



Co-funded by the  
Creative Europe Programme  
of the European Union



# Workshop

## New Business & Conceptual models

Conference Center of Campus Condorcet  
Paris-Aubervilliers

23 May 2022



INTERNATIONAL CENTRE  
FOR ARCHIVAL RESEARCH



## Table of contents

Introduction, <i>Yvan Corbat. DEX</i> .....	1
Self-copying service in the Spanish researching rooms, <i>Miguel Ángel Bermejo Alonso. Spanish State Archives</i> ....	3
Face to Face to e-Commerce: the Migration Experience, <i>José Furtado. General Directorate from the Book, Archives and Libraries</i> .....	5
Potential Avenues for New Revenue-Generating Business Models, <i>Leonard Callus. National Archives of Malta</i> .	7
Use of Artificial Intelligence, <i>Ole Gausdal. The National Archives of Norway</i> .....	11
Crowdsourcing, <i>Dorottya Szabó. National Archives of Hungary</i> .....	13
General Conclusions.....	15

## Introduction

### Yvan Corbat. DEX

The following report presents the results of the Workshop *“New Business and conceptual models”* coordinated by ICARUS in the framework of its #28<sup>th</sup> ICARUS convention in Paris that gathered 96 in person and an average of 25 in remote from its network of over 160 archives and scientific institutions.

One of the objectives of the project European Digital Treasures is to generate a greater added value, profitability, visibility and economic return of European archives, through the identification and implementation of new business models and activities.

In the first part of the workshop, *“Opening and introduction”*, two initial project activities that had been developed before the pandemic in order to catch this objective were presented:

- **Diagnosis & State of the Arts:** To identify, generate and transfer a series of homogeneous knowledge and to provoke cross-sectorial encounters.
- **Benchmark practices** made by institutions worldwide such as public and private Museums, Libraries, Heritage and Cultural Centres, to showcase innovative solutions in cultural management carried out by organisations that share the similar goals of protecting, enhancing arts and cultural heritage.

As far as the Pan European Diagnosis and State Arts is concerned, within its methodology, 42 questionnaires were obtained (21 at National levels / 21 at either local and/or regional levels) covering 21 countries.



It also gathered and analysed data from different European sources (networks, platforms, cultural institutions, etc. such as ICARUS, EBNA, EAG, Archive Portal Europe, etc.) and other parallel surveys being carried out, such as Identification of existing Good Practices or Innovative practices in the archives sector.

To complement the State of the Arts and inspire archival institutions to adopt innovative practices, the Benchmark identified 37 Good Practices (GPs) from other cultural institutions worldwide (Museums, Libraries, etc.) that were presented under 4 categories: 1) Management; 2) Audience Development; 3) Distribution/Communication; & 4) ICT/R&D/Cooperation), in order to generate a series of

conclusions and recommendations towards the potential transfer of GPs to the Archives sector.

Following these first recommendations and actions identified, the partners of the project agreed on a realistic objective to implement at least 4 new and alternative managerial action each (20 in total), that will contribute to improve their existing models, in terms of income generation, visibility or relation with existing and new users and visitors.

The present report refers to 5 of these 20 actions (one per National/State Archive) that were presented in the Workshop “*New Business and conceptual models*” .



- Self-copying service in the researching rooms of the Spanish State Archives
- e-commerce platform by the National Archives of Portugal
- Identification of potential for revenue generation in the National Archives of Malta
- Improve access to digitized archives using Artificial Intelligence in the National Archives of Norway
- Crowdsourcing activities in the National Archives of Hungary

Other managerial actions carried out during the project by the National /State Archives involved are:

- Collaboration with private institutions; Implementation of an online booking system for appointments at the researching rooms and Creation of an image bank for dissemination in the case of the *Spanish State Archives*.
- Opening the field for user generated contents; Participating in programmes that leads people to digitize their family memories and heritage; e-learning programmes for users (researchers, teachers, and students) and for the employees in the *National Archives of Hungary*.
- Broader test of HTR in combination with crowdsourcing; Make a plan to follow up the Adelaide declaration and improve the support of our users in the *National Archives of Norway*.
- Enrichment of data of the map collection; Collaborate with a government entity to popularise the archives; Saving an endangered archive or promoting volunteering in the *National Archives of Malta*.
- Info products model; Consulting and events and exhibitions in the National Archives of Portugal.

## Self-copying service in the Spanish researching rooms

### Miguel Ángel Bermejo Alonso. Spanish State Archives

The Spanish State Archives offer this service about their archival holding in order to achieve two objectives: provide the user with copies to work with and reduce pressure on saturated reprographics services of the Spanish State Archives.

The Spanish State Archives developed a pilot project in two archives of the system (Archive of the Royal Chancery of Valladolid and the General Archive of Simancas), to implement rules on self-copying for four months (October 2021 to February 2022) and to evaluate the appropriateness of extending this measure to all the archives of the network of state archives in Spain

This measure regulates the possibility for the beneficiaries of the archives to obtain reproductions of the documents of interest to them, using their own photographic camera, without having to wait for official copies from the corresponding reprographic service, which will save time and money.

This pilot project allows up to 50 captures with smartphones or tablets per user and study session. The users can take photographs about documents themselves, as long as there are no restrictions for access such as conservation, copyright and other aspects of intellectual property. There is a limit of 50 copies for user and work season the copies may be used for research purposes only, not for publication. Results of the experience:

ARCHIVE OF THE ROYAL CHANCELLERY OF VALLADOLID					
	Octubre	Noviembre	Diciembre	Enero	Total
Number of applications submitted	9	4	10	9	32
Number of approved applications	9	4	10	9	32
Number of rejected applications	0	0	0	0	0
Number of applicants for the service	7	4	8	6	25
Total number of documents included in the applications	15	11	15	13	54
Number of self-copied documents	14	11	15	11	51
Number of documents autocopied during more than one session	1	0	0	2	3
Number of applications reaching the limit of 50 self-copies	4	1	3	4	12
Total number of self-copies made	267	128	285	324	1004
Average number of self-copies per request	29	32	28	25	114
Average number of self-copies per user	38	32	35	54	159
Average number of self-copies per document	19	11	19	29	78

GENERAL ARCHIVE OF SIMANCAS					
	Octubre	Noviembre	Diciembre	Enero/Febrero	Total
Number of applications submitted	39	64	38	41	182
Number of approved applications	39	64	38	41	182
Number of rejected applications	0	0	0	0	0
Number of applicants for the service	19	25	15	16	75
Total number of documents included in the applications	85	119	87	73	364
Number of self-copied documents	85	119	87	73	364
Number of documents autocopied during more than one session	0	8	9	4	21
Number of applications reaching the limit of 50 self-copies	4	14	14		32
Total number of self-copies made	1112	1512	1357	1318	5299
Average number of self-copies per request	28	23	35	32	118
Average number of self-copies per user	2	60	90	1,7	153,7

According to our initial plan a proposal was presented to our Subdirector General in June 2021 in order to accomplish with all the procedure to get the permission form the General Directorate of the Fine Arts. This was the baseline to organize the coordination of the pilot



experience in the Archive of the Royal Chancery of Valladolid and the General Archive of Simancas from November and February 2021. Previously the preparation of infrastructures and materials and training of the staff involved in the archives involved.

Afterwards came the analysis of the experience by users and the archives involved. The users, for their part, have not ceased to express their absolute satisfaction with this service and consider it to be extremely useful for their work.

As far as the balance sheet for the archive is concerned, we consider it to be clearly positive. Although the volume of the pilot project is not sufficient to have a significant impact on the figures, our opinion after four months is that it offers clear advantages in terms of user satisfaction, reduction of waiting times, relief for the always overloaded reprographic services, savings in management costs.

Finally, we are evaluating the possibility to extend the experience to all archives under the General Subdirectorate of the Spanish State Archives and consequently to prepare the relevant administrative procedures needed.

## Face to Face to e-Commerce: the Migration Experience

### José Furtado. General Directorate from the Book, Archives and Libraries

From Face to Face to e-Commerce: The Migration Experience brief describes the General Directorate for the Books, Archives and Libraries experience on the transition from its previous approach and position online to E-commerce strategy adaptation, faced as an opportunity for paradigm change.

An original strategy for supported institutional image conceived 20 years ago was developed towards the idea that solely public who frequented the National Archive and its services, like the shop, would get the chance to buy or access the products.

However, on the course 2020 to 2022, directly or indirectly COVID 19 pandemic deeply affected the business model conducted. Abrupt decline in sales as well as of the clients due to confinement were observed; we had to deal with the difficulties that some of our human Resources felt to quickly adapt and response to the increased online demands; regarding logistics aspects we were confronted ourselves with the inefficient capacity of response from the webpage in use to the demands of our customers as well as a deficient stock management due to difficulties to access the main warehouse kept too far away from the business centre.

Part of the reason comes as a reflex of the incremental growth of online requests to services. As matter of fact, a 13% growth of on-line interactions were experienced between January and June comparatively to the last 6 months of 2019 that continuously kept to increase since then by users from 170 different countries that accessed to DGLAB's Archives databases to search on its more than 60,000,000 million images available online.

All the above make us believe we should focus on a shift of paradigm in this business area in order to better disseminate the institutional image of our archives, by introducing a different strategy, both at the physical shop and online service. The circumstances forces us to conclude that we should be where public is and delivered what they demand in the way they demand; to consider the adaptation of a more convenient way to communicate with our users/potential clients; to consolidate and reinforce the objective of use the DGLAB's shop as a privileged vehicle for the disseminating of the institutional image, both to national as international public that access National Archive and its services.

Decision were taken towards the professionalization of the DGLAB's Online shop presence by means of disruptive technology in such way that could force us to be also disruptive on our own work in order to optimize resources and process; aggregate other revenue-generating



DGLAB services on the same platform; increase efficiency in customer service both at the "front-office" as "back-office" level, through a more immediate control of "stocks"; reading of sales flow indicators and order requests; improved payment and introduce flexible calculation mechanisms on Shipping Functionalities; tax configuration per country; General Data Protection Regulation (GDPR) compliance at Customer support features; ability for customers to view their order history and status; Live-chat integration etc.

As disruptive as the implementation of a platform of this kind may be, one should not forget E-commerce platform is a mean, not an end, it is a tool to help us keep in path with the strategy defined which hopefully will introduce significant operational improvement gains.

Be where the users are (or be where public is) recall us the need to use ways of communication to facilitate the proximity of organizations to citizens.

Along with the consolidation of the strategic actions in the technological and operational field, start the path towards the creation of synergies that help us more easily identify and meet the needs of the user/customer/citizen by creating synergies with other entities is of crucial importance.



Encouraging easier access to archives heritage targeting public from other fields of research like design of products; magnifying the role of archives as a rich and unexplored source of resources for cultural commercial market; incentivizing public-private cooperation; generating product diversity to meet a broad range of public, from youngest to elderly and attract other "consumers" to the Archive; promoting internships in real work context focus towards web design, marketing or communication; promoting initiatives focused on creating opportunities for artistic production, fostering identity expression as a right of citizenship that the State should support, should be the way to change.

In that sense, the challenge for next years will be to face e-commerce platform as a laboratory of experiences and the archives as a source of inspiration.



## Potential Avenues for New Revenue-Generating Business Models

### Leonard Callus. National Archives of Malta

#### Introduction

A Pan-European Diagnosis and State of the Arts analysis, seeking to gain a deeper understanding of the existing archives management model, was carried out at the initial stage of the European Digital Treasures project.

The findings indicate that archives have considerably varying financial resources, depending on the country, their competences and level of responsibility. However, the prevalent scenario is one where a very large share of their budget is dedicated to operating expenses (personnel and general expenses), leaving archives with a minimum margin of economic manoeuvre to invest in new products and services such as facilities, equipment, promotion, external experts and research.

The situation in Malta mirrors the general trend in that the National Archives' funding is almost exclusively public and that the lion's share goes to salaries. Moreover local financial regulations require that any revenue generated by the institution has to be transferred to the consolidated fund.

#### Objectives

Against this background, the National Archives of Malta undertook to identify and analyse new business models to generate revenue that may be adopted by the National Archives of Malta. Although the applicability of such models in the local set-up was an requirement, these models could be adopted in other scenarios, maybe with some adaptations.

#### Activities

Within the European Digital Treasures project, the National Archives of Malta commissioned the Management Efficiency Unit (MEU), a business consultancy agency within the Office of the Prime Minister, to carry out research into potential avenues for revenue-generating business models, including the possibility to levy fees to the public administration and for selected services.

An analysis of current NAM activities and inherent revenue generation was undertaken, international practices were reviewed and internal brainstorming was carried out.

## Findings

The legal framework establishing and empowering the National Archives of Malta allows for the adoption of revenue-based financing framework. It was observed that the law itself envisages the attainment of financial self-sufficiency.

It is evident that the Maltese law recognizes the linear progression of the documentation created by a public entity throughout its active timeline, from creation up to the point it becomes an archived public record. The legislator designed the role and responsibility of the National Archives around this conceptual framework allowing the National Archivist to establish himself as the regulator of this continuum.



Embracing this approach, the National Archives/Archivists shifts from being the custodian of archived collective memory to its guarantor, ensuring that all public records of enduring historical value are immediately and suitably identified, classified, preserved and eventually archived and made public.

On this basis, the National Archives may legally generate revenue streams from fees levied to public entities for its regulatory / supervisory governance / oversight. This is a scheme of possible fees:

FACILITIES / PRACTICES	PRACTITIONERS	ENFORCEMENT
<ul style="list-style-type: none"> <li>• compliance</li> <li>• audits</li> <li>• inspections</li> <li>• certification</li> <li>• advisory</li> <li>• registration</li> </ul>	<ul style="list-style-type: none"> <li>• accreditation</li> <li>• training</li> <li>• professional development</li> <li>• advisory</li> <li>• registration</li> </ul>	<ul style="list-style-type: none"> <li>• non-compliance</li> <li>• fees/fines</li> </ul>

Such initiatives would require the publication of necessary secondary legislation.

The adoption of such regulatory-styled business model could have significant impact of the National Archives set-up. Consideration of the model's impact on the design and organisation of the functions, set-up, operations and resource deployment needs to be made.

As an alternative to the direct levying of fees, the corresponding fee structure may be encapsulated in the annual budgetary allocation of the National Archives. This would reduce the corresponding administrative burden. The research identified also several potential revenue-generation models that could be adopted by the National Archives.

These include:

AREA	TARGET	PROGRAMME
Advisory Services	Private operators Business NGOs Families	- cataloging - preservation - restoration - good archiving practices
Training	Private operators Public entities Business NGOs Families	- certified practitioner training - educational tours
Storage services	Private operators Public entities Business NGOs Families	- rental of storage space (physical, virtual)
Cultural tourism	Niche tourist segments	- themed programmes - attraction development
Funding (private)	Private sector Donors EU programmes International support programmes	- crowdfunding - Sponsorships - donations - Patreons
Research support services	Research institutions Third party organisations	- commissioned research - publications - internship programmes
Merchandising	Visitors Corporates	- Designer merchandise - Shop rental
Lab service	Private operators Public entities Business NGOs Families	- conservation - restorations - condition assessment
Partnership exhibitions	Museums Local council Specialised entities	- PPP - one-time collaboration agreement - loans

**Lessons learned**

There is scope for revenue-generating business models at the National Archives although, most probably, the operations will never become self-sustaining.

These activities are not free and the selected revenue streams should be aligned to the strategic and operational model to ensure that they don't become a net cost.

Besides the financial aspect, value of these business models is also in the fact that they are community engagement exercises in line with our mission to “enrich our knowledge of human society, to promote democracy, to protect citizens' rights and to enhance the quality of life.” (International Declaration on Archives, UNESCO).

## Use of Artificial Intelligence

### Ole Gausdal. The National Archives of Norway

The National Archives of Norway has started to use Artificial Intelligence (AI) in order to improve our user services. We have large volumes of paper-based archives with limited amount of metadata. As for today's standard, they are also quite poor categorized. The result has been that use of the material requires quite a lot of manual work. We are now working to improve this situation by use of AI. In order to use AI, the material have to be digitized. And that is of course a limitation. So far, we have digitized between 2-3 percent of our total holdings.

The goal for our use of AI is to make historical data searchable and to improve metadata in order to automate our services and improve the user experience and self-service. Artificial Intelligence has many branches. We are mainly working in the field of natural language processing (content extraction and classification), machine learning as well as image recognition. To follow up our AI ambitions, we have employed three AI experts and two experts to follow up handwritten text recognition (HTR).

#### **Examples of projects we have worked on:**

##### Redaction of Personal ID numbers from The Land Registers

We have 5,6 million scanned pages of the land registers, that are a much-used archive series. Many of the pages contain Personal ID Numbers of living persons. Due to Norwegian GDPR regulations, these pages cannot be published online without restrictions. We don't know which pages contain ID numbers. In order to be able to publish all the pages online, we have started a process to redact the ID numbers. Some 63,500 have been redacted manually. By using AI, another 32,000 ID numbers have been redacted in an automated process. We are quite satisfied with the results and the error rate is very low.

##### Data Extraction from the National Population Register

We have tens of millions of index cards that are part of the National Population Register. The series contains sensitive information and is closed to the public. But we have many requests for information from the index cards. Thus, we have started to create a name index using AI. The series spans over more than hundred years, with a variety of templates, different typewriters and handwritings. So far, we have processed more than 10 million cards and identified the names with satisfactory results. The results are available for researchers and case handlers at Tax Norway and the Norwegian Welfare and Labor Agency.

### Automatic detection of metadata

In a couple of projects, we have worked with automated detection of metadata. The results are quite good. In order to detect the information, we have preprocessed the files to improve the OCR results. We have captured metadata types like heading of the letter, name of the sender and receiver and the date it was sent. In one case, we have also categorized the metadata to be able to extract concepts from the documents. Examples of this are location, persons, organizations, and derived concepts. In one of the projects, we have used trained models to recognize case files that should be categorized as sensitive information – based on training data with typical sensitive content.

### **Conclusions**



We are happy to say that we have reached some of the goals we set out. And some of the results have also been put in production. By using AI, we have been able to make more historical data searchable. And we have improved metadata, user experience and self-service. One of the things we have learned so far, is how important it is to prepare the source material in order to optimize the results. Another one is that the results can be good enough, even if they're not perfect. We have also experienced the value of Nordic and international cooperation in this field, so we can learn from each other and don't have to invent everything ourselves.

## Crowdsourcing

### Dorottya Szabó. National Archives of Hungary

In autumn 2021, the *European Digital Treasures (EDT)* consortium's five national archives – National Archives of Norway, Malta, Portugal and Hungary as well as the State Archives of Spain – announced a unique palaeographical volunteering opportunity for the general public. Together with my colleagues, [Ildikó Szerényi](#) and [Anna Palcsó](#) we already wrote articles on the crowdsourcing activity, which are available at the website of the *EDT* project.

The five archives selected records as useful data for family historians and silver generation persons looking for family roots. This was essential to provide (enough) motivation for the volunteers.

- ❖ Norway chose the National Population Register consisting of registration cards of the permanent residents, from the time period of 1906-1914.
- ❖ The Portuguese DGLAB choose King John VI's General Register of Mercies (land grants, mayor's wards, rents, jurisdictions, recommendations, privileges, etc), between 1792-1826.
- ❖ The choice of the National Archives of Malta was immigration records from the Customs Department Fonds, from 1905-1966, with the data of arrivals and departures to/from Malta.
- ❖ Spain selected the collection of Series of Passport Issuance Register by the Consulate General of Spain in Buenos Aires, covering emigration to Argentina, between 1936 and 1939.
- ❖ In Hungary, we worked with the National Census of 1828, to assess the tax capacity of the population, covering the entire territory of Greater Hungary.

All of these collections were not only valuable source for genealogical researchers but they were well suited for indexing and retrieving information, when using Artificial Intelligence.

The archives' teams created a training set with transcriptions of personal data, from the collections mentioned above. The training set was sent to the validation software developed by [tranScriptorium](#), an IT company linked to the Valencian University of Technology. They created the algorithm that provided proposals of transcriptions of the handwritten texts and a secure web-interface, accessible from any location and at any time.

Every volunteer had its own "package" with digitized records, with photos of the records plus suggestions given by the algorithm for each name or word. The software provided 5-20 transcription suggestions, sorted by probability.



Parallel with the tranScriptorium's model-building and surface-creating activity, a crowdsourcing preparation started also, to find skilled volunteers and create a proper supporting background for their validation work. At the validation interface, volunteers were asked to select the correct transcription or, if necessary, to type it in independently and finalise the solution.

As the algorithm is self-learning, adapts to corrections, so it was crucial that the validation should be carried out by people with palaeographical experience. However, the volunteers' correction was supervised by archivists. Through this work, the handwritten text recognition capability of the software has been improved (as the results were sent back again to the Transkriptorium) and the final results became more accurate.

And now let me give you a short insight onto the Hungarian experience. The publication of the call for volunteers was very popular. However, the applicants had to test their knowledge on palaeography before they participated, so an online test was created for them. They were supported by a personal kick-off meeting and with a downloadable palaeographic example book, discussions and quizzes too. A Facebook group was also set up, where professional discussions and useful tips were exchanged, on a daily basis. (We planned to have more personal meeting, but the pandemic situation stopped these attempts.)



During the six-week programme in autumn of 2021, 66 volunteers worked from all over Hungary and from abroad too. They had to transcribe 50 pages of the census, but actually many of them processed significantly more. After closing the activity, volunteers responded to a questionnaire evaluating the programme, where they praised it and mainly said they would like to participate in similar activities in the future.

As a token of appreciation, the National Archives of Hungary presented commemorative certificate, awards and a merchandising product of the *EDT* project as a gift.

The archival partners could participate in a model-building activity and expand their activities with volunteers.

Through this work, not only the HTR capability of the software was improved and the final output became more accurate, but also a community building was initiated. And although this depends more on the experience of the archival partners, some of them will continue the work with volunteers for sure, including National Archives of Hungary.

Link to the presentation: <https://prezi.com/p/edit/0pzgbdffww5yr/>

## General Conclusions

These are some conclusions of these activities carried out in the National/State Archives involved in the project that will contribute to improve their existing models, in terms of income generation, visibility or relation with existing and new users and visitors:

- ❖ The European archives count on highly qualified and specialized teams.
- ❖ In general there is a lack of margin of economic manoeuvre to be able to invest in new products and services (facilities, equipment, promotion, external experts, research, etc.).
- ❖ There is also little experience in generating external income, particularly by private entities (e.g. sponsorships, financial donations, merchandising, etc.)
- ❖ It is important to highlight the tradition and commitment in international institutional cooperation (within the sector). Cooperation with other sectors could be a solution to their lack of resources (diversification, digitalization, visibility).
- ❖ The greater use and exploitation of new technologies, particularly in the matter of automatic recognition and treatment of texts and images, is not only a clear wish of the sector (particularly interested in topics such as Artificial Intelligence or Smart Data management), but also a necessity, since only a minimal part of the heritage kept on the shelves of the Archives is known and is available online.
- ❖ The European Archives must develop strategies to attract and manage new users and new audiences, without underestimating their social and educational responsibility, including the promotion of activities for groups at risk of social exclusion.
- ❖ It would be positive to increase the engagement with association of professionals, club of friends and crowdsourcing activities.
- ❖ Other fields of interest to income generations could be merchandising products, crowdfunding campaigns and sponsorship with private companies.

After the presentations of the EDT partners, Ancestry and Teklia presented themselves and their ongoing projects, followed by a discussion about how companies and institutions could include the strategies and activities presented before, within their daily business.