

1.08 Objavljeni znanstveni prispevek na konferenci  
1.08 Published Scientific Conference Contribution

Lajos Körmendy\*

## **BUILDING AN AUTOMATED REPOSITORY ASSESSMENT SYSTEM IN THE NATIONAL ARCHIVES OF HUNGARY**

### **Abstract:**

*Repository is the key factor of archival preservation: keeping the records in appropriate conditions is the most effective preservation job. Assessing the repositories is a prerequisite to any decision concerning long term preservation. When assessing one must take into consideration several hundred wide-ranging and interrelated facts and aspects as well as a huge amount of data. To do it well an IT system is needed. This system is intended to be a web based user friendly tool which enables an average archivist or librarian to make an objective, reliable and comprehensive assessment of any archival or library repository. The basis of the system is an electronic questionnaire containing 170 questions. When filling in the questionnaire the computer associates risk points to every answer, and - on the basis of these points - displays the risk level in security, fire protection, climatic conditions, etc. with comments. The National Archives of Hungary is looking for partners to finish the project.*

### **Key words:**

*preservation, repository assessment, web based IT system*

### **Izvleček:**

#### **Vzpostavlanje avtomatskega sistema za oceno skladišč v Madžarskem državnem arhivu**

*Skladišča so ključni dejavnik materialnega varstva arhivskega gradiva; hramba gradiva v primernih pogojih je najučinkovitejša zaščita pred uničenjem. Predpogoj za kakršnokoli odločitev glede dolgoročne hrambe je ocena skladišča. Pri tej mora ocenjevalec upoštevati na stotine širokih in medsebojno povezanih dejstev ter vidikov kakor tudi veliko količino podatkov. Za dobro opravljeno delo potrebujemo IT-sistem. Slednji je kot spletno, uporabnikom prijazno orodje namenjen temu, da arhivistom ali knjižničarjem omogoča objektivno in zanesljivo ocenjevanje kateregakoli arhivskega ali knjižničnega skladišča. Osnova sistema je elektronski vprašalnik, ki vsebuje 170 vprašanj. Ob vnašanju odgovorov računalnik oceni točke tveganja za vsak odgovor, na podlagi teh točk izpiše nivo tveganja pri varnosti, požarni varnosti, klimatskih pogojih itn. ter doda tudi opombe. Madžarski državni arhiv trenutno išče partnerje za končanje tega projekta.*

### **Ključne besede:**

*materialno varstvo, ocena skladišča, spletni IT-sistem*

## **THE ANTECEDENTS**

Repository is the key factor of archival preservation. Just an example: if you postpone the disintegration of acid papers for 70-80 or 100 years by assuring cool preservation conditions in the repository, you do not need to spend many millions of Euros for paper deacidification project now. We can say that keeping the records in

---

\* Dr. Lajos Körmendy, head of the Department of Preservation, National Archives of Hungary, 54-56 Úri street, 1014 Budapest, Hungary.

appropriate conditions in repositories is the most effective preservation job. But before doing anything one must be aware of the state of the repositories, i.e., one must assess them.

In 2009 the Hungarian Archives Advisory Board asked me to revise the repository requirements, which are parts of the Recommendation for the Preservation of Archival Material issued by the National Archives of Hungary in 2005. At the beginning I focused on the appropriate climatic values (temperature, relative humidity), and I made a survey of international literature, especially the standards. Looking through the standards I realized how *different* they are. For example the recommended parameters for long term keeping of paper records are as follows:<sup>1</sup>

	Germany	Hungary	UK	Denmark	ASRAE	Australia	IFLA	NARA	Netherlands
Temperature (°C)	2-18	19	13-16±1	12±1	-20 or 10	20±2	18	18	17±1
Relative Humidity (%)	35-50	50	45-60±5	45%±5	40 or 30-50	50±5	45-55	35-45±5	52±3

Another problem is that the standards are *inflexible*, which means that values beyond the recommended parameters are regarded as substandard. Some standards get around this difficulty by using from-to or plus-minus values. But it makes difference how big the deviation is, and how long it takes. For instance, is 24 °C for 3 months equal with 30 °C for 4 weeks? Briefly: the data should be *weighted* when we assess them. But it is not easy at all to make this calculation by a calculator if you have many stacks. The National Archives of Hungary, for example, has 60 stacks.

Many archivists think that assessing a repository is not a complicated job. I do not agree with them. Accessing a repository, especially the old ones, which were not built for archival purposes, is a very difficult job. A repository is a complicated and complex thing, and its assessment must be done on the basis of *many facts and values*, not only of temperature and RH parameters. There are much more than 100 facts and aspects belonging to security, fire protection, building, climate, etc. to be taken into consideration. *Objectivity* is also a dilemma because of the variety and the great number of data and aspects. For example lack of appropriate containers (boxes) must be faced with the lack of a smoke/fire detector and alarm system. What is more risky? Different facts often correlate with each other, and strengthen or weaken the harmful impacts.

After taking into account, comparing and combining several hundred data and information you must get the diagnose, and you also need a therapy: what to do differently, what to change, where to intervene.

Traditionally a repository assessment is made as follows: a preservation expert makes a survey, makes notes or fills in a questionnaire, then looks through the registered data, and finally takes stock of the situation, and makes a report. But such a traditional assessment can hardly overcome the problems mentioned above. The data are so numerous and wide-ranging that processing and weighting them needs a computer, i.e., we need to develop an IT-system.

<sup>1</sup> See for example ISO 11799:2003; DIN ISO 11799:2005-06; Levéltári állományvédelmi ajánlás 2005, Budapest; [http://webworld.unesco.org/safeguarding/en/all\\_envi.htm](http://webworld.unesco.org/safeguarding/en/all_envi.htm); BS5454:2000; [http://www.naa.gov.au/Images/standard\\_tables\\_tcm16-47303.pdf](http://www.naa.gov.au/Images/standard_tables_tcm16-47303.pdf); <http://www.archives.gov/foia/directives/nara1571.pdf>.

The alpha version of such an IT-system is already ready in the National Archives of Hungary. In this paper I would like to show where we are and what we are going to achieve.

## THE IT SYSTEM

### The goal

This system is intended to be a web based, user friendly tool, which facilitates the assessment of any kind of archival or library repository. With the help of this system an average archivist is able to make an objective, reliable and comprehensive assessment.

### The electronic questionnaire

The core of the system is an electronic questionnaire based, first of all, on the National Archives of Hungary's Recommendation for the Preservation of Archival Material, the chapters of the (Hungarian) Handbook for Archives relating to preservation as well as the Document Storage Requirements for Archive and Library Materials (ISO 11799:2003). The questionnaire contains about 170 questions divided into 11 subjects (groups of questions). Predetermined answers belong to most of the questions, others must be filled by the user.<sup>2</sup>

The screenshot shows the 'Electronic Repository Assessment' web application. On the left, a sidebar lists 11 subjects, with '4. Wall, roofing, doors & windows, hygienics, order' highlighted. The main content area displays questions 4.1 through 4.19. Question 4.1, 'Wall material of the repository', has a dropdown menu open showing options: Concrete, Synthetic resin, Artificial stone, Floor tile, and Concrete. Other questions include floor material, wall rendering, roof load capacity, and cleaning frequency. The interface includes a navigation bar with 'Home', 'New repository', 'Amend/erase data', 'Data assessment', 'Parameters', and 'Exit' buttons. The browser address bar shows the URL '193.224.149.7:81/arkar/index.php?site=obekereses'.

Figure 1: Subjects (left side) and detailed questions/answers of one of the subjects

<sup>2</sup> A number of archives apply a questionnaire in order to facilitate collecting and assessing repository data - see for example Stéphane Ipert, Laure Méric, Gilbert Le Guen, Benoît de Tapol, *Une méthode d'évaluation des pratiques de conservation préventive dans un service d'archives*, Février 2001, <http://www.archivesdefrance.culture.gouv.fr/static/1126> or the national questionnaire issued by the National Archives of UK <http://www.nationalarchives.gov.uk/information-management/our-services/self-assessment-results.htmv> - but as far as I know no one has a Web based automated repository and risk assessment system.

When filling in the questionnaire the computer associates risk points to every answer. The range of points is from 0 to 10. Zero means no risk, ten indicates an acute danger. For example the answers to the question about the floor material of the repository (4.2.) may be

Synthetic resin	0 risk point
Artificial stone	0 risk point
Floor tile	0 risk point
Concrete	1 risk point (because it powders away)
Lino	3 risk points (because of danger of fire)
Parquet or strip floor	3 risk points (because of danger of fire)
Other flammable material	4 risk points (because of danger of fire and uncertainty)

After filling in the whole questionnaire, the computer calculates all the risk points, and shows the general risk level of the repository (general assessment). Seven risk categories have been stated: *optimal*, *appropriate*, *average*, *emergency*, *risky*, *unsuitable* and *particularly dangerous*. Depending on the given risk points every subject is assigned to a risk category:

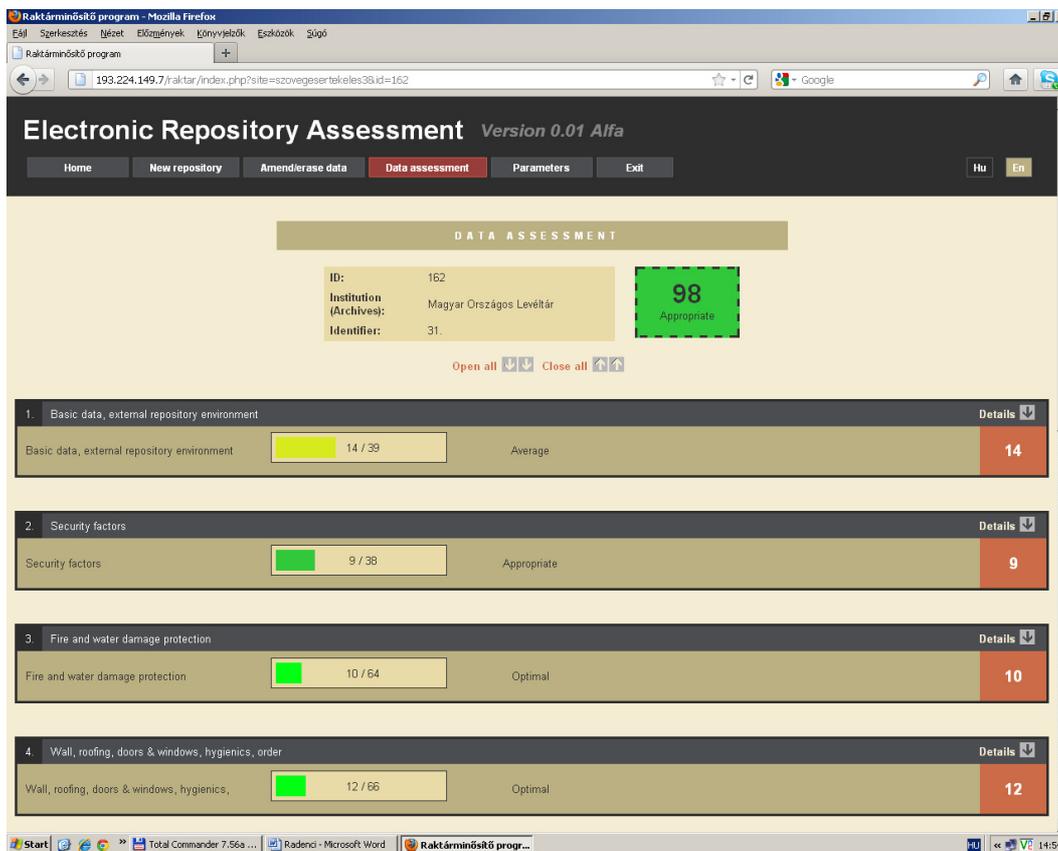


Figure 2: General assessment with risk points and categories

The system also makes a detailed assessment subject by subject. It is particularly important because almost everything is interrelated with almost everything. Let us see an example.

Questions 5.10-15 relate to the record containers (bundles, boxes). The question 5.18 asks about the light damage of the record. If the answer is yes to the questions 5.11 (the majority of the records are kept in bundles) and 5.18 (the records are lightly damaged), it means that the records kept in bundle are not well protected against light, and likely this is the reason of the damage. Therefore the computer displays a warning: „Damage caused by light is worsened by storage without container or in bundles.”

Figure 3: Detailed assessment with risk points, warnings and remarks

### Assessing climatic data

Answering all the questions is good but not enough to get a reliable assessment, the evaluation of climatic data of the repository is also needed. Temperature and relative humidity are maybe the most important preservation markers; for this reason it is strongly recommended to regularly log such data in the repository. The system is able to import CSV (Comma separated Valuables) files which is a common format, and most of the data loggers available on the market can create such files.

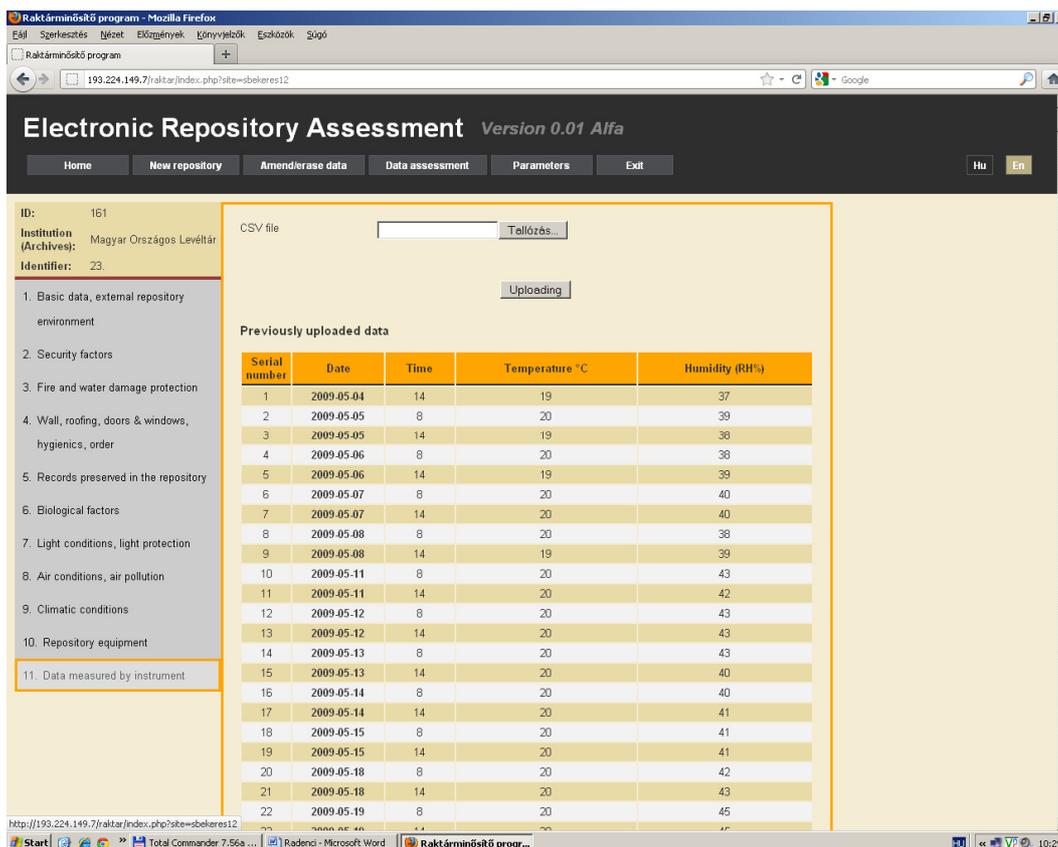


Figure 4: Uploading climatic data

Assessing climatic data means, first of all, that the computer weights them. Predetermined temperature and RH bands have been established and risk points associated to each band (see Figure 5). There are separate band tables for different types of archival material: paper, parchment, black and white/color photographic material, magnetic/optical and mixed material. (The questions 5.1 to 5.7 relate to the distribution of material [in percentage] stored in the repository.)

	Data	Risk points		Data	Risk points
<b>PAPER</b>			<b>PARCHMENT</b>		
Temperature °C	<2	1	Temperature °C	<2	1
	2-14	0		2-14	0
	15-18	2		15-18	2
	19-23	3		19-23	3
	24-26	5		24-26	5
	27-29	7		27-29	7
	30-33	9		30-33	9
	>33	10		>33	10
Daily fluctuation °C	<2	0	Daily fluctuation °C	<2	0
	2	1		2	1
	3	3		3	3
	4	5		4	5
	5-6	7		5-6	7
	7-8	9		7-8	9
	>8	10		>8	10
RH %	<20	9	RH %	<20	10
	20-30	7		20-35	9
	31-39	5		36-40	6
	40-45	1		41-45	4
	46-55	0		46-49	2
	56-58	1		50-55	0
	59-60	4		56-60	2
	61-64	5		61-64	5
	65-69	9		65-69	9
	>69	10		>69	10
Daily fluctuation RH	<4	0	Daily fluctuation RH	<4	0
	4	1		4	2
	5	2		5	3
	6	4		6	4
	7-8	5		7-8	5
	9-10	7		9-10	7
	11-14	9		11-14	9
	>14	10		>14	10
<b>PHOTO - B&amp;W</b>			<b>PHOTO - COLOUR</b>		
Temperature °C	<2	2	Temperature °C	<2	0
	2-14	0		2-5	2
	15-18	1		6-8	3
	19-22	3		9-12	5
	23-25	5		13-15	7
	26-27	7		16-20	9
	28-30	9		>20	10
	>30	10			

Figure 5: Temperature and RH parameters with risk points

## OTHER FEATURES

The system is web based, i.e., that a number of people (archives) can simultaneously use it. An access control will be developed which means that only authorized people may have access to their own data.

The system is multilingual. Now it is working in English and Hungarian, but any new language can be added to it. You can create your national (language) version on the basis of the English text.

## THE FUTURE

Although the most difficult part of the job, the professional one, has been done but the system is not ready yet. Some functions are missing (e.g., access control, maintenance, help), some others do not work well (calculation of risk points, relations), and the whole system is not enough flexible and dynamic. I mentioned above how different the standards are. It will be feasible that registered users may define their own parameters (climatic bands), which will be valid for their assessment. For example an archival institution may consider that a temperature of 18-20°C is ideal for long term preservation, another institution may prefer 12-15°C. The distribution of the risk points can be done separately, according to the preferred parameters.

The National Archives of Hungary is looking for partners to finish the project. An international consortium could finance and manage the project. I think that the achievement of such project would be considerable: the consortium could offer to any archives or library a handy tool through internet to solve the complex problem of the repository assessment.

## HOW TO TEST

You can test the system as follows

<http://193.224.149.7/raktar/>

User name: a

Password: a

If you want to look through the questions and answers click on *Amend/erase data*, than choose a repository. You can see the distribution of risk points by clicking on *Data assessment* (choosing the repository to be assessed is done in the same way). You can also test the system by adding your own repository data (*New repository*). You can upload and assess climatic data in format of CSV files separated by semicolons having the same data structure you can see in the test data of the National Archives of Hungary (except the Serial numbers which are given by the system).

## POVZETEK

### VZPOSTAVLJANJE AVTOMATSKEGA SISTEMA ZA OCENO SKLADIŠČ V MADŽARSKEM DRŽAVNEM ARHIVU

Sistem je osnovan na Priporočilu za hrambo arhivskega gradiva, ki ga je izdal Madžarski državni arhiv, na poglavjih (madžarskega) Priročnika za arhive, ki se nanašajo na to temo, ter Zahtevah za hrambo dokumentov arhivskega in knjižničnega gradiva (ISO 11799:2003).

Sistem je namenjen lažjemu ocenjevanju arhivskih skladišč. Skladišča so zelo komplicirana in kompleksna, pri njihovi oceni pa moramo upoštevati, primerjati in kombinirati na stotine podatkov ter informacij. Avtomatiziran sistem lahko to nalogo dobro opravi in tudi primerja podatke. Tako lahko na primer odgovori na vprašanja, kot so: "Kaj je bolj škodljivo za gradivo: temperatura 32 °C, ki traja en teden v juliju in drug teden v avgustu, ali 26 °C od začetka julija do konca avgusta?" S pomočjo sistema lahko posamezen arhivist poda objektivno in zanesljivo oceno skladišča.

Osnova sistema je elektronski vprašalnik, ki vsebuje okoli 170 vprašanj, razdeljenih v 10 skupin. Pri večini vprašanj so odgovori vnaprej nastavljeni, druge mora vpisati uporabnik.

Pri izpolnjevanju vprašalnika se vsakemu odgovoru dodeljujejo točke tveganja. Njihov razpon je med 0 in 10. Nič pomeni brez tveganja, 10 kaže na hudo nevarnost. Sistem upošteva medsebojno povezana vprašanja in odgovore. Na primer, vprašanja 6.1 in 6.2 se nanašata na plesen. Podana so tudi vprašanja, povezana z vodnimi madeži [3.26], vlažnostjo sten [4.5], prezračevanjem [8.2, 8.7] itd. Za plesen obstaja

veliko vzrokov, sistem pa opazuje vsak odgovor ter pomaga priti do diagnoze; po potrebi tudi opozarja.

Ko odgovorimo na skupino vprašanj, sistem sešteje točke tveganja ter poda splošno oceno za to skupino vprašanj. Po izpolnitvi celotnega vprašalnika računalnik izračuna vse točke tveganja in pokaže splošno oceno tveganja za skladišče.

Sistem je spleten, torej ga lahko istočasno uporablja več uporabnikov (arhivov). Kmalu bo razvit tudi nadzor dostopa, kar pomeni, da bodo do podatkov lahko dostopali samo avtorizirani uporabniki. Sistem je tudi večjezičen ter trenutno dostopen v madžarskem in angleškem jeziku. Možno pa je dodati tudi druge jezike. Trenutno sistemu manjka nekaj funkcij (npr. vzdrževanje, pomoč), ki so šele v razvijanju, nekatere funkcije še ne delujejo dobro (seštevanje točk, relacije), celoten sistem pa še ni dovolj fleksibilen in dinamičen. Najtežja naloga, strokovni del sistema, pa je vendarle opravljena.