AI in Archival Holdings and Digital History: Interrelational, Dialogical and Cooperative Interdisciplinary Partnerships

Pepita Raventós Pajares
Archives and Records Management
University of Lleida
Lleida. Spain
pepita.raventos@udl.cat

Meritxell Simon Martin
Pedagogy Department
University of Lleida
Lleida. Spain
meritxell.simon@udl.cat

Josep Ballber Botey
PhD candidate in Engineering & IT
University of Lleida
Lleida. Spain
josep.ballber@udl.cat

Celio Hernández Tornero
Archives and Records Management
University of Lleida
Lleida. Spain
celio.hernandez@udl.cat

Roberto García González
Associate Professor in Computer Science
University of Lleida
Lleida. Spain
roberto.garcia@udl.cat

Glòria Jové Monclús
Senior Lecturer in Teacher Education
University of Lleida
Lleida. Spain
gloria.jove@udl.cat

***Abstract*—This poster will map out the outcome of the interrelational, dialogical, and cooperative Digital Humanities partnership that is mutually enhancing not only the investigation carried out by the History of Education team but also the agenda of the Lleida University Archive and Records Management Service applying Semantic Technology.**

Keywords—Computational Archival Science, Archival holdings, Education History, Interoperability, Semantic Technology

Poster explanation

In a context of 21st century knowledge societies, archives are today very eager to provide digital data finding aids. Mass digitization of holdings is the precondition to do so. The Archive and Records Management Service of the University of Lleida (Spain) is currently working towards these goals: to digitize its holdings and to explore the functionalities of different digital tools, including AI, in order to offer a performative finding aid.

The case study is based on The Lleida Teacher Education College Holding (1841 – 1950), the oldest at the Universitat of Lleida and held at the Archive and Records Management Service, Lleida University. The Lleida Teacher Education College was a state-funded educational institution established in the mid-nineteenth century to train primary school teachers. It is today our Faculty of Education, Psychology and Social Work in Lleida University, which offers teacher education programs.

Seeking to promote this holding, we are currently carrying out a Digital Humanities project, intitled "The Lleida Teacher Education College during the Second Spanish Republic (1931-1936)", which consists of 3 transdisciplinary objectives:

1. Archival Studies: to clean, restore, classify, digitize, describe and make available to users the Lleida Teacher Education College Holding (1841-1950). This objective involves the use of technology of Handwritten Text Recognition (HTR) [1] and Computational resources of Semantic Technology, for the (A) processing of records and large-scale data, (B) the analysis, description and storage for long-term preservation, and (C) the accessibility of this Holding. This aim is conducted within the framework of a Management System for Records informed by ISO 30301.

2. History of Education: to study a part of this Holding: the Lleida Teacher Training College during the Second Spanish Republic (1931-1936). Education was indeed one of the priorities of the Second Spanish Republic in Spain. It was to be public, co-educational, active, creative, aimed at children, but also at adults and, in short, a driving force for social change. To carry out this project, the Provisional Republican Government built numerous new schools and refurbished the old ones. At the same time, it devised a new way of recruiting civil servant teachers: it introduced new public examinations to replace the presumably outdated rote-learning ones. This aim seeks to tease out, from a gender perspective [2], the materialization of these new public examinations in the case of the Lleida Teacher Education College.

3. Teacher Education: to transform the archival and historical outcome into a knowledge-transfer scheme intended for the educational community, especially teachers-in-training [3].

It is within the framework of this Digital Humanities project that we are conducting our goal: the digitization of our holdings and the design of a performative finding aid. In a previous case study, entitled "Building Interoperability in Archival Management Systems, Between Data and Records Long-term Digital Preservation for Joint Research. Challenges in a Faster Changing World" [4], we carried out the first step to do so. Taking the student assessment records of our university (managed by the Archives and Records Management Service), we designed an interoperable system for transferring digital objects and uploading them into the system of records and digital archiving applications of our university.

This interoperable system permits us now to offer authentic, reliable, integrated, and usable information. In order to transfer information from a digital archiving application to another, our interoperable system has to capture the values of the metadata and add those elements that are necessary and missing. In our study case, we find different profiles of end users with different enquiry needs. By way of example, our

users in objectives 2 and 3 in our Digital Humanities project demand different functionalities in their exploitation of the digital data of the holding.

In order to cater for them, our next step is to design a new way of describing the records by means of establishing links in a less hierarchical way than before, applying the entities model of ISO 23081, which establishes relationships between entities (e.g., records, agents, activities, mandates and relations). To set up the manner these links are established [5], we opt for applying semantic technology. In doing so, we design our records by defining them according to RDF. This requires an ontology. In our case, it is a Records in Context ontology: OWL.

Finally, to explore the resulting knowledge graph describing the entities of interest using semantic technologies, we use the Rhizomer [6] tool for interactive semantic knowledge graphs exploration. This tool provides and overview of the resulting graph, mechanisms for filtering based on facets and views to explore the details of each of the modelled entities.

ACKNOWLEDGMENT

We would like to thank the company Piql Iberia for their collaboration in the project.

REFERENCES

[1] P. Raventós, Hernández, C., M. Simon. “AI and Archive Handwritten Text Recognition Applied to Patrimonial Holdings: An Example of 10 Diaries Written by Spanish Republican Teachers in 1932”. AI collaboratory. Osaka: 2022.

https://ai-collaboratory.net/wp-content/uploads/2022/12/Pepita_Raventos-Pajares.pdf

P. Raventós, C. Hernández, M. Simon. “IA y archivo. La tecnología de reconocimiento de textos manuscritos en fondos patrimoniales: un ejemplo de 10 diarios en la formación de maestras y maestros en el año 1932“, Tábula, Number 25, pp. 83-98

https://doi.org/10.51598/tab.937

M.Simon, G. Jové. ‘Life-writing as an innovative learning method in teacher-education during the first biennium of the Second Spanish Republic (1931-1933)’. ”IABA 2023 “Life-Writing in Times of Crisis”, University of Warsaw, Poland, 2023

M. Simon. ‘History of Education as Digital History: the use of Handwritten Text Recognition (HTR) technologies in the study of the Normal School of Lleida during the Second Spanish Republic (1931-1936)’, “International Centre for Historical Research in Education (ICHRE) Summer Conference”, Institute of Education, University College London, United Kingdom, 2023.

[2] M. Simon, G. Jové, H. Murua, M. A. Goicoetxea. ‘The Contribution of Female Educationists to the New Public Examinations Designed by the Provisional Republican Government in Spain (1931-1933)’. “ISCHE 44 Histories of Education and Reform: Traditions, Tensions and Transitions”, University of Budapest, Hungary, 2023

[3] M. Simon, G. Jové, T. Segarra. ‘History of Education Meets Teacher Education: How the Educational Project of the Second Spanish Republic Inspires Innovative Teaching Today ’. ”ISCHE 44 Histories of Education and Reform: Traditions, Tensions and Transitions”, University of Budapest, Hungary, 2023.

[4] J. Ballber, P. Raventós., R. García. “Building Interoperability in Archival Management Systems, Between Data and Records Long-term Digital Preservation for Joint Research. Challenges in a Faster Changing World”. International Coincil on Archives. Rome, 2022.

https://www.ica.org/sites/default/files/a07_-_438_ballber_building_interoperability_in_archival_management_systems_841_x_1189_mm_6_compressed.pdf

[5] W. Duff, J. Haskell. “New Uses for Old Records: A Rhizomatic Approach to Archival Access”. The American Archivist. Vol. 78:1. Spring/Summer, 2015, pp.38-58.

https://americanarchivist.org/doi/pdf/10.17723/0360-9081.78.1.38

[6] R. García, J.-M. López-Gil, R. Gil, “Rhizomer: Interactive semantic knowledge graphs exploration”. SoftwareX. Vol. 20, 2022, p. 101235.

https://doi.org/10.1016/j.softx.2022.101235