
Integrating public and private archives on an international level

Kerstin Arnold*¹ and Marta Musso²

¹Archives Portal Europe Foundation – Germany

²Archives Portal Europe Foundation – United Kingdom

Abstract

Archives Portal Europe: from ingestion to automated topic detection

Archives Portal Europe (www.archivesportaleurope.net) is an online repository bringing together descriptive metadata from archival institutions from more than 30 countries, in more than 20 languages (and 5 different alphabets), and from a variety of institutions: national, regional and local archives, business archives, archives of schools and universities, archives of private persons and families, etc. While the Archives Portal Europe Foundation, established to manage the portal, is mostly controlled by institutional actors, private archives (from museums and galleries to corporate archives) actively collaborate to the initiative by providing their materials for ingestion, and by preparing them according to the same criteria as public institutions, so that all the metadata can be ingested and made searchable by users in Archives Portal Europe with a horizontal approach.

This paper aims to present the challenges and solutions adopted by Archives Portal Europe as a multi-level international aggregator, so that very different archives, from different countries and different traditions, can be merged in the portal for research. The paper will first focus on the portal's use of international standards for interoperability in the ingest, processing and access to such varied material. In this context, selected examples of private archives will showcase the established workflows and processes, which provide a general framework, but still allow for flexibility when required. Secondly, the paper will present some of the functionalities that Archives Portal Europe has worked on especially over the more recent years to facilitate archival research on part of the users, particularly with reference to extending over the usual keyword searches. Two projects will be looked at in order to illustrate this: the first is the possibility of crowdsourcing to users the tagging of relevant documents according to specific topics, which will become available as part of the portal's redesign in May 2022; the second is an automated topic detection tool that enhances and expands the keyword searches conducted by the users semantically in a multilingual environment. The latter is available online, but currently runs on a platform separate from the portal's main user interface.

Through the presentation of the new platform, the paper will explore the problems and opportunities provided by archival collaborations that go over national and international borders, to also embrace institutional diversity, in the creation of what is currently the largest archival repository in the world.

Keywords: Aggregation, Discovery, Automated Topic Detection, Interoperability, Standards

*Speaker