Introduction

Data Sharing, Holocaust Documentation and the Digital Humanities: Best Practices, Case Studies and Benefits

Laura Brazzo and Kepa J. Rodriguez

Abstract. This issue of Umanistica Digitale is dedicated to he workshop *"Data Sharing, Holocaust Documentation and the Digital Humanities: Best Practices, Case Studies and Benefits"*, that took place at the Università Cà Foscari in Venice on June 29–30, 2017, in the framework of the EHRI (European Holocaust Research Infrastructure) Project. The workshop was organized by the Centro di Documentazione Ebraica Contemporanea (CDEC), with the support of the Cà Foscari University's Master in Digital Humanities Program. The workshop coincided with the 4th edition of the LODLAM (Linked Open Data in Libraries, Archives and Museums) Summit. The aim of the workshop was to present the state of the art of data sharing practices and technologies, starting from the experiences and results obtained at the EHRI project; to discuss the usability and potential of data sharing in the Humanities; to investigate the possible connections between the EHRI project and other research infrastructures and digital humanities projects.

Il presente numero prende il titolo dal workshop "Data Sharing Holocaust Documentation and the Digital Humanities: Best Practices, Case Studies and Benefits" svoltosi nell'ambito del progetto EHRI il 29-30 giugno 2017. Il workshop è stato organizzato dalla Fondazione Centro di Documentazione Ebraica Contemporanea (CDEC), con il supporto del Master in Digital Humanities Program dell'Università Cà Foscari di Venezia. Scopo del workshop è stato quello di presentare lo stato dell'arte sulle pratiche e le tecnologie per la condivisione dei dati sullo specifico tema della storia della Shoah. Si è discusso delle opportunità e dei benefici derivanti dalla condivisione dei dati, a partire dalle esperienze e dai risultati ottenuti nell'ambito del progetto EHRI e delle possibilità di connessione fra EHRI ed altre infrastrutture di ricerca e progetti nel campo delle Digital Humanities. Gli articoli che presentiamo in questo numero speciale di Umanistica Digitale rispecchiano, e in alcuni casi integrano, i lavori presentati in sede di Workshop. Il numero include anche le tre presentazioni che hanno aperto le giornate di lavoro, ovvero i paper dedicati al progetto di digitalizzazione dei documenti sulla Shoah conservati dall'Archivio di Stato di Venezia e l'introduzione generale del progetto EHRI. I curatori desiderano ringraziare la rivista Umanistica Digitale e il suo comitato di redazione per aver accolto la proposta di pubblicazione; un particolare ringraziamento va a Marilena D'Aquino per il sollecito e costante supporto all'editing dei testi.

Preface

The workshop *"Data Sharing, Holocaust Documentation and the Digital Humanities: Best Practices, Case Studies and Benefits"* took place at the Università Cà Foscari in Venice on June 29–30, 2017, in the framework of the EHRI (European Holocaust Research Infrastructure) Project.[1]

The workshop was organized by the Centro di Documentazione Ebraica Contemporanea (CDEC), with the support of the Cà Foscari University's Master in Digital Humanities Program. The workshop coincided with the 4th edition of the LODLAM (Linked Open Data in Libraries, Archives and Museums) Summit.[2]

The aim of the workshop was to present the state of the art of data sharing practices and technologies, starting from the experiences and results obtained at the EHRI project; to discuss the usability and potential of data sharing in the Humanities; to investigate the possible connections between the EHRI project and other research infrastructures and digital humanities projects.

A Program Committee of the workshop was joined by four experts in fields related with data sharing and digital humanities:

- Laura Brazzo (PhD), Chairperson, Centro de Documentazione Ebraica Contemporanea (CDEC)[3], Milan, Italy

- Vladimir Alexiev (PhD), Ontotext Corp,[4] Sofia, Bulgaria

- Silvia Mazzini (Dr), regesta.exe,[5] Rome, Italy

- Kepa J. Rodriguez (PhD), Yad Vashem,[6] Jerusalem, Israel

21 submissions were collected, from which the program committee accepted 14 for presentation and discussion at the Workshop. Eleven of the presented submissions are represented by articles in this journal.

The selection process of the submissions showed to us a very heterogeneous landscape in conceptions and degree of development concerning data sharing. In some cases the focus of the projects was the online presentation of content of the databases; in other cases strategies to share the data among different agents in a more efficient way or strategies to integrate shared data in the own infrastructure.

The workshop was organized in four work sessions and opened with a general overview of the EHRI Project (*Veerle Vanden Daelen*) as well as a presentation of the recent State Archive of

---

1   http://www.ehri-project.eu
2   https://summit2017.lodlam.net/
3   http://www.cdec.it
4   https://ontotext.com/
5   https://www.regesta.com
6   http://www.yadvashem.org

Venice's project to digitize Holocaust materials (*Raffaele Santoro* and *Andrea Pelizza*).

The four sessions progressed from presentations of projects based on data integration (with potential for data sharing) to projects with an advanced data-sharing framework.

The three presentations scheduled for Session I were focused on different cases of data integration made at the most relevant institutions for Holocaust materials, the United States Holocaust Memorial Museum, in Washington (*Michael Levy* and *Megan Lewis)* and Yad Vashem in Jerusalem(*Olga Tolokonski*).

Session II was devoted to oral testimonies and audio-visual materials. Case studies on integration of materials and strategies for sharing were provided by representatives of the Fortunoff Video Archive for Holocaust Testimonies at Yale University (*Stefan Naron – Kevin Glick*), the Center for Digital Systems at Freie Universität Berlin (*Cord Pagenstecher*) and the USC Shoah Foundation's Visual History Archive (*Marta Stroud*).

Session III showcased two on-going digital humanities projects. The first project presented (*Paris C. Papamichos – Giorgos Antoniou*) is based on Digital Archive of the Greek Shoah and aimed to reconstruct the Greek Holocaust survivors' social networks; the second presentation concerned the Nuremberg Trials Project - a digital document migration case study carried out by the Harvard University (*Lidia Santarelli*).

In Session IV the scheduled presentations showed five advanced cases of data sharing and use of semantic approaches and technologies: the EHRI Project (*Vladimir Alexiev – Ivelina Nikolova – Nevi Hatela*); the JudaicaLink Project (*Kay Eckert – Maral Dadvar*); the Holocaust and WW2 LOD developments in the Netherlands (*Annelies Van Nispen – Lizzy Jongma*); and the LOD Navigator (*Giovanni Moretti - Rachele Sprugnoli – Sara Tonelli*). The presentation of some innovative ways to work with the archival standard (*Laurent Romary – Charles Riondet*) was also included in this final session.

Given the heterogeneity of the participants and their common interest in advanced methodologies to enable efficient data sharing, we a included in the Program a special panel devoted to Linked Open Data (LOD). Contributions were provided by two experts coming from the LODLAM Summit. Alessio Melandri (Synapta) provided an overview of LOD and queries SPARQL; Simon Cobb (Leeds University) introduced Wikidata and its potential for integration and indexing of cultural heritage information.

Organizers and Program Committee

Organizing committee:

Laura Brazzo, Fondazione CDEC, Milan;   Matteo Perissinotto, University of Trieste – Fondazione CDEC, Milan; Simon Levis Sullam, Università Cà Foscari, Venice

Program Committee:

Laura Brazzo (PhD), Chairperson, Fondazione CDEC, Milan; Vladimir Alexiev (PhD), Ontotext Corp; Sofia Silvia Mazzini (Dr), Regesta.exe, Rome; Kepa J. Rodriguez (PhD), Yad Vashem, Jerusalem.

Program

Thursday, 29th June 2017 9.00 Registration

9.30 Greetings Simon Levis Sullam (Università Cà Foscari)

Introduction

9.45 Raffaele Santoro - Andrea Pelizza (State Archive of Venice)

The archival series of Venetian Prefettura and Questura – Archivio di Stato di Venezia

10.15 Veerle Vanden Daelen (Kazerne Dossin)

Data Sharing, Holocaust Documentation and the Digital Humanities: Introducing the European Holocaust Research Infrastructure (EHRI)

11.15 Coffee break

Session I

11.30 Michael Levy (United States Holocaust Memorial Museum - USHMM)

Sharing Collections Data: One Perspective on the Practice at USHMM

12.00 Megan Lewis (United States Holocaust Memorial Museum - USHMM)

Using Names Lists for Social Network Analysis

12.30 Olga Tolokonsky (Yad Vashem)

Integration of Heterogeneous Shared Data: Yad Vashem's Perspective as a Data Aggregator

13.00 Light Lunch

Session II

14.00 Stephen Naron, Fortunoff Video Archive for Holocaust Testimonies, Yale University Library, New Haven *Fortunoff Video Archive for Holocaust Survivors Testimonies*

14.30 Cord Pagenstecher (Center for Digital Systems, Freie Universität Berlin) *Audiovisual Testimony Collections. Digital Archiving and Retrieval*

15.00 Martha Stroud (USC Shoah Foundation Center for Advanced Genocide Research)

Digital Testimonies and Research Access: USC Shoah Foundation Visual History Archive

15.30 Coffee break Discussion: *Semantic Archive, Data and Media Integration* 16.45
Conclusions

Friday, 30th June 2017

LOD TUTORIAL

9.00 Alessio Melandri (Synapta)

Getting New Knowledge from the Web of Data: LOD & SPARQL

10.00 Simon Cobb (Leeds University Library)

*An Introduction to Wikidata for Sharing and Visualising Cultural Heritage Resources 11.00 Coffee
break*

Session III

11.15 Paris Papamichos Chronakis (University of Illinois at Chicago)
*From Individual Survival to Social Networks of Survivors: Rethinking the Digital Archive of the
Greek Shoah*

11.45 Lidia Santarelli (Princeton University)

*Life and Death of the Archive. The "Nuremberg Trials Project": A Case Study in Digital Document
Migration*

Discussion: *Holocaust Research Questions and Use Cases Enabled by Data Sharing 12.45 Light
Lunch*

Session IV

14.00 Vladimir Alexiev (Ontotext)

Semantic Archive Integration for Holocaust Research: the EHRI Research Infrastructure

14.30 Charles Riondet (INRIA Research Center of Paris)

Towards Multiscale Archival Digital Data

15.00 Coffee break

15.15 Kai Eckert (Stuttgart Media University)

Judaica Link: A Knowledge Base for Jewish Culture and History

15.45 Annelies Van Nispen - Lizzy Jongma (Institute for War, Holocaust and Genocide Studies
- NIOD) *Holocaust and WW2 LOD Developments in the Netherlands*

16.15 Rachele Sprugnoli (Bruno Kessler Foundation) *LOD Navigator: Tracing Movements of Italian Shoah Victims*

Discussion: *Global Archive and Data Integration - Enlarging EHRI Data Connections to other Datasets and Projects*

16.45 Conclusions of the Workshop

Table of Contents

Preface Invited Talk

Full Papers