENK has been working in an integrated library system for more than 30 years. It records incoming documents and performs their services. However, the system we use (Corvina Integrated Library System) does not have a management information system from which, for example, we can regularly monitor indicators of processing legal deposit copies. In addition, there are ongoing activities that are invisible to the Corvina system because they are not performed there (e.g., manual works such as equipping, anti-theft protection, etc.). That is why the cart sheet, which accompanies the process, has never worn out.

In this study, we present the system that has been created by processing the electronic data of the cart sheets and which can be used to answer questions that concern library managers, such as

- How much did / do the colleagues deal with the legal deposit copies from the time the documents arrived at the Library to the time of shelving them?
- How are legal deposit copies distributed among each library collection?
- What types of documents do our colleagues deal with the longest / the shortest period of time?
- What types of documents do the most / least of our librarians deal with?
- Which activities are done by how many people and for how long?
- What activities can slow down the whole process that prevents the bookcarts from reaching their desired destination in time?
- How much do book carts wait between each activity?
- Is the central processing location “emptied” (from one shipment to another)?

The system’s indicators are presented on executive dashboards using various visualizations that are clear and understandable to everyone. A lot of innovation had to be introduced to carry out this task, as there has been no example of data being processed in this way since the creation of the cart sheet in 1952. We use data warehouse technology for data processing.

Defining the system and the system design is difficult to implement, it is only possible to know approximately what it covers. Many people have come up with many different definitions for it. One of them is as follows: "In a system thus we mean a whole that consists of elements, things, objects, element relations, relationships, relations, and connections that have certain properties." (Szenteleki, 2007, p.10)

There is also no uniform definition of data warehouse. The two best known figures in data warehouse design are Bill Inmon and Ralph Kimball. A possible definition of the latter from his book The Data Warehouse Toolkit: Data Warehouse: "The conglomeration of an organization's data warehouse staging and presentation areas, where operational data is specifically structured for query and analysis performance and ease-of-use." (For the Hungarian translation see: Szatmári, 2009, p.7)

Operational support systems within a “company” can be divided into two major groups: transaction-oriented OLTP (On-Line Transaction Processing) and analysis-oriented OLAP (On-Line Analytical Processing) systems. The difference between the two types is that in transaction-oriented source systems, employees do their daily work (that is, we record transactions in them), in these we track daily events, while analytical processing, which aims to collect long-term information and support decisions, takes place in OLAP systems. In the latter, the historical data of the company is stored and presented. Between the two of them there is the ETL (Extract-Transform-Load) process, in which data is extracted from transaction systems at regular intervals, cleaned if necessary, and transformed so that reports can be made from them, and finally loaded into the data warehouse structure. Transactional systems are characterized by a large number of users, i.e. employees who document daily events in the systems, while the users of OLAP systems are decision-makers and managers who have access to information. During the design of the OLTP, an individual-relationship model is built, and the resolution of the data is detailed, while OLAP deals with the processing of object-oriented, previously saved data. (Csepei, 2009)

Two methodologies for classical data warehouse design are known. One of them is a data-driven (Inmon), the other one is a requirement-following (Kimball) method. Both approaches have their drawbacks. The data-driven method may make it impossible to answer many questions from a larger, more complex data source due to tangled data. With the requirement-following method, there is a real danger that a cube will be created that can answer only one question.

Both systems therefore have advantages and disadvantages, which is why in the last approx. one and a half decades, hybrid design methodologies have evolved. These try to take over the advantages of the two methods and eliminate the disadvantages if possible. We have used such a hybrid methodology when redesigning the system (Takács et al., 2019).

1. Requirement Analysis. We collect questions with metrics and dimensionality of the problem also in this phase with the required visualizations from the decision makers via interviews. Then we formalize these specifications with our special structured stenography that is based on the terminology corresponding to the current problem. The output of this step is a set of formalized questions.
2. Deriving minimal granularity. Based on the set of formalized questions we specify the required minimal granularity for every indicator. The output of this step is the set of indicators with minimally detailed dimensions.
3. Deriving ideal schemata. We map the dimensional attributes and values to keys, produce the initial conceptual schemata. The output of this step contains ideal dimensions (keys, attributes and hierarchies) and ideal facts (with dimension keys for join), independently from the sources.
4. Source Analysis. The main question of this step is: What kind of transactions can we get them from? We decompose ideal facts into potential elementary transactional attributes and identify them in the source systems. The output of this step is the derived potential schemata.
5. Integration. Ideal schemata from the requirement analysis are compared with potential star schemata. Match occurs, when the two schemata contain the same fact, and, both have the same dimensionality in the same granularity level. In this step we define required transformations and calculate fact tables and common dimensions with attributes.
6. Multidimensional modeling. We build the cube(s) with dimensions, dimension hierarchies and measures.

3 RESULTS

3.1 Requirements analysis

In the requirements analysis phase, we collect and formalize managerial questions. The question and its description are defined in a textual and a formal way. A “manager/leader” is a person whom the system provides with information. As a result, he/she expects the report to appear in the form of visualized data on an IT interface. Reports and dashboards must be structured according to the requirements of access rights (it is not all the same, what level of managers can access certain reports).

In the requirements analysis, we “analyze” management issues based on the following aspects of Table 1:

What is the indicator? In which aggregation? What is the unit of measure? What kind of visualization would we like to see? At what resolution? What kind of slicers can be applied?

Table 1: Management question analysis

Indicator unit(s) aggregate function(s)	$\boldsymbol{I}^{\{u[,u]\}}_{\{af[,af]\}}$	the indicator ***I*** to be produced with *u* unit(s) in the upper right index and af aggregate function(s) in the bottom right index,
visualization slicer(s)	$\left(\begin{matrix} vt \\ \left[\left\{\begin{matrix} s \\ s \end{matrix}\right\}\right] \end{matrix}\right)^{v}$	the *v* visualization with the type *vt* (table, line diagram, bar graph, etc. ...) and optional *s* slicers (values can be $\boldsymbol{D}_{\{a\}}$ dimensional attribute, $\boldsymbol{D}_{\{v\}}$ subset of concrete values, or a $\boldsymbol{D}_{\{a\}}$ dimensional attribute in the *d* detail of another ***I*** indicator on the same dashboard)
detail(s)	$\left[\left(\begin{matrix} \boldsymbol{D}_{\sum\{a\}} \\ \left[\boldsymbol{D}_{\sum\{a\}}\right] \end{matrix}\right)^{\{d\}}\right]$	*d* details with $\boldsymbol{D}_{\{a\}}$ dimensional attribue(s), with optional $\sum\{a\}$ aggregation. *d* values e.g.: *row*, *col*umn, *cat*egory, *y* indicator

Formally: $\boldsymbol{I}^{\{u[,u]\}}_{\{af[,af]\}}\left(\begin{matrix}vt\\ [\{^{s}_{s}\}]\end{matrix}\right)^{v}\left[\left(\begin{matrix}\boldsymbol{D}_{\Sigma\{a\}}\\ [\boldsymbol{D}_{\Sigma\{a\}}]\end{matrix}\right)^{\{d\}}\right]$

The formal definition above covers the configuration and the display of indicators and these are related to one diagram. Several different visualizations could arise in the manager, related to the requirement specification, so many formal descriptions will be made in this section.

Table 2. Management questions and dashboards.

Logical description	Formally
Where is the bottleneck of the process? Waiting activity is where the average waiting time is the most. Sliced by year. Average waiting time by Previous and Next activity and by year.	$dA_T\left(\begin{matrix}card\\ \{TOP(\uparrow\boldsymbol{T}^{\boldsymbol{day}}_{\boldsymbol{avg}}(\mathbf{dA}_{\mathrm{T}}),1)\\ \mathbf{dA}_{\{\mathrm{wait}\}}\\ \mathbf{dD}_{\mathrm{Y}}\end{matrix}\right)^{v}$ $\boldsymbol{T}^{day}_{\mathrm{avg}}\left(\begin{matrix}matrix\\ \mathbf{dA}_{\{\mathrm{wait\ time}\}}\end{matrix}\right)^{v}(\mathbf{dA}_{\mathrm{T}})^{r}(\mathbf{dD}_{\mathrm{Y}})^{c}$ $\boldsymbol{T}^{day}_{\mathrm{avg}}\left(\begin{matrix}matrix\\ \mathbf{dA}_{\{\mathrm{wait\ time}\}}\end{matrix}\right)^{v}\left(\begin{matrix}\mathbf{dA}_{\mathrm{T}_{source}}\\ \mathbf{dDA}_{\mathrm{Y}}\end{matrix}\right)^{r}(\mathbf{dA}_{\mathrm{T}_{dest}})^{c}$
Average lead time of bookcarts by main types and/or month and year. Sliced by pandemic and department.	$\boldsymbol{AT}^{day}_{\mathrm{avg}}\left(\begin{matrix}column\\ \{\begin{matrix}pandemic\\ department\end{matrix}\}^{s}\end{matrix}\right)^{v}(\mathbf{dD}_{\mathrm{M}})^{y}(\mathbf{dD}_{\mathrm{Y}})^{sm}$ $\boldsymbol{AT}^{day}_{\mathrm{avg}}\left(\begin{matrix}column\\ \{\begin{matrix}pandemic\\ department\end{matrix}\}^{s}\end{matrix}\right)^{v}(\mathbf{dBC}_{\mathrm{type}})^{y}(\mathbf{dD}_{\mathrm{Y}})^{sm}$
Processed legal deposit distribution by libraries, category, librarians. Sliced by main types of bookcarts	$\boldsymbol{KM}^{\%}_{\mathrm{sum}}\left(\begin{matrix}pie\\ legal\ dep\end{matrix}\right)^{v}(\mathbf{dDest}_{\mathrm{library}})^{l}$ $\boldsymbol{KM}^{pcs}_{\mathrm{sum}}\left(\begin{matrix}treemap\\ \mathbf{dBC}_{\mathrm{type}}\end{matrix}\right)^{v}(\mathbf{dBC}_{\mathrm{cat}})^{cs}(\mathbf{dE}_{\mathrm{librarian}})^{r}$
Individual performance (avg time and sum processed document) per employee. Sliced by department, year and pandemic	$\boldsymbol{AT}^{day}_{\mathrm{avg}}\left(\begin{matrix}bar\\ \{\begin{matrix}pandemic\\ department\\ year\end{matrix}\}^{s}\end{matrix}\right)^{v}(\mathbf{dE})^{y}(\mathbf{dT}_{\mathrm{TT}})^{l}$ $\boldsymbol{ASz}^{pcs}_{\mathrm{sum}}\left(\begin{matrix}pie\\ \{\begin{matrix}\mathrm{proctime}\\ pandemic\\ department\\ year\end{matrix}\}^{s}\end{matrix}\right)^{v}(\mathbf{dE})^{l}$
Department performance (avg time and sum processed document) per employee. Sliced by year and pandemic	$\boldsymbol{AT}^{day}_{\mathrm{avg}}\left(\begin{matrix}column\\ \{\begin{matrix}pandemic\\ year\end{matrix}\}^{s}\end{matrix}\right)^{v}(\mathbf{dDept})^{y}(\mathbf{dT}_{\mathrm{TT}})^{l}$ $\boldsymbol{KM}^{pcs}_{\mathrm{sum}}\left(\begin{matrix}column\\ \{\begin{matrix}\mathrm{proctime}\\ pandemic\\ year\end{matrix}\}^{s}\end{matrix}\right)^{v}(\mathbf{dDept})^{l}$

3.2 Source analysis

In this project, our source is clearly the cart sheet and its electronic version. It shows the lead-time of a book. This is the time interval after the books arrive from the OSZK and are distributed and then at the end they are shelved in the appropriate library unit. The books arrive at the Kassai Street Campus Library (KCK). They are sorted according to the decree, and they are put on the carts according to the

destination. This is done thematically, by library units and by subject specialist librarians within that unit. The colleagues receive the cart sheet that accompanies the cart from the initial station to the destination. From 2020 the “path” of the cart has been recorded electronically, too. The cart number is indicated in the first line, followed by the type of material with the related note and destination. In addition to signing the monogram of the person starting the cart, he/she also indicates the number of books. During the process the books are equipped (stamp, barcode, RFID), assigned with accession numbers, cataloged. All of these happen at the central cataloging and acquisition place at the KCK. The colleague responsible for equipping the books and documents packs them into separate boxes per cart sheet. After this the delivery takes place. The starting and the ending date of any activity with the monogram of the person performing it are indicated on the cart sheet, together with the number of books and documents processed. If there are documents removed from the cart (for any reason), that must also be marked on the cart sheet (both electronically and on paper). The so-called “fast processing” documents are equipped and stocked by the designated colleagues. In the end a call number is given to every single document, and the “In transit” code is removed from them. (This code ensures that the reader can see in the online catalog during the processing activity that the book is not yet available.) In case of “fast carts” the subject analysis takes place during the bibliographic description, in case of the so-called “normal carts” later on, after delivery in the appropriate library, where the subject specialist awaits the material. In other words, parallel activities can take place. In the case of normal carts, it is the subject specialist who directs the books either into the open-shelf area or to the stacks. The books are then given their call numbers.

KOCSIKÍSÉRŐ LAP						
Kocsiszám	Anyag megnevezése fő kategória	Anyag megnevezése megjegyzés	Igényhely	Indítási dátum	Kocsit indító kézjegye	Darabszám indításkor
Leírás kezdete	Leírás befejezése	Leíró kézjegye	Feldolgozott könyvek darabszáma	Feldolgozott mellékletek darabszáma	Átadás oka	Másnak átadott dokumentumok száma
Leltározás kezdete	Leltározás befejezése	Leltározó kézjegye			Átadás oka	Másnak átadott dokumentumok száma
Szakozás kezdete	Szakozás befejezése	Szakozó kézjegye	Raktárba irányított dokumentumok száma	Szabadpolcra irányított dokumentumok száma	Átadás oka	Másnak átadott dokumentumok száma
Szállítás dátuma						
Helyrajzi szám adás kezdete	Helyrajzi szám adás vége	Helyrajzi számot adó kézjegye				
Mozgásban kód levétel dátuma	Mozgásban kódot levevő kézjegye					Mozgásban kód levétel darabszáma

Figure1: Cart sheet

Source entities based on the cart sheet

eBC: BookCart entity

- CartID: unique identifier attribute in format: {yyyy}-nr
- MainCategory: Hungarian equivalent of the main classification from {gift, obliged, purchased, exchanged, retrospective}
- SubCategory: Hungarian equivalent of the secondary classification from {-, fast, normal, local}
- Librarian: initials of the name of the librarians
- Destination: values from {faculties, libraries and their sub organizational elements)
- PCS: number of documents
- Comment: freestyle text

eA: Activity entity

- CartID: BookCart identifier attribute in format: {yyyy}-nr
- Date: date of the activity in {yyyy}.{mm}.{dd} or {yyyy}-{mm}-{dd} format
- Employee: initials of the name of the employees
- Activity: recorded step of the process mentioned above
- State: Hungarian equivalent from {start, end}

Additional entities to classify employees

- eEmployee(ID, DeptID, Position)
- eDepartment(ID, DepartmentName)
- ePosition(ID, Position)

Extra dimensions

dDate: classic date dimension with only the needed attributes and hierarchy levels

- Date: date of the activity in {yyyy}.{mm}.{dd} format
- Year: in {yyyy} format
- Month: number from {1, …, 12}
- MonthName: Hungarian name of the months
- MonthShortName: Hungarian name of the months (3-4 letters)
- QuarterCode: number from {1, …, 4}
- Quarter: roman number from {I, II, III, IV}
- PandemicTF: logical value
- Pandemic: ‘Waves of pandemic’ and non-pandemic classification of dates

dHoliday: holiday and transferred workday calendar

- Date: date of the activity in {yyyy}.{mm}.{dd} format
- DayName: Hungarian equivalent of the {Monday, …, Sunday}
- Type: Hungarian equivalent of the {holiday, transferred workday}
- Value: derived from the type respectively {1, 0}

4 CONCLUSIONS

In the DEENK, for the first time since 1952, it is possible to show by using electronic data how much time librarians deal with legal deposit copies. Which departments and which colleagues spend the most time with it. It shows exactly how long it takes for each cart to be "emptied", which type of book description takes longer or shorter time, which activities require carts to be stationed the most, which books go where, in what distribution. And managers also can get an idea of, for example, how the pandemic affected the life of the Library, the everyday work in terms of processing legal deposits. There are a lot of innovations that need to be introduced to accomplish this task, because since the creation of the “cart sheet” in 1952, there has been no example of anyone processing data in this way.

Constructing the system fits in very well with DEENK's upcoming tasks of quality assurance work. The review of the work processes and the definition of the indicators and target values are now on the agenda. The KPIs set by the Associate Director of Content, Collections and Discovery in this process are that incoming books should be available to the reader within 2 weeks. The system we have created helps us to monitor this. If the process was stuck somewhere and the achievement of the target value was hindered, it would be possible to intervene immediately, especially at the activity where the jam occurs.

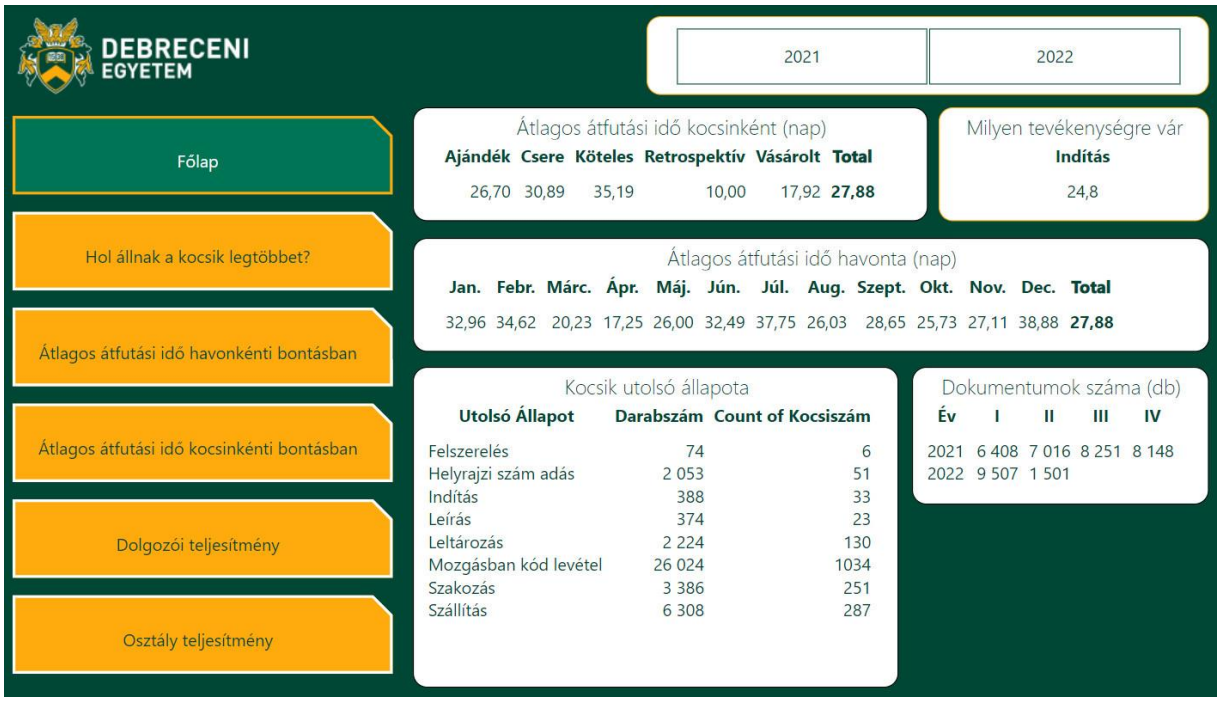

Figure2: Executive dashboard summary

The system will allow us to process not only the current data, but also the data that will be recorded electronically later on. This allows for comparison. It is a serious achievement that librarians have begun to monitor their own work and understand why it is important to accurately record their activities. This is an important change, as so far attempts by the management have mostly failed.

Of course, there are still further opportunities in the system. It is mostly possible to form a basis for return calculations by adding costs. Of course, this part may already require authorization management, as operational and strategic management typically have different access to financial data.

Anyway, the Library has taken another step towards modernization by processing and implementing some of their operational data with this technology. At the same time, we must emphasize once again that the underlying cart sheet, together with the legal deposit collection, is 70 years old. The concept already existed then. In today's words, our predecessors already had the “IT thinking”, they just did not have the right technology to extract valuable information from the huge amount of data.

By now, Clive Humby’s mathematical phrase, “Data is the new oil,” has become a proverb. Then, Palmer complete: “It’s valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc., to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value.” (Palmer, 2021)

REFERENCES

17/1986. (VIII. 20.) MM rendelet a sajtótermékek köteles- és tiszteletpéldányairól (In Hungarian)

717/2020. (XII. 30.) Korm. rendelet a kiadványok kötelespéldányainak szolgáltatásáról, megőrzéséről és használatáról (In Hungarian)

Bátfai, E., Lakatos, R., Molnár, G., & Takácsné Bubnó, K. (2016). BPMN 2.0 a könyvtári minőségbiztosításban. *Tudományos és műszaki tájékoztatás*, 63(8), pp. 305-316. (In Hungarian)

Bázing, Zs. (2013). Virgács Húsvétra – avagy támadás a szellemi tulajdon ellen. Jogász az új kötelespéldány rendeletről. *Tudományos és műszaki tájékoztatás,* 60(7), pp. 318-320. (In Hungarian)

Debreceni Magyar Királyi Tisza István Tudományegyetem Könyvtárának jelentése az 1930-1953. tanévről. Debrecen, Hungary: Magyar Királyi Tisza István Tudományegyetem. (In Hungarian)

Di Tria, F., Lefons, E., & Tangorra, F. (2017): Cost-benefit analysis of data warehouse design methodologies*. Information Systems* 63, pp 47–62.

Gomba, Sz. (1994): A kötelespéldány-szolgáltatásról. A kötelespéldány szolgáltatás debreceni tapasztalatai. *Könyv, könyvtár, könyvtáros*. 3(5), pp. 13- 19. (In Hungarian)

Institutional organogram from the webpage of Debreceni Egyetem Egyetemi és Nemzeti Könyvtár. Retrieved from https://lib.unideb.hu/hu/organogram

Kovács, M. (ed.). *A debreceni Kossuth Lajos Tudományegyetem Könyvtárának évkönyve (1954-1958)*. Debrecen, Hungary: Kossuth Lajos Tudományegyetem Könyvtára. (In Hungarian)

Majoros, E. (2012): *A kötelespéldány-szolgáltatás gyakorlata, jogi keretei itthon és külföldön*. Szakdolgozat. Debrecen, Hungary: Debreceni Egyetem Informatikai Kar. (In Hungarian)

Legal deposit informations from the webpage of Országos Széchényi Könyvtár. Retrieved from https://www.oszk.hu/koteles_szolg

Összefoglaló jelentés a Debreceni Egyetemi Könyvtár 1958-1976. évi működéséről. (1958-1976) Debrecen, Hungary: Kossuth Lajos Tudományegyetem Könyvtára. (In Hungarian)

Palmer, M. (2021): Data is the New Oil. *ANA Marketing Maestros*. Retrieved 16 May 2021.

Szenteleki, K. (2007): *Információs rendszerek.* Debrecen, Hungary: Debreceni Egyetem Agrár- és Műszaki Tudományok Centruma, ISBN 978-963-9732-67-4. (In Hungarian)

Takács, V. L., Bubnó, K., Ráthonyi G. G., Bácsné Bába, É., & Szilágyi, R. (2020): Data Warehouse Hybrid Modeling Methodology. *Data Science Journal*. 19:38. pp. 1–23. doi: 10.5334/dsj-2020-038

Varró, T. (2006): *A Debreceni Egyetemi Könyvtár története 1949-1956-ig*. Szakdolgozat. Debrecen, Hungary: Debreceni Egyetem Informatikai Kar. (In Hungarian)