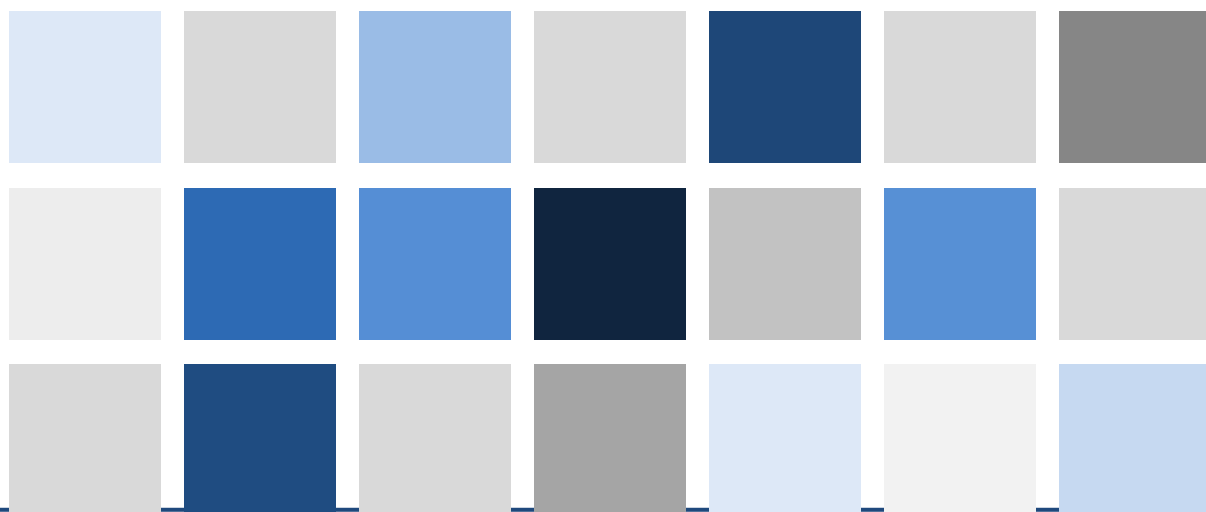


Long-term data for Europe

EURHISFIRM

D11.1: Best Practices for the Preservation and Access to Digitized Material of Interest for European Cultural Heritage



AUTHORS:

Abe de Jong (Monash University, Melbourne/ Erasmus University Rotterdam)

Joost Jonker (University of Amsterdam/International Institute for Social History, Amsterdam)

Joanna Sławatiniec (Erasmus University Rotterdam)

APPROVED IN 2021 BY:

Jan ANNAERT (*Universiteit Antwerpen*)

Wolfgang KÖNIG (*Goethe-Universität Frankfurt*)

Angelo RIVA (*École d'Économie de Paris*)

ACKNOWLEDGMENTS:

We are indebted to the following colleagues for their willingness to share their knowledge and insights with us:

Bertrand COÜASNON (*Institut national des sciences appliquées de Rennes*)

Thierry PAQUET (*Université de Rouen Normandie*)

Contents

Executive summary	4
1. Introduction: aim, objectives, partners, deliverables, and deadlines of WP11	5
2. WP11's general approach.....	6
3. The issues at hand	6
4. The complications of Category (3) documents.....	9
5. Options for presenting Category (3) materials.....	13
6. The long way ahead.....	15
7. Conclusions.....	16
8. Recommendations.....	16
Annex 1, Questionnaire for the EURHISFIRM WP11 roundtable, 10 November 2020	18
Annex 2, Websites surveyed	19



Executive summary

EURHISFIRM “Historical high-quality company-level data for Europe” is a design study to build a world-class research infrastructure (RI) compliant to the FAIR (findable, accessible, interoperable, reusable) data principles. The project aims to increase the accessibility and usability of historical company-level data (financial, governance, and geographical) and to expand the available pool of this data.

Though quantitative data form the EURHISFIRM’s main concern, the project consortium recognizes the imperative of adding digitized material documenting the cultural dimension of the European corporate and financial experience to the RI. That dimension shows to best effect Europe’s defining characteristic as a region, as a continent: unity in diversity. This first of two deliverables surveys the cultural heritage material relating to European corporations and finance.

The fundamental problem concerning the cultural heritage material that interests us is twofold. First, the material is highly diverse; second, unlike the quantitative material, it is randomly preserved, scattered over numerous collections, and poorly catalogued, so difficult to find. At the same time the material is of great importance to economic historians and it has great outreach potential to other disciplines and even the wider public.

Moreover, cultural heritage material poses additional, serious IT complications if they are to be accessible for quantitative databases and other disciplines at the same time. Consequently EURHISFIRM needs to find a partner with which to develop ways of presenting the cultural heritage materials. WP11’s survey of 48 websites identified Europeana and Clarin as desirable partners. The second report will analyze them in more detail.

At the same time EURHISFIRM should take initiatives, on the one hand, to develop a standard classification of relevant material, so national projects can start identifying and cataloguing it; and on the other, to bring institutions with big collections together for collaborating on themed exhibitions.

1. Introduction: aim, objectives, partners, deliverables, and deadlines of WP11

According to the overall EURHISFIRM project design, Work Package 11:

‘explores concepts and tools to stimulate the lasting conservation of the digitized material and provides guidelines for making those materials publicly accessible. It also explores innovative ways to use digitized images as documentation for the data extracted from them and evaluates alternative strategies to use digitized material’.

This task is given three broad objectives:

- (1) The use of digital images to document data and inspire further research and identify sources of interest for cultural heritage;
- (2) The promotion of Europe’s cultural heritage by facilitating digital preservation and online accessibility of sources with a unique historical value;
- (3) The mobilization of digitized images of historical sources as an exceptional additional documentation for the data (including the exploration of ways to make materials accessible and connected to EURHISFIRM data).

The WP11 objectives are to be reached by executing seven specific steps:

- (1) To analyze sources of interest for the cultural heritage regarding stocks, bonds, corporate finance, the securities trade, and stock exchanges;
- (2) To analyze best practice cases of the preservation and access to digitized material relating to the cultural heritage of stocks, bonds, corporate finance, the securities trade, and stock exchanges;
- (3) To explore innovative tools to use digitized material to enhance data documentation;
- (4) To analyze how this data and information are currently provided;
- (5) To stimulate the proper conservation of materials found;
- (6) To explore ways to make those materials accessible and connected to EURHISFIRM data, whenever needed and practical in digital form;
- (7) To define strategies and practices of the already existing RIs (e.g. ERIC DARIAH) that provide platforms and processes enhancing European cultural heritage.

WP11 therefore differs from EURHISFIRM’s other work packages in two key respects. First, rather than helping to build a research infrastructure for corporate governance and finance data, WP11 concerns itself with an entirely different dimension, cultural heritage. As will be argued more fully below in Section 4, the nature of the cultural heritage material requires to step outside of the EURHISFIRM framework and seek collaboration with other, existing infrastructures with cultural heritage dedicated facilities. Second, WP11

does not concern itself with issues concerning the care for and preservation of the digital data itself, that remains the responsibility of the data storage institutions concerned. Nor does WP11 aim to develop or provide facilities for storage and/or retrieval of documents and sources identified. In accordance with EURHISFIRM's federated structure, such facilities also remain the responsibility of the national institutions concerned in the seven consortium countries Belgium, France, Germany, the Netherlands, Poland, Spain, and the United Kingdom. All WP11 aims to do is provide best practice guidelines for the national institutions to follow with respect to its objectives. We request the EURHISFIRM community to turn those guidelines into imperatives so as to make our chosen federated structure work.

Four institutions partner WP11 leader KNAW/IISH in carrying out the work: Paris School of Economics (EEP-PSE), Goethe Universität (GU), Université Rouen (UR), and Institut National des Sciences Appliquées de Rennes (INSA). For practical reasons Erasmus University Rotterdam (EUR) has become a fifth WP11 partner institution, KNAW/IISH and EUR having decided to employ one postdoc stationed at EUR, Joanna Kinga Sławatiniec, to work on both EUR's WP9 and on WP11.

WP11 needs to deliver two products: a report on best practices for the preservation and access to digitized material of interest for European cultural heritage plus a position paper on the best practice institutions' relevance as partners for EURHISFIRM. The intended contents of these two products are discussed in more detail below. The first product was scheduled for April 2020, the second one for October 2020. For COVID-19 reasons both deadlines have been postponed indefinitely; our intention now is to produce both no later than March 31, 2021.

2. WP11's general approach

To tackle its tasks, the WP11 team proceeded as follows. In October 2020 we organized a roundtable with our four partner institutions, represented by Angelo Riva, Wolfgang König, Bertrand Couäsnon, and Thierry Paquet. Structured around a short position paper/questionnaire concerning WP11's plan of action and deliverables (see Annex 1), the highly inspiring round table greatly helped to structure the WP11 team's subsequent efforts in executing its task programme. The team then set out to research and write the present report, partly based on a survey of 48 cultural heritage websites presenting collections of documents and objects. In January 2021 the draft report was circulated amongst the representatives of our partner institutions and following revisions prompted by their comments submitted to the steering committee for approval.

3. The issues at hand

EURHISFIRM builds a European Research Infrastructure (RI) to 'connect, collect, collate, align, and share detailed, reliable, and standardized long-term company-level data'. The RI enables researchers, policy makers, and the public at large to access, download, analyze, and compare those data from all participating countries. At its core the RI will give access to national databases with, on the one hand, securities prices



from the countries' stock exchanges, and on the other financial statements and securities' issues data plus names of managers, directors, and other executives. Those databases have been compiled from a variety of paper sources. WP4 produced highly useful, comprehensive survey of documents and document collections in the consortium countries listed by country and type.

For our purposes, this list may be reduced to two document categories:

- (1) Various stock exchange-related publications, dedicated newspapers, and investors' guides and handbooks;
- (2) Additional supporting publications, such as corporate annual reports or reports from credit rating agencies, commercial publishers, or investor service firms.

Databases such as the Antwerp SCOB and Paris DFIH were originally built around (1) and added (2) when funds permitted or links could be made to scans made and held by other institutions, in the DFIH case the French Bibliothèque Nationale. Ideally users should be able to consult scans of the original documents occasionally, frequently, or continuously, depending on their intentions and research interests. The Paris DFIH already does this, allowing users to move from a particular security to both the original page in the price current and to the relevant page in the stock exchange yearbook. This has been done by linking the prices database to the relevant documents scanned by the Bibliothèque Nationale and available through that library's website Gallica. It is desirable for EURHISFIRM to provide guidance on how best to preserve and facilitate access to them, so the RI's users switching between national databases may also easily find and retrieve the relevant supporting documentation.

In addition a third source category relevant to stock exchange and corporate history must be distinguished, which we summarize non-exhaustively as:

- (3) Corporate articles of association; press releases and coverage; publicity posters; flyers and folders; paper shares and bonds; prospectuses for securities issues; pamphlets; rules and regulations concerning stock exchanges; paintings; engravings; photos of people, buildings, and trading floors; architectural and other drawings; cartoons; newspaper or magazine articles; films; give-aways like keyrings, pens, ties, caps, ashtrays, and cigarette or cigar lighters; decorative plates and other crockery; gadgets and knickknacks commemorating particular events; flags; stationery; office equipment like ticker tape machines; trading booths and counters.

This category poses particular problems which we discuss at length in Section 4.

WP11 identified five key issues concerning the preservation of and access to digitized documents, common to all three categories:

1. Digitized documents stored must remain integrated with the databases in such a way as to guarantee the persistence of links between them when databases are updated or migrated;
2. Digitized documents stored must be arranged in such a way as to allow later, correct insertion of additional, similar documents, such as missing volumes of serial works;



3. Copyright of digitized documents stored must have been secured so they can be shared freely;
4. Digitized documents must be stored in a durable way;
5. The original paper sources must be kept in appropriate preservation conditions.

Issues 1, 2, and 3 need not concern WP11 as being the responsibility of both the national database builders and managers and of other EURHISFIRM WPs: WPs 6 and 7 (Data Connecting and Matching respectively Data Extraction and Enrichment System) in the case of No. 1 and 2, and WP3 (Legal and Ethical Design) for No. 3. No.'s 4 and 5 are also the responsibility of national projects, but with most of those in statu nascendi it makes sense to survey best practices of preserving digitized and original paper sources.

Issue 5 is the simplest one to tackle. Contrary to widespread opinion, digitization is a means of widening access to paper sources, not an alternative way of preserving them. Digital files are considered more perishable than the paper originals from which they were made. Indeed some archivists are of the opinion that even digital born archives are best conserved in print-on-paper form. Consequently following digitization the paper originals should be carefully preserved and not destroyed. That also means that the integrity of such originals should not be compromised in any way by the digitization process. On the plus side digitization does enable the libraries and archives concerned to remove those paper sources from publicly accessible storage to cheaper, long-term safekeeping facilities. Issue 5 has greater relevance for document Category (3), the ephemera, for reasons to be explored in more detail below: they tend to be randomly collected, poorly identified in archive catalogues and, in the case of old securities, even destroyed in large numbers following the switch to electronic securities.

Issue 4 is a central one because the format of digitized documents determines the size and quality of storage space needed. The machine-reading software being developed by our partners in the EURHISFIRM syndicate generates very large amounts of data that need to be carefully preserved so users can link particular items to the exact location in the documents processed. Simple scanning of documents commonly results in very large, high-resolution TIFF format files, which are durable in the sense that TIFF has been industry standard for a long time and looks likely to remain so for the foreseeable future. The downside of large TIFF files is their slow loading and need for lots of storage space. For that reason they are usually converted into formats like JPEG or PDF, smaller, lower resolution, less detailed, and with a risk of conversion errors, but faster to load and lighter on storage space. Libraries, archives, and museums with lots of publicly accessible digitized documents therefore tend to keep the original TIFFs on cheap, slow, back-office storage systems and offer public access only to JPEGs or PDFs from more expensive and faster front-office storage. This allows such institutions also to update their publicly accessible documents from the original, full digital copies should the JPEGs or PDFs become corrupted or obsolete with the arrival of new digital image formats. Summing up, storing digitized images durably therefore requires a dual storage system.

That dual storage requirement looks likely to complicate a durability question which, though strictly speaking not part of WP11's brief, needs mentioning here. The EURHISFIRM-connected national databases, large and highly complex as they are, pose a formidable challenge when it comes to durability, that is to say, the funding of long-term maintenance and upkeep. The need for dual storage systems adds



another dimension to that cost problem which national project teams need to factor in. Storing digitized images durably therefore adds two layers of complexity: managing high-res and low-res image files, and running two separate storage and retrieval systems. Some countries possess storage and maintenance organizations to care for databases created by academic research projects. DFIH, for instance, uses the French national HUMA-NUM system for mass data storage. Other countries cannot rely on such facilities and are dependent on universities now willing to host databases, but which may change their minds. Moreover, national projects may become orphaned when their initiators move on or retire. This set of data storage-related problems needs to be thoroughly considered by the EURHISFIRM consortium if the RI is to achieve durability.

4. The complications of Category (3) documents

As indicated above Category (3) documents, the ephemera, form a case apart. Produced as they were by stock exchanges, market participants, commercial publishers, or business corporations themselves, this category is not just highly voluminous, it also consists of a huge variety of different materials. For ease of reference we repeat our non-exhaustive summary given above:

Corporate articles of association; press releases and coverage; publicity posters; flyers and folders; paper shares and bonds; prospectuses for securities issues; pamphlets; rules and regulations concerning stock exchanges; paintings; engravings; photos of people, buildings, and trading floors; architectural and other drawings; cartoons; newspaper or magazine articles; films; give-aways like keyrings, pens, ties, caps, ashtrays, and cigarette or cigar lighters; decorative plates and other crockery; gadgets and knickknacks commemorating particular events; flags; stationery; office equipment like ticker tape machines; trading booths and counters

Unlike Categories (1) and (2), Category (3) materials tend to be:

- (a) Incidental rather than serial;
- (b) Randomly preserved;
- (c) Scattered over libraries, archives, private collections, and museums;
- (d) Randomly catalogued and therefore difficult to find;
- (e) Of great interest to financial historians, but generally neglected by them, partly because of (d);
- (f) Of great outreach potential to other disciplines, including cultural historians and historians of art and design, and to a wider public of people nursing an interest for financial history ephemera, such as collectors and news media.

The key issue here is (d). Because the ephemera are generally difficult to find and access, they are underused even by financial historians, let alone other disciplines or a wider public. That is an important missed opportunity, because many Category (3) materials, though definitely not all, hold information now

often difficult to access and therefore important additions to national price database systems. Prospectuses, for instance, hold valuable information about the conditions of securities issued, about the evolution of corporate governance notions, and about how over time bankers sought to persuade investors to part with their money. Popular with collectors, old paper securities also provide vital information about issuing conditions and their often attractive look should be welcome fodder for historians of art and design for what it tells us about the world in which they were issued. Rules and regulations, and some pamphlets, give insight into the market microstructure, how the market worked and evolved over time. Architectural drawings and depictions of trading floors are beginning to find interest from cultural historians exploring the relationship between the built environment and the people who use it. Similarly, stationery, office equipment like ticker tape machines, trading booths and counters are important relics for re-creating the material world of stock exchange trading. Plays, poems, pamphlets and cartoons lampooning or lambasting the securities trade generally and speculation in particular show the reaction of contemporary opinion to what happened on the market.

Category (3) materials are therefore of primordial, if neglected, interest to financial historians while at the same time offering great outreach potential for national digitization projects and for EURHISFIRM beyond their primary constituency of financial historians and economists. In addition high-resolution images of some Category (3) materials, for instance old securities and cartoons lampooning speculation, have potential revenue value for national database projects and/or EURHISFIRM.

That primary goal, widening access to these neglected materials, requires eliminating bottleneck (d). The WP11 team recommends using the EURHISFIRM framework to develop a classification system for Category (3) documents, based in as far as possible and practicable on existing classification systems of national libraries, archives, and dedicated institutions like banks or stock exchanges and their archives or museums. Identifying, describing, and cataloguing them similarly across Europe will at the same time widen access, facilitate discovery, enable the correct insertion into national database systems, and ensure similar items can be found and linked across the EURHISFIRM RI.

However, there's more, much more. Taken together, Category (3) materials represent a dimension EURHISFIRM must take into account if it is to succeed: the socio-cultural dimension of European business and finance. The EURHISFIRM RI links national database systems of securities prices and corporate finance data so as to enable researchers and policy makers to analyze economic developments and trends across the continent over long periods of time. If these developments and trends are to be understood, their social and cultural embeddedness must be properly valued and fully taken into account: from market microstructure via prospectus regulations and corporate governance norms to trading floor lay-out, the design of exchange buildings, and the look of securities, precisely the kind of information to be found in Category (3) materials.

In their bewildering variety and number, these materials show Europe's defining characteristic as a continent: unity in diversity. All countries experienced similar, interconnected phenomena: modern economic growth, the appearance of big corporations, and the rise of financial markets. But they did not experience them at the same time, in the same way, with the same attitude, or with the same consequences and effects. Sometimes they embraced change, sometimes they resisted or even fought it. Sometimes they learned from each other, for good or ill; sometimes they chose their own way, discovering

new horizons or a cul-de-sac. New ideas, techniques, and practices travelled from one end of the continent to the other, remaining recognizably the same but, subtly tweaked to suit local circumstances along the way, rarely the same in substance.

In addition to being important sources for financial history research and items with key outreach potential, the Category (3) materials therefore represent an unmissable opportunity for giving EURHISFIRM a social and cultural dimension vital to it becoming a success as an all-European project. To sum up, these materials should be added to national database systems, so users of those databases can click through to see the original prospectus or security. As a matter of course EURHISFIRM must connect collections of Category (3) materials and unlock the potential for exploring this particular, underused dimension of the European economic, social, and cultural heritage. If at all feasible this should be done in a way to maximize EURHISFIRM's outreach, to other academic disciplines, but more especially to as wide a public as possible.

That said Category (3) materials pose a huge challenge. The WP11 team highlights two serious complications. First, the volume, variety, and diversity of Category (3) materials is very large, perhaps limitless even for individual countries. Core items of interest to financial historians, for instance paper securities or prospectuses, already number in the tens, perhaps even hundreds of thousands. Before the switch to entirely digital securities stock exchanges used to preserve paper specimens, as did banks, other financial institutions and some dedicated financial or economic history archives and museums, more rarely libraries. Old paper securities are avidly collected, spawning sometimes large private collections. However, the preservation and cataloguing of paper securities and prospectuses varies from institution to institution and from country to country. Some institutions discarded or even destroyed their hoards of paper securities. At other institutions the securities form collections of their own, at still others they have to be located and dug out more or less by hand from a mass of other documents. The other Category (3) types of material, equally numerous but considerably more diverse, are even harder to locate, catalogue, and digitize.

Adding a comprehensive collection of Category (3) materials to national database systems therefore will require major efforts to classify, locate, catalogue, describe, and digitize them. Given the difficulty of obtaining sufficient funding for digitizing prices it is hard to see this being done on a large scale anytime soon. It is worth mentioning here that at present neither SCOB nor DFIH, the two oldest and most advanced national price database systems, also hold large collections of Category (3) materials. The WP11 team therefore emphasizes that building a comprehensive digital images collection of these materials at EURHISFIRM level remains a distant ideal. The proposed uniform classification system will bring that ideal a little closer, but in our view EURHISFIRM could develop other initiatives towards that end, discussed in Section 6.

Secondly, adding Category (3) materials to price database systems raises the latter's IT complications a notch or two. Of course any Category (3) image collection should be linked to the main price database so researchers can directly retrieve and view low-resolution images of the items relevant to their enquiry. Prospectuses, corporate statutes, or share specimens, for instance, should be given identifiers so their images can be easily matched to corresponding issuing entities and prices. That will be harder to do for many if not most objects, but the current trend towards linked data systems may make it easier by adding signifiers to the object description. The obvious way forward, discussed at greater length in Section 6, is

for EURHISFIRM to collaborate with an existing European RI such as CLARIAH, CLARIN, or Europeana which have already developed systems for most, if not all, types of cultural heritage materials.

However, the real issue lies elsewhere. Given the number, variety, poor cataloguing, and dispersion of Category (3) materials even major efforts to build a comprehensive image collection are unlikely to unearth all of them. Such collections should therefore have a facility allowing third parties, say stock and bond collectors, librarians, or archivists, to add images and descriptions of specimens in their possession, which necessitates either a stand-alone system or, perhaps more practical, a waiting room for outside contributions cleared weekly by a systems administrator.

The very nature of Category (3) materials presents a further IT complication. The great attraction of many, perhaps most ephemera resides in their design, their look, and their materiality. Databases holding them should therefore be designed with reverse priorities from the price databases: not data first and metadata, scans and photos at the back, but scans and photos upfront and the linked data and metadata at the back. If those images are to be monetized, the system needs corresponding features ensuring payment, copyrighting, and retrieval of high-resolution images. In that case it needs to be kept in mind that monetizing Category (3) images and a facility for third party contributions sit uneasily together, probably to the point of excluding each other. Images of Category (3) materials also need to be interoperable with other cultural heritage infrastructures such as CLARIN and/or Europeana. In addition WP11 recommends exploring within the framework of WP7 (Data Extraction and Enrichment System) and/or WP9 (Infrastructure Policy and Architecture) the possibility and desirability of extending the interoperability between EURHISFIRM and CLARIN/Europeana to the Category 1 and 2 documents, i.e. the scanned price lists and yearbooks, so as to widen their visibility and accessibility.

Finally, any Category (3) system design should facilitate drawing on the collection for thematic web exhibitions at both national and EURHISFIRM levels so as to generate website traffic and raise public awareness of and interest in financial history as embodied by the collection. It should be noted, however, that properly fulfilling this vital form of outreach requires national database projects and/or EURHISFIRM possessing a team member able and willing to act as curator for such exhibitions or else funding to employ guest curators. Simply presenting pretty images without context or connection should be discouraged and wider educational purposes favoured.

Summing up, an IT system holding Category (3) materials must:

- (1) Be capable of holding images, descriptions, links to location, and metadata of highly diverse documents and objects;
- (2) Link images, descriptions, links to location, and metadata to items in the main price databases;
- (3) Link images, descriptions, links to location, and metadata to items in other EURHISFIRM databases;
- (4) Facilitate continuous updating with new images, descriptions, links to location, and metadata by both insiders and outsiders;
- (5) Prioritize images over description;



(6) Be fully interoperable with other cultural heritage infrastructures;

(7) Facilitate web exhibitions at national and EURHISFIRM levels.

Therefore the Category (3) materials require, in addition to considerable time and expense for their collection, description, and digitization, fairly complex IT systems and facilities for access and presentation. Consequently the best we can do for now is at the same time what is most necessary, i.e. explore the available options for presenting Category (3) materials and develop recommendations for both the EURHISFIRM syndicate and for national database projects to follow.

5. Options for presenting Category (3) materials

As argued above, Category (3) materials should be:

- a. Added to national database systems;
- b. Connected through EURHISFIRM;
- c. Used to unlock Europe's economic, social, and cultural heritage and maximize EURHISFIRM's outreach;
- d. Made interoperable with other European cultural heritage infrastructures.

In order to find out how that might best be done, we conducted a brief internet survey to see how archives, document collections, websites, and museums in Europe and around the world present their subjects and/or materials. We focused primarily on those concerning economic, social, financial, and business history, but included a few general archives, websites, libraries, and museums to get a wider perspective. For now we concentrated on the front end of websites, their look and functions, and left their technical back end with the underlying systems requirements for imaging and cataloguing software for later consideration. Annex 2 presents the list of 48 sites analyzed with their characteristics. We classified sites by functionality into four types, in ascending order of complexity:

(1) Thematic or institutional websites with or without collection catalogues and with or without images;

(2) Thematic or institutional websites giving access to particular, searchable collections of documents and/or objects, with images of selected items and/or webexhibitions;

(3) Websites with a meta portal function linking collections and/or database systems, searchable by keyword, with or without images;

(4) Meta portals linking collections and/or database systems combining virtual exhibitions with presentations of individual collections or museums plus a search engine to locate objects in the collections of member institutions.

Unsurprisingly type (1) turned out to be by far the most common website: fairly simple, though often attractive, presentations of particular collections or institutions, nearly always with pretty images (<https://history.bnpparibas/>). Type (1) sites offer more rarely searchable catalogues or digitized documents, so for that reason alone the type is unsuitable for EURHISFIRM purposes. The Rothschild Archive (<https://www.rothschildarchive.org/>) and Baring Archive (<https://www.baringarchive.org.uk/>) present good examples of the more sophisticated type (2) with attractive web exhibitions, digital collection catalogues and, in the case of the latter, digitized documents from the archive.

That brings us to a first observation. Type (1) websites offer digitized documents and images of objects rarely if at all, type (2) websites more often, but then mostly if they belong to public institutions such as the Bibliothèque nationale de France (<https://www.bnf.fr/>), the Bundesarchiv (<https://www.bundesarchiv.de/>), or the National Archives Kew (<https://images.nationalarchives.gov.uk/>). Type (1) and type (2) websites are also mostly dedicated to presenting the collections of the institutions which own them and, links apart, no more. Those two characteristics rule types (1) and (2) out as examples for a collection of Category (3) materials to follow. After all, such collections will consist almost entirely of digitized documents and objects owned not by the institution running the particular website, but by a variety of other, linked institutions.

However, some institutional websites looking like type (2) double as type (3) in the sense that they combine their primary function, presenting an institution's collection, with a meta portal function providing users with keyword hit access to collections outside of the institution. National libraries tend to provide such a service (<https://www.bnf.fr/>, <https://www.dnb.de/>, <https://www.bl.uk/>), though as a rule only for documents, photos, and pictures, rarely if ever for objects. There are also a number of meta portals proper, websites offering users keyword access to the collections of member institutions. European public archives possess such a meta portal (<http://www.archivesportaleurope.net/>), for instance, as do institutions dedicated to preserving movies (<http://www.filmarchives-online.eu/>). We also classified as type (3) the European research infrastructures like DARIAH (<https://www.dariah.eu/>) or CESSDA (<https://www.cessda.eu/>), meta portals facilitating users access to database systems and metadata of member institutions. Amongst them, CLARIN (<https://www.clarin.eu/>) is a little different and, devoted as it is to language resources and technology, at first sight not really a match for EURHISFIRM. However, CLARIN's structure of family nodes grouped under thematic umbrellas might be just the ideal way of storing and presenting the very heterogeneous material gathered by EURHISFIRM. WP11 explores this issue further in its second report.

Partnered as it is by CESSDA, the EURHISFIRM research infrastructure will also likely be a type (3) meta portal for the price data, scans of supporting documents, and metadata of its member database projects. However, the one type (4) website, Europeana (<https://www.Europeana.eu/>), looks the better, indeed the obvious partner for the Category (3) materials. Europeana aims provide, as the website puts it, ways for sharing 'cultural heritage for enjoyment, education and research'. The site gives fully searchable access to the collections of a very large number and a wide range of museums, archives, and libraries, providing pictures about any conceivable subject and of just about anything imaginable. Europeana presents numerous web exhibitions, provides distance learning programmes, and also allows third-party contributions ranging from photos to personal testimonies and stories. In short, it looks the ideal website

for showcasing Europe's corporate and financial history through the Category (3) materials, either on its own or in combination with CLARIN.

6. The long way ahead

As indicated above, EURHISFIRM is unlikely to be in a position to start showcasing soon. Category (3) documents need to be identified, located, catalogued, scanned or photographed, and entered into national databases before anything can be shown at all at national levels, let alone the European level. Given the priority of digitizing stock exchange prices and corporate finance data, and given the funding difficulties for doing that, national database projects will probably and understandably focus on digitizing supplementary documentation like corporate articles of association, prospectuses, and stock exchange regulations because such documents are important for the context of securities prices. Other attractive ephemera are thus likely to be included last if at all, reducing the outreach potential and the possibility of inviting a wide public to explore the cultural dimension of Europe's corporate and financial history.

At this point in time the best EURHISFIRM can do is also the most necessary: stimulate and streamline that long process by laying the groundwork. WP11 recommends five closely related steps to be taken:

- (1) Establish a working group of financial historians and librarians and archivists of financial institutions to develop an elaborate, international classification of Category (3) materials;
- (2) Ask WPs 7 and/or 9 to set up a working group for choosing a standard software and website format suitable for connecting Category (3) materials to national data projects;
- (3) Explore the options and potential for collaborating with CLARIN/Europeana for rendering Category (3) materials, and perhaps Cats. (1) and (2) as well, accessible;
- (4) Join CLARIN/Europeana and promote web exhibitions showcasing Europe's corporate and financial history;
- (5) Use the classification, software and website standards, and CLARIN/Europeana membership, as levers for mobilizing institutions with large Category (3) materials to catalogue, digitize, and show them through both national websites and Europeana.

These five steps are simple in the sense that they require initiative, coordination, and perseverance, more than staff or money. They also rely on something EURHISFIRM is good at, that is to say, bringing people and institutions together for achieving common, complex goals. However, rendering Category (3) materials accessible for a wide public exceeds EURHISFIRM's time horizon, and exceeds it by far. The WP11 team recommends to work around that by using the remaining months to convince a couple of relevant institutions of the importance, indeed the need to develop a common classification of Category (3) materials with a view to showcasing them via Europeana. The French Bibliothèque Nationale already publishes a variety of Category (3) documents on Gallica. Institutions like AEX Paris, Brussels, and Amsterdam, for instance, have big collections, as do some of the banks. If approached via top officials

and/or sectoral organizations with a good concept for a number of joint exhibitions themed around Europe's corporate and financial history and heritage, they might be willing to start collaborating on rendering Category (3) materials more accessible through a common classification system and digital images. Linking those images across websites and to national database systems then becomes an easier and obvious next step. In addition national associations or foundation with large ephemera collections, such as the Belgian, Polish or Dutch association of securities collectors, may be approached to join the consortium and contribute images of their collections to the central one.

7. Conclusions

There exists a wealth of material vital for financial and corporate history research and therefore important as add-ons to the kind of price databases linked through the EURHISFIRM research infrastructure. That same material holds great, unused potential for outreach to other disciplines and to a wider public. However, the material is randomly preserved, poorly catalogued, seldom digitized, and scattered far and wide. Linking it to price databases and unlocking its outreach potential therefore depends on major classification and digitization efforts that will both outstrip the resources of national price database projects and outlast EURHISFIRM's duration. To get around that problem the WP11 team proposes to get major financial institutions interested in showcasing their collections via themed exhibitions about Europe's joint financial and corporate history on the Europeana website. At the same time EURHISFIRM should explore the options and potential of a close collaboration with Europeana and/or CLARIN for making Category 3 materials and perhaps also the Cats. 1 and 2 accessible to wider public. This exploration will be the subject of WP11's second report.

8. Recommendations

- The data storage-related problems of national database projects need to be thoroughly considered by the EURHISFIRM consortium if the RI is to achieve the durability desired.
- The EURHISFIRM framework should be used to develop a classification system for Category (3) documents, based in as far as possible and practicable on existing classification systems of national libraries, archives, and dedicated institutions like banks or stock exchanges and their archives or museums.
- EURHISFIRM should use the remaining months to convince a couple of relevant financial institutions of the importance, indeed the need to develop a common classification of Category (3) materials with a view to showcasing them via Europeana.
- EURHISFIRM should establish a working group of financial historians and librarians and archivists of financial institutions to develop a good, international classification of Category (3) materials.



- EURHISFIRM should ask WPs 7 and/or 9 to set up a working group for choosing a standard software and website format suitable for connecting Category (3) materials to national data projects.
- EURHISFIRM should explore the options and potential for collaborating with CLARIN/Europeana for rendering Category (3) materials, and perhaps Cats. (1) and (2) as well, accessible.
- EURHISFIRM should use the classification, software and website standards, and Europeana membership as levers for mobilizing institutions with large Category (3) materials to catalogue, digitize, and show them through both national websites and Europeana and/or CLARIN.
- EURHISFIRM should enter into a close collaboration with Europeana and/or CLARIN based on criteria to be developed in the second WP11 report.



Annex 1, Questionnaire for the EURHISFIRM WP11 roundtable, 10 November 2020

According to the overall EURHISFIRM project design, Work Package 11:

‘explores concepts and tools to stimulate the lasting conservation of the digitized material and provides guidelines for making those materials publicly accessible. It also explores innovative ways to use digitized images as documentation for the data extracted from them and evaluates alternative strategies to use digitized material’.

This task is given three main objectives:

- (1) The use of digital images to document data and inspire further research and identify sources of interest for cultural heritage;
- (2) The promotion of Europe’s cultural heritage by facilitating digital preservation and online accessibility of sources with a unique historical value;
- (3) The mobilization of digitized images of historical sources as an exceptional additional documentation for the data (including the exploration of ways to make materials accessible and connected to EURHISFIRM data).

Taken together these texts suggest a dual task: (1) develop guidelines for the preservation of and access to digitized material, i.e. scans of price currents, stock exchange yearbooks, corporate documentation, etc.; and (2) identify further sources of cultural heritage interest for addition to EURHISFIRM databases.

The WP11 team invites your opinion on these tasks so as to best perform them. To structure our discussion, we have formulated the following four questions. Please read them and take a little time to ponder over them.

Question 1

In your view, what are the key issues concerning these tasks that WP11 should address?

Question 2

In your view, which of those issues touches areas other WPs are set to deal with?

Question 3

In your view, which of those issues has already been solved by similar database projects?

Question 4

In your view, to what extent should EURHISFIRM’s federated structure determine the outcome of WP11?

Question 5

Do you have any other suggestion related to the promotion of Europe's cultural heritage in the area of financial markets?

Annex 2, Websites surveyed

	Country	Website	Type	Remarks
1	A	https://zedhia.at/	1	Thematic website, research portal for Central European economic history. No images. Access by subscription only
2	A	https://www.oenb.at/	1	Bank site, searchable collections, web exhibitions, images objects
3	A	https://www.onb.ac.at/	2	Austrian national library. Keyword searchable. Large numbers of digitized documents and images, web exhibitions
4	B	http://www.scob.be/	1	Thematic website with price database and additional documentation
5	B	https://www.bnpparibasfortis.com/	1	Bank site with archive catalogues, some images
6	B	https://www.avae-vvba.be/	1	Association for business archives

7	B	http://www.crisp.be/	1	Centre for social and political history
8	B	https://www.finarcheo.org/	1	Institution for financial history, some images, no collection catalogue
9	B	https://www.etwie.be/	1	Centre for technology and industrial heritage. Images, film clips, no collection catalogue
10	B	https://www.nbb.be/	1	Bank site, no images or archive catalogue
11	CA	https://www.bankofcanadamuseum.ca/	2	Bank site, web exhibitions
12	CH	https://www.finanzmuseum.ch/	1	Museum site, small web exhibitions, large non-searchable world securities collection
13	D	https://www.bundesarchiv.de/	2	Archive website, searchable catalogues, web exhibitions, digitized documents, images of objects
14	D	http://www.macrohistory.net/	1	Thematic website with extensive and downloadable macro-economic data
15	D	https://digi.bib.uni-mannheim.de/aktienf%C3%BChrer/data/index.php	1	Thematic website with scans and digitized German stock exchange manuals 1870-1999

16	D	http://ykawabat.sakura.ne.jp/statutenTopEng.htm	1	Thematic website with scans of German company statutes
17	D	https://datacenter.safefrankfurt.de/	1	Price database with links to other German databases, some of limited accessibility. No images,
18	D	https://www.dnb.de/	3	German national library also serving as metaportal with links to other collections. Keyword searchable, some documents, images, no objects
19	EU	http://www.archivesportaleurope.net/	3	Meta portal European archives, mostly public ones. Keyword searchable, documents, no objects, few images
20	EU	https://www.Europeana.eu/	4	Meta portal European museums, archives, libraries. Keyword searchable access to collections of documents, images of objects, video clips and sound clips. Web exhibitions. Allows third-party additions



21	EU	https://www.cessda.eu/	3	Meta portal promoting datasharing and datastorage for social sciences and humanities. No images, catalogues, or collections
22	EU	https://www.dariah.eu/	3	Research infrastructure for the arts and humanities
23	EU	https://www.erih.net/	1	Thematic website industrial heritage. Some images, no collections or catalogues
24	EU	http://www.filmarchives-online.eu/	3	Meta portal European film archives, keyword searchable. Some images
25	F	https://dfih.fr/	1	Attractive thematic website with database of stock exchange prices, voluminous additional data and documentation, digitized documents
26	F	http://www.entreprises-coloniales.fr/	1	Thematic website French colonial companies. Downloadable administrative data

27	F	https://www.banque-france.fr/	1	Bank site, small web exhibition, no catalogues or searchable collection
28	F	https://history.bnpparibas/	1	Bank site, web exhibitions, searchable collection, documents and photos, no objects
29	F	https://www.bnf.fr/	3	French national library and meta portal with extensive links to other collections. Keyword searchable. Large numbers of digitized documents and images, web exhibitions, audio and video clips
30	I	The historical archive of Italian Joint Stock Companies	1	Thematic website with references, no documents, data, collection, images
31	NL	https://art-heritageabnamro.nl/historische-collectie/historische-archieven	2	Bank site, searchable collection documents and objects, web exhibition
32	NL	http://www.vvof.nl/	1	Thematic website collectors old securities. No images, but downloadable catalogue
33	NL	https://financieelerfgoed.wordpress.com/	1	Thematic website Dutch financial history

				with images, maps, documents from various collections
34	NL	https://iisg.amsterdam/	2	Library and archive website, keyword searchable, many digital resources, images, audio and video clips
35	P	http://starepapiery.pl/	1	Old Polish shares, English-language translation under construction. Images and documentation, no digitized documents
36	SE	https://data.houseoffinance.se/	1	Thematic website Swedish financial history. Database with company reports and Stockholm price currents. No images
37	UK	https://images.nationalarchives.gov.uk/	2	Archive website, searchable catalogues, web exhibitions, digitized documents, images of objects
38	UK	https://www.rothschildarchive.org/	2	Extensive and lively bank site, non-searchable guides to collections, richly illustrated range of web exhibitions

39	UK	https://www.baringarchive.org.uk/	2	Comprehensive bank site, catalogue of collections, digitized document series, web exhibitions
40	UK	https://www.bankofengland.co.uk/museum	1	Bank museum, web exhibitions, no searchable collections
41	UK	https://www.bl.uk/	3	British national library with links to other collections. Keyword searchable. Some digitized documents and images, small web exhibitions, audio and video clips
42	USA	https://som.yale.edu/faculty-research/our-centers-initiatives/international-center-finance/data/historical-financial	2	Will Goetzmann's ICF historical datasite. Reference to various price databases and other sources, some images of original sources
43	USA	https://www.moaf.org/	2	Museum website with short description of past exhibitions. Images of documents and objects, videos. Searchable collection.
44	USA	https://www.library.hbs.edu/	2	Library website with searchable collection catalogues, high-res

				images of individual items, research portals, and web exhibitions
45	USA	https://www.openicpsr.org/	1	Searchable data repository social sciences
46	USA	https://loc.gov/	2	American national library. Keyword searchable. Large numbers of digitized documents and images, web exhibitions, audio and video clips
47	USA	https://www.nypl.org/	2	New York Public Library. Keyword searchable. Large numbers of digitized documents and images, web exhibitions, audio and video clips
48	EU	https://www.clarin.eu/	3	Research infrastructure for language resources and technology with an interesting facility for creating families of related items, databases, or images grouped together under a particular, thematic umbrella

