
Monasterium and ERC "From Digital to Distant Diplomatics" Strategy Workshop - Planning a Call for Digitisation Proposals

Georg Vogeler*¹ and Karl Heinz*²

¹University of Graz – Austria

²ICARus – Austria

Abstract

Monasterium.net has served the archival and research community for the past 20 years. Still the largest portal for mediaeval and early modern charters, it is less dynamic in attracting further charter data than it was 10 years ago. The Workshop invites archivists, archival users, technicians and researchers to consider how this can be changed.

Indeed, the ERC project "From Digital to Distant Diplomatics" (DiDip, <https://didip.eu>) attempts to develop tools that facilitate large-scale diplomatics work based on state-of-the-art machine learning based methods. This will include open source contributions to handwritten text recognition, object detection on the image (e.g., seals, notarial signs), natural language processing (e.g., named entity recognition), indexing and searching by visual features or textual features. The project can rely on the huge data set collected in monasterium.net and will make its results available in a modernised MOM-CA application in the portal.

Machine learning relies on high-quality data to train the appropriate models. To achieve a solid ground for this kind of methods applied to the domain of mediaeval and early modern charters, Monasterium.net provides a good basis. However, it has an obvious bias towards material from Central Europe. To mitigate this, the DiDip project plans to hand out financial grants to support archives in digitising their charter collections and making them available online.

The aim of the workshop is to discuss the best methods how to frame the grant call. We will discuss questions like:

- What kind of digital content concerning mediaeval documents are there to be digitised?
- What are the obstacles to having it online?
- Which methods do you prefer to be applied for the online publication? (e.g., iif, data standards ...)
- What can attract archivists to share data via Monasterium.net?
- Which legal consideration should be addressed in the call?
- What kind of practicalities should be made clear for this call (e.g., costs, scanning process, time frame)?
- What are the long term perspectives of contributions to Monasterium.net?

*Speaker

Keywords: Monasterium.net, Charter, digitisation, archives