Memoriaali – An Online Platform for Digital Reception and Enrichment of Archived Materials

Siiri Simpanen; Muisti Centre of War and Peace; Mikkeli, Finland, Anssi Jääskeläinen, South-Eastern Finland University of Applied Sciences; Mikkeli, Finland, Mika Kokkonen, Elka Central Archives for Finnish Business Records; Mikkeli, Finland

Abstract

People are participating in the digital environment more than ever before and have gotten used to sharing their lives and expertise through social networks. There is a great potential in users sharing private archival materials and their knowledge digitally, but they need adequate guidance. At the same time, the archival sector is struggling to process the fast-growing amount of digital material. Reception and its processes are slowed down by insufficient technological development and lack of resources. What is needed to obtain valuable private archives and enriching metadata from the users, but not increase the workload of archival organizations? In this paper, we are introducing Memoriaali, an easy-to-use platform that provides tools for archival donors and professionals.

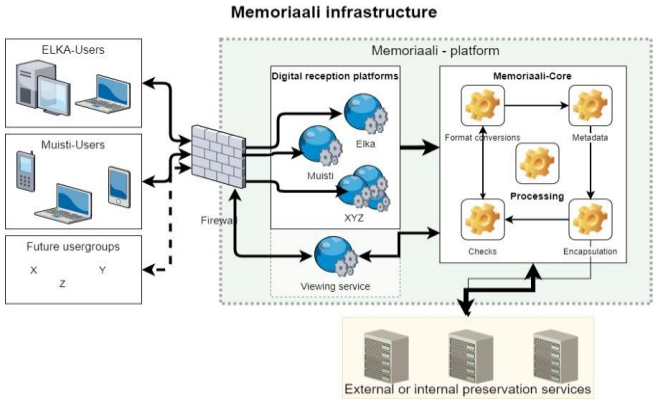

Figure 1. *Upper-level Memoriaali infrastructure*

Motivation

People possess great amounts of historically valuable physical documents. Stuffed in boxes and drawers, these documents may easily end up in the trash. With good luck, the documents will be digitized, but this does not mean that those will be preserved and become part of a wider cultural-historical horizon and research. Digitalization has changed the way people produce and store their photographs, letters, records, and other items, but as the physical media of information is becoming marginalized, the potential threat of losing existing content arises.

There exist various examples on using professional digital archiving services in creation of family archives. One of such cases is described by Uotila [8]. Uotila states that creating a family archive is a straightforward task but takes a lot of time with professional tools. Furthermore, there have been experiments in creating a digital archive for citizens [2], As noble as the idea is the biggest obstacle in creating such a service is the lack of paying customers. The Memoriaali aims to ease these cases and obstacles by creating a simple to use data receiving and enriching platform which utilizes existing archival service providers. Figure 1 presents the designed upper-level infrastructure of Memoriaali platform.

The digital era has also shifted societal norms and values from hierarchical structures towards networked models, changing society's expectations of the archival field [3]. Democratization has already been adapted in platform economy and in many service applications like Uber, Wolt and AirBnB. In these services applicable data is open for everyone and there is an opportunity to participate in them as customer or service provider. The archival sector is yet to take part in this trend.

In archival science, the concept of participatory archive acknowledges that rights, responsibilities, and needs, as well as stewardship and expertise regarding archives are shared, but research has only begun to explore the meaning of participatory practices such as interaction, radical user orientation and contextualization [1],[5]. These practices can be seen as a key to democratizing the archive and knowledge production in it.

A special characteristic of the Finnish archival field is its strong connection to the activities of the authorities. At the same time, private archives receive less attention, and there is little tradition of community archives in Finland. Archival tasks are also carried out by private sector organizations, but their activities also often focus on preserving the records of companies and larger non-governmental organizations. In Finland, it is difficult for an individual to have their archival material preserved unless the archive contains material about a historically significant person or is otherwise assessed as invaluable. This creates a situation where preserved documents tell the history of governance and notable characters, while the perspective of common people is difficult to find from the archives.

By making the archival transfer process accessible, easy to use and available for private donors, it is possible to widen the way we see history through archives. If even larger number of people are seen as stakeholders, knowledge production around archived materials can permanently change and allow participation of more users than before.

For institutions to meet new expectations and utilize the potential of users, new kind of technical solutions are needed. Developing digital solutions for reception and access make archives more accessible to both input and output users. The demand for remote access has increased, and future needs can be even greater, as the COVID-19 pandemic has shown.

Problem

The professional discussion inside the Finnish archival sector indicates that receiving private records is a problem waiting to be

solved. This paper addresses three problems around archiving that has been the motivation of Memoriaali project.

Firstly, many archives do not accept donations from private citizens because they do not fit inside the preservation strategy. Even if the value of documents is understood and the donation could be accepted, the process might still be too time consuming with the available resources in the archival organization.

Secondly, although people have better devices for digitizing historically valuable materials at home or for example in nearby library, there is no easy-to-use service for private archival donors who wants to submit their materials to the local archive. Available solutions (such as RODA-in) are more suited for professional use and require understanding of archival process and technological skills beyond average archival donor. Current solutions also do not enable users to enrich the archived materials and the important information remains uncovered.

Finally, archival organizations struggle with processing analogical material and the fast-growing masses of digital material. A big part of the digital and analogical information being handled is not machine-readable and requires human labor before becoming a part of the archive. This manual labor is expensive, time-consuming and it slows down the preservation work, especially for smaller archiving organizations. Technical requirements, such as validations and migrations of the submissions, are making it difficult to reduce the work of professionals.

In spite the fact that most of the private documents will never be archived, there is still a clear need for a technical solution that enables it. Memoriaali platform is a possible solution, as it simplifies the archival submissions and makes enriching the archived material possible with crowdsourcing [7].

The Memoriaali platform for archival submissions

Memoriaali project approaches the problems described above by developing a new kind of easy-to-use platform for receiving archival submissions and retrieving data from the archive. Acting as a simplified receiving and retrieving functionality between the donor and the repository, Memoriaali stores the received data in local database until it is processed, and after acceptance procedures transfers the data into a true digital repository which handles the storing, indexing and retrieval tasks. Figure 2 presents the preservation workflow from the donor to the actual preservation phase. Organization using Memoriaali will have a possibility to utilize any compatible digital repository as their final archive, into which Memoriaali connects via series of API calls.

Technical solutions and tools developed in this project supports user participation in pre-ingest and ingest phases of the archiving process and reduce the manual workload of archival professionals. Memoriaali environment guides the user at each stage of the submission process, providing room for users with different abilities to take part in the process [7].

Memoriaali also takes care of required technical aspects such as format conversions, validations, virus checks etc. so the donor does not have to consider those at all.

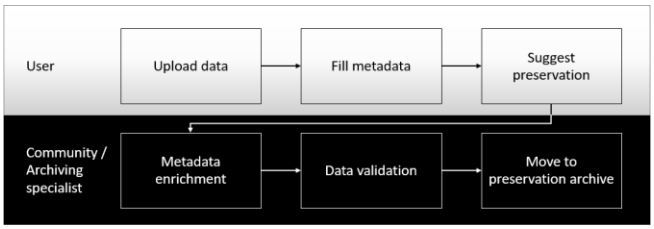

***Figure 2**. Simplified preservation workflow*

Collecting metadata in Memoriaali

The platform allows archival donors to submit the materials remotely and give the descriptive metadata and some of the technical metadata themselves. Memoriaali will allow the organization using the system to modify the metadata scheme or metadata fields collected from the user. This way Memoriaali will serve multidisciplinary organizations. In the developing and piloting phases Dublin Core and EAD2002 are used. The main reason for the selection of these two descriptive metadata options is commons-ip by Keeps which Memoriaali uses for generating SIPs (Submission Information Package). E-ARK SIP specifications on the other hand supports the newer EAD3, but a reference implementation was not available when the project was started.

In the UI the basic view of collecting the metadata is narrowed down to just a few essential fields with the possibility to expand the view and input additional metadata. If could be argued that a lack of standardization via authority systems could lead to a data which cannot be compared in the future. However, from the perspective of an average data donator just some mandatory fields are more likely to be filled than full set of mandatory descriptive metadata. This view is backed up the guidelines by Nielsen Norman Group, they state that “Eliminating unnecessary fields requires more time, but the reduced user effort and increased completion rates make it worthwhile”[6]. To obtain this mandatory information, the organization using Memoriaali needs to define the required fields that the donor cannot leave blank [7].

In addition to the metadata collected from the user, Memoriaali utilizes open-source tools and libraries such as Exiftool and Apache Tika to collect technical metadata automatically. Later, a sophisticated content analysis is embedded into the workflow, adding more automatically created content related metadata. These analysis tools will enable reading and understanding the content of both text and image files.

Enriching materials with crowdsourcing

Crowdsourcing is one of the focal points of Memoriaali project [7]. By developing tools for expert users, such as researchers, genealogists, and history enthusiasts, platform will gather additional information from these users to be added to the public materials in Memoriaali. This allows high level user participation and opens the archival processes not only for the donor, but also to users that may have unrecognized knowledge. Especially the descriptive metadata can be enriched by allowing users with knowledge to add information.

One of the piloting organizations, Muisti Center for War and Peace, is launching a collection of Finnish war history using the Memoriaali platform. The purpose is to bring the materials and knowledge of individuals and small communities alongside the

archives of authorities. Photographs, letters, interviews, and other materials gathered from private archival donors are of great interest to the field of Finnish war history enthusiasts. At the same time these people have wide knowledge on Finnish war history and might be able to tell information about the materials that the donor may not know. By allowing these expert users to submit proposals for additional metadata, Memoriaali allows participation in ways that are not common in the archival field.

To avoid spam and incorrect additions, Memoriaali also uses crowdsourcing and participation of volunteer expert users, to approve proposed additions, alongside automatic inspection tools such as reCAPTCHA by Google. Later, it is possible to build a trust-based user ranking system which automatically allows trusted users to do corrections and additions.

Security and legal matters

Memoriaali ensures compliance with copyright and privacy policies, such as GDPR, by directing the user to sign a submission agreement. Donors must guarantee that they possess the needed privileges to publish the data and that it does not contain regulated material. The legality of the doner is verified when the account is created with either a strong authentication, electronic signature or manual signature which is scanned into the system.

One of the Memoriaali platform focal points is crowdsourcing, but it is up to the donor if they want their collections to be open to public or stay private. If donor chooses the data to stay private, Memoriaali platform ensures that data will not show up on searches and browsing. Hiding the data can be done multiple ways. The donor can choose if all the metadata or some of the metadata will be published while the uploaded documents are kept private. These private materials remain visible to those who have the correct permission for it (for example the donor and the archival specialists of the organization administrating the Memoriaali platform).

Along private citizens, Memoriaali will allow business organizations and communities to transfer their materials to archival organization. Pilot of Elka Central Archives for Finnish Business Records focuses on importing materials of corporations, making different levels of usage restrictions particularly significant. In Elka pilot the workflow consists of donating and sharing documents in private corporate environments. Emphasis on this pilot is to assure that working with materials with different levels of public or private metadata is fluent, even if the user needs to restrict the access level of the documents and metadata. Crowdsourcing can still be applied within users of this private environment.

Usage of archived material in Memoriaali

The purpose of Memoriaali is not only to make the archive available for the donor, but also to make the materials available for other users. The submission process described in this paper aims to ensure that enough information about the materials is gathered to make them usable and understandable. The required metadata fields are used to retrieve the minimum of pivotal information about the content, but Memoriaali instructs the user to hand over all the information they know and submit it to the correct metadata field. Involving the donor in description process and filling the metadata in a format that archival systems can read, reduces the work of archival organizations.

All the materials submitted to Memoriaali might not be immediately archivable. If the donor uploads an old photo but has no knowledge about its content, the platform will not allow archival transfer. In such cases Memoriaali uses crowdsourcing and knowledge of archival specialists, which allows materials to be enriched before the repository transfer. This enriching will eventually make the materials understandable and therefore usable.

Memoriaali platform also includes its own internal search which allows the user to browse or search the materials. This search functionality is presented on the right side of Figure 3, where the Finnish word for skiing is searched. Left side of the Figure 3 presents the file upload phase, where four different filetypes are being uploaded into Memoriaali. As stated earlier the actual archived data is preserved into another system which is linked to the Memoriaali via APIs. In addition, interfaces which will make the materials available in national collection services (such as Finnish Finna.fi -search service) or internationally (for example, Europeana) are being developed.

Implementation of the Memoriaali platform

To reduce the bottlenecks of archival organizations, Memoriaali also provides tools and technical solutions, such as automated error checks, analysis and cleaning of external hard drives. Creating OAIS compliant E-ARK SIP information packages is also supported to speed up the ingest process with OAIS

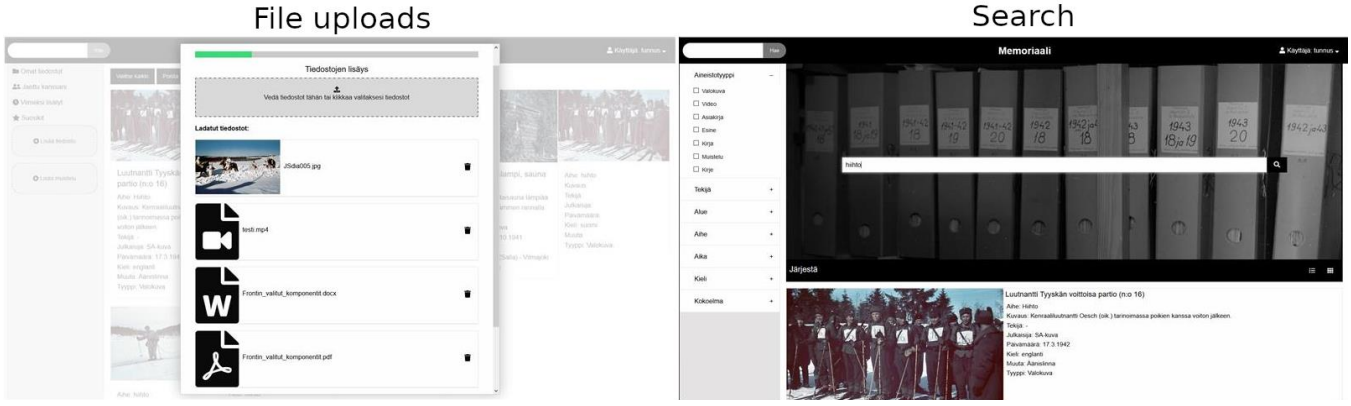

Figure 3. *File upload and search views of Memoriaali platform*

compliant archives. This is one of the most important functions provided by Memoriaali. A customized workflow is created by utilizing the open-source solutions provided by the E-ARK research project (RODA Commons IP). This structure adds a lot of overhead to the packages, but it ensures that the packages created in Memoriaali can be ingested into any E-ARK-compatible digital repository. This way Memoriaali will truly be a platform-independent solution. However, it is also possible to bypass the SIP creation phase and submit the payload and metadata into the preservation archive without encapsulation, this way OAIS incompatible repositories can also be used as the final archive.

Another important functionality of Memoriaali is related to linked open data. Memoriaali project co-operates with the Finnish Semantic Computing Research Group, and the data from their Sampo portals (https://seco.cs.aalto.fi/projects/) is linked with the Memoriaali data and vice versa. This way data that already exists somewhere does not need to be filled again and semantic meaning is ensured.

Memoriaali development is based on the utilization of open-source code and tools. This practice speeds up the development and allows the developers to focus on sections that need modification the most. For the front end, a comparison between Vue, React and Angular was made, and React was chosen due to wide extensibility, being more popular than Vue and being less HTML or CSS heavy than Angular. Using React is also supported by the wide community support that comes with the popularity [4].

The infrastructure of Memoriaali uses modular design, so archival organizations can easily extend functionalities without affecting the core functions of Memoriaali.

Conclusions

By enabling participation in the processes of archiving, Memoriaali allows users to submit their materials to archives with ease. With crowdsourcing metadata, Memoriaali democratizes knowledge production and enriches archived materials. This will improve receiving, understanding and accessing of private archives. Even if the donor wants to keep his data private, Memoriaali offers a solid way for the donor and archival specialist to communicate and complete the data for it to be moved to long lasting archiving system.

Offering easy-to-use, yet customizable tools and solutions, Memoriaali will be relatively simple to add to the workflow of archival organizations. It will make the ingest process easier by offering automation to time-consuming tasks. Memoriaali project will promote digital expertise at the local level in Finland, but also have wider meaning by producing open-source code and ready-to-use online platform [7].

The versions of the Memoriaali platform developed in the project remain the property and control of the parties involved. However, according to funding regulations and open source principles, all applicable codes will be published in the GitHub service after the project is finished [7].

Although Memoriaali is still a work in progress, both users and professionals of the Finnish archival sector have expressed interest in Memoriaali, showing the necessity and potential of the project. Discussion around the Memoriaali platform and tools has made it clear, that the development aims of the project are vital for the archival sector in Finland. As most of the problems experienced in Finnish archives are universal, the project will contribute to the development of the archival sector internationally.

Memoriaali is the joint EU-funded project of Mikkeli Development Miksei (the development company of City of Mikkeli), Elka Central Archives for Finnish Business Records, Muisti Centre of War and Peace, and South-Eastern Finland University of Applied Sciences Xamk. The project started on the 1st of June 2020 and will end on the 31st of December 2021 [7]. The final results of the project will be available in December 2021. The development of Memoriaali platform can be viewed at memoriaali.xamk.fi, although the platform will only be available in Finnish at the time of publication of this paper.

References

[1] A. J. Gilliland & S. McKemmish, “The Role of Participatory Archives in Furthering Human Rights, Reconciliation and Recovery” (Atlanti: Review for Modern Archival Theory and Practice, 24. 2014) Available at: https://escholarship.org/uc/item/346521tf

[2] A. Jääskeläinen, L. Uosukainen (2018) Citizen Archive: My Precious Information, New Review of Information Networking, DOI: 10.1080/13614576.2018.1537800

[3] C. Findlay, “Participatory cultures, trust technologies and C. Findlay, “Participatory cultures, trust technologies and decentralisation: innovation opportunities for recordkeeping” (Archives and Manuscripts, 45:3, 176-190, 2017), pg. 178. DOI: 10.1080/01576895.2017.1366864

[4] E. Wohlgethan. Supporting Web Development Decisions by Comparing Three Major JavaScript Frameworks: Angular, React and Vue.js. (2018) avail. https://reposit.haw-hamburg.de/bitstream/20.500.12738/8417/1/BA_Wohlgethan_2176410.pdf

[5] I. Huvila, “Participatory archive: towards decentralised curation, radical user orientation and broader contextualisation of records management.” (Archival Science, 2008, 8 (1), 15-36. Springer).

[6] K. Whitenton. Website Form Usability: Top 10 Recommendations, available. https://www.nngroup.com/articles/web-form-design/

[7] Mikkelin kehitysyhtiö Miksei Oy. Hankesuunnitelma, Memoriaali – Aineistojen digitaalisen vastaanoton, käytettävyyden ja verkkopalvelun kehittäminen Memory Campuksessa. (2020). In Finnish.

[8] P. Uotila, Using a professional digital archiving service for the construction of a family archive, Proc Archiving 2014. pg. 188 (2014).

Author Biography

Siiri Simpanen is the archival specialist of Memoriaali project. They have a Bc in Culture and Arts (2011) and they are finishing a MSc in Information Science at Tampere University at the moment.

Anssi Jääskeläinen has an IT MSc. (2005) and a PhD (2011) from the Lappeenranta University of Technology. He has an extensive knowledge of user experience, usability and programming. He is currently working as a Research Manager.

Mika Kokkonen has a Business IT Bc (2008) and he is MSc Student (Development of Data-based Wellbeing Services) at South-Eastern Finland University of Applied Sciences. He has history of working as software engineer. He is currently working as project specialist in Memoriaali project (Technology scope).