

Digital contributions to a 300 years old methodology: Diplomatics & DH

Luger, Daniel

daniel.luger@uni-graz.at
Universität Graz, Austria

Nicolaou, Angelos

angelos.nicolaou@uni-graz.at
Universität Graz, Austria

Decker, Franziska

franziska.decker@uni-graz.at
Universität Graz, Austria

Atzenhofer-Baumgartner, Florian

florian.atzenhofer-baumgartner@uni-graz.at
Universität Graz, Austria

Lamminger, Florian

florian.lamminger@uni-graz.at
Universität Graz, Austria

Vogeler, Georg

georg.vogeler@uni-graz.at
Universität Graz, Austria

Aoun, Sandy

sandy.aoun@uni-graz.at
Universität Graz, Austria

Kovács, Tamás

tamas.kovacs@uni-graz.at
Universität Graz, Austria

Diplomatics research has developed a rich set of features observable in medieval charters which can inform us about the making of a charter, the norms of chancery work and the people producing charters for a prince or an institution.

Typical features of charters are writing material, script, seals or other forms of authentication, chancery marks and other traces of the chancery workflow, ornamental elements such as initials, but also abbreviations, punctuations or signatures. The original method in diplomatics is the search for “Kanzleimäßigkeit”, i.e., the study of a certain chancery practice. The diplomatist tries to define the characteristic features of a charter and find them in other contemporary charters in order to deduce the practice of a so-called chancery and to gain information on their historical context (ranging from forgery detection to the political situation of their creation) (Breßlau 1912, Vogtherr 2008, Kölzer 2011, Lackner 2013, Luger 2020).

However, establishing the “typical charter” in a certain chancery becomes hard when you have to deal with a number of documents beyond the capacity of a single researcher, as is obviously the case in the archives of Central and Southern Europe in the 14th and 15th century (and even earlier in Italy). Detecting features of charters manually is not possible in a large-scale dataset as we would need for later Middle Ages (Hlaváček 2006). Mass digitization makes available large numbers of documents, as for instance Monasterium.net includes more than 660'000 charters, with a special focus on Austria and the surrounding Central and Eastern European countries, including Hungary, Slovenia, Croatia, North Macedonia, Romania, the Czech Republic and Slovakia.

In the ERC-Advanced Grant Project “DiDip – From Digital to Distant Diplomatics” (PI Georg Vogeler) we want to show which charter features are identifiable automatically and test which other features are recognized by the machine that can be useful for Diplomatics research.

One of the newer methods we want to use for analyzing late medieval chanceries is Computer Vision (CV). The recent works of Christlein 2018 and Leipert et al. 2020, for example, have clearly shown that CV can provide quantified stylistic attributes of all graphical features of a charter (e.g. decorations, notarial signs). There are also indications that CV can infer the date of a historical document (Cloppet 2017, Seuret 2021) and can classify the handwriting style from a paleographical perspective.

To test the potential of these methods for our source material, we have built a ground truth of a random selection of 1000 charters from the Monasterium.net corpus and made experiments with object detection.

By reformulating layout analysis as an object detection problem instead of the typical image segmentation approach, we drastically economize the effort for annotating charters. For the annotation process we use a tool developed by Angelos Nicolaou called FRAT.¹ It allows fast drawing of rectangles and labeling them. We have defined a group of visual elements of charters in 11 different classes which will be of interest in our future studies: “Seal”, “Writable Area”, “Ornament”, “Old Text”, “Old Note”, “New Text”, “New Other”, “Fold” (i.e. Plica), “Calibration Card”, “No Class” and “Ignore”. We have been able to annotate the 1000 charters with all these 11 classes within 25 hours. With this data we could train a simple Yolo5 (Jocher 2022) model with significant detection quality in (95% accuracy model at a 50% intersection over union (IoU) threshold), which can already be used as a base classification to find occurrences of the selected features in the documents.

With object detection and other computational methods, we may be able to revolutionize the study of late medieval charters. Possible concrete research questions would be, for example, the influences of public notaries on the authentication of charters in the late Middle Ages (Weileder 2019) made possible by selecting charters with a notarial sign, or the spread of humanistic charter script in the European context and an analysis of the humanistic networks behind it (Luger 2016) made possible by automatic script classification.

Notes

1. <https://github.com/angelos/frat>

Bibliography

Breßlau, Harry (1912): *Handbuch der Urkundenlehre für Deutschland und Italien*. Berlin.

Christlein, Vincent (2018): „Automatic Detection of Illuminated Charters“. In *Illuminierte Urkunden. Beiträge aus Diplomatie, Kunstgeschichte Und Digital Humanities*, herausgegeben von Gabriele Bartz und Markus Gneiß. Archiv für Diplomatik, Schriftgeschichte, Siegel- und Wappenkunde, Beihefte 15. Wien: Böhlau: 45–51.

Id. (2018): „Technical Tools for the Analysis of High Medieval Papal Charters“. In *Papstgeschichte im Digitalen Zeitalter. Neue Zugangsweisen zu einer Kulturgeschichte Europas*, herausgegeben von Klaus Herbers. AfD Beiheft. Wien: Böhlau Verlag GmbH & Cie: 45–53.

Cloppet, Florence / Eglin, Véronique / Helias-Baron, Marlène / Kieu, Cuong / Vincent, Nicole / Stutzmann, Dominique (2017): „ICDAR2017 Competition on the Classification of Medieval Handwritings in Latin Script“. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*: 1371–76.

Hlaváček, Ivan (2006). „Das Problem der Masse: Das Spätmittelalter“. *Archiv für Diplomatik, Schriftgeschichte, Siegel- und Wappenkunde* 52: 371–93.

Jocher, Glenn / Chaurasia, Ayush / Stoken, Alex / Borovec, Jirka / Kwon, Yonghye TaoXie, et al. (2022): *ultralytics/yolov5: v6.1 - TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference (Version v6.1)*. Zenodo. <https://doi.org/10.5281/ZENODO.3908559>.

Kölzer, Theo (2011): „Diplomatics.“ In *Handbook of Medieval Studies - Terms – Methods – Trends*, edited by Albrecht Classen, 405–24. Berlin et al.: De Gruyter.

Lackner, Christian (2013): Möglichkeiten und Perspektiven diplomatischer Forschung: zum Privileg Herzog Albrechts III. für die Universität Wien vom Jahre 1384. Vol. 4. Stabwechsel. Antrittsvorlesungen aus der historisch-kulturwissenschaftlichen Fakultät der Universität Wien. Wien u.a.: Böhlau.

Leipert, Martin / Vogeler, Georg / Seuret, Mathias / Maier, Andreas / Christlein Vincent (2020): „The Notary in the Haystack – Countering Class Imbalance in Document Processing with CNNs“. In *Document Analysis Systems*, herausgegeben von Xiang Bai, Dimosthenis Karatzas, und Daniel Lopresti. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020: 246–61.

Luger, Daniel (2016): *Humanismus und Humanistische Schrift in der Kanzlei Kaiser Friedrichs III. (1440-1493)*. Vol. 60. MIÖG Erg. Bd. Wien u.a.: Böhlau, 2016.

Id. (2020): „Über Sonderformen und Fälschungen von Urkunden Kaiser Friedrichs III.“ In *FS Christian Lackner*, 127–48, 2020.

Seuret, Mathias / Nicolaou, Angelos / Rodríguez-Salas, Dalia / Weichselbaumer, Nikolaus / Stutzmann, Dominique / Mayr, Martin / Maier, Andreas / Christlein, Vincent (2021): „ICDAR 2021 Competition on Historical Document Classification“. In *Document Analysis and Recognition – ICDAR 2021*, ed. by Josep Lladós, Daniel Lopresti, und Seiichi Uchida, Cham: Springer International Publishing: 618–34.

Vogtherr, Thomas (2008): *Urkundenlehre*. Vol. 3. Hahnsche Historische Hilfswissenschaften. Hannover: Hahn'sche Buchhandlung.

Weileder, Magdalena (2019): *Spätmittelalterliche Notarsurkunden: Prokuratorien, beglaubigte Abschriften und Delegatensurkunden aus bayerischen und österreichischen Beständen*. Vol. 18. Archiv für Diplomatik. Beiheft. Köln, Wien: Böhlau.