King's Research Portal

Citation for published version (APA):
Blanke, T., Bryant, M., Frankl, M., Kristel, C., Speck, R., Vanden Daelen, V., & Van Horik, R. (2017). The European holocaust research infrastructure portal. *ACM Journal on Computing and Cultural Heritage*, *10*(1), 1-18. Article 1. https://doi.org/10.1145/3004457

The European Holocaust Research Infrastructure Portal

Tobias Blanke, King's College London
Michael Bryant, King's College London
Michal Frankl, Jewish Museum Prague
Conny Kristel, Koninklijke Nederlandse Akademie van Wetenschappen
Reto Speck, King's College London
Veerle Vanden Daelen, Centre d'Etudes et de Documentation Guerre et Sociétés Contemporaines
René van Horik, Koninklijke Nederlandse Akademie van Wetenschappen

Over the course of the last century there have been significant changes in the practices of archives driven by the massive increase in the volume of records for archiving, a larger and more diverse user base, and the digital turn. This paper analyses work undertaken by the European Holocaust Research Infrastructure project (EHRI) to develop heritage archives into research infrastructures by connecting their knowledge and making it relevant for research. In the paper, we focus on EHRI's work on an integrated collection portal; acting as a central gateway to the rich information on Holocaust-related sources. At the time of writing, the portal contains over 150,000 descriptions of over 1,850 institutions that hold Holocaust-related archival material in 51 countries. In addition, it hosts concise reports that provide in-depth per-country information about the Holocaust history and archival situation in 47 countries, topic-focused research guides, and a range of other services. The paper presents how the EHRI portal work connects to the state-of-the-art of heritage portals and the novel solutions we had to develop to align the portal with the requirements of a research infrastructure.

1. INTRODUCTION — FROM THE MILK CAN TO THE EHRI PORTAL

In 1940 in the Warsaw Ghetto the historian Emanuel Ringelblum established a secret archive, code named Oneg Shabbat.[1] He engaged 50-60 people who contributed descriptions of events on a regular basis and wrote reports on different aspects of daily life in the ghetto. Emanuel Ringelblum was a professionally trained historian, but he worked closely together with people from different backgrounds. The majority of the participants, called "fighter-historians" by Levi [1988], were teachers, writers and economists, but they were also joined by a number of workers and craftsmen. Ringelblum and his team collected documents (diaries, drawings, songs, newspaper clippings, etc.) illustrating the victims' experience; or as Primo Levi wrote "how the ghetto lived and died day by day" [Levi 1988]. In 1942 and 1943, they buried the underground archive in three different hiding places in metal boxes and milk cans in and close by the ghetto. Two parts of the underground archive were dug up after the war.

This brief story about Oneg Shabbat underlines that history is not made in a laboratory, but tells the story of events in the real world. Holocaust studies — although it was not yet called that way at the time

---

[1]The archive was called Oneg Shabbat or in Yiddish Oyneg Shabes ("joy of the Shabbat"), because its team met weekly on Saturday afternoon.

— did not start within the walls of a university. The story also shows that early research on the Holocaust engaged a wide range of experts and colleagues. Ringelblum brought this diverse group together not just because the difficulties of the ghetto demanded it, but also because the subject required the input from several disciplines. Holocaust research has remained an interdisciplinary challenge and it has never become an exclusively academic undertaking.

The need for broad engagement with the public as well as the requirement to bring together different types of expertise are clearly reflected in the European Holocaust Research Infrastructure (EHRI) [EHRI 2016d]. EHRI became active in 2010 and has started its second four-year phase in 2015 as part of the Horizon2020 programme. It is a joint undertaking of Holocaust historians, archivists and specialists in the digital humanities. It is driven by the objective to make the documents that were often produced and collected under bleak circumstances available and accessible for future generations and to develop innovative techniques and methodologies for analysis and interpretation of such documents. Sources related to the Holocaust are extremely fragmented and dispersed across many public repositories and private collections. In order to connect material scattered around Europe and beyond and to shed light on hidden archival sources, EHRI undertakes a programme of trans-national networking and research to integrate local infrastructures, knowledge and expertise into a shared infrastructure of global dimension.

Some of the documents, which the Oneg Shabbat group put in a milk can, can now be found on the Internet, whereas the entire collection can be studied in a digital format in several research institutions in Warsaw, Jerusalem and Washington, DC. For EHRI, it was clear from the beginning that the online availability of reliable and properly contextualised Holocaust material has relevance well beyond the walls of academia. The evident significance of collections such as Oneg Shabbat emphasises the cultural and heritage dimension of archives in general. The Oneg Shabbat collection is undoubtedly a human creation intended to inform; an idea that is central to all definitions of cultural heritage [Feather 2006]. Like all archival material, Holocaust related collections are today not only considered as records that have been kept to serve as evidence, but just as much as sources of memory about the past; that are 'about history, heritage, and culture, about personal roots and family connections, about who we are as human beings and about glimpses into our common humanity through recorded information in all media' [Cook 2000].

Given the continued challenges to the memory of the Holocaust in some European countries, gathering evidence remains an important part of all work on the Holocaust. In this respect, it should be mentioned that producing and safeguarding evidence was also a key driving force for Oneg Shabbat. Ringelblum and his colleagues feared that the unprecedented crime would not be believed if there would be a lack of adequate evidence. This illustrates perfectly that evidence, as used by historians, and memory, as exhibited by heritage institutions, should not be seen as irreconcilable and mutually exclusive concepts, but that they do exist side by side.

EHRI's effort to integrate Holocaust-related documentation shows that archivists can no longer be considered neutral custodians or simple record-keepers but increasingly enact and perform social memory through their practices. As a new public space the World Wide Web creates new opportunities for public engagement and involvement for all heritage organisations, mainly through online portals that offer the chance to present their work to a large audience. This paper reports on the specific challenges to develop the EHRI Online Portal [EHRI 2016d]; acting as a central gateway to the rich information on Holocaust-related sources. At the time of writing, the portal contains descriptions of archival institutions that hold Holocaust-related archival material in 51 countries. On average, 200-300 research sessions a day take place on the EHRI portal. In addition, it hosts a range of other services, concise reports that provide in-depth per-country information about the Holocaust history and archival situation in 47 countries as well as research guides.

EHRI is at its heart an integration project. By connecting collections from different archives and countries with research, EHRI aims to make a significant contribution towards more integrated research on the Holocaust, as we analyse in Section 3. To this end, in the first phase of the project we have developed an

environment to provide high-quality research descriptions for relevant collections including many that were previously hidden. Section 4 presents the novel technical environment that we had to develop in order to make it easier to integrate collections that are relevant to research but not easily accessible. The technical architecture of the portal (Section 4) is developed around this challenge. In Sections 5 and 6, we present how users can work together through the portal; finding collections and sharing them; developing research guides; building virtual collections; and presenting research documentation blogs.

2. BACKGROUND - THE CHANGING NATURE OF PORTALS

According to Barber [2006], portals have two key functions: the integration of content, and the presentation of that content to users. To this end, all portals incorporate services and data; make use of the federation and integration of data; and provide to varying degrees access control, customisation, and personalisation. Historically speaking, portals first emerged as Enterprise Information Portals (EIP) in order to share information in enterprises using web-based interfaces in the late 1990s [Firestone 1999]. However, from the beginning EIPs supported complex workflows for users to engage with the portal's contents [Firestone 1999]. After the initial enthusiasm for portals as a single point of access in the enterprise that contains everything, it was soon realised that this ambition might overload portals. Rather than being built around heavy standards, modern leaner portals concentrate on essential functionality and make use of web technologies: HTTP-based APIs, the social web, etc.

Most if not all heritage organisations operating digitally today have some kind of portal, following on from the success of EIPs. Through portals the institutions promise their online visitors a view "at culture as a whole instead of just gazing at it in thin slices" [Mäkelä et al. 2012]. In the domain of the Holocaust two existing large scale heritage portals stand out: Yad Vashem in Israel [Yad Vashem 2016] with the Central Database of Shoah Victims' Names holding over 4 million information items and the US Holocaust Memorial Museum [USHMM 2016]. EHRI builds upon their success.

Heritage portals have become a distinct research topic concentrating on several themes. A first research theme looks into enabling specialised searching of heritage collections through portals. "Semantic searching", for instance, groups results by related concepts, as Schreiber et al. [2008] discuss for large virtual collections of cultural-heritage objects. An extension and specification to semantic search is the spatio-temporal search, as suggested by Kauppinen et al. [2008]: "The idea is that regions offer one view to a content and they can be used to select a subset of the content by specifying constraints (. . .)." We will come back to semantic searching when we discuss the features of the EHRI portal in Section 5.

A second active research topic around portals is related to the distribution and integration of diverse material, as it is common in heritage organisations [Blanke and Hedges 2008]. Advanced computational reasoning to integrate vocabularies and metadata, as presented by Chortaras et al. [2014], is nowadays part of many portal applications; set up as distributed applications and services. The heritage portal Europeana [Isaac et al. 2012] experiments currently with a service structure, where the central Europeana.eu portal is only one type of front end to its content. Europeana will become a data provider brokerage service, which other heritage organisations can plug into and develop their own applications upon. Advanced portals should thus not just harvest content from organisations but allow these to present the material through their own interfaces [Mäkelä et al. 2012]. In Section 4, we present our work on integrated APIs that will deliver the EHRI material to remote digital services.

In addition to searching and integrating material at a central point of access, a third major research theme around portals is the integration of user generated content [Mäkelä et al. 2012]. While the representation of heritage content needed for computer-to-computer communication can be complex, simplicity supports human understanding of material and subsequent interaction with it (see Section 5.2). For Mäkelä et al. [2012] simplicity of access leads to new innovative user interactions with the content, as they create exhibitions and

combine material with narrative elements. We present in Section 6 our attempts to create dynamic research guides which help users understand how to access material across archives.

In the context of research applications, portals have evolved to become so-called Virtual Research Environments (VREs) or "web-based, community-oriented, comprehensive, flexible, and secure working environments" [Candela et al. 2013]; connected to some of the largest scientific infrastructures. VREs were supposed to be to research what virtual learning environments are to teaching. Several specialised VREs for humanities were created; mainly to work with textual documents. Sarwar et al. [2013] integrate high-performance computing resources to support linguistic work, while Hedges et al. [2013] engage and support researcher working on critical editions. In the original vision, however, VREs were supposed to cover generic use cases and span all the arts and sciences. They thus relied on relatively heavy portals to develop a generic set of tools and services that were supposed to serve all communities. While VREs have been successfully created for a particular "community of practice" [Candela et al. 2013], it is less clear whether a generic set of VRE tools and services will ever be developed. Research processes are compared to e-learning activities much less controlled and well-defined and even in the world of e-learning new more interactive environments emerge contionously.

In this paper, we present how EHRI builds upon the above-mentioned three research trends for portals and has developed its own version of searching and accessing heterogeneous material that is enriched with user-generated content. EHRI departs from traditional heritage portals, however, by adding a focus on research to accessing, integrating and presenting digitally available material [Blanke and Kristel 2013]. This is expressed in our workflows, which start with a focus on relevant but often hidden research material (Section 3) and end with an idea of the complex narrative integration of specialised materials as presented in Section 6. From other humanities VRE efforts, we depart in a renewed focus on simplicity for any user interaction as our audience is wide-spread across several research communities. While we support some of the standard features of VREs, as presented in Section 5.2, we do not rely on heavy work-flow models.

3. INVESTIGATION WORK

A central component of both EHRI's first and second phase was its investigation work identifying research archives and collections relevant to the Holocaust. At the time of writing, over 1,850 institutions with Holocaust-relevant collections (spread over 51 countries) and over 150,000 archival descriptions were formally described according to an International Council of Archives (ICA) standard. Where possible, the descriptions are connected with a multilingual thesaurus. In our investigation work we concentrate on all archives that are relevant to research, whether they are easy to access or - just as likely - closed or restricted in nature. Because Holocaust-related sources are dispersed across Europe and beyond and held by a wide variety of institutions, the portal furthermore hosts country reports which provide concise information on the Holocaust history, archival situation, and the status of EHRI's research on data identification and integration work relevant countries. Such reports are available for 47 countries at the time of writing.

3.1 Identifying the Archives

When EHRI started its data identification and integration work in October 2010, the project first began integrating information that was already available. Here, EHRI did not have to start from scratch, because organisations throughout the world have already done excellent work in collecting and saving documents, objects, photos, film and art related to the Holocaust. Identifying widely dispersed Holocaust-related sources, however, remains a major challenge. Prior to EHRI, large heritage institutions in the field such as the United States Holocaust Memorial Museum (USHMM) and Yad Vashem have spent several decades surveying archives to discover sources on the Holocaust and are far from concluding their work.

The first step for EHRI consisted of merging information on institutions which hold Holocaust-related archival material, the collection-holding institutions (CHIs), into a consolidated overview that took into ac-

count the focus of previous surveys. The EHRI identification work's starting points were the Directory of Holocaust-related Archives of the Conference on Jewish Material Claims Against Germany [Claims Conference 2016][2] and the Guide des archives sur la Shoah of the Mémorial de la Shoah in France [Memorial de la Shoah 2016]. Finally we worked with a list of institutions which Yad Vashem has worked with and copied material from, as well as Yad Vashem's internal overviews of surveyed institutions. This compiled list was further completed with information from USHMM as well as aggregators such as national research portals on archival sources on the Second World War[3], national archival guides, experts and published studies on the subject.

The next step in the identification work was to add information on the collections held by these CHIs which contain Holocaust-relevant material. All sources related to the Holocaust could and should in principle find their place on the EHRI portal and the project works to provide an overview, as complete and inclusive as possible. However, for pragmatic reasons it was necessary to establish some priorities. In order to structure the identification work and provide the consortium with a common framework, EHRI developed a concise internal working definition of the Holocaust. EHRI's surveying work starts with Germany, its allies and the countries occupied by the Axis Alliance, but special attention is paid to identifying Eastern (and Central) Europe sources, as these are the places where most victims perished and where sources for research have been less inventoried and made accessible in comparison with Western Europe.

Even though EHRI's main focus for active surveying goes to the aforementioned countries, others are included as well, and the project integrates relevant information where ever feasible. As such, an overview of restitution and compensations archives has been integrated in the portal because EHRI received it via the Conference on Jewish Material Claims Against Germany. In this way, EHRI works with a pragmatic and inclusive approach to bringing in as much relevant content as possible into its portal.

In the identification work, we have encountered a large variety of Holocaust CHIs, as relevant materials are situated in research archives as well as traditional memory institutions. The surveyed institutions include public and private archives, memorial sites and museums, etc. The CHIs likewise have myriad ways of describing and providing access to their data; from a complete absence of descriptions, to highly detailed descriptions in an analogue and/or digital environment, using archival standards and/or a standardised database with or without export facilities. In practice, however, very few institutions work in a standardised and open way, so that EHRI can publish, add and update their data in the EHRI portal without manual preparatory work. EHRI, however, aims to have as many institutions as possible join in via standardised digital connections. Therefore tools are currently being developed and tested using the ResourceSync [Haslhofer et al. 2013] framework in order to assist institutions in the sustainable digital publishing of their material.

3.2 Integrating Material in the EHRI Portal

The integration of material in the EHRI portal follows a well-defined workflow. After initial intermediary steps (using Excel and consequently in the ICA-AtoM software [Bushey 2012]), all the collected information on CHIs and their holdings have been entered into the EHRI portal, which is open source and complies with the relevant ICA standards. ISDIAH (International Standard for Describing Institutions with Archival Holdings) has been used for the CHIs and ISAD(G) (General International Standard Archival Description) for the collections. All CHI descriptions are provided in English as are collection descriptions if they are available in English. They are entered into the portal and connected to the CHIs via two pathways, via (semi-)automated ingest or manual data entry (see Section 4).

Once the material is in the portal, EHRI invests in contextualising it in several ways. Because of the trans-national aspect of the Holocaust the documentation is highly dispersed and highly context-dependent.

---

[2]The site is not active anymore but we received the underlying database from the Claims Conference directly.
[3]e.g. http://www.archievenwo2.nl/ in the Netherlands or http://www.ns-quellen.at/ in Austria

A typical example is the manner in which collections can be fragmented due to changes in regional political regimes during the Second World War, resulting in material with shared provenance being organised under different names before and after a given date. EHRI aims at providing information for users to make these connections via linking such collections to each other.

In order to contextualise EHRI's identification work within a systematic and structured framework, country reports on the state of Holocaust research are a key part of the portal. All EHRI country reports follow the same general structure. First, in two short paragraphs a general overview of a country's history during the Second World War is given. The first paragraph covers questions of statehood as well as German rule and influence, while the second paragraph focuses on Holocaust history and also includes information on the size of the pre-war Jewish community as compared to the total population of a country and an estimated number of Jewish victims. In a second section the reports describe briefly the archival situation: a first paragraph deals with the archival culture of a country and how its archives are organised (centralised system, role of the state, and legislation can be addressed here); the second paragraph gives more information about which archives are most relevant for Holocaust research. The third part of the reports provides information on EHRI's research in the country. This results in a short summary per country. If further research was undertaken by EHRI, the reports can be longer. In consulting the summaries, the reader should get a clear and concise overview of why the country is being researched and what the current state of knowledge and access to Holocaust-related archives in this country is, including EHRI's own identification efforts.

Not only the changes during the Second World War have made it necessary to contextualise the material in the portal. Equally, the multitude of documentation projects in the post-war period has resulted in multiple copy collections being available for research in a variety of institutions worldwide. This means that the same documents are described by multiple CHIs, both original- and copy-holding ones. Here too, EHRI is exploring novel ways to linking these descriptions. We aim to provide the researcher not only with information on the locations where she can consult the materials, but also with information on the available metadata and its languages. This can open up sources for researchers that were previously difficult to discover. Perpetrator sources on the Holocaust, for instance, are most often written in German, but the metadata of their host institutions can be in languages the German-reading researcher might not master. Hence, making the connection with, for example, an English-language copy-collection description can open up original sources for researchers spread throughout Europe and beyond. Similarly, victim sources can be in Yiddish and Hebrew and the metadata in various other languages; their content varying because of the different context, focus, and areas of expertise of their describer.

In summary, the material we needed to integrate into the EHRI portal is therefore marked by multiple complex relationships [Bryant et al. 2014]. Such relationships may be of a variety of different types; for instance, materials may be related because they pertain to common research themes, because they refer to a specific event or because they relate to a common geographic feature and so on. Many archival materials created by one agency have furthermore become fragmented and are currently held in multiple repositories that have internally organised these materials according to diverse principles. Finally, as just discussed, large institutions dedicated to Holocaust research, archiving and memory such as Yad Vashem, USHMM and the International Tracing Service (ITS) have over the years copied Holocaust-related sources from repositories across the world and thus created large copy archives. This means that a given archival collection may be available from multiple repositories and may have been described and organised according to a variety of paradigms. The technical environment of the portal has to answer to all these challenges.

Through contextualising Holocaust material, EHRI aims to support new research and to enable historiographical progress by providing access and by sharing information on sources that were hitherto relatively unknown. But the diversity of institutions with Holocaust material [Speck and Links 2013] is further complicated by the number of different approaches to archival description employed by those institutions. The next section will present our approach to dealing with the richly varied data that is the result of this heterogeneity.

4. TECHNICAL ARCHITECTURE OF THE PORTAL

As discussed in the previous section, the information contained in the EHRI portal has different provenances. While all country reports and descriptions of archival institutions have been authored by EHRI, descriptions of archival materials have been sourced and integrated via a variety of methods: when available, existing descriptions authored by the partner repositories or other documentation projects have been imported into the portal; however, for particularly important collections that have so far not been described in a format suitable for import, EHRI has authored its own descriptions [Vanden Daelen et al. 2015].

This multi-pronged approach is needed to fulfil EHRI's central mission to integrate existing Holocaust resources while paying particular attention to hitherto hidden collections. However, it also points to one of the major difficulties we had to overcome, namely the need to integrate data of highly heterogeneous nature. While we have followed applicable international standards for all EHRI authored descriptions, descriptions sourced from elsewhere are typically not standard-compliant. Mäkelä et al. [2012] report on the typical problems faced by heritage organisations. The underlying databases from which the information is to be integrated are often legacy systems that pre-date the applicable standards, run on obsolete or unsupported systems, and bear the mark of many years of ad-hoc evolution in business practices.

In dealing with this, the approach of standardising institutionally diverse descriptions to a common denominator was rejected, since it would have entailed an unacceptable loss of information for specialised research. Moreover, institutionally idiosyncratic descriptions can often reveal much about the complicated archival histories these collections have undergone, and as such constitute in themselves a valuable information resource (Links et al. [2015]; see also Duff and Harris [2002]). Unlike other large-scale archival integration projects such as ApeX [Sugimoto 2013], therefore, we decided to keep standardisation at the source to a minimum, and follow a "take it as it comes" approach to integrating data.

The EHRI portal's technical architecture answers these specific challenges using a fairly simple design, which is largely centred around a Neo4j graph database [J.Sadalage and Fowler 2013] to hold both archival and administrative metadata. Access to the Neo4j instance is mediated via a web-service which performs several higher-level functions:

—domain-object serialisation and deserialisation

—enforcing read and write access according to a flexible role-based system

—maintaining an audit log of actions which modify data batch operations, such as data ingest and bulk modifications

In addition to this Neo4j-based store there are several other components which perform in the overall system and support the kind of access required by modern heritage portals, as discussed in Section 2:

—an RMDBS system (PostgreSQL) containing user account data and other non-archival content

—a search engine (Apache Solr) for free-text queries

—a suite of small, single-purpose tools to transform data streams from the metadata store for search and semantic indexing

—a semantic repository (Ontotext GraphDB), currently under development, that will provide a SPARQL endpoint for a subset of publicly-accessible data and native storage for EHRI's SKOS-based resources

4.1 The EHRI Graph Database as a Rich Data Source

The decision to use Neo4j as the reference data store deserves attention, since often more traditional databases are preferred here, while other more advanced research applications seem to prefer semantic web store as reference stores [Blanke et al. 2013]. Initially, the main considerations were:
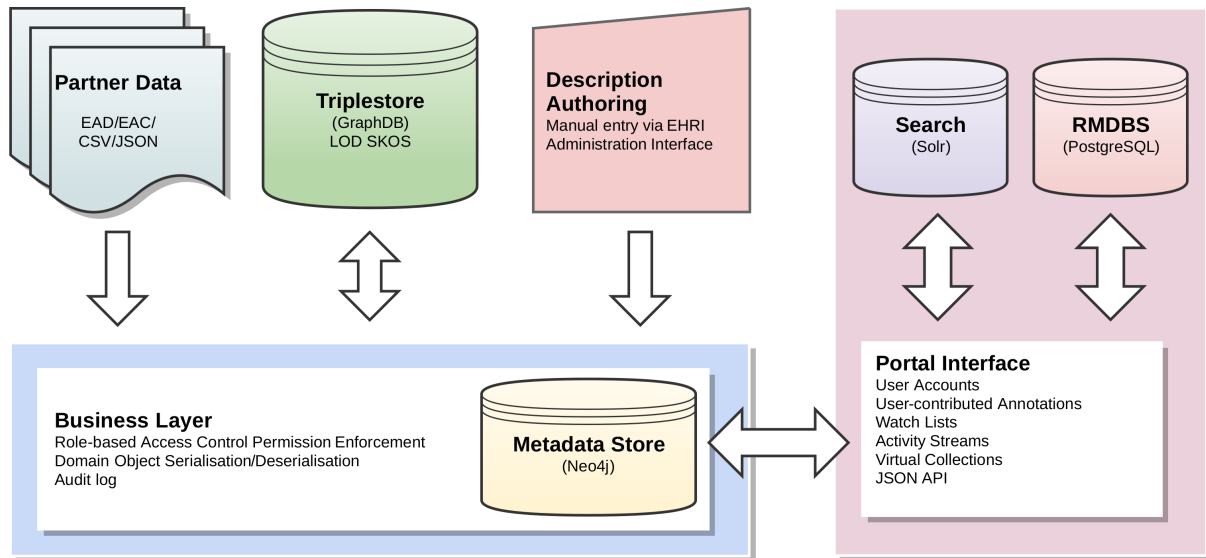
Fig. 1.   EHRI Architecture Model

—EHRI dealt primarily with archival metadata and enriching this metadata with additional contextual information was a key part of EHRI's mission. The entirety of this metadata was therefore considered *in-use* rather than (in data-centre parlance) *at-rest*, so an online system catering to dynamic manipulation was favoured over a more static repository-like system. Moreover, where XML data such as ICA's Encoded Archival Descriptions (EAD) were used, it was predominantly as a serialisation format, exported (often in an ad-hoc manner) from legacy systems in use at partner archives and not intended as a reference for public consumption.

—As presented in Section 4, archival collections are described in hierarchies of non-uniform shape and thus require hierarchical queries to navigate. Neo4j, and many other native graph databases, optimise for the use-case of traversing unbounded paths (recursive, self-referential joins in SQL parlance) without the need for structural book-keeping such as adjacency lists or nested sets. There were other cases in which the *relationship-first* approach of a graph database (and of query languages such as Neo4j's Cypher) was considered to be a good fit for EHRI's data; in particular a preponderance of join-heavy queries.

—Due to the integration of data from partner archives proceeding in parallel to the development of the EHRI portal, a "schema-less" database system was deemed desirable, allowing us to easily modify data types, attributes, and relationships as the requirements (and our understanding of the domain) developed.[4]

Secondary, but still important, considerations were:

—The property-graph paradigm of Neo4j, in which data is represented as nodes and relationships, both of which can possess arbitrary data attributes, appeared to allow considerable simplifications in modelling the heritage domain of archives. Most notably, it allowed interconnections between material of different types and a degree of object polymorphism not possible in a typical RMDBS schema without the introduction of undue complexity.

---

[4]Newer versions of Neo4j than were available at the beginning of EHRI provide various schema and data integrity features which we make use of in our newer production instance.

—Neo4j was both free and developed in-the-open under an open-source license; albeit by a commercial company. This made it easy for EHRI developers to obtain it and, if necessary, browse the source code.

—An extensive ecosystem of third-party tools and APIs exist for Neo4J; including some expressly designed for bridging between different graph databases, reducing the degree of vendor lock-in to which we felt exposed.

—Neo4j was developed in, and extensible by, the Java programming language, with which developers on the project were familiar and comfortable.

Using graphs, we have developed in EHRI our own rich data source that is servicing the portal but can also be used from the outside, and were the first digital heritage project to use a graph database to this end. In the portal, the backend web service provides an interface to a business layer that sits atop Neo4j; translating between the raw database primitives (nodes, relationships, properties) and domain objects (archival descriptions, holding repositories etc). It also enforces access control for read and write operations and maintains a log of data changes. EHRI's need to support both aggregation via automated harvesting and manual data entry by a varied group of archivists and researchers as well as user-generated content from historians in multiple countries, meant that a large focus of development effort was ensuring that the provenance of the information was maintained as it evolved through these different means. This required furthermore fine-grained access control.

The primary unit of exchange in the EHRI backend is a subgraph, comprised of multiple connected graph nodes, which represent domain objects of certain known types. When a domain object is created, updated or deleted, the business layer ensures that all nodes which comprise it are added, modified or removed in a transactional manner. Moreover, a subgraph is always updated in an idempotent fashion, meaning that operations which would not result in logical changes to the data do not result in changes to either the audit log or the underlying database. Supporting idempotent, side-effect-free repetition of data update and ingest operations in the business level allows us to support periodic re-harvesting and integration of disperse material; modifications will only be applied if the logical data (as opposed to the source file timestamp, for example) has actually changed. This system is used generically across all data types.

Unlike many otherwise comparable heritage infrastructures which aggregate material that is already in the public domain, EHRI had a requirement to provide restricted access to potentially sensitive data. The requirements of the VRE (see Section 5.2) also mandated the existence of access control mechanisms to ensure a user's data and her generated content remained private in the portal. Moreover, the relatively large and diverse group of (often non-technical) specialists manually curating different aspects of EHRI's data demanded a flexible role-based system of permissions that could cover many scenarios: from managing all EHRI's archival metadata for a given CHI; to allowing editing of records without also permitting their creation or deletion; and to managing the description of a single repository.

Several of the access scenarios emerged from experience with existing metadata management tools, while others came from discussions with working archivists. It was clear at an early stage that the interactions between hierarchically-structured data forming scopes in which permissions could apply (type of item, individual item, country, institution, fonds, etc.), and the hierarchy of roles to which a given user could belong would be complex. To counter this complexity, the system was implemented as a core part of the backend and made as generic as possible, so that rules could be enforced on many different data types in a uniform manner and be consistently and efficiently used by higher-level components. In practice, the EHRI portal makes use of the backend's permission system in many different contexts, such as mediating access to archival data modification, user-generated content such as annotations and updates to a user's own profile when they log-in via a third-party authentication system such as OAuth2.

4.2 Data Integration Approaches

The complexity of various aspects of cross-institution data integration, touched on above, has led us to adopt a higher-level, more abstract approach to several aspects of domain modelling for data integration in comparison to that permissible in a less heterogeneous environment. This allowed us to implement a novel approach to the portal research challenge of integrating material with a particular focus on the EHRI domain's archival access points and dispersed copy sources.

Linking of archival descriptions via so-called archival "access points" (subject terms and authority file references) is well supported by both the existing conceptual and technical standards (ISAD(G) and EAD, respectively.) Our experience as data integrators, however, was that there is relatively little use of shared controlled vocabularies across archives that could be used for access points; in large part due to the diverse range of institution types and varied internal working practices. EHRI is also constrained from being able to change data provided to us by third-parties, both to avoid clouding the provenance of that data and to prevent issues with repeatable re-harvesting, since our changes would create conflicts with the original data.

EHRI's framework for description interlinking, touched upon in Section 3, therefore operates on a higher level of abstraction than access points used for *intra*-archive linking, elevating the link itself to a full-fledged entity, rather than just a pointer from one item to another, as it typically implemented. In cases where we unambiguously know the item to which an access point refers a link is created binding the two items together, with an access point reference (in the form of a simple text label) providing its *body*. This decoupling allows us to integrate harvested, third-party material without having to resolve access point references before or during ingest, while also preserving the ability to do so afterwards without affecting the original data.

A particular challenge to Holocaust research is that it is common for the same *physical* archival material to have been described multiple times by different institutions. In some cases, multiple descriptions can simply represent the same text in multiple languages. More frequently, however, we have found that alternate descriptions are created not just in different languages but at different times and for different purposes, and with the intention of addressing different questions [Links et al. 2015]. EHRI therefore takes the approach of modelling the conceptual item (representing, for example, an archival fonds) in a manner distinct from an individual description of that item.

4.3 Structured Data Interfaces

While our priority in EHRI's first phase was on delivering functionality for non-technical users (see Section 5), the provision of machine-readable interfaces to the EHRI store is an important consideration and a key focus of our efforts in the second phase of the project. The approach is a pragmatic one which gradually expands the scope and functionality of the structured data interfaces according to LOD best practices, starting with that which is easiest to align with external sources of knowledge.

Currently, all of EHRI's descriptions of archival material, CHIs, and authority files are available in XML format as EAD, EAG and EAC respectively. This data is also continuously indexed, on a selective basis, to a semantic repository (an instance of Ontotext's GraphDB [Dimitrov et al. 2015]) to provide a SPARQL interface, with the quantity of data available as RDF being incrementally expanded as challenges with semantic alignment are overcome. In addition, we are currently testing a JSON-based API[5] (based on the JSON API specification[6]) that serves as an easy-to-use structured interface to the EHRI portal's primary data on archival descriptions, institutions, country reports, and authority files.

---

[5] Available at: https://portal.ehri-project.eu/api/v1/. Currently this API serves as an integration point with EHRI's tools that use third-party platforms, such as the Wordpress-based Document Blog described in section 6.2.
[6] http://jsonapi.org/

5. PORTAL END-USER FUNCTIONALITY

The information integrated into the portal can be searched for and explored by users in a variety of ways. The implementation of specific search functionality is an important research topic for digital heritage, as outlined in Section 2. We have also spent a lot of effort to understand how users access archives, and to then develop corresponding services. Our user requirement work indicated that Holocaust researchers have two broad information retrieval requirements: an ability to perform precise searches that allow for the location of the famous pin in the haystack; as well as support for information retrieval methods that allow for serendipitous discovery (Anderson and Blanke [2015]; see also Duff and Johnson [2002]).

5.1 Searching and Browsing

In our experience, serendipity in access to material in the portal is best supported by offering users opportunities for browsing with self-defined entry points; thereby allowing them to freely leaf through the portal's information content. The portal's material is structured in a hierarchical manner which provides users with one possible browsing method. The top level are the countries, which EHRI has detailed in its country reports and which provide one possible entry-point into the portal. From these reports, users can browse the CHIs within each country and from each institution. Researchers can then proceed to a list of the collection descriptions held therein. Collection descriptions are similarly hierarchical, typically consisting of a top-level description with a varying number of child levels. Across these hierarchies, users can further follow links to content that is related along several axes: provenance, subject, geography, individuals, organisations, etc.

Browsing is supplemented in the portal by a universal free-text search facility, as described in Section 4 that returns results for any types of material matching a user's query, on the basis that the country reports and descriptions of institutions may be just as relevant in answering a researcher's question as descriptions of collection material. If a user has a more focused idea of what she is looking for, any of the hierarchical scopes discussed above (country, institution or material) can likewise be searched; allowing cases such as searching document descriptions within a given institution, or searching child items within a single collection.

Where a specific item type is the target of a user's query, search results can be further narrowed via the application of facets to filter the data. Facets consist of coarse categories into which data can be clustered; such as language, source and level of detail and provide an important narrowing mechanism where a textual search query results in an overly broad set of matches. An additional way for users to refine their search queries is by adding a field constraint, restricting matching to those where the query applies to a specific part of the target records.

5.2 Virtual Research Environment

One of the findings of EHRI's user requirements investigation was that Holocaust researchers originate from diverse backgrounds and possess a correspondingly wide set of working methods. The VRE features of the EHRI portal are therefore designed to complement a user's exploration and research practices without attempting to impose upon them a particular work-flow or set of tools. As discussed in Section 2, we do not claim to develop a generic VRE but one that is designed for our particular communities of practice; following the principles developed by Anderson and Blanke [2012].

The VRE's services are available to all users who have opted to create a free account. Once they have created an account a user can opt to provide a photograph and details such as their location, interests and areas of research. These details are available to other registered EHRI users and are searchable and browsable. While a user's e-mail address is not publicly visible, an account by default enables a user to be contacted by other registered users via a messaging form.

The first VRE service intended to facilitate keeping track of relevant material in the EHRI portal is the ability for registered users to create notes on collection materials. Notes can be created on both item descriptions and individual fields within a description and are visible to that user both in the portal's

browsing interface and on the user's personalised notes page, where they can also be searched and exported in JSON, CSV or plain text format. Notes are, by default, private to the user who created them and invisible to others. If a user wishes to make a note publicly visible, she can indicate this and add it to a moderation queue. Members of EHRI's moderation group can then vote on notes submitted for publication. Given the sensitivity surrounding Holocaust research avoiding the publication of inappropriate user-generated content is of utmost importance, and only notes with a positive voting score become publicly visible.

Another VRE service for keeping track of material is the ability to "watch" items. In addition to adding the item to a searchable list that persists across browser sessions, watching also means that updates to the chosen items such as edits and the addition of public notes made by other users will appear in the user's activity stream. The items in a user's watch list can be exported in the same manner as notes. It is finally possible to follow other users and be followed in return. Following a user means subscribing to her activity in the portal's personalised activity stream, adding to the time-line notices when she adds material to her watch list or creates publicly visible notes. It is intended, therefore, that the portal's personalised activity stream serves to facilitate serendipitous discovery by alerting users to the existence of material that is of interest to others.

All these are services, common to many Web 2.0 applications, complement the above-described standard functionality of the portal to provide access to content and display it. We are, however, also working on tools to give researchers the ability to define their own view on content by providing end-user access to functionality currently only available to EHRI's data curators. As touched on in Section 4, the heterogeneity of archival data (and Holocaust-related archival material in particular) presents significant challenges in the development of a coherent integrated portal. While a primary goal of the portal is to represent as accurately as possible the material available across numerous dispersed archives, the complex composition and provenance of this material is itself a major barrier to the undertaking of research in the field. The use of *virtual collections* in the EHRI portal was therefore developed as a secondary and more flexible approach to representing archival material in a browsable and searchable form [Bryant et al. 2014]. Virtual collections have two primary use cases within the EHRI portal: to represent physically-dispersed collections of shared provenance and to represent thematic collections of mixed provenance. In both instances, the virtual collection can contain additional structure and descriptive information that ties together, extends, and lends context to the material sourced directly from partner archives. In this way, virtual collections can play a role in how users curate their own content.

6.   CONTENT CURATION AND ONLINE STORIES

EHRI has concentrated on advanced techniques for content curation of archival material. While there has been much progress in the area of digitisation and creation of digital finding aids, the methods and practices of presentation of archival material on the Internet are so far limited, and, as a result, many digital archives are difficult to use. In 2012, EHRI organised the very successful Public History of the Holocaust conference [EHRI Public History 2016] in Berlin; with funding coming directly from the German Ministry of Research and Education. One of the findings was that Holocaust historians and archivists often lack the tools and training to engage in the public history mission and to create easy-to-use and powerful online representations.

Current approaches to digitisation of archives often do not allow easy access and release of digital material.This failure has in turn been invoked by conference participants as evidence that digital databases are not relevant to research communities. EHRI's work on advanced content curation assists both professional and non-professional research audiences to familiarise themselves with the typical types of documents frequently encountered in Holocaust research. Archives have known content curation practices for a long time under the name of research guides [Valenza et al. 2014]. They provide narratives contextualising and integrating material from various CHIs on a particular topic such as the Terezín concentration camp. The new methods of content curation and presentation tested by EHRI are part of a broader process of rethinking the role of

Fig. 2.   EHRI Research Guide on Terezín

the archive as an infrastructure, and of archivists who want to open up the material they store, conserve and protect. We explore ways to build upon the rich data available in the EHRI portal and to provide researchers of all types and levels of expertise with new tools and approaches. Whereas a large proportion of the records in the EHRI portal focus on higher level descriptions (predominantly that of entire collections), the research guides seek to highlight, contextualise and connect the digital material itself.

6.1   Research Guides

The two research guides developed in the first phase of EHRI were meant to overcome the high degree of fragmentation of Holocaust-related archival material on a document level and to promote comparative Holocaust research. Collating together detailed descriptions of archival material from several countries, however, brought significant challenges. The heterogeneity we encountered for material in other parts of the portal, was exacerbated here by the fact that we wanted to integrate the actual material and not just its descriptions.

The Terezín (Theresienstadt) Research Guide [EHRI 2016c] connected the largest archival collections (in Yad Vashem, the Terezín Memorial, Beit Theresienstadt and in the Jewish Museum in Prague) documenting the history of this ghetto into which Jews from the Protectorate Bohemia and Moravia and from other European countries were deported. More than thirty thousand died there due to the horrible living conditions while most of other prisoners were murdered following their deportation to extermination camps in the East. The main aim of the Terezín Research Guide was to help researchers overcome the fragmentation of the surviving fraction of original documents from the war-time and of the post-war accounts. The guide was built by developing a story and connecting it to more than 20,000 descriptions in the EHRI portal. Linking descriptions to controlled vocabularies not only enabled searching using faceted browsing, but also provided

the skeleton of new integrated finding aids listing Terezín archival material from different archives together. A thesaurus of key concepts was used to construct a thematic guide through the documents. The Terezín "Council of Elders" (or "Jewish council") helped build a finding aid organised around the structure of the Terezín Jewish leadership subordinated to the SS command. Finally, we used the most complete database of Terezín prisoners developed by the Terezín Initiative Institute in Prague as an authority set; thus linking the archival material to almost ten thousand prisoners.

Detailed geo-coding made it possible to place the Terezín documents on the map of the former ghetto (see Figure 2). Geo-coding enables the location of Holocaust related events in space, by virtually mapping them onto sites of memory. An authoritative list of locations makes it, for instance, possible to connect locations in a ghetto with documents in the archives. EHRI partners created a comprehensive gazetteer of locations in the concentration camp that describes barracks and crematorium, etc. The research guide links these locations to the archival evidence supporting them.

A second research guide deploys a similar methodology and combines information about collections of the "Jewish councils" in Vienna, Prague and Budapest [EHRI 2016b] with the goal of fostering comparative research into this significant subject. In particular, it highlights the dramatic archival history of these collections which all carry traces of war-time destruction and were collected, combined, moved, constructed and reconstructed after the war. In the future, EHRI plans to extend this research guide with information about further collections from different countries, thereby enhancing its potential for comparative research.

6.2 EHRI Document Blog

The EHRI Document Blog [EHRI 2016a] reacts to the challenge of digital archives by looking at the archival document itself and its history, archival and other contexts and readings. Launched in January 2016, the blog is developing quickly into a space to share ideas about Holocaust-related documents as well as their presentation and interpretation using digital tools.[7] EHRI partners and fellows, amongst others, have the opportunity to document their activities and experiment with different ways to explain and show digital archival content. Using visualisation and storytelling tools, the EHRI bloggers can, for instance, plot the text of a particular document on a map allowing the reader to follow the narrative in space. Another blog post, for instance, visually annotates a page of a death certificate from Terezín and allows users to understand and interpret a set of more than 20 thousand such documents published online.

Such a shared exploration can be the first step to a full-scale scholarly edition. EHRI will create a toolkit to support such scholarly thematic online editions of documents connected to the EHRI portal. During the current second stage of EHRI, two such editions will be prepared: an edition of early Holocaust testimonies focused on comparing their content, form and method and an edition of diplomatic reports about Nazi Germany and the Holocaust (1939-1945), which will complement an earlier edition of reports of foreign diplomats about Nazi Germany [Bajohr and Strupp 2011].

EHRI's content curation work thus produces services to develop presentations based on combinations of narratives with data from the EHRI portal, to geo-code records and to integrate other EHRI services such as thesauri in order to finally publish full-scale scholarly editions.

7. CONCLUSION AND FUTURE WORK

Selecting and integrating dispersed and fragmented archives of relevance to the specialised area of Holocaust research has the potential to stimulate new topics of enquiry and generate new research questions. The establishment of large research infrastructures and their integrating activities will provide a platform on

---

[7]See for instance, Magdalena Sedlická: Testimony of Valerie Straussová, https://blog.ehri-project.eu/2016/03/11/testimony-of-valerie-straussova/ or Wolfgang Schellenbacher: Death Certificate of Gabriel Frankl from the Terezín Ghetto, https://blog.ehri-project.eu/2016/02/18/death-certificate-of-gabriel-frankl-from-the-terezin-ghetto/.

which researchers can follow an investigative trail across archives, connect and link dispersed material and discover new material currently hidden away. The portal work is a key component of this, as we have shown. We have demonstrated how multifaceted the portal work of a research infrastructure is.

For the portal of the research infrastructure to be effective, it needs to bring together the expertise of archivists who understand the collection requirements, researchers who can identify hidden archives and collections and finally digital humanities experts who can develop a virtual bridge connecting the material. The EHRI portal includes all this expertise and has been the focus point of our work in the first phase of the project.

In the second phase of EHRI, we plan to pay closer attention to the trustworthiness and long-term availability of the institutions we connect through the portal. We are thus developing services to transform all Holocaust CHIs into trusted digital repositories and ensure that EHRI can be sustained as a decentralised infrastructure. The goal is to keep the institutions' digital material trustworthy and usable in the long run. To this end, we are developing data management strategies and services around the way the EHRI information is used by researchers.

We thus strive to implement policies that strengthen the sustainability of all parts of the infrastructure but in particular of the Holocaust related digital content. Currently, the data storage facilities of all CHIs are assessed in order to gain and share expertise on the way the EHRI information might be archived and accessed in the long term. The repository certification framework used by the EHRI project are the criteria of the Data Seal of Approval (DSA) [Harmsen 2008]. At the end of EHRI's second phase, we aim to have a governance structure in place that will ensure the reliability and sustainability of all Holocaust material held within EHRI.

Thus, for the second phase of EHRI we would like to complement the centralised portal structure that we have concentrated on in the first phase with a decentralised data service that helps surface the valuable material its partners hold. This should lead to increased levels of collaboration across Europe and beyond as well as, ultimately, to more integrated research on the Holocaust.

REFERENCES

Sheila Anderson and Tobias Blanke. 2012. Taking the long view: from e-science humanities to humanities digital ecosystems. *Historical Social Research/Historische Sozialforschung* (2012), 147–164.

Sheila Anderson and Tobias Blanke. 2015. Infrastructure as intermeditation–from archives to research infrastructures. *Journal of Documentation* 71, 6 (2015), 1183–1202.

Frank Bajohr and Christoph Strupp. 2011. *Fremde Blicke auf das Dritte Reich: Berichte ausländischer Diplomaten über Herrschaft und Gesellschaft in Deutschland, 1933-1945*. Vol. 49. Wallstein Verlag GmbH.

D. Barber. 2006. *Portal Building*. Lulu Enterprises Incorporated.

Tobias Blanke, Michael Bryant, and Mark Hedges. 2013. Back to our data—experiments with nosql technologies in the humanities. In *Big Data, 2013 IEEE International Conference on*. IEEE, 17–20.

Tobias Blanke and Mark Hedges. 2008. Providing linked-up access to Cultural Heritage Data. In *Proceedings of the ECDL 2008 Workshop on Information Access to Cultural Heritage. Aarhus, Denmark*.

Tobias Blanke and Conny Kristel. 2013. Integrating holocaust research. *International Journal of Humanities and Arts Computing* 7, 1-2 (2013), 41–57.

Mike Bryant, Linda Reijnhoudt, Reto Speck, Thibault Clerice, and Tobias Blanke. 2014. The EHRI Project-Virtual Collections Revisited. In *Social Informatics*. Springer, 294–303.

Jessica Bushey. 2012. ICA-AtoM: open-source software for archival description. *Archivi & Computer article* (2012), 1–16.

Leonardo Candela, Donatella Castelli, and Pasquale Pagano. 2013. Virtual research environments: an overview and a research agenda. *Data Science Journal* 12, 0 (2013), GRDI75–GRDI81.

Alexandros Chortaras, Nasos Drosopoulos, Ilianna Kollia, and Nikolaos Simou. 2014. Web Reasoning for Cultural Heritage. In *Web Reasoning and Rule Systems*. Springer, 23–28.

Claims Conference. 2016. Claims Conference Archives. (2016). http://www.claimscon.org/archivist_forum/archive_search.asp

Terry Cook. 2000. Beyond the screen: the records continuum and archival cultural heritage. *Beyond the Screen: Capturing Corporate and Social Memory* (2000), 9.

Marin Dimitrov, Alex Simov, and Yavor Petkov. 2015. Low-cost Open Data As-a-Service in the Cloud.. In *ESWC Developers Workshop*. 35–40.

Wendy M Duff and Verne Harris. 2002. Stories and names: archival description as narrating records and constructing meanings. *Archival Science* 2, 3-4 (2002), 263–285.

Wendy M Duff and Catherine A Johnson. 2002. Accidentally found on purpose: Information-seeking behavior of historians in archives. *The Library Quarterly* (2002), 472–496.

EHRI. 2016a. EHRI Document Blog. (2016). https://blog.ehri-project.eu/

EHRI. 2016b. EHRI Jewish Councils Research Guide. (2016). https://portal.ehri-project.eu/guides/jewishcouncil

EHRI. 2016c. EHRI Terezín Research Guide. (2016). https://portal.ehri-project.eu/guides/terezin

EHRI. 2016d. European Holocaust Research Infrastructure. (2016). https://portal.ehri-project.eu

EHRI Public History. 2016. European Holocaust Research Infrastructure Public History Conference. (2016). http://www.ehri-project.eu/public-history-holocaust

John Feather. 2006. Managing the documentary heritage: Issues from the present and future. *Preservation management for libraries, archives and museums* (2006), 1–18.

Joseph M. Firestone. 1999. Defining the Enterprise Information Portal. (1999).

JH Harmsen. 2008. Data Seal of Approval, assessment and reviewing of research data repositories. In *Proceedings of The Fifth International Conference on Preservation of Digital Objects*. British Library.

Bernhard Haslhofer, Simeon Warner, Carl Lagoze, Martin Klein, Robert Sanderson, Michael L Nelson, and Herbert Van de Sompel. 2013. ResourceSync: leveraging sitemaps for resource synchronization. In *Proceedings of the 22nd international conference on World Wide Web companion*. International World Wide Web Conferences Steering Committee, 11–14.

Mark Hedges, Heike Neuroth, Kathleen M Smith, Tobias Blanke, Laurent Romary, Marc Küster, and Malcolm Illingworth. 2013. Textgrid, textvre, and dariah: Sustainability of infrastructures for textual scholarship. *Journal of the Text Encoding Initiative* 5 (2013).

Antoine Isaac, Robina Clayphan, and Bernhard Haslhofer. 2012. Europeana: Moving to linked open data. *Information Standards Quarterly* 24, 2/3 (2012).

Pramod J.Sadalage and Martin Fowler. 2013. Distilled, NoSQL — A Brief Guide to the Emerging World of Polyglot Persistence. (2013).

Tomi Kauppinen, Jari Väätäinen, and Eero Hyvönen. 2008. *Creating and using geospatial ontology time series in a semantic cultural heritage portal*. Springer.

Primo Levi. 1988. *The drowned and the saved*. Vintage.

Petra Links, Reto Speck, and Veerle Vanden Daelen. 2015. Who holds the key to Holocaust-related sources? Authorship as subjectivity in finding aids. *Holocaust Studies* (2015), 1–23.

Eetu Mäkelä, Eero Hyvönen, and Tuukka Ruotsalo. 2012. How to deal with massively heterogeneous cultural heritage data–lessons learned in CultureSampo. *Semantic Web* 3, 1 (2012), 85–109.

Memorial de la Shoah. 2016. Memorial de la Shoah Archives Guide. (2016). http://www.memorialdelashoah. org/index.php/en/archives-and-documentation/guides-to-archives/international-guide-to-shoah-archives

Muhammad S Sarwar, Thomas Doherty, John Watt, and Richard O Sinnott. 2013. Towards a virtual research environment for language and literature researchers. *Future Generation Computer Systems* 29, 2 (2013), 549–559.

Guus Schreiber, Alia Amin, Lora Aroyo, Mark van Assem, Victor de Boer, Lynda Hardman, Michiel Hildebrand, Borys Omelayenko, Jacco van Osenbruggen, Anna Tordai, and others. 2008. Semantic annotation and search of cultural-heritage collections: The MultimediaN E-Culture demonstrator. *Web Semantics: Science, Services and Agents on the World Wide Web* 6, 4 (2008), 243–249.

Reto Speck and Petra Links. 2013. The missing voice: archivists and infrastructures for humanities research. *International Journal of Humanities and Arts Computing* 7, 1-2 (2013), 128–146.

Go Sugimoto. 2013. Archives Portal Europe: A New Frontier in Digital Space. *Uncommon Culture* 3, 5/6 (2013), 93–101.

USHMM. 2016. United States Holocaust Memorial Museum. (2016). https://www.ushmm.org/

Joyce Kasman Valenza, Brenda L Boyer, and Della Curtis. 2014. Curation in academic libraries. *Library Technology Reports* 50, 7 (2014), 45–50.

Veerle Vanden Daelen, Jennifer Edmond, Petra Links, Mike Priddy, Linda Reijnhoudt, Václav Tollar, and Annelies Van Nispen. 2015. Sustainable Digital Publishing of Archival Catalogues of Twentieth-Century History Archives. In *"Open History: Sustainable digital publishing of archival catalogues of twentieth-century history archives"*.

Yad Vashem. 2016. Yad Vashem — The Holocaust Martyrs' and Heroes' Remembrance Authority. (2016). http://www.yadvashem.org/