

The use of digital tools for the processing of late medieval papal letters

Adam Zapała (History Institute Polish Academy of Sciences, Cracow)

The medieval papal letters are an unique collection of documents on a European and world scale. Their uniqueness stems from the wide range of issues (religious, political, legal and career-related) they concern, their universal scope (they affected the entire Latin civilisation), and their very good state of preservation. Although most of the original letters have been lost, the majority of papal documents from the late Middle Ages have been preserved as copies in papal chancellery registers. Only from the 15th century, several hundred thousand documents (not including supplications or sources on the economic aspect of exchange) have survived to our times.

Therefore querying, editing and publishing papal material has become one of the most important works of medieval historiography. Since the beginning of the 20th century, projects in many European countries try to make the papal letters available for scholars. The scope and volume of the material resulted in a situation where projects ongoing on for decades are still largely unfinished, and no one has ever ventured to collect the 15th century material in its entirety (Zajchowska 2017). Until recently, such an idea seemed impossible; however, the use of digital tools is changing this perspective. This paper aims to discuss how digital tools are used to compile late medieval papal letters in the project "Monumenta Vaticana res gestas Polonicas Illustrantia" and more generally what opportunities they offer in that field.

In the project, we are working on a comprehensive digital workflow to improve the efficiency of processing papal letters. Our work is carried out on three levels: digitisation and harmonisation of the existing output of historiography, indexing of papal documents, and full-text editions. The first one means primarily the digital processing of printed editions (layout analysis, OCR), automatic extraction of metadata from the text, and its normalisation and the creation of a database (in XML). On the second level, the aim is to obtain basic information from the documents (such as people, places, institutions, types of privileges, types of documents present). For this purpose, image indexing software using GIS technology - INDXR - is used (Borek et al. 2020).

The third and most time-consuming part is the full-text edits. Here as well, the work of historians can be facilitated by digital tools, especially in two areas: the automation of the work and improving its quality. For the former, tests are underway to use artificial intelligence to automatically transcribe and process medieval texts (HTR, NLP). Although the results are still not satisfactory, these solutions have great potential. The latter, the use of

XML/TEI encoding gives great possibilities for analysis and visualisation of texts. Moreover binding tags to external identifiers (Handle.net, Wikibase) allows a objectivity of the work, which gives hope for the harmonisation of different projects in the future. Digital collating of text can speed up work and prevent many mistakes (as a large proportion of papal letters were written according to strictly respected forms), digital tools can also help to automatically date documents, etc.

The author defended his PhD thesis at the Jagiellonian University of Kraków in 2018. After having participated in the projects “Monumenta Vaticana res gestas Polonicas illustrantia” (2013–2018) and “Atlas historyczny Polski XVI wieku” [Historical atlas of Poland in the 16th century] 2018–2020, he became principal investigator of DARIAH-PL – Digital Research Infrastructure for the Arts and Humanities project at the Historical Institute of the Polish Academy of Sciences. His research interests include ecclesiastical history in the late Middle Ages; Polish contacts with the Holy See in the 15th century, prosopography of late medieval Poland and digital humanities.

Bibliography:

A. Borek, T. Związek, M. Słomski, M. Gochna, G. Myrda, M. Słoń, *Technical and methodological foundations of digital indexing of medieval and early modern court books*, “Digital Scholarship in the Humanities”, Vol. 35, No. 2, 2020, p. 233-253.

A. Zajchowska, „*Monumenta Vaticana*” — koncepcje edycji późnośrednio-wiecznych źródeł papieskich. *Przyczyny sukcesów i porażek*, in: *Editiones sine fine*, ed. K. Kopiński, W. Mrozowicz, J. Tandecki, Toruń 2017, p. 143-162.