BUILDING TIES BETWEEN ARCHIVAL CREATORS AND PUBLIC ARCHIVES – AMBITIONS AND REALITY

Tarvo Kärberg
Tiiu Kravtsev
Ülle Mägin
The National Archives of Estonia

Introduction

The main task of the National Archives of Estonia (NAE) is to ensure preservation and usability of written memory of the society, documented cultural heritage for nowadays and future generations. Therefore NAE collects archival records of public agencies. In order to document the society in all its diversity, NAE collects records created by some selected private and non-profit data producers as well. NAE signed cooperation agreements with some private bodies in 2004, at the time of developing its appraising policy. Currently, NAE's collecting field includes about 75 private institutions, such as the Open Estonia Foundation (OEF).

OEF, that is a non-governmental foundation, was established in 1990 with the support of the philanthropist George Soros. The main goal of OEF is to help people work on behalf of a free and open society. Over the past 20 years the foundation's initiatives are driven by a vision of democracy, civil society, social responsibility and equal opportunities in democratic decision-making processes. OEF has ranged from school reforms to contemporary art and legal reforms. Today their focus is on strengthening open governance and civil society and reinforcing the principles of participatory democracy. OEF runs and supports initiatives to advance openness and democracy in decision-making processes, and inclusive civil dialogue in society. The organisation helps to shape and advocate public policies and actions that assure greater justice in society and promote fundamental rights. OEF also shares experiences with countries in democratic transition and build cross-border partnerships in the European Union and across the globe. Over the years, OEF has supported grants and projects with more than 30 million Euros.

OEF has created several series of records during their business activity which have been appraised as having archival value. NAE chose for permanent preservation the following series of OEF's records: minutes of board and council meetings; director decrees, newsletters, annual reports, applications for projects, contracts and reports. The largest series of records were related to the projects of the OEF (applications for projects, contracts and reports). The previously mentioned series were chosen to be transferred to NAE.

The procedure of transferring records from public agencies to public archives requires a transferor to arrange and describe documents in accordance to the requirements of NAE. The Archives Act of Estonia states that the transferor should prepare the archival records for transfer and bear the expenses of the transfer, including expenses incurred during re-arrangement, descriptive work and transport of archival records according to the requirements (Archives Act of Estonia[1]). Therefore producers may need to do quite a lot of work for preparing the records for preservation and they are very keen on optimizing their transfer processes to overcome the relatively high financial cost that the preparation of archival material may bring along.

NAE has constantly helped all archival records creators in meeting the requirements specified by regulations, no matter whether they are governmental or non-governmental organisations, but it has been still very resource-consuming. Fortunately there has been an innovative jump in this collaboration in the recent past. NAE has created and applied a special pre-ingest software tool, the Universal Archiving Module (UAM) which helps archivists in the agencies to meet the

---

[1] http://www.legaltext.ee/et/
andmebaas/tekst.asp?loc¼text&dok¼2012X02&keel¼en&pg¼1&ptyyp¼RT&tyyp¼X&
query¼arhiiviseadus.

criteria set by the archives. The UAM can be used for describing the archival creator, funds, series, files (cases) and even records (items) and computer files manually or automatically. This tool was used for the OEF transfer project as well.

Case study

OEF submitted the request for transferring of records to NAE in 2013. Before preparing the transfer, OEF had earlier created document-level descriptions (reference code, title of the program and project, date, cost) in an Excel worksheet. There were no descriptions of the organisation (OEF) and archives.

When OEF turned to NAE with its request for transferring these records, it seemed expedient to use UAM in the creation of descriptive data and transferring the digital documents for its predictable efficiency. The main reason was the great amount of descriptions already existing in digital form (in Microsoft Excel files) and also some digitally born documents which needed to be described and transferred. In addition to that, currently about 80% of OEF's records created in everyday business are in digital form and future transfers are also planned to be performed digitally.

Based on the analysis of the institution's functions, the following archival scheme (Table 1) was set up and approved by NAE.

Table 1: Series of the OEF

Series number	Series name in Estonian	Series name in English*
1	Üld- ja personalikäskkirjad	General and personal decrees
2	Nõukogu ja juhatuse koosoleku protokollid	Minutes of Board and Council meetings
3	Infolehed "Teataja"	Newsletter "Teataja"
4	Projektide dokumendid	Projects' documentation
5	Programmi Hea Algus projektid	A Good Start program projects
6	Programmi Avatud Eesti Raamat projektid ja lepingud	Open Estonian Book program projects and contracts
7	CIVICUSe uuringu dokumendid	CIVICUS research documentation
8	Aastakokkuvõtted	Annual reports

* free translation of the series' names from Estonian

Building the archival arrangement and descriptions

UAM provides functionality for designing the archival arrangement by creating various hierarchical levels (fond, sub-fonds, series, sub-series, files, items/records). It can be done in two ways: importing data from XML files, thus creating archival classification scheme automatically or adding levels manually in the UAM graphical interface. If needed, the classification tree can be flexibly modified by adding new levels and deleting or modifying existing ones. It is very comfortable as actions which are not allowed (i.e. adding files directly under a fond, if you have already some series under it) are hidden from the user, so it is difficult to make any logical mistakes.

All levels can be properly described in an automated (importing metadata from XML files) or manual way (inputting metadata into the fields manually). Even the archival creator and transfer project details can be described in the UAM.

All metadata elements are described in the guidelines produced by the NAE. The metadata elements described in the guidelines are very strongly influenced by the General International

Standard Archival Description (ISAD(G))[2] and International Standard Archival Authority Record for Corporate Bodies, Persons and Families (ISAAR(CPF))[3] standards.

The archival classification scheme of OEF was inputted to the UAM manually and enriched with metadata according to the guidelines set by the archives.[4] An example screenshot of the metadata tab "Content and Structure" of a series can be seen in Figure 1.
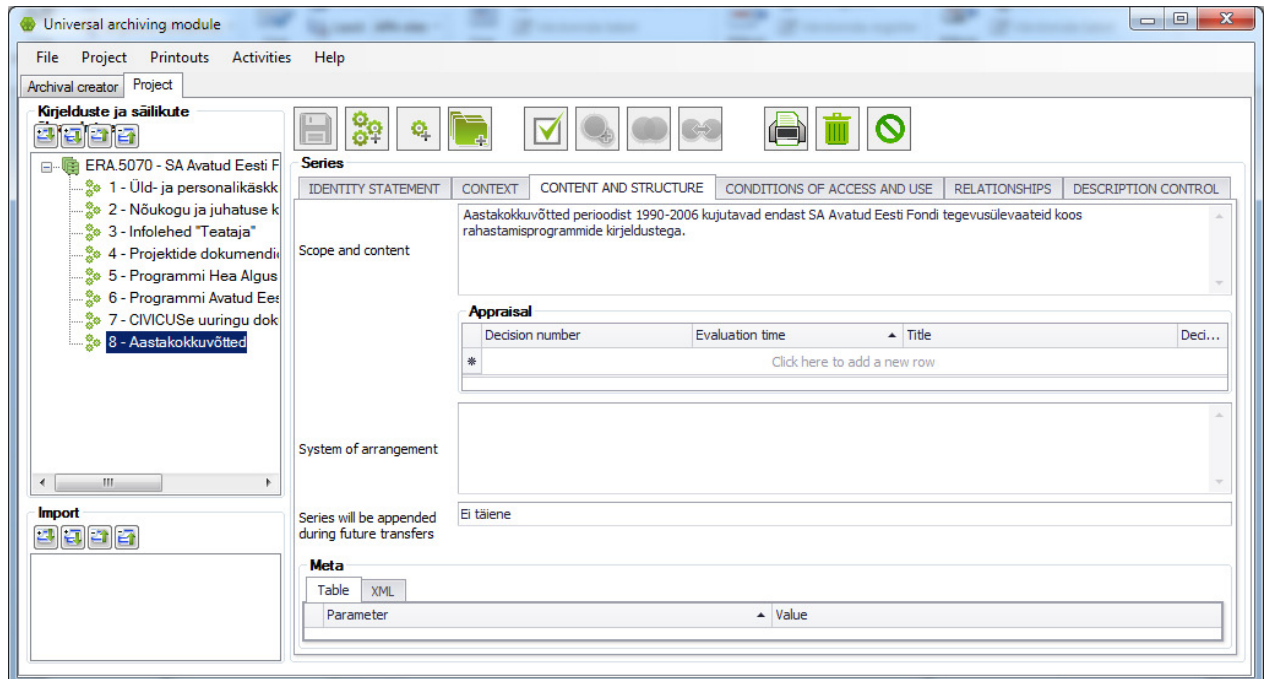

Figure 1: Descriptive data of a series in the UAM

The work was preceded by the process of describing the organisation (archival creator) and fond level. This step can be very time-consumable as it requires in-depth information about archival creator's current situation, past and relationships and acquiring that information from several sources could be complicated. Fortunately this description usually changes only slightly in time and can be used over and over again for future transfers. Therefore the archivists took more time and created a very comprehensive description for OEF. They added information to 7 tabs in total – each containing several metadata fields. An example view of archival creator's "History" metadata tab can be seen in Figure 2.

---

[2] http://www.ica.org/10207/standards/isadg-general-international-standard-archival-description-second-edition.html
[3] www.ica.org/10203/standards/isaar-cpf-international-standard-archival-authority-record-for-corporate-bodies-persons-and-families-2nd-edition.html
[4] NAE's future ambition is to use the same archival classification scheme in next upcoming transfers (send the scheme to the agency if they haven't preserved it), so that the next deliver projects could take less time, thus be even more effective and cost-optimal.
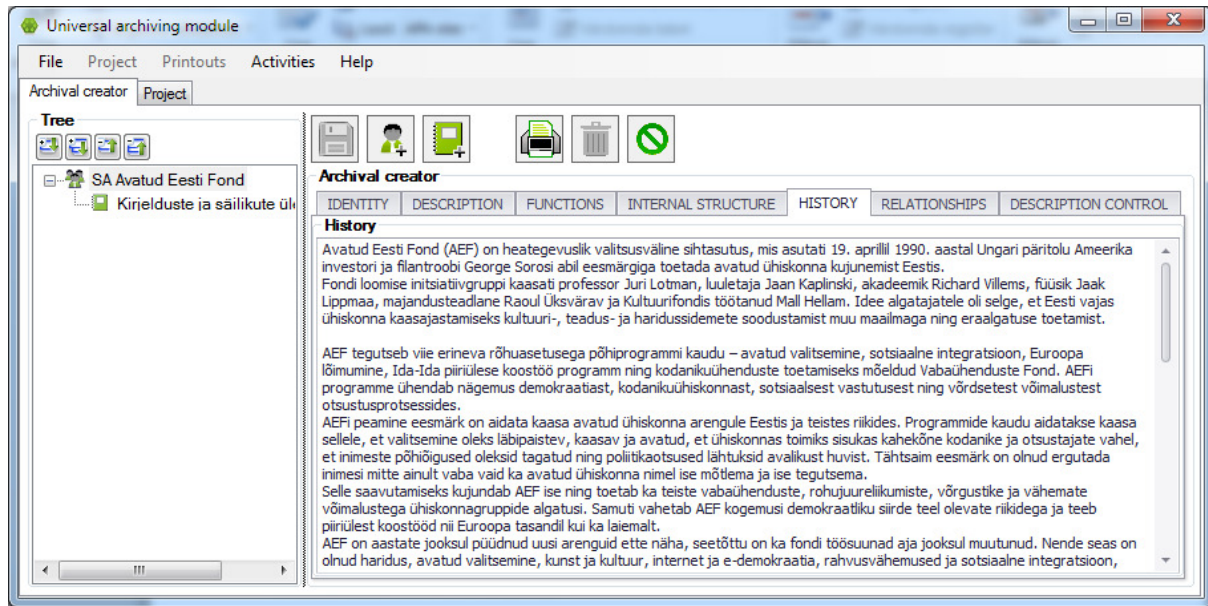
Figure 2: View of the Archival Creator in the UAM

Metadata for describing the actual transfer project was also added as the UAM groups each transfer to a separate project. The description of a transfer project was very simple and consisted only of the project's number and name, author's first and last name, author's position, location (i.e. city), agency's name, archival list number (finding aid number), first folder number (file number). Dates of creation, modification and transfer were automatically added by UAM.

The next step was getting the approval to the archival classification scheme. As the UAM doesn't have any automated procedures for getting approval, it was obtained externally. When the archival classification scheme was in place, work continued with describing the files (dossiers).

As the descriptions of file-level items were on Excel spreadsheets (created already previously in the agency), the archivists copied the descriptions to another pre-defined Excel spreadsheet to ensure that the descriptions contain all mandatory metadata elements and that they can be automatically mapped to the XML Schema set by the archives. Some example cells are seen in Table 2.

Table 2: Example of data on pre-defined Excel spreadsheet

Säiliku nr / File No*	Sarja tähis / Series No	Pealkiri / Title	Alg-daatum / Date from	Lõpp-daatum / Date to	Kogus / Amount	Ühik / Unit	...
ns1:KyViit**	ns1:KyVanemViit	ns1:KyPealkiri	ns1:Vaartus	ns1:Vaartus 2	ns1:Kogus	ns1:Yhik	...
1	1	Üld- ja personalikäskkirjad / General and personal decrees	1991	1993	15	lehte / pages	...
2	2	Nõukogu ja juhatuse koosolekute protokollid / Minutes of Board and Council meetings	1990	1992	15	lehte / pages	...
3	2	Nõukogu ja juhatuse koosolekute	1993	1993	15	lehte	...

		protokollid / Minutes of Board and Council meetings				/ pages	
…	…	…	…	…	…	…	…

* free translation to English added after "/" mark in cells
** element names in XML Schema

This Excel file was in compliance with the UAM input as the XML output of the Excel file used the same metadata elements. The XML file was produced by a simple Save-As dialogue in Excel and the file was imported to the UAM. All descriptions were automatically placed in to proper places in the archival classification hierarchy according to series numbers.

The import was possible because of the pre-defined Excel spreadsheet which placed all metadata into the right tags in XML files, so that the UAM could easily identify them during import.

As the descriptions of the item-level were in Excel spreadsheets, created by the agency as well, the archivists copied them also to another pre-defined Excel spreadsheet and imported to the UAM in a similar manner.

NAE's future ambition is to use the same Excel table templates for other agencies as well, so that they could organize their descriptions very effectively before using the UAM.

Computer files

Negotiation processes with data producers have shown that agencies have often some valuable information encapsulated in computer files which have been not put into the ERM systems, but are on some network drives. These computer files are usually organised in some way for better finding purposes, but they are not automatically related to the classification scheme used in ERMS. Therefore the UAM is designed to help to create relations between computer files and appropriate records or (case) files. This can be done in two ways. One option is that the archivist selects computer file(s) or some catalogue in the operation systems and the UAM imports the computer file(s) to the indicated place in the software and automatically characterizes and migrates them, if needed. This is useful when there is a relatively small amount of computer files to import as it requires quite a lot of manual intervention. The second way is to use XML files for the list of available computer files during import. The XML file can be created from the console of the operation system by printing out the listing of contents of a directory. This is extremely useful when computer files are organised by the agency's series or functions on the network drives. It is important to note that all actions are logged and it is possible to check whether some computer file is a result of a migration or it is the original file. As no computer files are being deleted during the migration process, it is possible to repeat the migration (into some other file format) later if needed.

As the OEF project contained also some computer files, they were imported to the UAM using the first described way of import. All descriptions of computer files were created automatically. During the import process the computer files were checked whether they are in an archival format or not. All of them were in PDF format, so no migration was needed. But in the case of encountering some non-archival formats, the UAM would have tried to use a OpenOffice or LibreOffice converter to migrate computer files into an archival format. As the migration capability lies on OpenOffice and LibreOffice and their main functionality is not the migration, then it would be very reasonable if the agency could already create the computer files in archival formats or migrate them into the archival format before using the UAM.

Validation

The UAM has a built-in multi-step validation functionality which helps to ensure that all actions and transfer activities are performed according to the rules set by the archives. First step includes validation on manual input – when the data is inputted to some metadata field it will be automatically validated. Existence of mandatory values is checked during saving the UAM views and additional validation (correctness of the classification tree etc.) is done when "Validation"

button is pressed. The final validation is performed during transfer. The validation levels duplicate some rules, but they are mainly complementing each other. For example some metadata are not mandatory when the archivist starts the describing process and the absence of it will be not treated as a problem until the archival classification scheme has been marked as approved by the NAE. There are also some rules which only indicate some mistakes and are treated as warnings, thus they are not compromising the transfer, but will be highlighted in the delivery agreement later.

The OEF project was successfully validated in all stages. Although some messages about a missing metadata element value were received during the validation, they were treated as warnings and the final decision was left to be done in the archival ingest module where additional checks are made before archival storage. The ambition is to continue to use the multi-step validation and perform most of the checks on both (archives and agency) sides to ensure that the problems are being identified as early as possible and that no invalid information is being placed to the archival repositories.

<u>Transfer</u>

Descriptions and computer files can be sent to the archives over the national secured channel X-Road[5] and DEC (document exchange centre)[6] or the information packages can be saved to some external carriers and brought physically to the archives.

When sending information packages over X-Road and DEC the packages will be automatically split to smaller pieces for more efficient performance. The system checks that no pieces get lost during transfer and informs the user about the result of the process. If the user decides to save information packages to some external media instead of using DEC then it is possible to choose the location where to save and indicate the maximum capacity of the media in the dialog box (Figure 3).
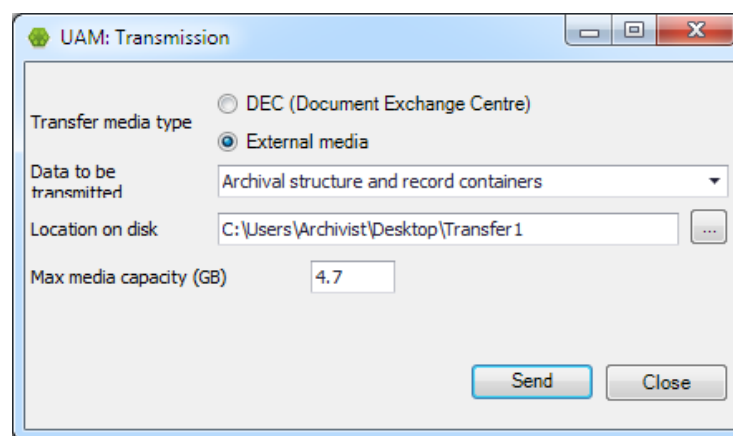

Figure 3: Sending information packages to archives

It is very important to mark the maximum media capacity in cases when the size of the transfer project exceeds the available media size. For example, if we want to save information packages to a single-layer DVD then we should put the size of 4.7 GB in the field of maximum media capacity. The UAM splits the information in a way that all DVDs will contain index XML files which gives information about the information package(s) on that DVD and the total amount of them.

---

[5] The data exchange layer X-Road is a technical and organisational environment, which enables secure Internet-based data exchange between the state's information systems.
https://www.ria.ee/x-road/
[6] DEC is an infrastructure for the transmission of documents (i.e. a mediation layer for document exchange services of information systems) relying on the X-Road as a transport-level infrastructure.
https://www.ria.ee/dec/

It was chosen to use the X-Road for transferring computer files, digital descriptions for both digital and paper documents of OEF and the transfer was successful.

The information packages were received by NAE's ingest module where the further processing continued. The paper documents were brought to the archives and put to a storage room as usually.

As this transfer project was quite small in the number of megabytes, it was not possible to encounter any performance issues, but NAE is aware of the fact that transferring and processing large XML files can be resource-consuming and that additional measures for optimizing the process could be needed in the future when dealing with big amount of data.

Conclusion

Most of Estonian private associations and companies were founded after the restoration of independence in 1991, thus civil initiatives that currently operate have a history of about 20 years. So, year by year, the need to solve the problem of how to preserve archives of private organisations is increasing. Until today they have not had an urgent need for transferring their records. This problem has become more acute as the period of twenty years seems to be a maximum time for storing records by such organizations. Although NAE may demand public bodies to follow guidelines and regulations set by the Archives Act, they cannot do the same with private organisations. On the contrary, NAE when collecting private records must reach a mutual agreement by motivating the owners to describe their archives on the required level. Having comfortable and user friendly tools NAE will have a greater chance to achieve this goal.

The UAM tool, created by NAE, has proven itself well by providing multiple functionalities to some regular problems (e.g. missing metadata) encountered in transfers. It has helped the agency to prepare digital documents and their metadata (including also metadata for paper based documents) for archiving and to create the schema of hierarchy and offers a flexible way to change it. It is important for both, the archives as well as the agency, that the UAM encompasses a built-in system of validation, which helps to ensure that the descriptions and the internal structure of records (structure of archives) was compiled in accordance to the requirements of the NAE, in a less time and labour consuming way. By using the built-in validation system inadequacies in the data were automatically detected and easily eliminated.

UAM has also helped to gather and organize computer files from outside of the ERMS and comprise them to the transfer project of OEF.

It was a very significant project for NAE as OEF was the first private (non-profit) organization which records and descriptions were transferred by using the pre-ingest tool UAM. So, the first experience of taking over the records of a private organisation, occurred just in time and was of significant practical value for both sides.