ACTA HISTORICA TALLINNENSIA

32/2

2026

ACTA HISTORICA TALLINNENSIA

Eesti Teaduste Akadeemia rahvusvaheline ajalooajakiri,
mis ilmub koostöös Tallinna Ülikooliga

International history journal of the Estonian Academy of Sciences
published in collaboration with Tallinn University

PEATOIMETAJA
Marek Tamm *Tallinna Ülikool*

TOIMETAJAD
Liisi Keedus *Tallinna Ülikool*
Mart Kuldkepp *University College London*
Anu Mänd *Tartu Ülikool*
Kristo Nurmis *Tallinna Ülikool*

TOIMETUSKOLLEEGIUM
Karsten Brüggemann *Tallinna Ülikool*
Liisi Esse *Stanfordi Ülikool*
Jörg Hackmann *Szczecini Ülikool*
Andres Kasekamp *Toronto Ülikool*
Jüri Kivimäe *Toronto Ülikool*
Eva Piirimäe *Tartu Ülikool*
Aivar Põldvee *Tallinna Ülikool*
Elina Räsanen *Helsingi Ülikool*
Lars Fredrik Stöcker *Nordost-Institut Hamburgi Ülikoolis*
Tõnu Tannberg *Tartu Ülikool*
Kristjan Toomaspoeg *Salento Ülikool*
Aro Velmet *Lõuna-California Ülikool*
Bradley Woodworth *New Haveni Ülikool*

EDITOR-IN-CHIEF
Marek Tamm *Tallinn University*

EDITORS
Liisi Keedus *Tallinn University*
Mart Kuldkepp *University College London*
Anu Mänd *University of Tartu*
Kristo Nurmis *Tallinn University*

ADVISORY BOARD
Karsten Brüggemann *Tallinn University*
Liisi Esse *Stanford University*
Jörg Hackmann *University of Szczecin*
Andres Kasekamp *Universty of Toronto*
Jüri Kivimäe *University of Toronto*
Eva Piirimäe *University of Tartu*
Aivar Põldvee *Tallinn University*
Elina Räsanen *University of Helsinki*
Lars Fredrik Stöcker *Nordost-Institute at the University of Hamburg*
Tõnu Tannberg *University of Tartu*
Kristjan Toomaspoeg *University of Salento*
Aro Velmet *University of Southern California*
Bradley Woodworth *University of New Haven*

TOIMETUSE AADRESS
Tallinna Ülikool
Humanitaarteaduste instituut
Narva mnt 25, 10120, Tallinn, Eesti
marek.tamm@tlu.ee

EDITORIAL ADDRESS
Tallinn University
School of Humanities
Narva mnt 25, 10120, Tallinn, Estonia
marek.tamm@tlu.ee

Published since 1997
The journal is published biannually

kirj.ee/historica

New Approaches to Medieval Livonia in the Digital Age

Ed. by Marek Tamm

127 *Marek Tamm, Hembo Pagi*

Medieval Studies after the Digital Turn: Methods, Infrastructures, and the Making of Digital Livonia

169 Digipööre keskajauuringutes: meetodid, taristud ja „Digitaalse Liivimaa" kogemus. *Resümee*

173 *Nicolas Bourgeois*

The Historian and Modern Statistical Methods: Applications to Thirteenth-century Livonian Chronicles

202 Ajaloolane ja tänapäevased statistilised meetodid: rakendusi 13. sajandi Liivimaa kroonikate uurimisel. *Resümee*

205 *Rūta Bruževica*

From Manuscript to Visualisation: A Digital Approach to Medieval Livonian Craft Ordinances

233 Käsikirjast visualiseeringuni: digitaalne lähenemine keskaja Liivimaa käsitööliste skraadele. *Resümee*

235 *Anu Mänd*

Digital Methods and the Study of Fragmentary Medieval Grave Slabs

259 Digitaalsed meetodid fragmentaarsete keskaegsete hauaplaatide uurimisel. *Resümee*

263 *Juhan Kreem*

Imaging Medieval Livonian Seals: Techniques and Technologies in the History of Sigillography

277 Kujutise loomise tehnikad keskaja Liivimaa sigillograafia ajaloos. *Resümee*

279 *Tiina Kala, Kaspar Kolk*

The Tallinn *Ex quo* Vocabulary: Textual Tradition and Historical Context

310 „Ex quo" sõnastiku Tallinna käsikiri: tekstitraditsioon ja ajalooline taust. *Resümee*

313 *Kristjan Toomaspoeg*

"We Are from the North": Livonica in the Vatican Archives and the Special Case of Aleksander Kurtna

341 „Me oleme Põhjast": Livonica Vatikani arhiivis ja Aleksander Kurtna erijuhtum. *Resümee*

Abstracted/indexed in: Central and Eastern European Online Library (CEEOL), EBSCO Publishing database, Gale's Academic OneFile database, Index Copernicus International World of Journals, SCOPUS (Elsevier), Web of Science (Arts & Humanities Citation Index)

ISSN 14062925 (print)
ISSN 17367476 (electronic)

Acta Historica Tallinnensia 32/2 https://doi.org/10.3176/hist.2026.2.01

Medieval Studies after the Digital Turn: Methods, Infrastructures, and the Making of *Digital Livonia*

Marek Tamm

Tallinn University, School of Humanities, Institute of History, Archaeology and Art History, Narva mnt 25, 10120 Tallinn; marek.tamm@tlu.ee

Hembo Pagi

Tallinn University, School of Humanities, Institute of History, Archaeology and Art History, Narva mnt 25, 10120 Tallinn; hembo@tlu.ee

Abstract. This article examines the digital turn in medieval studies through a methodological survey and a reflective study of Digital Livonia, a regional research platform. It reviews developments in digital collections, editions, transcription, computational analysis, networks, GIS and artificial intelligence, then examines their application to dispersed sources, heterogeneous datasets and established research practices. It argues that integration requires selective formalisation: shared structures must enable comparison while preserving source distinctions, provisional identifications and the provenance of scholarly assertions. Digital Livonia demonstrates how data models, researcher workflows and institutional responsibility shape the possibilities and limits of digital historical enquiry.

Keywords: digital medieval studies, medieval Livonia, research infrastructures, digital editions, handwritten text recognition, computational text analysis, source criticism

© 2026 Authors. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International License CC BY 4.0 (http://creativecommons.org/licenses/by/4.0).

INTRODUCTION

Over the past two decades, the digital turn has been one of the most significant transformations in medieval studies. It has, however, not been the only major change to reshape the field. Global, material, environmental, postcolonial and gender-oriented approaches have altered its geographical horizons, its objects of study and many of its fundamental questions. Digital technologies have increasingly become part of the infrastructure through which such research is pursued. Their effects extend across almost every stage of the research process, from the discovery and reproduction of sources to their transcription, analysis, interpretation and publication. Even medievalists who would never describe themselves as digital humanists now work in an environment shaped by online catalogues, digitised manuscripts, searchable corpora, bibliographical databases and electronic communication. In this broad sense, digital resources have become part of the ordinary practice of medieval studies, including research that makes no use of computational analysis.[1]

The term 'digital turn' nevertheless risks suggesting a rupture where the historical development has been cumulative and uneven. Computers entered medieval research long before the World Wide Web, while printed catalogues, critical editions, archival visits and close reading continue alongside databases, digital facsimiles and machine learning. The digital also encompasses practices that involve very different degrees of formalisation and methodological intervention, from viewing a manuscript image or consulting an online repertory to training a model to identify scribal hands.[2] For analytical purposes, we distinguish three overlapping operations: digitisation makes existing sources and scholarly products available electronically; digital structuring represents them

1 For some general discussions, see D. J. Birnbaum, S. Bonde, M. Kestemont. The Digital Middle Ages: An Introduction. – Speculum, 2017, 92, S1, S1–S38. https://doi.org/10.1086/694236; R. Bleier, F. Fischer, T. Hiltmann, G. Viehhauser, G. Vogeler. Digitale Mediävistik und der deutschsprachige Raum. – Das Mittelalter, 2019, 24, 1, 1–12. https://doi.org/10.1515/mial-2019-0001; B. Burrichter, B. Gebert, C. Mackert, G. Viehhauser. Digitale Mediävistik. – Das Mittelalter, 2021, 26, 1, 101–117. https://doi.org/10.17885/heiup.mial.2021.1.24312; Medieval Manuscripts in the Digital Age. Ed. by B. Albritton, G. Henley, E. Treharne. Routledge, London, New York, 2021; Digital Medieval Studies: Practice and Preservation. Ed. by L. K. Morreale, S. Gilsdorf. Arc Humanities Press, Leeds, 2022. https://doi.org/10.1017/9781802700152; Digital Medieval Studies: Experimentation and Innovation. Ed. by S. Gilsdorf, L. K. Morreale. Arc Humanities Press, Leeds, 2024. https://doi.org/10.1515/9781802702330.

2 On identifying scribal hands, see, for instance, C. De Stefano, F. Fontanella, M. Maniaci, A. Scotto di Freca. Reliable Writer Identification in Medieval Manuscripts through Page Layout Features: The "Avila" Bible Case. – Engineering Applications of Artificial Intelligence, 2018, 72, 99–110. https://doi.org/10.1016/j.engappai.2018.03.023.

as searchable, linkable and reusable objects; computational analysis applies formal, statistical or algorithmic procedures to texts, images and historical data.[3] These distinctions concern what researchers do with evidence, rather than separate branches or successive stages of medieval studies. Search itself involves computation, and the production of a digital facsimile already requires choices about representation and description. They nevertheless provide a useful framework for examining how digital approaches have reorganised medieval research and how sources move from images and transcriptions to structured records and analytical corpora.

This introductory article develops that examination in two parts. The first part surveys selected areas of digital research in medieval studies, including digital collections and editions, manuscript transcription, computational textual analysis, relational and spatial methods, and artificial intelligence. The second examines the construction of Digital Livonia, a regional platform bringing together digitised sources, editions, thematic databases and research services. The survey establishes the methodological possibilities against which the choices and limitations of the platform can be assessed. Our argument develops from a practical problem encountered while building Digital Livonia: what should a regional research platform integrate, and what should remain specific to its constituent collections? Shared structures make evidence comparable, but they can also obscure differences between documentary attestations, uncertain identifications and modern reconstructions. Conversely, preserving every collection in its existing form limits the connections that researchers can establish between them. We examine how this tension shaped the platform's data model, technical architecture and working procedures. These choices form part of what we call a hybrid practice of medieval research, in which digital procedures are developed and assessed through sustained engagement with philology, palaeography, codicology and historical source criticism. Collaboration between historians, technical specialists and collection custodians helps to maintain an identifiable relationship between historical assertions and the evidence on which they depend.

The other contributions to this special issue examine these questions through different bodies of Livonian evidence. Nicolas Bourgeois assesses statistical and machine-learning approaches to the *Chronicle*

3 On the distinction between digitisation, digital methods and digital history, see C. Brennan. Digital Humanities, Digital Methods, Digital History, and Digital Outputs. – History Compass, 2018, 16, 10, e12492. https://doi.org/10.1111/hic3.12492.

of Henry of Livonia and the *Livonian Rhymed Chronicle*, considering how computational analysis can support verifiable interpretation and collaboratively curated metadata. Rūta Bruževica follows the transformation of craft ordinances from Riga and Tallinn into a digital corpus, exploring how textual analysis and visualisation depend on transcription, classification and the interpretation of normative sources. Turning to material evidence, Anu Mänd demonstrates how reflectance transformation imaging and digital photomontage can improve the reading of worn grave slabs and reconnect dispersed fragments, including through comparison with historical drawings. Juhan Kreem places digital imaging within the longer history of reproducing medieval Livonian seals, showing how successive techniques have shaped their study and why the proliferation of images needs to be accompanied by adequate scholarly metadata. The final two articles examine the manuscript and archival contexts on which digital research also depends. Tiina Kala and Kaspar Kolk investigate the textual composition, binding and liturgical fragments of the *Ex quo* vocabulary manuscript Cm 16 in the Tallinn City Archives, distinguishing its connections with the Dominican friary from the unresolved identity and affiliation of its scribe. Kristjan Toomaspoeg reconstructs the history of research on Livonica in the Vatican Archives, with particular attention to Aleksander Kurtna, and identifies the need to bring dispersed editions, transcriptions and research papers into a common database. Together, these studies show how the possibilities of digital medieval research are shaped by the transmission of texts and objects, the history of their documentation and the scholarly decisions through which they become available for comparison.

FROM HUMANITIES COMPUTING TO DIGITAL MEDIEVAL STUDIES

The history of digital medieval studies is commonly narrated through the Italian Jesuit Roberto Busa and his *Index Thomisticus*. From his doctoral research in the mid-1940s, Busa sought a means of indexing every occurrence of every word in the works of Thomas Aquinas and related authors; during a visit to the United States in 1949, he secured the support of IBM. Yet, as Steven E. Jones has shown, this familiar account functions as a founding myth and should not be mistaken for a simple

point of origin.[4] Busa did not step directly into modern computing. His early work depended on punch-card machines developed for business data processing and is more accurately described as humanities data processing than as digital humanities in the present sense. The project grew out of a particular post-war conjunction of Jesuit scholarly networks, IBM's technical support, engineers such as H. Paul Tasman, and extensive, often unacknowledged, data labour, much of it performed by women.[5] Its importance lies both in the unprecedented scale of the resulting eleven-million-word concordance and in this socio-technical form of scholarship: text was decomposed into units, encoded according to explicit rules and reassembled, while the available machinery helped to determine what could be asked and produced.

Busa's project was exceptional in scale, but it should not obscure the multiple beginnings of computer-assisted medieval research. Edgar Lejeune's recent research on European medievalists in the 1960–1990 period shows that the field emerged in several rather different settings. At the Centre de traitement électronique des documents (CETEDOC) in Louvain, Paul Tombeur and his collaborators developed automated concordances of medieval Latin texts. From 1966, David Herlihy and Christiane Klapisch-Zuber organised a transatlantic project to turn the Florentine *catasto* of 1427 into a database for demographic and social analysis. In the 1970s, Jean-Philippe Genet and his colleagues created the ALINE program and the MEDITEXT corpus for lexicometric analysis. These enterprises represented their sources in different ways – as full texts, selected variables, documentary descriptors, concordances or programs – but all were collective, bringing historians together with demographers, linguists, computer scientists, punch-card operators and institutional computing centres. Lejeune's central contribution is to locate the transformation in what he calls the 'practices of sources': their selection and criticism, the constitution of corpora and the operations by which evidence was synthesised. The Florentine *catasto* project, initially described as an 'edition for the computer', is exemplary. Its data were not simply present in the manuscript waiting to be extracted; they were produced through decisions about variables,

4 S. E. Jones. Roberto Busa, S. J., and the Emergence of Humanities Computing: The Priest and the Punched Cards. Routledge, New York, 2016, 1–29, 107–173.

5 See J. Nyhan, A. Flinn. Computation and the Humanities: Towards an Oral History of Digital Humanities. Springer, Cham, 2016, esp. 59–62, 79–80, 90. https://doi.org/10.1007/978-3-319-20170-2; M. Terras, J. Nyhan. Father Busa's Female Punch Card Operatives. – Debates in the Digital Humanities 2016. Ed. by M. K. Gold, L. F. Klein. University of Minnesota Press, Minneapolis, 2016, 60–65. https://doi.org/10.5749/j.ctt1cn6thb.9.

categories, omissions, correction and uncertainty. Early computer use thus demanded a new degree of formalisation and made particularly visible that data are constructed forms of evidence. It also prompted collective organisation beyond individual projects. A meeting in Rome in 1975 led to the foundation in 1979 of the review *Le Médiéviste et l'ordinateur*, whose debates about workflows, open or closed research instruments, common standards, reuse and technological obsolescence remain strikingly current.[6]

By the 1990s, projects such as the Electronic Beowulf and the Piers Plowman Electronic Archive were testing what an electronic scholarly environment could offer that print could not.[7] Manuscript images, transcriptions, descriptions and apparatus could now be brought into direct relation, while the Text Encoding Initiative (TEI), founded in 1987, supplied a shared framework for representing textual structures. Personal computing, optical media and especially the web moved such work beyond institutional computing centres and specialised laboratories.[8] Yet many of the questions raised by the earlier generation (standards, reuse, collaboration and technological obsolescence) remained unresolved.[9]

Since the 2010s, the growth of digitised collections, interoperable image services, handwritten text recognition, computer vision and machine learning has expanded the possibilities for connecting research operations. Earlier projects also sought to integrate databases, corpora and editions, while many contemporary resources remain isolated. What has changed is the availability of shared standards and services through which materials can move between digitisation, description,

6 E. Lejeune. Médiévistes et ordinateurs: Organisations collectives, pratiques des sources et conséquences historiographiques (1966–1990). Thèse de doctorat d'histoire et de philosophie des sciences, Université de Paris, 2021. https://theses.hal.science/tel-03682773v2. See also E. Lejeune. 'Open' or 'Close' Research Instruments? Conflicting Rationales in the Organization of Early Digital Medieval History in Europe (1960–1990). – Digital Humanities Quarterly, 2024, 18, 2. https://doi.org/10.63744/f7ddtgjkbjyw; E. Lejeune. From Sources to Data: Medievalists, Demographers, and Computer Scientists Producing a Database from a Quattrocento Census (1964–1978). – History of Humanities, 2025, 10, 2, 397–425. https://doi.org/10.1086/737038.

7 A. Prescott. The Electronic Beowulf and Digital Restoration. – Literary and Linguistic Computing, 1997, 12, 3, 185–196. https://doi.org/10.1093/llc/12.3.185; The Piers Plowman Electronic Archive, Vol. 1: Corpus Christi College, Oxford, MS 201 (F). Ed. by R. Adams, H. N. Duggan, E. Eliason, R. Hanna III, J. Price-Wilkin, T. Turville-Petre. University of Michigan Press for the Society for Early English and Norse Electronic Texts, Ann Arbor, 2000.

8 J. Unsworth. Medievalists as Early Adopters of Information Technology. – Digital Medievalist, 2012, 7. https://doi.org/10.16995/dm.34.

9 P. Robinson. Current Issues in Making Digital Editions of Medieval Texts – or, Do Electronic Scholarly Editions Have a Future? – Digital Medievalist, 2005, 1. https://doi.org/10.16995/dm.8; S. P. McCormick. A Guide to Digital Medieval Studies in North America. – Perspectives médiévales, 2016, 37. https://doi.org/10.4000/peme.9655; P. Sahle. What Is a Scholarly Digital Edition? – Digital Scholarly Editing: Theories and Practices. Ed. by M. J. Driscoll, E. Pierazzo. Open Book Publishers, Cambridge, 2016, 19–39. https://doi.org/10.11647/OBP.0095.02.

transcription, annotation, analysis and publication.[10] Donald J. Waters has described this as the emergence of an 'end-to-end' digital infrastructure for the humanities.[11] Such an infrastructure remains incomplete and unevenly distributed, but the concept draws attention to the connections between research operations and the work required to sustain them.[12] The developments briefly outlined here did not replace one another, and their boundaries remain porous. The history of the field is better understood as a cumulative and uneven attempt to integrate changing research instruments with the older practices of source criticism and historical interpretation.[13]

DIGITAL COLLECTIONS AND RESEARCH INFRASTRUCTURES

For working medievalists, the most immediate effect of digitisation has been access. Manuscripts, charters, seals, inscriptions and other sources held in different countries can now be compared within the same working session. This has changed the geography of research as well as its speed. Materials once accessible mainly to scholars with substantial travel budgets and institutional connections can increasingly be consulted from elsewhere.[14] The International Image Interoperability Framework (IIIF) has pushed this development beyond the provision of images. Because compatible repositories can expose images through a common protocol, a manuscript held in one collection can be viewed, compared and annotated alongside material held elsewhere. Codices, fragments and documentary series dispersed by the history of collecting can therefore be provisionally reunited in a common research environment.[15]

10 C. L. Borgman, A. Scharnhorst, M. S. Golshan. Digital Data Archives as Knowledge Infrastructures: Mediating Data Sharing and Reuse. – Journal of the Association for Information Science and Technology, 2019, 70, 8, 888–904. https://doi.org/10.1002/asi.24172.

11 D. J. Waters. The Emerging Digital Infrastructure for Research in the Humanities. – International Journal on Digital Libraries, 2023, 24, 87–102, here 88–90. https://doi.org/10.1007/s00799-022-00332-3.

12 U. Pawlicka-Deger. Infrastructuring Digital Humanities: On Relational Infrastructure and Global Reconfiguration of the Field. – Digital Scholarship in the Humanities, 2022, 37, 2, 534–550. https://doi.org/10.1093/llc/fqab086; C. L. Borgman, A. Scharnhorst, M. S. Golshan. Digital Data Archives as Knowledge Infrastructures.

13 P. Svensson. The Landscape of Digital Humanities. – Digital Humanities Quarterly, 2010, 4, 1. https://doi.org/10.63744/5dhf3ymfs2gg.

14 A. Cappellotto. From Codex to Apps: The Medieval Manuscript in the Age of Its Digital Reproduction. – Umanistica Digitale, 2020, 4, 9, 1–18. https://doi.org/10.6092/issn.2532-8816/11459; T. Burrows. The History and Provenance of Manuscripts in the Collection of Sir Thomas Phillipps: New Approaches to Digital Representation. – Speculum, 2017, 92, S1, S39–S64. https://doi.org/10.1086/693438.

15 R. Sanderson, B. Albritton, R. Schwemmer, H. Van de Sompel. SharedCanvas:

Yet access remains uneven. Digitisation is expensive and selective, and well-funded or visually attractive collections are more likely to appear online than poorly catalogued materials. The image itself is also a transformation of the object: scale, weight, texture, sequence and traces of handling are difficult to reproduce on screen.[16] Advanced imaging can also reveal features unavailable in conventional photographs. Multispectral imaging, reflectance transformation imaging and 3D modelling can disclose erased writing, surface relief and traces of pigment that may be difficult or impossible to perceive through ordinary inspection.[17] For seals, inscriptions and other material sources, lighting direction and image processing can make a worn letter or shallow impression more legible, while also shaping its appearance. Researchers therefore need documentation of the conditions of capture and processing and, where possible, access to alternative views of the same surface. These allow another scholar to assess a proposed reading and consider how the imaging procedure may have enhanced, obscured or distorted the feature under examination. Catalogue metadata require the same scrutiny as images. Inherited classifications and uncertain attributions can acquire a misleading precision when presented as database fields; documenting their origin and revision allows users to distinguish institutional descriptions from subsequent scholarly interpretations.[18]

Our experience with Digital Livonia has made clear how much this critical work depends on infrastructure. Preserving images, documenting their production and keeping research data available for scrutiny and reuse require repositories, standards, software maintenance and sustained institutional commitments. FAIR provides a useful vocabulary – findable, accessible, interoperable and reusable – but achieving these qualities requires continuing scholarly work.[19]

A Collaborative Model for Digital Facsimiles. – International Journal on Digital Libraries, 2012, 13, 1, 3–16. https://doi.org/10.1007/s00799-012-0098-8; L. F. Davis. Framing the Fragment, from Parchment to Pixel. – Digital Medievalist, 2024, 17, 1, 1–23. https://doi.org/10.16995/dm.10343.

16 H. Bamford, E. Francomano. On Digital-Medieval Manuscript Culture: A Tentative Manifesto. – Digital Philology, 2018, 7, 1, 29–45. https://doi.org/10.1353/dph.2018.0002; L. F. Davis. Life On – and Off – the Continuum. – Digital Medieval Studies: Practice and Preservation, 99–111. https://doi.org/10.1017/9781802700152.006.

17 B. Endres. Digitizing Medieval Manuscripts: The St. Chad Gospels, Materiality, Recoveries, and Representation in 2D and 3D. Arc Humanities Press, Leeds, 2019.

18 P. Andrist. Toward a New Generation of Databases and Database Applications for Describing Ancient Manuscripts. – Digital Scholarship in the Humanities, 2021, 36, Supplement 2, ii9–ii16. https://doi.org/10.1093/llc/fqab022.

19 M. D. Wilkinson et al. The FAIR Guiding Principles for Scientific Data Management and Stewardship. – Scientific Data, 2016, 3, 160018. https://doi.org/10.1038/sdata.2016.18.

DIGITAL EDITIONS AND THE MODELLING OF TEXTUAL TRADITIONS

Scholarly editing is among the areas in which medieval sources have posed particularly productive challenges to digital methods. Medieval works often survive in witnesses that differ in wording, order, language, layout and material context. Printed editions have developed sophisticated ways of representing this plurality, including synoptic texts and extensive apparatuses, but the relations between these elements remain constrained by the organisation of the page and volume. Digital editions permit different arrangements of images, diplomatic and normalised transcriptions, translations, variants and commentary. Their value depends on which relations they allow readers to inspect and how clearly they disclose the editorial decisions behind them.[20]

A digital edition constitutes an explicit model of the textual tradition. The editor must decide what counts as a textual unit, how abbreviations and corrections are represented, whether uncertain readings are retained, how witnesses are aligned and which relations between text and material object are made visible. Encoding languages such as TEI render many of these decisions explicit and machine-readable. They allow the same data to generate different views: a close transcription beside the facsimile, a normalised reading text, a synoptic comparison of witnesses or a corpus for linguistic analysis. The scholarly argument therefore lies not only in the introductory essay and reconstructed text, but also in the structure of the data.[21] For historians working with charters, inventories or account books, another question concerns the relationship between an edited text and the information extracted from it. Georg Vogeler's concept of the 'assertive edition' addresses this relation by linking textual evidence to structured statements about people, objects, events and their connections.[22]

Digital editions are particularly well suited to what has been called the *mouvance* or variance of medieval textuality.[23] Instead of treating differences between witnesses mainly as corruption on the

20 E. Pierazzo. Digital Scholarly Editing: Theories, Models and Methods. Ashgate, Farnham, 2015, 5–24. https://doi.org/10.4324/9781315573564.

21 P. Sahle. What Is a Scholarly Digital Edition?

22 G. Vogeler. The 'Assertive Edition': On the Consequences of Digital Methods in Scholarly Editing for Historians. – International Journal of Digital Humanities, 2019, 1, 309–322. https://doi.org/10.1007/s42803-019-00025-5.

23 P. Zumthor. Essai de poétique médiévale. Seuil, Paris, 1972, 164, 507; B. Cerquiglini. Éloge de la variante: Histoire critique de la philologie. Seuil, Paris, 1989, 54.

way to an original, they can represent textual change as an object of study in its own right. Gabriel Viehhauser's survey of digital editing in medieval German studies shows both the potential and the limits of this development. Projects such as the digital edition of *Der arme Heinrich* combine manuscript images, several levels of transcription, translation and information about the codex, thereby restoring a degree of visibility to the plurality of the tradition. At the same time, the number of fully developed digital medieval editions remains smaller than repeated predictions of a digital revolution might lead one to expect.[24] These limitations arise in large part from the institutional, financial and technical demands of digital scholarly editing. Digital editions require sustained collaboration between editors, libraries and research software specialists; they are costly to develop and demanding to maintain. Interfaces age faster than printed books, links decay and project-specific software may become unusable. Moreover, academic evaluation still tends to recognise the finished monograph or critical edition more readily than the creation of data, code or an evolving scholarly resource. The digital edition has consequently revealed an institutional problem as much as a technical one: scholarship increasingly depends on collective and continuing labour, while systems of authorship, funding and credit remain oriented towards individual and final products.[25]

MAKING MANUSCRIPTS MACHINE-READABLE

Digitisation has created an abundance of manuscript images, but images are not searchable textual corpora. For most medieval sources, transcription remains the principal bottleneck. Optical character recognition (OCR) developed for modern print performs poorly on handwriting, abbreviations and irregular layouts. Handwritten text recognition (HTR) addresses the problem through supervised machine learning: a model is trained on images paired with corrected transcriptions, learns recurring visual patterns and predicts characters on pages that were not included in the training material. As Peter Stokes

24 G. Viehhauser. Zum gegenwärtigen Stand der digitalen Editorik in der Altgermanistik: Einige Bemerkungen. – Computer im Musenhain: Von träumenden Büchern und der Aura des Digitalen. Hrsg. von S. Bartsch, L. Borek, P. Hegel. Technische Universität Darmstadt, Darmstadt, 2023, 15–25. https://doi.org/10.26083/tuprints-00027905.

25 L. K. Morreale, S. Gilsdorf. Introduction: The Medievalist, Digital Edition. – Digital Medieval Studies: Practice and Preservation, 1–10. https://doi.org/10.1017/9781802700152.001; D. J. Waters. Emerging Digital Infrastructure, 96–99.

and Benjamin Kiessling emphasise, the performance and range of such systems depend fundamentally on the data available to train them.[26] Progress has consequently been fastest for modern, widely used languages and scripts, whereas many historical writing systems lack corpora of sufficient size and consistency. The selection and preparation of training data already determine which materials become amenable to automated study and which remain comparatively inaccessible.[27]

The practical importance of HTR is considerable. Output need not be publication-ready to support full-text search, the identification of names and formulae, or the selection of passages for correction. A relatively small set of manually transcribed pages can make a much larger collection searchable. General models for Latin and vernacular scripts are also reducing the quantity of project-specific training material required, although performance varies sharply between hands, genres, layouts and states of preservation. The increasing circulation of reusable models and training datasets may reduce duplicated labour, particularly for comparatively rare scripts. Their reuse nevertheless requires common standards, or at least explicitly documented conventions, for transcription and data exchange.[28]

HTR has returned an old philological question to the centre of attention: what should a transcription record? Expanding abbreviations, modernising punctuation or normalising spelling may aid reading but erase features relevant to scribal practice. Conversely, recording every glyph can produce data that are difficult to reuse. Estelle Guéville and David Joseph Wrisley argue that transcription schemes must follow the research question. Their work on Paris Bibles preserves abbreviation and orthographic variation because these features help distinguish scribes and regional habits. The 'ground truth' used to train a model is therefore already an interpretation: transcription determines which features of the manuscript become available for subsequent computational analysis and

26 P. A. Stokes, B. Kiessling. Sharing Data for Handwritten Text Recognition (HTR). – The Companion to Digital Humanities in Practice. Ed. by C. Crompton, L. Estill, R. J. Lane, R. Siemens. Routledge, London, New York, 2025, 95–104. https://doi.org/10.4324/9781003327677-8.

27 G. Mühlberger et al. Transforming Scholarship in the Archives through Handwritten Text Recognition: Transkribus as a Case Study. – Journal of Documentation, 2019, 75, 5, 954–976. https://doi.org/10.1108/JD-07-2018-0114; A. Pinche, P. A. Stokes. Historical Documents and Automatic Text Recognition: Introduction. – Journal of Data Mining and Digital Humanities, 2024. https://doi.org/10.46298/jdmdh.13247.

28 A. Pinche, T. Clérice. Reconnaissance automatique d'écriture et sources historiques: Limites et nouvelles perspectives. – Studi Francesi, 2025, 206, 357–366. https://doi.org/10.4000/15gi8; D. J. Wrisley, E. Guéville. Medieval Manuscripts and the Computational Humanities: Big Data, Scribes, and the "Paris Bible". Arc Humanities Press, Leeds, 2026, 43–82. https://doi.org/10.17302/MFSK1586.

which disappear through normalisation.[29] This problem has also been examined by Elena Pierazzo and Peter Stokes, who argue that manuscript transcription should preserve the relationship between text and codex rather than extract a purely linguistic sequence from its material setting. Pages, columns, lines, additions, erasures and spatial arrangements are not incidental containers of textual content; depending on the question being asked, they may form part of the evidence that the transcription is intended to represent.[30] Transcription is consequently a model of the manuscript, created through decisions about which dimensions of the object are retained and which are omitted.

Ariane Pinche's research on early French vernacular legendaries illustrates how improvements in HTR can extend the corpus available for philological investigation. Her account moves from an initial experiment on a single manuscript towards the use of recognition models on further witnesses. The resulting transcriptions provide material for comparing the composition of hagiographical collections and investigating the possible circulation of smaller groups of saints' lives before their incorporation into larger compilations.[31] A comparable expansion of the available evidence is occurring in visual and material analysis, where digitisation increasingly means producing new forms of evidence from reflected light, surface structure and comparison across images. Computer vision can identify page regions, measure proportions of text and decoration, compare letterforms or search for recurring motifs. The eCodicology project analysed around 170,000 images from manuscripts associated with the Benedictine abbey of St Matthias in Trier, extracting layout features in order to identify patterns across the collection.[32] DigiPal and related projects led by Stokes have similarly used digital images and structured descriptions of letterforms to support comparison, scribal attribution and the study of scripts. Such approaches render

29 E. Guéville, D. J. Wrisley. Transcribing Medieval Manuscripts for Machine Learning. – Journal of Data Mining and Digital Humanities, 2024. https://doi.org/10.46298/jdmdh.9805.

30 E. Pierazzo, P. A. Stokes. Putting the Text Back into Context: A Codicological Approach to Manuscript Transcription. – Kodikologie und Paläographie im digitalen Zeitalter 2. Codicology and Palaeography in the Digital Age 2. Ed. by F. Fischer, C. Fritze, G. Vogeler. Books on Demand, Norderstedt, 2010, 397–430; P. A. Stokes. Holistically Modelling the Medieval Book: Towards a Digital Contribution. – Anglia, 2021, 139, 1, 6–31. https://doi.org/10.1515/ang-2021-0002.

31 A. Pinche. Comprendre la composition des premiers légendiers en langue vernaculaire à l'aide des méthodes numériques: Acquisition automatique du texte (HTR) et analyses stylométriques. – Bulletin du centre d'études médiévales d'Auxerre, 2022, Collection CBMA. https://doi.org/10.4000/cem.19452.

32 H. Busch, S. Chandna. eCodicology: The Computer and the Mediaeval Library. – Kodikologie und Paläographie im digitalen Zeitalter 4. Hrsg. von H. Busch, F. Fischer, P. Sahle. Books on Demand, Norderstedt, 2017, 3–23; B. Endres. Digitizing Medieval Manuscripts.

selected palaeographical features measurable and comparable, while their identification and the interpretation of resulting patterns continue to depend on palaeographical judgement.[33]

The possibilities opened by HTR and related forms of image analysis are accompanied by problems of uneven reliability. In HTR, errors are not random: models may perform well on regular main text but fail on marginalia, unusual abbreviations, changes of hand or damaged passages – precisely the features that may matter most. Accuracy percentages also conceal the difference between common and historically significant errors.[34] To make these limitations assessable, published transcriptions should document their provenance, identify the recognition model and transcription conventions used, and distinguish uncorrected output from text reviewed by researchers, with a record of subsequent corrections. Where recognition accuracy is reported, the evaluation sample and its relation to the wider collection should also be specified. The dependence of recognition quality on the material condition of the source is especially visible in the Rescapé project on the fire-damaged Turin manuscript BNU L.II.14. Severe scorching initially prevented automatic segmentation because the contrast between ink and parchment was insufficient, while tears, cuts and warped text lines produced further recognition problems. The project therefore combined specialised imaging with complementary tools for region and line detection and adapted its transcription conventions to represent damaged and illegible text. The case locates HTR within a longer chain extending from conservation and imaging through segmentation and transcription to interoperable data.[35]

33 P. A. Stokes. Scribal Attribution across Multiple Scripts: A Digitally-Aided Approach. – Speculum, 2017, 92, S1, S65–S85. https://doi.org/10.1086/693968; S. Brookes, P. A. Stokes, M. Watson, D. Marques de Matos. The DigiPal Project for European Scripts and Decorations. – Writing Europe, 500–1450: Texts and Contexts. Ed. by A. Conti, O. Da Rold, P. Shaw. D. S. Brewer, Cambridge, 2015, 25–58.

34 M. J. Hill, S. Hengchen. Quantifying the Impact of Dirty OCR on Historical Text Analysis: Eighteenth Century Collections Online as a Case Study. – Digital Scholarship in the Humanities, 2019, 34, 4, 825–843. https://doi.org/10.1093/llc/fqz024; T. Clérice, M. Vlachou-Efstathiou, A. Chagué. CREMMA Medii Aevi: Literary Manuscript Text Recognition in Latin. – Journal of Open Humanities Data, 2023, 9, 4. https://doi.org/10.5334/johd.97.

35 G. Kagan. Numériser des manuscrits brûlés. Techniques et méthodes à l'aide de l'intelligence artificielle: nouvelles perspectives. – Studi Francesi, 2025, 206, 335–340. https://doi.org/10.4000/15gi5; P. O'Connor. Handwritten Text Recognition Technology and MS Turin, BNU, L.II.14 (T). The "Rescapé" Case Study. – Studi Francesi, 2025, 206, 367–376. https://doi.org/10.4000/15gi9.

READING TEXTS AT SCALE

Once texts are machine-readable, they can be compared across corpora beyond the limits of ordinary linear reading. Computational textual analysis includes word frequencies, stylometry, authorship attribution, text-reuse detection, topic modelling and semantic methods. Their common premise is that patterns too dispersed or repetitive for unaided reading may become visible when textual features are counted and compared systematically.[36] An earlier experiment in Pinche's research on vernacular legendaries, conducted jointly with Jean-Baptiste Camps and Thibault Clérice, combined HTR and stylometry to test Paul Meyer's hypotheses about the grouping of saints' lives in Paris, BnF, fr. 412. Working with texts copied by a single hand helped to limit scribal variation, although recognition errors and differences between the texts still complicated the analysis. The resulting groupings provided evidence for assessing the proposed compilations and possible shared authorship, with their interpretation requiring comparison with the manuscript tradition.[37]

Stylometry is one form of this work. It measures recurrent features such as frequent words, character sequences, vocabulary or syntax to compare texts and has been applied to anonymous authorship, scribal intervention, genre and dialect. Medieval transmission complicates the inference from statistical similarity to authorship because the surviving text may preserve interventions by compilers, translators, scribes and modern editors.[38] A comparison of letters by Hildegard of Bingen, Guibert of Gembloux and Bernard of Clairvaux suggested that Guibert substantially reworked two visionary texts transmitted under Hildegard's name.[39] Quantitative spelling analysis of *The Canterbury Tales* has confirmed palaeographical identifications of manuscripts copied by the

36 See, for instance, F. Stella. Digital Philology and Quantitative Criticism of Medieval Literature: Unconventional Approaches to Medieval Latin Literature II. Brepols, Turnhout, 2021. https://doi.org/10.1484/M.USML-EB.5.119334; J. De Gussem. Computational Stylistics and Medieval Texts. – Routledge Resources Online: Medieval Studies. Ed. by H. Klemettilä, S. Niskanen, J. Willoughby. Routledge, 2022, 1–12. https://doi.org/10.4324/9780415791182-RMEO391-1.

37 J.-B. Camps, T. Clérice, A. Pinche. Noisy Medieval Data, from Digitized Manuscript to Stylometric Analysis: Evaluating Paul Meyer's Hagiographic Hypothesis. – Digital Scholarship in the Humanities, 2021, 36, Supplement 2, ii49–ii71. https://doi.org/10.1093/llc/fqab033.

38 D. J. Wrisley, E. Guéville. Medieval Manuscripts and the Computational Humanities, 89–116; S. Dows-Miller. Medieval Authorship and Statistical Analysis: Reconsidering Jehan de Saint-Quentin's 'œuvre' in the Digital Age. ARC Humanities Press, Leeds, 2026, 17–28.

39 M. Kestemont, S. Moens, J. Deploige. Collaborative Authorship in the Twelfth Century: A Stylometric Study of Hildegard of Bingen and Guibert of Gembloux. – Digital Scholarship in the Humanities, 2015, 30, 2, 199–224. https://doi.org/10.1093/llc/fqt063.

same scribes,[40] while word embeddings applied to the DEEDS corpus have traced changes in administrative vocabulary across the Norman Conquest. The questions differ, but each depends on converting textual variation into features that can be compared systematically.[41] Corpus analysis and close reading operate most effectively in succession. The first locates anomalies and tests representativeness; the second evaluates what a pattern may mean. Often the process is recursive.

Such analysis is complicated by the particular character of medieval corpora. They are frequently small, multilingual, editorially heterogeneous and unrepresentative; spelling, dates and attributions are unstable. Latin and canonical authors remain easier to model than many vernacular or documentary traditions. Differences in transcription, editorial intervention and corpus composition must therefore be examined before statistical patterns are attributed to medieval authors or scribes.[42]

RELATIONS AND PLACES: PROSOPOGRAPHY, NETWORKS AND GIS

Charters, registers, letters, chronicles and administrative records repeatedly connect persons, institutions, offices, events and places. Digital prosopography makes it possible to identify such entities across sources and investigate them collectively, while network analysis and geographic information systems (GIS) provide ways of examining the relations and spatial patterns that emerge from these data.

These relations, however, are not simply given by the sources. A line in a network graph is already an interpretation. Co-witnessing a charter, kinship, correspondence, monastic membership and denunciation in an inquisitorial record arise from different documentary practices and cannot be treated as interchangeable evidence of 'connection'. A graph is not a picture of medieval society, but a model constructed from selected

40 P. M. W. Robinson, T. Tresoldi. Analyzing Spelling Patterns in the Manuscripts of the *Tales of Canterbury*. – Digital Scholarship in the Humanities, 2026, 41, 1, 347–363. https://doi.org/10.1093/llc/fqaf123.

41 Y. Liu, G. Tilahun, X. Gao, Q. Wen, M. Gervers. A Comparative Study of Static and Contextual Embeddings for Analyzing Semantic Changes in Medieval Latin Charters. – Proceedings of the First Workshop on Language Models for Low-Resource Languages. Association for Computational Linguistics, Abu Dhabi, 2025, 182–192. https://aclanthology.org/2025.loreslm-1.14/.

42 On the inseparability of corpus construction and interpretation, see K. Bode. A World of Fiction: Digital Collections and the Future of Literary History. University of Michigan Press, Ann Arbor, 2018, 17–44. https://doi.org/10.3998/mpub.8784777; D. J. Wrisley, E. Guéville. Medieval Manuscripts and the Computational Humanities.

sources and decisions about nodes and ties. Measures such as centrality, clustering or brokerage acquire historical meaning only in relation to the evidence from which those relations were derived.[43] This dependence on documentary form makes source criticism part of the network model itself. An apparent difference between the network positions of men and women, for example, may reflect the composition of surviving testimonies, the interests of investigators or the circumstances under which information was recorded.[44] Such factors can sometimes be encoded and tested rather than left as external qualifications. Zbíral, Sikk and Shaw have proposed precisely this kind of integration of digital methods and source criticism in their work on inquisitorial records.[45] Courtney Luckhardt's study of women's networks in Carolingian Septimania shows a complementary possibility: relational analysis can recover actors and connections obscured by narratives organised around office, dynasty or male political agency.[46] Studies of inquisitorial material likewise reconstruct dissident communities while asking how interrogation, denunciation and documentary procedure produced the network preserved in the source.[47]

Mapping Medieval Chester reconstructed the urban landscape around 1500 and set it beside English, Latin and Welsh textual representations of the city. The comparison made it possible to examine how the same urban space was perceived and narrated within a multilingual frontier community.[48] Justin Colson confronted a different problem in late medieval and early modern London, where deeds and wills located properties through relational descriptions such as 'next to', 'opposite'

43 N. Ruffini-Ronzani, S. de Valeriola, E. Cuvelier. Dealing with the Heterogeneity of Interpersonal Relationships in the Middle Ages: A Multi-Layer Network Approach. – Digital Medievalist, 2022, 15, 1. https://doi.org/10.16995/dm.8070; S. de Valeriola. Can Historians Trust Centrality? Historical Network Analysis and Centrality Metrics Robustness. – Journal of Historical Network Research, 2021, 6, 1, 85–125. https://doi.org/10.25517/jhnr.v6i1.105; K. Sikk, R. Välimäki, D. Zbíral. How Inquisitive Was Medieval Inquisition? A Network-Analytical Approach to Information Flow in the Trials for Brandenburg-Pomeranian Waldensians (Late 14th c.). – Digital Scholarship in the Humanities, 2026, 41, 1, 396–411. https://doi.org/10.1093/llc/fqaf138.

44 D. Salihović, J. L. Estévez. Gender Bias in Medieval Inquisitions and Its Place in Shaping Knowledge about the Heterodox. – Social Science History, 2026, 50, 1, 171–201. https://doi.org/10.1017/ssh.2025.10099.

45 D. Zbíral, K. Sikk, R. L. J. Shaw. Bridging Digital History Methods and Source Criticism: A Research Agenda for the Study of Inquisition Records. – Acta Historica Tallinnensia, 2025, 31, 1, 46–72. https://doi.org/10.3176/hist.2025.1.02.

46 C. Luckhardt. Mapping Women's Relational Networks in Carolingian Septimania, 795–850. – Early Medieval Europe, 2025, 33, 3, 412–435. https://doi.org/10.1111/emed.12779.

47 D. Zbíral, K. Riccardo, T. Hampejs, Z. Brys. Gender, Kinship, and Other Social Predictors of Incrimination in the Inquisition Register of Bologna (1291–1310): Results from an Exponential Random Graph Model. – PLOS ONE, 2025, 20, 2, e0315467. https://doi.org/10.1371/journal.pone.0315467.

48 K. D. Lilley. Urban Mappings: Visualizing Late Medieval Chester in Cartographic and Textual Form. – Mapping the Medieval City: Space, Place and Identity in Chester, c. 1200–1600. Ed. by C. A. M. Clarke. University of Wales Press, Cardiff, 2011, 19–41.

and 'between'. Translating these descriptions into GIS-compatible data produced a relative urban topography from evidence that had never been organised according to modern cartographic coordinates.[49] These projects illustrate a recurring problem in historical GIS: spatial analysis requires the formalisation of historical descriptions without imposing a precision foreign to the sources. Geocoded data can reveal density, proximity, accessibility and change over time, but coordinates are themselves scholarly propositions.[50] David Joseph Wrisley's mapping of place names in medieval French literature extends the same principle from reconstructed physical space to textual geography, allowing the geographical horizons of genres, circulation and literary production to be examined comparatively.[51]

Neither relational nor spatial formalisation removes uncertainty. A medieval name may identify several people or places, dates may remain ranges, and a proposed location may be only approximate. Data models need to preserve these alternatives and degrees of uncertainty. Historical gazetteers, linked data and knowledge graphs can connect variant names and competing identifications across projects, provided that interoperability does not erase local distinctions or conceal the evidential basis and degree of certainty attached to each assertion.[52]

ARTIFICIAL INTELLIGENCE, RETRIEVAL AND THE LIMITS OF GENERAL-PURPOSE MODELS

Artificial intelligence (AI) belongs within the longer history of computational medieval studies. Machine learning already underlies HTR, image classification, named-entity recognition and linguistic analysis. Generative AI has changed the terms of engagement by allowing

49 J. Colson. Reinterpreting Space: Mapping People and Relationships in Late Medieval and Early Modern English Cities Using GIS. – Urban History, 2020, 47, 3, 384–400. https://doi.org/10.1017/S0963926820000164.

50 See R. L. J. Shaw, K. Sikk, D. Zbíral. Toponymic Surnames and the Spatiality of Heresy Prosecutions: Peter Seila's Register of Sentences from the Quercy Region (Languedoc), 1241–1242. – Humanities and Social Sciences Communications, 2024, 11, 195, 1–14. https://doi.org/10.1057/s41599-024-02689-z.

51 D. J. Wrisley. Locating Medieval French, or Why We Collect and Visualize the Geographic Information of Texts. – Speculum, 2017, 92, S1, S145–S169. https://doi.org/10.1086/694300.

52 R. Simon, L. Isaksen, E. Barker, P. de Soto Cañamares. The Pleiades Gazetteer and the Pelagios Project. – Placing Names: Enriching and Integrating Gazetteers. Ed. by M. L. Berman, R. Mostern, H. Southall. Indiana University Press, Bloomington, 2016, 97–109. https://doi.org/10.2307/j.ctt2005zq7.12; A. Mertel, D. Zbíral, Z. Stachoň, H. Hořínková. Historical Geocoding Assistant. – SoftwareX, 2021, 14, 100682. https://doi.org/10.1016/j.softx.2021.100682.

scholars to interact with systems that produce fluent text, translations, summaries and code. Large language models can support exploratory searching, language work, information extraction and the creation of research tools, especially for scholars without advanced programming skills.[53]

For medieval studies, however, fluency is a poor proxy for historical reliability. Relevant languages and textual traditions are unevenly represented in training data, manuscript variation and editorial mediation are typically obscured in general-purpose systems, and plausible answers may contain fabricated references or silently modernised readings. Translation can smooth ambiguity and reproduce assumptions inherited from modern editions. Commercial systems create an additional problem because their provenance and reproducibility are difficult to establish.[54] Specialised language technologies can offer greater control over these procedures. Frameworks designed for pre-modern languages may document their models, preprocessing choices and linguistic assumptions and permit evaluation against defined corpora. Specialisation alone, however, guarantees neither transparency nor reproducibility. These depend on access to the relevant components, stable versions, adequate documentation and an evaluation appropriate to the task. The same criteria should guide the assessment of general-purpose models, whose usefulness for a particular operation must be established through testing rather than inferred from the fluency of their output.[55]

Retrieval over a curated corpus offers a more readily assessable experimental use of generative AI. A system can search a delimited collection of transcriptions, catalogue records or editions and return identifiable passages alongside a generated synthesis. This makes some errors easier to diagnose, although a cited passage may still fail to support the claim attached to it, and relevant evidence may be missed

53 Y. Fox. Using Generative AI in Historical Practice. Cambridge University Press, Cambridge, 2026. https://doi.org/10.1017/9781009740067.

54 On the fundamental limits of large language models as sources of knowledge, see E. M. Bender, T. Gebru, A. McMillan-Major, S. Shmitchell. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? – Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. ACM, New York, 2021, 610–623. https://doi.org/10.1145/3442188.3445922; W. H. Walters, E. I. Wilder. Fabrication and Errors in the Bibliographic Citations Generated by ChatGPT. – Scientific Reports, 2023, 13, 14045. https://doi.org/10.1038/s41598-023-41032-5; É. Ollion, R. Shen, A. Macanovic et al. The Dangers of Using Proprietary LLMs for Research. – Nature Machine Intelligence, 2024, 6, 4–5. https://doi.org/10.1038/s42256-023-00783-6.

55 K. P. Johnson, P. J. Burns, J. Stewart, T. Cook, C. Besnier, W. J. B. Mattingly. The Classical Language Toolkit: An NLP Framework for Pre-Modern Languages. – Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations. Association for Computational Linguistics, Online, 2021, 20–29. https://doi.org/10.18653/v1/2021.acl-demo.3.

during retrieval. For Digital Livonia we have so far kept general-purpose conversational AI outside the public interface. Our preferred direction is to investigate retrieval over curated collections, evaluating the selection of evidence separately from the accuracy of any generated answer and retaining links to records, passages and source images.

A similar emphasis on assessability appears in pedagogical uses of generative AI. Sophie Whittle's work on Chaucer asks students to compare AI-generated translations with manuscripts, editions and human-produced alternatives, identifying omissions, biases and anachronisms. Here, error itself becomes an object of criticism and a means of teaching editorial judgement. The model's output functions as material for comparison rather than as an authoritative interpretation.[56] A more research-oriented version of the same principle can be found in structured information extraction. When language models identify and classify evidence according to a closed annotation scheme, their performance can be evaluated systematically. A reproducible workflow records the model and its version, minimises stochastic variation and compares the results with a manually created gold-standard sample. Precision and recall then provide more meaningful measures than the apparent fluency of individual answers. Experiments with medieval inquisition records have used this procedure to distinguish testimony from procedural language, identify persons and groups, classify their grammatical and narrative roles, and assign reported actions to predefined semantic categories. Such evaluation leaves the model's internal opacity intact while placing empirical bounds on the reliability of its output.[57]

Across these examples, the location of interpretative work shifts. Decisions once concentrated in editions and historical arguments now also enter corpus construction, transcription schemes, data models and model evaluation. Computational methods alter the scale and organisation of comparison; their historical value depends on the evidential relations they make visible. For us, this shift requires criticism at three stages: the source itself, its transformation into data and the tools applied to those data. Zbíral, Sikk and Shaw's notion of 'source criticism 2.0' goes further by incorporating documentary production,

56 S. Whittle. Chaucer in the Age of AI: Critical and Collaborative Translation in Digital Medieval Studies. – Digital Culture & Education, 2026, 16, 3, 67–90. See also M. Schöffel, M. Wiedner, E. Garces Arias, P. Ruppert, C. Heumann, M. Aßenmacher. Modern Models, Medieval Texts: A POS Tagging Study of Old Occitan. – Proceedings of the 5th International Conference on Natural Language Processing for Digital Humanities. Ed. by M. Hämäläinen, E. Öhman, Y. Bizzoni, S. Miyagawa, K. Alnajjar. Association for Computational Linguistics, Albuquerque, 2025, 334–349. https://doi.org/10.18653/v1/2025.nlp4dh-1.30.

57 D. Zbíral, K. Sikk, R. L. J. Shaw. Bridging Digital History Methods and Source Criticism.

transmission, perspective and survival into computational analysis itself.[58] The Digital Livonia case that follows makes this redistribution of scholarly work visible in practice.

BUILDING A REGIONAL KNOWLEDGE INFRASTRUCTURE: THE DIGITAL LIVONIA PLATFORM

No single research platform can encompass all six domains outlined above. Digital Livonia concentrates on collections, editions, transcription, entity management and spatial access, while computational textual analysis and network analysis are principally downstream uses of exported data; general-purpose generative AI has not been incorporated into the public environment. Its contribution therefore lies less in embedding every analytical method than in creating documented and reusable points of passage between sources, structured data and external research tools.

Digital Livonia was established at Tallinn University in 2021 to create a common research environment for the history of medieval Livonia. Funded by the Estonian Research Council until 2025, the project brought together historians, art historians, archaeologists, manuscript specialists and researchers in digital technologies.[59] Its principal public outcome is a web platform available in both Estonian and English, combining digitised sources, digital source editions, research databases, practical applications, images of material objects and a growing set of services for searching, filtering, mapping and reusing them.[60] As participants in the conception and construction of Digital Livonia, we examine decisions, revisions and abandoned assumptions that are rarely visible in the finished interface. Our account draws on this experience rather than an independent evaluation of the platform's usability or research impact.

58 D. Zbíral, K. Sikk, R. L. J. Shaw. Bridging Digital History Methods and Source Criticism; R. L. J. Shaw, K. Suba, T. Hampejs, D. Zbíral. Syntactic-Semantic Capture of Historical Texts as a Platform for Source-Critical Analysis: Telling the Story of a Premodern Heresy Trial with Computer-Assisted Semantic Text Modelling (CASTEMO). – Digital Scholarship in the Humanities, 2026, 41, 2, 1001–1020. https://doi.org/10.1093/llc/fqag016.

59 The project (PRG1276) was led by Marek Tamm, and the core team included Rūta Bruževica, Inna Jürjo, Villu Kadakas, Tiina Kala, Kaspar Kolk, Juhan Kreem, Ivar Leimus, Anu Mänd, Hembo Pagi, and Karl Peeter Valk.

60 Digital Livonia: For a Digitally Enhanced Study of Medieval Livonia (c. 1200–1550), https://dl.tlu.ee/en/ (accessed 20 September 2026). Unless otherwise stated, all descriptions of the platform's current contents and functionality refer to the version consulted on that date.

The platform formed one component of a broader research programme organised around social and prosopographical studies, spatial and network analysis, intellectual and manuscript history, and economic history. The thematic databases and digital editions were developed within these research strands rather than produced according to a single platform-wide schema. This organisation helps to explain both the scholarly richness of the resulting collections and the differences between their structures, vocabularies and levels of completion.

A REGION CONSTITUTED BY DISPERSED EVIDENCE

Medieval Livonia presents a particularly revealing case for such an undertaking because the historical region no longer corresponds to a modern political, linguistic or institutional unit. Its territory is divided chiefly between Estonia and Latvia, while relevant sources are held in archives, libraries and museums across the Baltic region, Germany, the Vatican, Poland, Sweden, Denmark, Russia and elsewhere. Their dispersal reflects the political and ecclesiastical organisation of medieval Livonia, its connections with the Hanseatic world, the papacy and the religious orders, as well as the subsequent movement of archives, libraries and private collections. The study of the region has consequently depended on crossing modern state borders and combining source traditions organised according to different national historiographies and archival practices.

The surviving evidence is heterogeneous even within a single institution. The Tallinn City Archives, for example, preserves municipal registers, account books, correspondence, property records, guild statutes, wills and manuscript codices, whereas the National Archives of Estonia holds charters, seal impressions and various materials concerning landed property and ecclesiastical institutions. Museums preserve seals, coins, inscriptions, liturgical objects and works of art; libraries hold manuscript fragments and printed editions; archaeological records add buildings, urban plots, castles and excavated objects. Much of this material has been catalogued and/or digitised, but digitisation has usually proceeded according to the remit and technical system of the holding institution. A charter may therefore appear in one repository, its seal in another database, its printed edition in a nineteenth-century volume held by a

digital library, and the persons or places mentioned in it in a separately compiled scholarly repertory.

The fragmentation of sources affects both access and the kinds of questions researchers can readily pursue. Institutional catalogues generally reflect the boundaries of particular collections and source types, while historical research often requires connections across them. A study of a Tallinn merchant, for example, may draw on municipal registers, probate inventories, property transactions, correspondence, guild membership records, artworks and material held in foreign archives. Research into a cleric's intellectual life may similarly connect prosopographical information with manuscripts, university attendance, ecclesiastical offices and the circulation of particular texts. Where these sources can be searched only through separate institutional or disciplinary systems, such connections are harder to identify and their reconstruction demands considerable time and scholarly effort.

Digital Livonia was conceived in response to this problem. Its object was not to create a comprehensive digital copy of the 'medieval Livonian archive', which would have been technically, legally and intellectually unrealistic. The more practicable aim was to establish a common point of access where selected digital collections, newly prepared source editions and research data could be described according to compatible principles and, where possible, connected through shared historical entities. In this respect the project treated medieval Livonia as a field of evidence rather than a fixed territorial container. The coherence of the platform would derive from the relationships between people, places, institutions, events, documents and objects, even when the underlying resources remained in different repositories. Such an approach accords with a broader tendency in recent Livonian historiography to replace an image of the region as an isolated periphery with an analysis of the networks through which people, texts, objects and practices circulated.[61]

The platform's regional remit nevertheless exceeds its present coverage. Many of its collections concern Tallinn or the territory of present-day Estonia, and the public interface is available in Estonian and English. Although several datasets also include evidence from Latvia, this

61 See, for instance, M. Tamm. Deprovincialising Livonia: St. Catherine's Convent of Tallinn in Thirteenth-Century Dominican Networks. – Livland – eine Region am Ende der Welt? Forschungen zum Verhältnis zwischen Zentrum und Peripherie im späten Mittelalter / Livonia – a Region at the End of the World? Studies on the Relations between Centre and Periphery in the Later Middle Ages. Ed. by A. Selart, M. Thumser. (Quellen und Studien zur baltischen Geschichte, 27.) Böhlau, Köln, Weimar, Wien, 2017, 61–90; Making Livonia: Actors and Networks in the Medieval and Early Modern Baltic Sea Region. Ed. by A. Mänd, M. Tamm. Routledge, New York, London, 2020. https://doi.org/10.4324/9780429296000.

distribution reflects the location of the project team, its research interests and its institutional partnerships. Digital Livonia should therefore be understood as a growing contribution to the study of the region, with substantial geographical imbalances that need to remain visible to users. Extending coverage also requires decisions about collaboration, editorial responsibility and multilingual access.

FROM PROPOSAL TO PLATFORM

The first plans for Digital Livonia were ambitious. The platform was expected to aggregate digitised resources from several memory institutions, host new prosopographical and thematic databases, publish searchable source editions, support maps and network analysis, enrich imported metadata and provide a semantic search across different source types. The envisaged data model was to be compatible with the CIDOC Conceptual Reference Model, while Dublin Core and controlled vocabularies would provide a basis for metadata exchange. Several interfaces were imagined: a research environment for scholars, a more accessible public version and, eventually, educational or gamified applications. These aims were intellectually coherent, yet they assumed a degree of technical integration and standardisation that could not be achieved simply by selecting an existing content-management system.

An early formulation described Digital Livonia as a data project rather than a software project. The distinction expressed an important concern. Technologies age quickly, while historical data should remain usable after a particular interface or application has disappeared. A small research group cannot make the continued maintenance of a bespoke information system its principal scholarly task. The proposed solution was therefore to keep the data as independent as possible from the public website, to rely on open standards and established tools, and to store digitised materials in institutional repositories with stronger preservation infrastructures. These principles have remained central. The expectation that an adequate platform could be assembled largely from ready-made components, however, proved optimistic.

The difficulty became apparent when the project moved from general entities such as ‘person’, ‘document’ or ‘object’ to the structure of actual sources. A person mentioned in a municipal account book is not documented in the same manner as a knight brother whose career has been reconstructed from charters and office lists. A book may survive

as a physical codex, be known from a probate inventory, appear as a title reconstructed by a modern scholar, or combine all three conditions. Prices and wages introduce units, currencies, commodities, dates and conversion problems that bear little resemblance to the description of manuscripts or castles. A generic schema broad enough to accommodate every collection risked becoming too abstract for meaningful research, while a separate structure for each dataset would reproduce the fragmentation that the platform was intended to overcome.

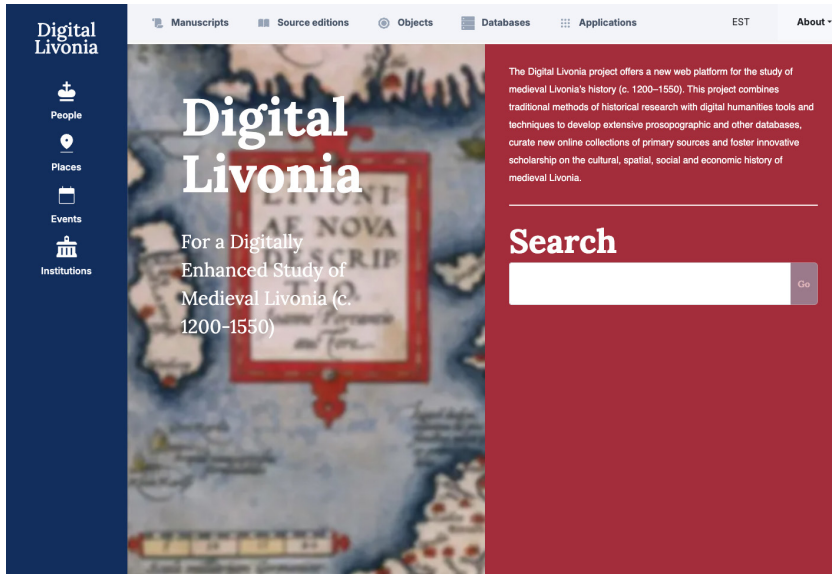
During development, we moved from planning comprehensive semantic integration to implementing connections incrementally. Digitised images, transcriptions, thematic databases, maps and entity indexes could become usable before every relation between them had been resolved. The present system relies principally on relational structures, controlled vocabularies and selected links between records. CIDOC CRM remains a conceptual reference for defining relations across collections, but a complete mapping was beyond the available resources. This was a decision about the extent and timing of implementation: relational storage can support source-specific assertions and subsequently expose selected information through CRM mappings or linked-data services.

This change reflects the need to accommodate developments in research as well as a move towards a simpler model. As scholarship advances, new sources may require unforeseen categories, identifications may be revised, and previously accepted relations may become uncertain. Fields designed to address one research question may also prove inadequate for another. A platform for historical research therefore needs enough structure to support comparison, together with the flexibility to revise its data model. The initial Digital Livonia model was understood as a working hypothesis, open to revision through further research. The resulting architecture aims to make existing collections usable and to support connections between them, without assuming that a single, comprehensive model can adequately represent the historical world to which they relate.

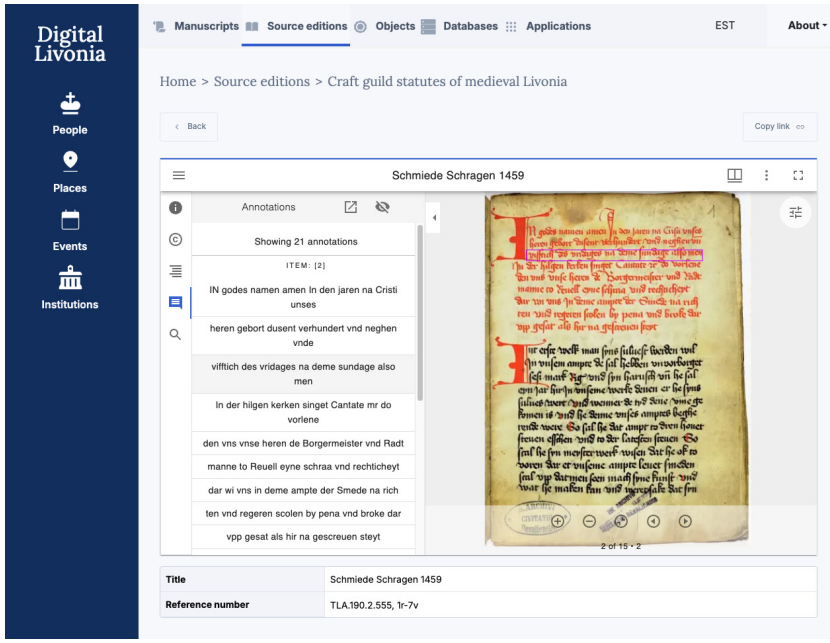
SOURCES, EDITIONS AND RESEARCH DATA

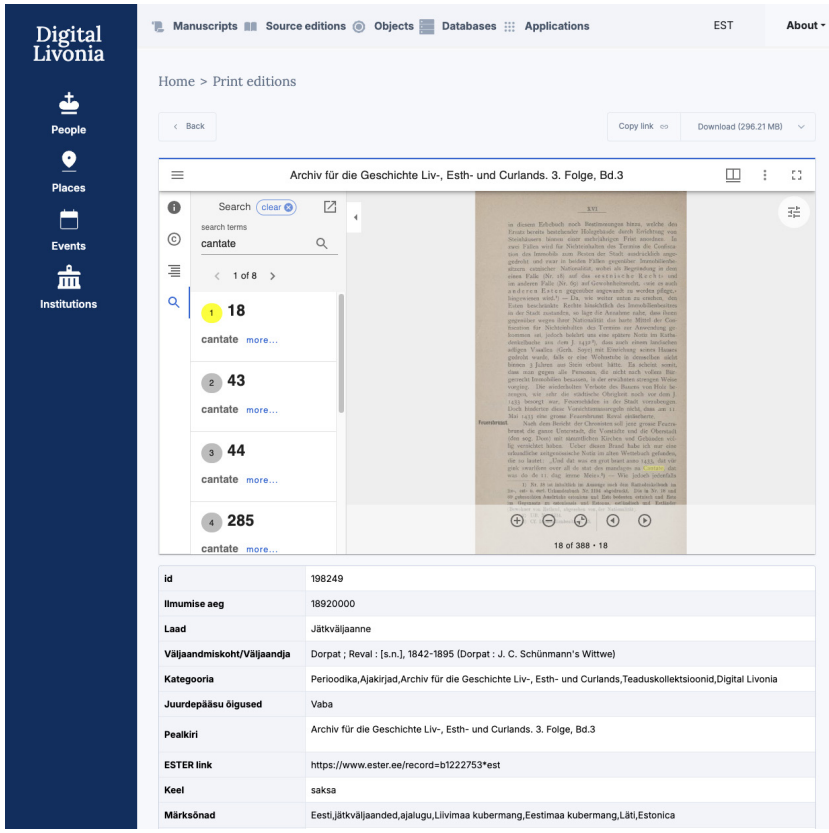
The public platform reflects this layered development. At the time of writing, its contents are organised into five principal sections: manuscripts, source editions, objects, databases and applications. The manuscript section contains seven collections, including the membership books of the Tallinn Brotherhood of the Black Heads, parchment charters and letters, manuscript fragments gathered under *Fragmenta membranea Estoniae*, Tallinn account books, registers of burghers, property and mortgage registers, and manuscript codices. The source-edition section comprises the probate inventories of medieval Tallinn, the statutes of craft guilds from Tallinn and Riga, the civic ordinances known as *burspraken* and digitised printed editions of Livonian sources. A separate object collection presents medieval Livonian seals.[62]

These categories describe different forms of digital representation. Some collections lead primarily to digitised images accompanied by catalogue information. Others combine images with manually prepared transcriptions and segmentation, while some editions use HTR in Transkribus to produce an initial transcription that is subsequently checked and corrected by researchers. The collection of printed source editions draws heavily on the digitisation infrastructure of the Tallinn



62 Digital Livonia, Manuscripts, https://dl.tlu.ee/en/parchments/; Source editions, https://dl.tlu.ee/en/source-editions/; Objects, https://dl.tlu.ee/en/objects/ (accessed 6 September 2026).


Digital Livonia
People
Places
Events
Institutions
Manuscripts
Source editions
Objects
Databases
Applications
EST
About
Home > Source editions > Craft guild statutes of medieval Livonia
Back
Copy link
Schmiede Schragen 1459
Annotations
Showing 21 annotations
ITEM: [2]
IN godes namen amen In den jaren na Cristi unses
heren gebort dusent verhundert vnd neghen vnde
vifftich des vridages na deme sundage also men
In der hilgen kerken singet Cantate mr do vorlene
den vns vnse heren de Borgermeister vnd Radt
manne to Reuell eyne schraa vnd rechticheyt
dar wi vns in deme ampte der Smede na rich
ten vnd regeren scolen by pena vnd broke dar
vpp gesat als hir na gescreuen steyt
2 of 15 • 2
Title
Schmiede Schragen 1459
Reference number
TLA.190.2.555, 1r-7v


Digital Livonia
People
Places
Events
Institutions
Manuscripts
Source editions
Objects
Databases
Applications
EST
About
Home > Print editions
Back
Copy link
Download (296.21 MB)
Archiv für die Geschichte Liv-, Esth- und Curlands. 3. Folge, Bd.3
Search
clear
search terms
cantate
1 of 8
1 18 cantate more...
2 43 cantate more...
3 44 cantate more...
4 285 cantate more...
18 of 388 • 18
id
198249
Ilmumise aeg
18920000
Laad
Jätkväljaanne
Väljaandmiskoht/Väljaandja
Dorpat ; Reval : [s.n.], 1842-1895 (Dorpat : J. C. Schünmann's Wittwe)
Kategooria
Perioodika,Ajakirjad,Archiv für die Geschichte Liv-, Esth- und Curlands,Teaduskollektsioonid,Digital Livonia
Juurdepääsu õigused
Vaba
Pealkiri
Archiv für die Geschichte Liv-, Esth- und Curlands. 3. Folge, Bd.3
ESTER link
https://www.ester.ee/record=b1222753*est
Keel
saksa
Märksõnad
Eesti,jätkväljaanded,ajalugu,Liivimaa kubermang,Eestimaa kubermang,Läti,Estonica

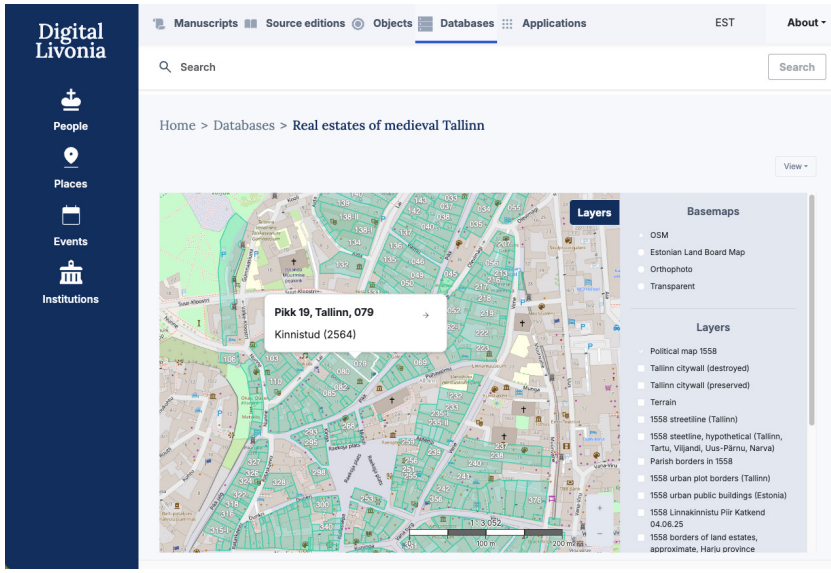


Figure 1. Selected interfaces of the Digital Livonia platform

University Academic Library and makes large nineteenth- and early twentieth-century publications available as searchable digital objects. The probate inventories, *burspraken* and guild statutes approach a digital edition more closely, since the images, transcribed text and structural divisions of the documents have been prepared together. The platform thus contains resources situated at several points between digital facsimile, searchable corpus, source edition and structured dataset. This diversity is appropriate to the sources, although it also means that the degree of searchability and analytical preparation varies from one collection to another.

Thirteen thematic databases are currently available. They concern real estate in medieval Tallinn, inhabitants of small medieval towns in present-day Estonia, inscriptions on Livonian artworks, prices, Tallinn scribes, books, castles, wages, the knight brethren of the Livonian branch of the Teutonic Order, Livonian clergy, manuscript codices, higher clergy and medieval Estonian manors.[63] Their subjects correspond to distinct fields of medieval research, yet together they illustrate the range of entities and relations that a regional platform must accommodate. The database of knight brethren reconstructs careers through offices

63 Digital Livonia, Databases, https://dl.tlu.ee/en/databases/ (accessed 20 September 2026). Since the individual databases continue to be supplemented, their descriptions should be understood as statements of their current coverage rather than claims to completeness.

and places, while the database of higher clergy records episcopal and capitular positions. The database of inhabitants of medieval towns contains prosopographical information on residents of towns in present-day Estonia, excluding Tallinn, covering eight chartered towns as well as Keila, Kuressaare and Lihula. The book database reconstructs Livonian book culture by combining surviving codices and early printed editions with documentary references to volumes that have since been lost, whereas the manuscript-codices database provides detailed codicological and textual descriptions of selected extant manuscripts. The inscription database relates verbal texts to material objects; the price and wage databases formalise amounts, currencies, commodities and labour; and the databases of castles, manors, and urban real estate link historical evidence to spatial locations and allow the data to be explored on an interactive map. The current scope and operational status of the databases are summarised in Table 1.

Table 1. The thematic databases currently available in Digital Livonia as of 20 September 2026

DATABASE	PRINCIPAL ENTITIES AND INFORMATION RECORDED	PUBLIC SCOPE AND PRESENT COVERAGE
Real estate of medieval Tallinn	Urban plots, owners and property transactions	Medieval Tallinn; coverage continues to expand
Inscriptions on medieval Livonian artworks	Inscribed texts, objects, languages and locations	Selected artworks from medieval Livonia; expanding corpus
Prices in medieval Livonia	Commodities, monetary amounts, currencies and dates	Coverage determined by the surviving source series
Wages in medieval Livonia	Forms of labour, payments, currencies and dates	Coverage determined primarily by surviving account books
The scribes of medieval Tallinn	Scribes, hands, documents and municipal offices	Documentary culture of medieval Tallinn
Knight brethren of the Livonian branch of the Teutonic Order	Persons, offices, careers and places	Livonian branch of the Teutonic Order

Table 1. Continued

DATABASE	PRINCIPAL ENTITIES AND INFORMATION RECORDED	PUBLIC SCOPE AND PRESENT COVERAGE
Medieval Livonian higher clergy	Bishops, canons, ecclesiastical offices and careers	Higher ecclesiastical officeholders in medieval Livonia; expanding corpus
Medieval Livonian clergy	Bishops, canons, priests, monks, friars, ecclesiastical offices and careers	Broader coverage of secular and regular clergy in medieval Livonia
Books in medieval Livonia	Surviving codices and early printed editions, together with documentary references to lost volumes	Medieval Livonian book culture; expanding corpus
Manuscript codices in medieval Livonia	Codicological and textual descriptions of surviving manuscripts	Selected Livonian manuscript collections; expanding corpus
Castles in medieval Livonia	Fortified sites, variant names, locations and historical information	Medieval Livonia; expanding corpus
Medieval Estonian manors	Sites, names, ownership and chronology	Territory of present-day Estonia; expanding corpus
Inhabitants of medieval Estonian small towns	Persons, origins, ownership, places	Towns in the territory of present-day Estonia; expanding corpus

Note: All databases listed were publicly operational on 20 September 2026; several remain under active expansion.

The platform currently includes two specialised applications that illustrate the formalisation required by digital historical work. A calculator for medieval Livonian currencies takes account of the fact that monetary values cannot be converted through a single stable ratio, since accounting systems and the silver content and relationships of units changed over time; it therefore distinguishes chronological periods and allows users to convert between the mark, ferding, schilling, öre, artig and pfennig according to historically defined equivalences.[64] A second

64 Digital Livonia, Medieval Livonian currency calculator, https://dl.tlu.ee/en/calculator/ (accessed 20 September 2026).

application, the weight calculator, translates between the principal weight units used in Livonian towns, using the Riga weight system as a common reference point while accounting for local variation in their actual values. It thus makes historically variable systems of measurement directly comparable across Riga, Tallinn, Tartu/Narva, Pärnu and Viljandi, turning a potentially cumbersome set of conversion rules into a reusable research tool.[65]

Several collections offer tabular interfaces through which users can search records, select columns, sort results, change views and download data. These functions are modest when compared with advanced analytical environments, but they mark an important transition from consultation to reuse. A fixed catalogue page presents the institution's description of an object; a configurable and downloadable table allows researchers to reorganise a corpus according to their own questions and carry it into statistical, network or geographical software. Digital Livonia does not perform every possible analysis within the platform. It seeks instead to provide sufficiently structured data from which further analysis can begin. The same principle informs our general indexes of people and places.[66] The people interface brings together records from several databases in a searchable and downloadable table, while the places section presents place records on a map organised by basemaps, layers and data sources. Our aim is to develop these interfaces into an entity-centred platform in which a person or place can provide a route between documents, objects and datasets. Cross-platform search for events and institutions remains under development, and links between the individual collections are being established progressively.

Digital Livonia currently supports parts of a source-centred workflow. The alignment of transcriptions with images allows readers to examine textual readings against the digitised document, while downloadable tables make selected research datasets available for further analysis. Movement between an individual source and related entries in other collections remains less consistent: some connections are supplied through links and indexes, whereas others require manual searching and historical identification. Following a charter through its image, description, named persons and associated objects remains a task shared between the platform and the researcher.

65 Digital Livonia, Medieval Livonian weight unit calculator, https://dl.tlu.ee/en/weight-calculator (accessed 20 September 2026).

66 Digital Livonia, People, https://dl.tlu.ee/en/databases/people/; Places, https://dl.tlu.ee/en/places/ (accessed 4 August 2026).

ARCHITECTURE AND INTEROPERABILITY

The technical architecture was designed around a separation between data management and public presentation. The principal data layer uses Directus, a headless content-management system placed above a relational database. A headless system stores and manages content without prescribing the form in which it must be displayed. The public frontend is a separate web application that requests data through an application programming interface. Directus generates REST and GraphQL API interfaces from the underlying database structure, including its relational fields.[67] The same data can consequently be used by the current website, a future interface, a map service or an external research application without being duplicated for each presentation. Figure 2 summarises the resulting division of responsibilities. Master files and repository metadata remain with memory institutions and ETERA; image conversion and delivery, transcription, structured data management and public presentation are handled by separate but connected services.

Separating these components allows the public interface to be redesigned without necessarily rebuilding the research databases, and permits other applications to use the same records. It also creates dependencies that require continuing maintenance: interfaces, permissions and software versions must remain compatible. Modularity therefore facilitates replacement and migration but does not itself

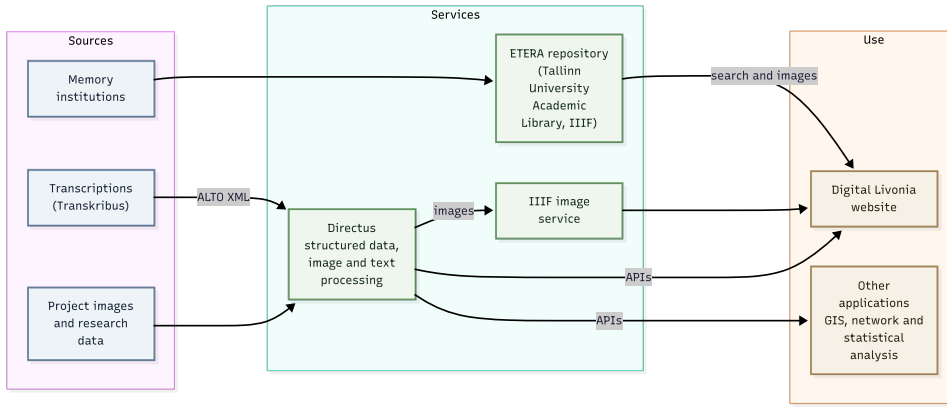


Figure 2. The modular architecture and principal data flows of Digital Livonia

67 Directus. Headless CMS. – Directus Documentation, https://directus.com/docs/getting-started/overview; Directus. API Reference. – Directus Documentation, https://directus.com/docs/api. REST – Representational State Transfer, https://en.wikipedia.org/wiki/REST; GraphQL – an open-source query language for APIs and a server-side runtime, https://graphql.org/ (accessed 20 September 2026).

preserve the data. Long-term reuse also requires durable exports, documented schemas and vocabularies, identifiable releases and an institution responsible for retaining them. These requirements apply especially to the project's original research datasets, whose preservation is distinct from the institutional custody of source images.

The data model had to represent sources that exist in several forms at once. A charter is a physical archival object with a shelfmark and material properties; it may have one or more digital images, a catalogue record, a transcription, a *regest*, a printed edition and relations to people, places or events. Each representation has its own provenance and may contain information absent from the others. Treating all of them as a single record obscures these distinctions, whereas separating them completely makes navigation difficult. Digital Livonia addresses the problem through relational structures that connect records while retaining their separate functions. Controlled vocabularies are used where stable comparison is required, although the treatment of variant medieval names and uncertain identifications remains uneven across collections. A fuller CIDOC CRM (Conceptual Reference Model) mapping would make the meanings of these relations more explicit beyond the local database. Its implementation requires each collection's categories to be examined, ambiguous relations to be represented and the provenance of imported information to be retained.[68] Our immediate priority has been to publish usable collections and establish selected connections between them. Further semantic alignment can proceed within this architecture, provided that the underlying records preserve the distinctions needed for subsequent mapping.

Images are handled through the International Image Interoperability Framework (IIIF). IIIF separates the delivery of an image from the application in which it is viewed. Its Image API allows a client to request a specified region, size, rotation, quality and format through a standardised address, while the Presentation API describes the sequence and structure of compound digital objects such as codices, registers or multi-page documents. Textual annotations associated with these objects can be searched through the Content Search API.[69] The practical advantage is that an image held by one institution can be displayed,

68 C. Bekiari et al. Definition of the CIDOC Conceptual Reference Model. Version 7.1.3. February 2024. https://cidoc-crm.org/sites/default/files/cidoc_crm_version_7.1.3.pdf (accessed 20 September 2026).

69 International Image Interoperability Framework Consortium. IIIF Image API 3.0, 2020, https://iiif.io/api/image/3.0/; IIIF Presentation API 3.0, 2020. https://iiif.io/api/presentation/3.0/; IIIF Content Search API 2.0, 2022. https://iiif.io/api/search/2.0/ (accessed 20 September 2026).

compared or annotated in another compatible environment without transferring the master file. For Digital Livonia, IIIF delivery draws on two sources. Printed source editions held in the Tallinn University Academic Library's ETERA repository are served through ETERA's own IIIF output, which supplies a durable institutional setting for this material. For images managed directly within Directus, the project assembled its own image-processing and delivery pipeline: a self-contained Docker service using IIPImage.[70] This supports local and S3-compatible storage, provides a IIIF Image API 3.0 endpoint and offers a REST interface for transcoding images. Since the service is independent of a particular content-management system, it can continue to deliver images if the Digital Livonia frontend or data-management layer changes. Two Directus extensions were developed to connect this infrastructure to the platform, one generating IIIF Presentation manifests and the other managing image delivery. Their release as open-source components allows the work undertaken for a regional project to be reused elsewhere.[71]

This architecture changes the status of the digital image. In a conventional website, an image is often an illustration embedded in a page and used only within that page. In an IIIF environment, the image becomes an addressable resource that can be cropped, enlarged, cited and brought into a different viewer. A manuscript page held by the Tallinn City Archives can be displayed beside a related fragment or parallel text served by another institution, provided both use the same standards. The larger promise is the virtual reconstruction of documentary contexts dispersed by the history of preservation. As the preceding discussion has emphasised, such reconstruction cannot restore the material object in all its dimensions, but it can create forms of comparison that would otherwise require repeated travel and the simultaneous cooperation of several holding institutions.

Transcription constitutes a further layer. Project texts are prepared in Transkribus, where document images can be segmented, transcribed, corrected and enriched with structural or textual tags. Transkribus permits export in formats ranging from plain text and searchable PDF to PAGE XML, ALTO XML, spreadsheets and TEI XML.[72] In the Digital

70 IIPImage is an advanced high-performance image server system for web-based streamed viewing and zooming of ultra high-resolution images, https://iipimage.sourceforge.io/ (accessed 20 September 2026).

71 IIIF Conversion and Image API End-point. GitHub repository, https://github.com/digital-livonia/iiif-convert-and-serve/ (accessed 20 September 2026).

72 Transkribus. Downloading: Formats Available to Store, Publish or Further Analyse Transcriptions. – Transkribus Help Center, https://help.transkribus.org/downloading/ (accessed 20 September 2026).

Livonia workflow, we convert ALTO XML into IIIF annotations, so that the resulting transcription is aligned with the corresponding images and prepared for IIIF presentation and content search. This procedure allows the manuscript or document to remain visible while its text becomes searchable and reusable. It also avoids making Transkribus itself the final publication environment. The transcription can move from the recognition and editing service into the platform, a digital edition or another analytical workflow.

Search operates across several technical layers. The platform queries its own structured datasets and selected content preserved in ETERA, the digital repository of the Tallinn University Academic Library. Through IIIF Content Search, annotations associated with the library's materials and texts transcribed within the project can also be searched. The original proposal envisaged a semantic search that would allow users to specify whether they were looking for people, places, events, institutions or other relations connected with a search term. Elements of this entity-centred approach have been implemented in the current people and places interfaces, but a uniform semantic search across all collections is still under development. Full-text search, field-based database search and semantic search operate at different analytical levels and therefore answer different questions. Finding the occurrence of a name in a transcription does not establish which historical person it denotes, whereas retrieving all records linked to an identified individual depends on prior scholarly work of disambiguation.

The public interface and the principal collection descriptions are available in Estonian and English, while the sources retain their historical languages and record-level translation remains uneven. Multilinguality therefore extends beyond the interface itself. The longer-term aim is to reconcile Latin, Middle Low German and vernacular forms with modern Estonian, Latvian, German, and English names without obscuring the form attested in the source. Variant labels are currently preserved within individual datasets, but a platform-wide multilingual authority layer has not yet been established. Developing such a layer will require stable entity identifiers, language-tagged labels, explicitly defined preferred and alternative names, and documentation of uncertain identifications.

The architecture is organised as a connected sequence of services rather than as a single self-contained system. Institutional repositories preserve and serve digital objects; IIIF standardises their delivery and presentation; Transkribus supports transcription; Directus manages structured research data and exposes them through APIs; the frontend

brings these services together for the user; and downloadable data can be transferred to GIS, network-analysis or statistical applications. In this respect, the architecture resembles the 'end-to-end' infrastructure described by Waters, although not every link in the chain has yet been completed and several transitions still depend on substantial manual mediation.[73] Its coherence lies in the movement of materials between services and in the standards that make this movement possible.

RESEARCHER WORKFLOWS AND THE HUMAN INFRASTRUCTURE OF DATA

The most instructive problems encountered during construction concerned neither APIs nor image servers, but the relationship between the data model and the working habits of researchers. The project team consisted primarily of historians, alongside art historians and archaeologists whose habitual tools were Word, Excel, personal notes and, in some cases, handwritten tables. Their expertise lay in reading difficult sources, identifying people and places, reconstructing chronologies and evaluating uncertain evidence rather than creating a database or a web application whose value depends on discrete, reusable and consistently structured records. They did not join the project in order to learn the internal logic of a complex database interface, and it gradually became clear that expecting them to adapt their work to such an interface would consume time better devoted to the sources.

The initial design assumed that researchers would enter information directly into forms reflecting the intended relational structure. In practice, some found the interface difficult and postponed data entry, while others used free-text fields or improvised conventions that circumvented the model. The resulting problem was not a lack of scholarly competence. The interface had exposed the technical structure at the point where researchers needed intellectual flexibility. A historian gathering information from an account book often does not know at the beginning which categories will prove significant, whether two names denote the same person, or how a doubtful date should eventually be represented. Premature formalisation can interrupt this exploratory process and encourage uncertain judgements to be entered as stable facts.

73 D. J. Waters. The Emerging Digital Infrastructure for Research in the Humanities, 88–92.

A more workable procedure began from the tools that researchers already used, especially spreadsheets. These already impose a model through their columns, rows and recording conventions, but allow that model to be revised during exploratory work. Their limitations became apparent when the data had to represent multiple relations, maintain consistent values or support collaborative revision. We increasingly used spreadsheets as provisional research environments from which records could be imported into relational collections through scripts and agreed procedures. This shifted the point at which stricter validation was required, while leaving historians and technical staff jointly responsible for checking the transformations. For this project, accommodating familiar working practices proved more effective than requiring researchers to adopt the database interface at the outset. It did not remove the need for common conventions or changes in how evidence was recorded. It also transferred work to those preparing imports, reconciling categories and checking transformed records. The arrangement therefore depends on procedures through which researchers can review those changes and return corrections to the maintained dataset. Our experience suggests that the timing of formalisation needs to be negotiated alongside responsibility for the work it creates.

The complexity of historical evidence need not be incompatible with the requirements of structured data. Organising evidence as data necessarily involves simplification, but it should preserve the distinctions on which historical interpretation depends.[74] A date may be exact, approximate or established within limits defined by two documented events; a place name may have several plausible identifications; and a person may appear under different names or be difficult to distinguish from a namesake. Absence from a source may likewise reflect a gap in the record rather than evidence of non-existence. Database fields need to accommodate these uncertainties without requiring researchers to claim greater precision than the evidence allows. The design of Digital Livonia benefited from a growing emphasis on how the system could represent the state of historical knowledge, including its uncertainties, which prompted reconsideration of how sources were accommodated within it.

This process also made visible the human infrastructure behind the platform. Every record depends on archival identification, transcription, normalisation, checking and decisions about correspondence between

74 J. Drucker. Humanities Approaches to Graphical Display. – Digital Humanities Quarterly, 2011, 5, 1. https://dhq-static.digitalhumanities.org/pdf/000091.pdf.

sources. Technical specialists must understand enough of the historical problem to avoid building formal relations that distort the evidence; historians must understand enough of the data model to recognise what their classifications will permit or exclude. Librarians and archivists contribute knowledge of collections, identifiers, rights and preservation. Open-source communities maintain the standards and software on which the platform relies. The finished interface conceals much of this labour, but its reliability is inseparable from it. Digital infrastructure consists as much of expertise, conventions, documentation and institutional trust as of code and servers.

DIGITAL LIVONIA AMONG MEDIEVAL RESEARCH PLATFORMS

Digital Livonia belongs to a wider ecology of portals and research environments created during the past three decades. We compare Digital Livonia with selected projects to clarify how its scale, architecture and scholarly functions relate to other approaches. Manuscriptorium, maintained by the National Library of the Czech Republic, represents the model of a large international aggregator. It brings together catalogue records and digitised documents from hundreds of institutions and supports interoperable access through IIIF.[75] Digital Livonia operates on a much smaller regional scale and combines aggregated material with databases and editions created through the project's own research; its emphasis therefore falls less on the breadth of discovery than on establishing relations between collections and historical entities.

Monasterium.net and Fragmentarium illustrate research environments organised around particular source types. Monasterium's collaborative archive provides access to more than half a million charters and related documents organised into archival fonds and research collections. Standardised metadata, the Charters Encoding Initiative and mechanisms for scholarly contribution allow documents from different institutions to be described, edited and reused within a common environment.[76]

75 Manuscriptorium: Gateway to Written Cultural Heritage, https://www.manuscriptorium.com/en/front-page-english/ (accessed 20 September 2026). On its new full-text edition module, see Manuscriptorium. Full-Text Edition Module, 29 May 2026, https://www.manuscriptorium.com/en/updates/-/full-text-edition-module/ (accessed 20 September 2026).

76 Monasterium.net: The Collaborative Archive, https://www.monasterium.net/mom/home/; Monasterium.net, Import Charters, https://www.monasterium.net/mom/import-charters/ (accessed 20 September 2026).

Fragmentarium addresses the particular problems posed by medieval manuscript fragments, supporting cataloguing, tagging, transcription and the virtual reconstruction of dismembered codices, with IIIF enabling images from partner institutions to be brought into a shared working environment.[77] Their concentration on particular source types supports specialised operations and shared descriptive conventions. Digital Livonia's regional remit requires it to connect more heterogeneous evidence, but both projects offer useful precedents for developing research functions within individual collections. Monasterium's collaborative architecture also suggests how controlled contributions, transparent revision histories and procedures for returning enriched information to holding institutions could support a continuing scholarly workspace.

Biblissima and the Finnish Sampo portals provide the closest comparisons for Digital Livonia's longer-term entity-centred ambitions. Biblissima aggregates catalogues and research databases concerning the transmission of texts and images from antiquity to the Renaissance, organising access around authority-controlled entities such as manuscripts, works, people, places, institutions and collections. Shared identifiers and IIIF services connect resources that retain their distinct institutional origins.[78] The Sampo model separates reusable linked-data services from the portals through which their contents are searched, visualised and analysed, using shared entities and semantic relations to connect heterogeneous datasets.[79] Mapping Manuscript Migrations offers a medieval example: it harmonises metadata on more than 220,000 medieval and Renaissance manuscripts, supporting the investigation of manuscripts, provenance events, owners and places through faceted search, analytical views and a reusable SPARQL endpoint.[80] These projects demonstrate the editorial and institutional investment required to connect independently constructed resources. Their scale exceeds Digital Livonia's present capacities, but their architecture provides a precedent for selectively aligning entities across its collections.

77 Fragmentarium. Digital Laboratory, https://fragmentarium.ms/pages/about/digital-laboratory/ (accessed 20 September 2026).

78 Biblissima: Observatory of Written Cultures, https://portail.biblissima.fr/en/; Biblissima, Data Sources, https://portail.biblissima.fr/en/data-sources/ (accessed 20 September 2026).

79 E. Hyvönen. Digital Humanities on the Semantic Web: Sampo Model and Portal Series. – Semantic Web, 2023, 14, 4, 729–744. https://doi.org/10.3233/SW-223034; Sampo Model, Data Services, and Series of Semantic Portals. Semantic Computing Research Group, Aalto University, https://seco.cs.aalto.fi/applications/sampo/ (accessed 20 September 2026).

80 E. Hyvönen, E. Ikkala, M. Koho, J. Tuominen, T. Burrows, L. Ransom, H. Wijsman. Mapping Manuscript Migrations on the Semantic Web: A Semantic Portal and Linked Open Data Service for Premodern Manuscript Research. – The Semantic Web – ISWC 2021. Ed. by A. Hotho et al. Springer, Cham, 2021, 615–630. https://doi.org/10.1007/978-3-030-88361-4_36.

People of Medieval Scotland offers an equally pertinent comparison from digital prosopography. Its factoid-based model represents statements about people, offices, places and relationships as source-specific assertions rather than merging them silently into a single modern biography. Conflicting or uncertain information can therefore remain attached to the evidence from which it derives, while the aggregated data can support maps, network analysis and linked-data reuse.[81] Its regional focus and source-centred treatment of identity are especially relevant to Digital Livonia, where people must frequently be reconstructed from variant names and fragmentary attestations distributed across independently created databases.

For Digital Livonia, these precedents suggest developing specialised functions within collections while connecting selected entities across them. The extent of integration will depend on the scholarly value of each connection and the editorial work required to establish and maintain it. The comparison also warns against assuming that digital availability is tied permanently to a particular technical implementation. The Digitized Medieval Manuscripts app (DMMapp), a widely used directory linking users to digitised manuscript collections, has announced that its current website will cease operation on 22 November 2026, although a successor version of the application has already been made available.[82] This case illustrates both the vulnerability and the potential resilience of digital scholarly resources: hosting costs and maintenance demands can render one implementation unsustainable, while open code, reusable data and careful documentation can facilitate migration to a new environment. Interoperable standards and open-source development therefore improve the prospects of continuity, although long-term sustainability still depends on the availability of technical maintenance, governance and institutional or community support.

81 A. Beam, J. Bradley, D. Broun, J. R. Davies, M. Hammond, N. Jakeman, M. Pasin, A. Taylor et al. People of Medieval Scotland: 1093–1371. Glasgow, London, 2019, https://poms.ac.uk/ (accessed 4 August 2026); M. Pasin, J. Bradley. Factoid-based Prosopography and Computer Ontologies: Towards an Integrated Approach. – Digital Scholarship in the Humanities, 2015, 30, 1, 86–97. https://doi.org/10.1093/llc/fqt037.

82 DMMapp: Digitized Medieval Manuscripts App, current site (https://digitizedmedievalmanuscripts.org/) and successor application (https://sexycodicology.github.io/Digitized-Medieval-Manuscripts-app/) (accessed 20 September 2026). The existing website announces that it will cease operation on 22 November 2026 and directs users to a newly available version of the application. See also G. Menna. The End of an Era: Saying Goodbye to the DMMapp after 12+ Years. – DMMapp Blog, 19 December 2025. https://blog.digitizedmedievalmanuscripts.org/the-end-of-an-era-saying-goodbye-to-the-dmmapp-after-12-years/ (accessed 20 September 2026).

LIMITS AND FUTURE DEVELOPMENT

The Digital Livonia platform now constitutes the first stable configuration of infrastructure that remains under development. Several of its present limitations reflect the history of its construction. Metadata varies in scope and granularity because the collections were assembled from sources created at different times and for different purposes. Some databases are comprehensive within a narrowly defined corpus, whereas others remain preliminary and explicitly anticipate the addition of further records. Relations between people, places, documents and objects have been established in selected collections but have not yet been normalised across the platform as a whole. A uniform semantic search, comprehensive CIDOC CRM mapping, and general indexes of events and institutions remain under development.

Evaluation has so far been formative rather than systematic. Feedback from project researchers shaped data-entry interfaces and workflows during development, but no published user study, citation analysis or public usage report is yet available. The next phase is intended to combine task-based testing with historians and collection custodians, structured feedback on failed searches and ambiguous records, and API and download metrics. Such measures would assess how effectively the platform supports reproducible research tasks rather than merely recording the number of visits to its pages.

These limitations are also methodologically revealing. A regional knowledge infrastructure does not become comprehensive simply through the accumulation of records. Its usefulness depends on decisions about identity, provenance, uncertainty and the meaning of relations. Before two occurrences of a name can be linked to a single person, their dates, offices, social connections and documentary contexts must be evaluated. Before a historical place can be mapped, variant names and possible locations must be reconciled without concealing doubt. Before data imported from another repository can be enriched, the original description must be distinguished from additions made within the project, and responsibility for each assertion must remain identifiable. The slow work of entity resolution and provenance is therefore central to the platform's further development.

The next phase will concentrate on connections that can be established and evaluated within the available resources. Selected people, places and institutions recurring across collections provide a starting point, but their identification must retain the supporting evidence and

any unresolved alternatives. Documentation of schemas, vocabularies, provenance and revision procedures is needed alongside this work so that exported records remain intelligible outside the platform. Multilingual labels and external authority links can then be developed without suppressing local forms. These improvements also require explicit responsibility for data correction, versioned releases, preservation and software maintenance beyond individual research grants. The FAIR principles provide a useful orientation for this work, since the platform aims to make its data findable, accessible, interoperable and reusable. Historical data complicate each of these terms. Findability depends on names that vary across Latin, Middle Low German, Estonian, Latvian and modern scholarly usage. Accessibility may be constrained by the rights attached to images held by different institutions. Interoperability requires mappings that preserve rather than erase differences between source traditions. Reuse depends on documentation sufficient to explain how a modern record was derived from medieval evidence. Maintaining these qualities requires continued work as records are corrected, collections expand and the platform changes.

Following the completion of the original research grant at the end of 2025, the platform has remained hosted by ETAIS, the Estonian Scientific Computing Infrastructure, operated as shared national infrastructure under the Estonian Research Infrastructures Roadmap, while the Academic Library continues to provide the ETERA repository and image-delivery layer. Its further development is currently supported by 'Digital Livonia 2.0: For a Digitally Enhanced Study of Medieval Livonia (c. 1200–1550)' (TUF25012), funded by Tallinn University. Its broader institutional setting now includes Tallinn University's participation in the Estonian Research and Cultural Data Infrastructure for 2025–2029. Further continuity will be provided by a new five-year research project funded by the Estonian Research Council, due to begin in early 2027 and devoted to the study of medieval Livonian sources in the Vatican Apostolic Archives through methods and services supported by the Digital Livonia infrastructure. Together, these arrangements strengthen the local environment in which the platform can be maintained and further developed, although a platform-specific governance policy still needs to be formalised.

Sustainability also requires a clear distinction between components that must be preserved and those that can be replaced. The public interface will age and should be expected to change. The underlying data, stable identifiers, documentation and links to source images have a longer

scholarly life. IIIF enables images to remain usable in other viewers; APIs allow structured data to be consumed by new applications; and institutional repositories provide a more durable home for master files than a project server. The modular architecture of Digital Livonia was designed with this distinction in mind. The longer-term measure of its success will be whether future researchers can understand, cite and reuse its materials after the present combination of software has ceased to be current.

CONCLUSION

The developments surveyed in this article show how digital methods have expanded access to medieval sources and altered the scale and organisation of their comparison. Their consequences extend into the constitution of evidence itself: imaging determines which material features become visible, transcription and encoding select what can be searched, and computational models establish the terms on which patterns are identified. Across these operations, source criticism must address both the historical formation of the surviving material and the procedures through which it becomes available for analysis. The history of digital medieval studies also shows that these questions have accompanied computer-assisted research from its beginnings, even as its instruments and institutional settings have changed.

The construction of Digital Livonia brings these methodological questions into the practical work of organising a regional research infrastructure. Our initial ambition for comprehensive semantic integration encountered differences between sources, datasets and researchers' working practices that could not readily be resolved within a common model. An incremental architecture allowed collections to become usable while their relationships remained under development. Accommodating familiar tools such as spreadsheets helped researchers contribute their material, but transferred further work to those preparing, checking and maintaining the data. These choices enabled publication and reuse while leaving substantial questions of entity identification, provenance, multilingual access and institutional responsibility unresolved.

Read together, the survey and the platform study suggest that a hybrid practice of medieval research depends on sustained engagement between digital procedures and philological, palaeographical,

codicological and historical judgement. Such engagement begins when sources are selected and represented and continues through modelling, analysis and publication. For regional infrastructures such as Digital Livonia, its effectiveness depends on whether researchers can follow the connections between sources and scholarly assertions, recognise uncertainty and revise earlier interpretations. Their lasting contribution will rest on keeping those connections intelligible and open to scrutiny as research questions, software and institutional arrangements change.

Acknowledgements

The research and writing of this article were supported by two projects devoted to the development of digital approaches to medieval Livonia: Digital Livonia: For a Digitally Enhanced Study of Medieval Livonia (c. 1200–1550) (PRG1276), funded by the Estonian Research Council, and Digital Livonia 2.0: For a Digitally Enhanced Study of Medieval Livonia (c. 1200–1550) (TUF25012), funded by Tallinn University.

DIGIPÖÖRE KESKAJAUURINGUTES: MEETODID, TARISTUD JA „DIGITAALSE LIIVIMAA" KOGEMUS

Marek Tamm, Hembo Pagi

Digipööre on viimase paarikümne aasta jooksul muutnud keskajauuringute peaaegu kõiki tööetappe, alates allikate leidmisest ja transkribeerimisest kuni nende analüüsimise, tõlgendamise ja publitseerimiseni. Digitaalsed keskajauuringud ei ole seetõttu enam üksnes kitsas eriala digihumanitaaria sees, vaid üha enam keskajauuringute üldine töökeskkond. Samas ei ole tegemist katkestusega senises uurimistraditsioonis. Digitaalsed meetodid on kujunenud pika ja ebaühtlase arengu tulemusena ning eksisteerivad kõrvuti trükikataloogide, kriitiliste väljaannete, arhiivitöö ja lähilugemisega. Artiklis eristatakse analüütiliselt kolme kattuvat tegevust: 1) digiteerimine teeb allikad ja uurimistulemused elektrooniliselt kättesaadavaks; 2) digitaalne struktureerimine võimaldab neid otsida, seostada ja taaskasutada; 3) arvutuslik analüüs rakendab tekstide, piltide ja ajalooliste andmete uurimisel statistilisi või algoritmilisi meetodeid.

Digitaalsete keskajauuringute kujunemislugu ulatub Teise maailmasõja järgsesse aega. Üheks selle sümboolseks alguspunktiks on Roberto Busa „Index Thomisticus", kuid oluline on silmas pidada, et arvutipõhiste keskajauuringute teke oli märksa mitmekesisem protsess. 1960.–1980. aastatel arendati mitmes Euroopa uurimiskeskuses keskaegsete ladina tekstide automaatseid konkordantse, ajaloolisi andmebaase ja leksikomeetrilisi korpusi. 1990. aastatel avasid personaalarvutid, internet ja teksti kodeerimise standardid uue etapi, mida esindasid näiteks projektid „Electronic Beowulf" ja „Piers Plowman Electronic Archive". Alates 2010. aastatest on esiplaanile tõusnud omavahel ühendatud töövood, kus digiteerimine, transkribeerimine, annoteerimine, analüüs ja publitseerimine moodustavad üha enam ühe tervikliku teadustaristu.

Digitaalsete kogude suurim vahetu mõju on olnud ligipääsu laienemine. Rahvusvaheline pildikoostalitlusraamistik IIIF võimaldab eri mäluasutustes hoitavaid käsikirju ühises keskkonnas vaadelda, võrrelda ja annoteerida. Digiteerimine ei kõrvalda siiski vajadust allikakriitika järele. Veebis kättesaadav materjal on paratamatult valikuline ja digitaalne kujutis muudab materiaalse objekti omadusi. Teisalt võivad multispektraalne fotograafia, pinnastruktuuri analüüs ja 3D-modelleerimine tuua nähtavale selliseid tunnuseid, mida palja silmaga on raske või võimatu märgata. Digitaalne allikakriitika peab seetõttu küsima nii kujutise loomise, metaandmete päritolu ja valikute kui ka identifikaatorite, kasutusõiguste ja säilitamise kohta.

Oluline osa digitaalsetest keskajauuringutest on seotud kriitiliste editsioonide ja käsikirjade masinloetavaks muutmisega. Digitaalne väljaanne ei ole pelgalt mahukam trükiväljaanne, vaid tekstipärimuse eksplitsiitne mudel, milles tuleb määratleda tekstiüksused ning lühendite, paranduste, ebakindlate lugemite ja eri käsikirjade omavahelised suhted. Käsikirjalise teksti automaattuvastus ehk HTR on omakorda märgatavalt laiendanud võimalusi suurte käsikirjakorpuste transkribeerimiseks, samas kujundavad transkriptsioonireeglid juba ette seda, millised käsikirja tunnused muutuvad hiljem arvutuslikult analüüsitavaks. Normaliseerimine võib lihtsustada teksti lugemist ja otsimist, kuid kaotada paleograafiliselt olulist teavet. Seetõttu on HTR kõige tulemuslikum inimese ja masina koostöös.

Masinloetavad tekstid võimaldavad uurida keskaja tekstikorpusi sellises ulatuses, mida klassikaline lähilugemine ei võimalda. Stilomeetria, autorluse määramine, teksti taaskasutuse tuvastamine, semantilised mudelid ja muud arvutuslikud meetodid toovad nähtavale tekstimustrid, mis on hajutatud sadade või tuhandete käsikirjade vahel. Nende

tulemused vajavad siiski hoolikat tõlgendust. Statistiline sarnasus võib tuleneda autorist, žanrist, kirjutajast, ümberkirjutamisest, toimetamisest või transkriptsioonivigadest. Sama kehtib võrgustikuanalüüsi ja ajaloolise GIS-i kohta. Võrgugraafi seos on juba tõlgendus: sugulus, kirjavahetus, üriku kaastunnistamine või inkvisitsiooniprotokollis mainimine ei ole samalaadsed suhted. Ka ajaloolise ruumi sidumine tänapäevaste koordinaatidega võib luua eksitava täpsuse. Hea andmemudel peab seetõttu säilitama ebakindluse ja võimaldama konkureerivaid tuvastusi.

Generatiivse tehisaru levik on toonud keskajauuringutesse uusi võimalusi ja riske. Suured keelemudelid võivad toetada otsingut, tõlkimist, infoeraldust ja töövahendite loomist, kuid väljundi ladusus ei ole usaldusväärsuse mõõt. Keskajal kasutatud keeled ja tekstipärimused on treeningandmetes ebaühtlaselt esindatud ning kommertsmudelite puhul jäävad sageli varjatuks kasutatud väljaanded, normaliseerimispõhimõtted ja mudeli täpne päritolu. Artiklis käsitletakse ühe katsetamisvõimalusena kureeritud korpustel põhinevat infootsingut, mis seob genereeritud vastuse kontrollitavate tekstikohtadega. Ka sellisel juhul tuleb eraldi hinnata leitud tõendite asjakohasust ja vastuse täpsust, sest viidatud tekstikoht ei pruugi väidet toetada ning oluline teave võib jääda leidmata. „Digitaalse Liivimaa" avalikku kasutajaliidesesse ei ole üldotstarbelist vestluslikku tehisaru seni lisatud.

Artikli teine osa analüüsib „Digitaalse Liivimaa" veebiplatvormi kui piirkondliku teadustaristu näidet. 2021. aastal Tallinna Ülikoolis algatatud projekti eesmärk on koondada ja seostada keskaja Liivimaa uurimiseks vajalikke hajutatud allikaid ja uurimisandmeid. Liivimaa on selleks eriti sobiv näide, sest ajalooline piirkond ei lange kokku ühegi tänapäevase riigi territooriumiga ning selle allikad paiknevad Eesti ja Läti kõrval Saksamaa, Vatikani, Poola, Rootsi, Taani, Venemaa ja teiste maade kogudes. „Digitaalse Liivimaa" avalik keskkond koondab digiteeritud käsikirju, digitaalseid allikaväljaandeid, materiaalseid objekte, kolmteist temaatilist andmebaasi ja kaks arvutuslikku rakendust. Andmebaasid hõlmavad muu hulgas kinnisvara, vaimulikke, Saksa ordu Liivimaa haru rüütelvendi, raamatuid, käsikirjakoodekseid, mõisaid, hindu, palku ja Eesti keskaegsete väikelinnade elanikke. Rahaühikute- ja kaalukalkulaator näitavad, kuidas ajalooliselt muutlikke mõõtühikuid saab formaliseerida ajaliselt ja piirkondlikult täpsete reeglite alusel. Platvormi arhitektuur on moodulipõhine: institutsionaalsed hoidlad säilitavad algfaile, IIIF tagab piltide standardiseeritud esituse, Transkribus toetab transkribeerimist, Directus haldab struktureeritud andmeid ning API-liidesed võimaldavad neid kasutada ka väliste GIS-i, võrgustiku- ja statistiliste töövahenditega.

Projekti üks olulisemaid kogemusi puudutab uurijate tööpraktikat. Algne eeldus, et ajaloolased sisestavad andmed otse keerukasse relatsioonilisse andmebaasi, osutus liiga piiravaks. Tõhusamaks kujunes töövoog, kus uurijad said esialgu kasutada tuttavaid vahendeid, eeskätt arvutustabeleid, ning andmed formaliseeriti hiljem tehniliste protseduuride ja spetsialistide toel. Sellest tuleneb üks projekti keskseid põhimõtteid: tööriistad tuleb kohandada uurijate tööviisidele enne, kui uurijatelt nõutakse kohanemist tööriistadega. Digitaalne taristu koosneb seega samavõrd inimestest, teadmistest, kokkulepetest ja institutsionaalsest vastutusest kui tarkvarast ja serveritest.

„Digitaalse Liivimaa" kogemus näitab, et piirkondliku teadustaristu loomisel tuleb otsustada, milliseid andmeid ja seoseid saab ühtlustada ning millised allikate ja uurimisküsimuste erinevused peavad säilima. Platvormi praegune sisu on geograafiliselt ebaühtlane ning kogude omavaheline sidumine alles jätkub. Edasise arenduse keskmes on valitud isikute, paikade ja institutsioonide tuvastamine eri andmekogudes, tõlgendusotsuste dokumenteerimine ning uurimisandmete säilitamise ja ajakohastamise vastutuse määratlemine.

Acta Historica Tallinnensia 32/2 https://doi.org/10.3176/hist.2026.2.02

The Historian and Modern Statistical Methods: Applications to Thirteenth-century Livonian Chronicles

Nicolas Bourgeois
Urban Radar, Fougères, France; nbourgeo@phare.normalesup.org

Abstract. The article examines the application of statistical and machine-learning-based tools to two thirteenth-century Livonian chronicles: the *Chronicle of Henry of Livonia* and the *Livonian Rhymed Chronicle*. It highlights both the potential and the limitations of these methods for the analysis of medieval texts. The article discusses key challenges, including sparse and multilingual corpora, difficulties of evaluation, and methodological complexity. The article argues that the principal value of increasingly complex computational methods lies less in autonomous historical interpretation than in their integration into transparent, verifiable, and collaboratively curated scholarly infrastructures. It therefore proposes principles for designing digital tools and metadata that are accessible, reliable, and open to continuous scholarly verification.

Keywords: medieval chronicles, medieval Livonia, lexicometry, clustering, topic modelling, embeddings, transformers

The pace of progress in statistical methods, particularly in the field of machine learning, often rebranded as artificial intelligence, has created a gap between the capabilities of these systems and researchers' ability to keep pace with their development, even when those researchers have a technical background. Many open systems are also presented through technical language, abstract concepts, or opaque interfaces that make

© 2026 Author. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International License CC BY 4.0 (http://creativecommons.org/licenses/by/4.0).

it difficult even for scholars with high levels of digital literacy to grasp their full potential – and their serious limitations. On the other hand, some tools now available to a general audience are remarkably easy to use, requiring little or no knowledge of their underlying mechanisms, while also producing highly convincing results. This combination creates both an urgent need for validation and a significant challenge, not only to determine how such outputs should be validated, but also to acquire and master the techniques required to do so effectively. As a result, researchers may underestimate what is possible, feel overwhelmed and avoid engaging, or, conversely, display excessive confidence in results they may not have the capacity to assess.

This article approaches these problems through two thirteenth-century Livonian chronicles: the *Chronicle of Henry of Livonia* and the *Livonian Rhymed Chronicle*. They provide useful test cases because they combine several difficulties characteristic of medieval textual corpora: limited size, different languages, unstable historical vocabularies, complex textual transmission, and substantial differences of genre, authorship, and communicative context. Rather than asking whether computational methods can replace close reading or historical interpretation, we ask more specifically: What can different levels of computational text analysis contribute to the study of medieval historical texts? Where do their evidential limits become apparent? And under what conditions can increasingly complex and opaque models be integrated responsibly into historical research?

Our central argument is that the principal value of such methods lies less in producing autonomous historical interpretations than in supporting the construction of transparent, verifiable, and collectively curated scholarly resources. Simpler techniques such as lexicometry, clustering, and topic modelling can provide useful exploratory perspectives when their assumptions remain visible to the researcher. More complex methods, including embeddings and large language models, may be harder to interpret directly, but they can still be valuable when used to generate, organise, or enrich metadata within controlled scholarly workflows. In this sense, computational models are most productive when understood not as substitutes for historical interpretation, but as part of a broader paratextual infrastructure comparable to indexes, regesta, critical apparatuses, and scholarly editions.

The article develops this argument in several stages. We begin with comparatively transparent methods based on word frequency, clustering, and dimensionality reduction, examining both their usefulness and their

limitations when applied to the two Livonian chronicles. We then turn to probabilistic topic modelling and embeddings, which allow more complex relationships between words and passages to be explored but also introduce greater interpretative and methodological difficulty. Finally, we consider the use of transformer-based language models for metadata generation and argue that their most promising role in historical research lies in collaboratively designed and carefully documented systems in which computational outputs remain open to scholarly verification and revision.

NATURAL LANGUAGE PROCESSING FOR TEXT ANALYSIS

Text analysis has evolved considerably over the past sixty years, as computer scientists have taken advantage of increasing computing resources to design algorithms that can treat texts less as bags of isolated words and more as complex structures in which each word derives its full meaning from its relation to others.[1]

The early stages of this evolution focused on incorporating the immediate context within a given text, from simple counts of co-occurrences to more in-depth analyses of grammatical structure. These methods have long been integrated into the toolkits used by historians familiar with quantitative methods.[2] In addition, statisticians and machine learning specialists have applied techniques such as supervised classification, clustering, regression, and dimensionality reduction to generate insights, metrics, and visualisations from large corpora of texts.[3]

By contrast, recent breakthroughs have focused on building statistical models of language itself, validated by their ability to perform some relatively easy-to-evaluate tasks such as sentence completion, rather than by any a priori understanding of linguistic structure. These advances were initially driven by broader tasks such as translation,

1 M. Piotrowski. Natural Language Processing for Historical Texts. (Synthesis Lectures on Human Language Technologies, 5). Springer, Cham, 2012. https://doi.org/10.1007/978-3-031-02146-6.

2 N. Bourgeois, M. Cottrell, S. Lamassé, M. Olteanu. Search for Meaning Through the Study of Co-occurrences in Texts. – Advances in Computational Intelligence. IWANN 2015. Lecture Notes in Computer Science, 9095. Ed. by I. Rojas, G. Joya, A. Catala. Springer, Cham, 2015, 578–591. https://doi.org/10.1007/978-3-319-19222-2_48.

3 N. Bourgeois, A. Pellet, M. Puren. Using Topic Generation Model to Explore the French Parliamentary Debates During the Early Third Republic (1881–1899). – Digital Humanities in the Nordic and Baltic Countries Publications, 2022, 4, 2, 35–51. https://doi.org/10.5617/dhnbpub.11259.

question answering, and image captioning, but were quickly applied to text analysis. The resulting tools, including recurrent neural networks and transformers,[4] have proven extremely powerful, but also increasingly difficult to interpret and evaluate.

As a result, historians now have access to a wide range of text analysis methods that are increasingly sophisticated, but also opaque and difficult to master. We argue that all of these techniques can be useful for researchers, though not in the same way. Statisticians should design tools that help historians make use of the most recent methods not only by making them easy to use, but, more importantly, by ensuring that they are easy to understand, evaluate, and safely integrate into scholarly workflows.

On the other hand, historians could benefit from developing a solid familiarity with earlier, more established techniques, since mastery of these provides the best safeguards for engaging with more recent approaches. We therefore advocate incorporating these technologies into a collective process aimed at building, proofreading, sharing, curating, and ultimately using secondary materials, rather than using them in stand-alone studies that are difficult for a wider scholarly audience to understand or replicate.

WRITTEN SOURCES FOR THIRTEENTH-CENTURY LIVONIA

In this article, we set aside non-literary evidence, such as archaeological or numismatic finds. This in no way means that such evidence cannot provide valuable material for automated analysis. This is the case not only because it often contains inscriptions, but also because image analysis can be fully integrated into language analysis. For instance, recent transformer-based vision–language models construct a latent space in which both text and images are jointly encoded. To give an intuitive example, an image of a sword would be assigned a representation close to

4 S. Hochreiter, J. Schmidhuber. Long Short-term Memory. – Neural Computation, 1997, 9, 8, 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735; A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin. Attention is All You Need. – Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, 6000–6010. https://dl.acm.org/doi/10.5555/3295222.3295349; J. Devlin, M.-W. Chang, K. Lee, K. Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. – Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Vol. 1 (Long and Short Papers). Association for Computational Linguistics, Minneapolis, 2019, 4171–4186. https://doi.org/10.18653/v1/N19-1423.

that of the word 'sword' itself.[5] However, incorporating these methods would add a significant layer of complexity to an article that is primarily intended as an introduction to the field.

With that restriction in mind, the documentary basis for studying thirteenth-century Livonia falls broadly into two categories: narrative accounts and normative documents. Some have been preserved in their original form, while others have come down to us only through later manuscript traditions and scholarly compilations, which may have introduced changes in vocabulary.

The most structurally informative sources are legal and administrative documents. These include charters, privileges, treaties, monastic or episcopal statutes, and, above all, an impressive corpus of papal letters, whose numerical prominence is undoubtedly related to the better preservation conditions of pontifical archives. Together, these sources form the backbone of our understanding of the early history of Livonia.[6] Most are written in Latin, although a significant number of German texts also survive. These documents vary considerably in their form, vocabulary, and scope, and thus constitute the closest, albeit much scarcer, equivalent to the corpora used to train large language models. For the same reason that we do not use archaeological sources, we do not include these texts in this preliminary study. Nevertheless, they should be regarded as valuable material for any serious quantitative text analysis and as absolutely necessary for any attempt to train a model on the language used in the chronicles.

This leaves us with narrative sources only. Among them, the *Chronicle of Henry of Livonia* remains foundational. Written in Latin in the 1220s by the priest Henry, it provides a detailed, near-contemporary account of missionary activity, conquest, and the establishment of ecclesiastical structures, as well as extensive information about the eastern Baltic societies of the time. It is undoubtedly the most detailed and best-studied text on medieval Livonia and, for these reasons, will also be the main focus of our study.[7]

5 A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, I. Sutskever. Learning Transferable Visual Models from Natural Language Supervision. – Proceedings of the 38th International Conference on Machine Learning. Vol. 139 of Proceedings of Machine Learning Research, 2021, 8748–8763. https://doi.org/10.48550/arXiv.2103.00020.

6 Liv-, est- und kurländisches Urkundenbuch nebst Regesten. Erster Band 1093–1300. Hrsg. von F. G. von Bunge. Kluge und Ströhm, Reval, 1853.

7 See Crusading and Chronicle Writing on the Medieval Baltic Frontier. A Companion to the Chronicle of Henry of Livonia. Ed. by M. Tamm, L. Kaljundi, C. S. Jensen. Ashgate, Farnham, 2011.

A complementary perspective appears in the *Livonian Rhymed Chronicle*, composed later in the thirteenth century in vernacular Middle High German poetic form. This text reflects the viewpoint of the successive military orders – the Brothers of the Sword and the Livonian branch of the Teutonic Order – and places greater emphasis on military events. Its narrative priorities and chronological precision differ considerably from those of Henry's chronicle, and individual episodes cannot always be reconciled with other surviving evidence. It is important to include it in our study because, while referring to many of the same or similar events, it employs a markedly different vocabulary, shaped not only by its vernacular language but also by its intended audience, purpose, and date of composition.[8]

For the purpose of this exploratory article, we use digital editions of the chronicles, which are based on the automated optical character recognition (OCR) of Leonid Arbusow's and Albert Bauer's 1955 edition of Henry's chronicle, referred to as HCL,[9] and Leo Meyer's 1876 edition of the *Livonian Rhymed Chronicle*, referred to as RC.[10] In the case of the RC, while the OCR quality is sufficient for a methodological introduction, any attempt to build reliable metadata should, of course, rely instead on a proofread digital edition, including an additional layer of linguistic disambiguation.

In the first part of this article, we focus on the possibilities and limitations of traditional lexicometric methods, including clustering and dimensionality reduction. These techniques can be used without an extensive mathematical background. In the second and third parts, we introduce more sophisticated tools based on robust statistical models, including latent Dirichlet allocation, latent space models, and transformers, which may be difficult for scholars to apply independently with a high degree of confidence. In the final part, we explore how the automated construction of metadata in a controlled context may help bridge this gap.

8 See A. V. Murray. The Structure, Genre and Intended Audience of the Livonian Rhymed Chronicle. – Crusade and Conversion on the Baltic Frontier, 1150–1500. Ed. by A. V. Murray. Ashgate, Aldershot, 2001, 235–251; A. V. Murray. Formulaic Language in the *Livonian Rhymed Chronicle*: Set Phrases and Discourse Markers in Middle High German History Writing. – Amsterdamer Beiträge zur älteren Germanistik, 2019, 79, 86–105; A. V. Murray. The *Livonian Rhymed Chronicle* as a Transitional Text: Formulaic Language in Middle High German Verse History. – Rilce: Revista de Filología Hispánica, 2020, 36, 4, 1324–1343. https://doi.org/10.15581/008.36.4.1324-43.

9 Heinrichs Livländische Chronik = Heinrici chronicon Livoniae. Bearb. von L. Arbusow, A. Bauer. Hahnsche Buchhandlung, Hannover, 1955.

10 Livländische Reimchronik, mit Anmerkungen, Namenverzeichnis und Glossar. Hrsg. von L. Meyer. F. Schöningh, Paderborn, 1876.

FROM WORD FREQUENCY TO CONTEXT ANALYSIS

At the most basic level of automated analysis lies lexicometry, an approach that treats a text as a collection of measurable units. In this framework, words are counted, compared, and statistically analysed without regard to grammatical structures or more sophisticated local statistical patterns, although the approach still has many advantages. It is easy to understand with a minimal mathematical background, and the errors it produces due to its rudimentary nature (homonyms, non-standard spelling, lack of contextual information) are equally easy to interpret, if not to prevent. It can be very useful at the exploratory stage for gaining insight into a large corpus, or at a more advanced stage to test a given hypothesis.

The combination of lexicometry with close reading and traditional textual analysis provides a compelling example of how historians can incorporate quantitative methods into the construction of an argument. Wojtek Jezierski has previously examined the vocabulary of emotions across the two chronicles, using word frequencies both to illustrate and to support his interpretation.[11] Rather than relying solely on raw counts, he carefully contextualises the data, thereby avoiding the common pitfall of drawing conclusions from frequency alone. Notably, his article also suggests that topic modelling could extend this line of analysis even further.

Table 1. Most frequent words in the two chronicles, excluding stopwords: HCL before lemmatisation (top), HCL after lemmatisation (middle), and RC without lemmatisation (bottom).

castrum	castro	episcopus	lyvones	terram	milicie	exercitum
236	146	140	133	131	130	122
castrum	episcopus	exercitus	frater	terra	christianus	magnus
451	404	296	240	207	162	161
meister	brûdere	komen	sprechen	heide	manche	landen
478	321	205	205	204	203	197

11 W. Jezierski. Politics of Emotions and Empathy Walls in Thirteenth Century Livonia. – Making Livonia: Actors and Networks in the Medieval and Early Modern Baltic Sea Region. Ed. by A. Mänd, M. Tamm. Routledge, London, New York, 2020, 113–142. See also W. Jezierski. Emotional Bonding and Trust during Sieges, Twelfth–Thirteenth Centuries. – Risk, Emotions, and Hospitality in the Christianization of the Baltic Rim, 1000–1300. Brepols, Turnhout, 2022, 175–205, esp. 192–203.

If we perform a raw count of words in the HCL, we quickly encounter a well-known problem: different conjugated or declined forms of a single root are counted separately. This can be addressed relatively easily through lemmatisation, that is, by grouping together the different inflected forms of a given word, as long as the spelling remains consistent (cf. Table 1). However, lemmatisation has many shortcomings: it handles variations in spelling poorly, struggles with words that are not included in dictionaries such as *lyvones*, and performs less well with computationally under-resourced historical languages, including Middle High German. More importantly, it does not group synonyms together or distinguish between homonyms. For instance, the words *dominus* and *frater* can have different meanings throughout Henry's chronicle, as becomes evident from the following:

> *Factum est autem eodem tempore, ut augeret Dominus de die in diem numerum et familiam fratrum milicie.* (HCL XI.3, pp. 48–49)
>
> It happened also at the same time that the Lord increased from day to day the number and retinue of the Brothers of the Militia.[12]

> *Signatis itaque per Theuthoniam multis hominibus signo crucis tandem domnus episcopus ad naves revertitur, assumpto secum fratre suo Rothmaro de claustro Sygeberge.* (HCL IX.6, p. 28)
>
> Having signed many men throughout Germany with the sign of the cross, at length the bishop returned to the ships, taking with him his brother Rothmar from the monastery of Segeberg.

TOPICS BASED ON FREQUENCY

However, with a chronicle such as Henry's, written over a short period by a single author, lexicometry can be useful for comparing different parts of the text. While reliance on a single word may be risky, statistics based on frequently co-occurring groups of words may prove more robust; the combination of several elements among *frater*, *miles*, and *magister* leaves little doubt, as does the association of *mare*, *navis*, and *tempestas*. Figure 1 shows that the prevalence of one topic or another is a more reliable indicator of where the action takes place than the frequency of a single word.

12 This and all subsequent translations from Henry's chronicle are taken from *The Chronicle of Henry of Livonia*. Trans. by J. A. Brundage. The University of Wisconsin Press, Madison, 1961.

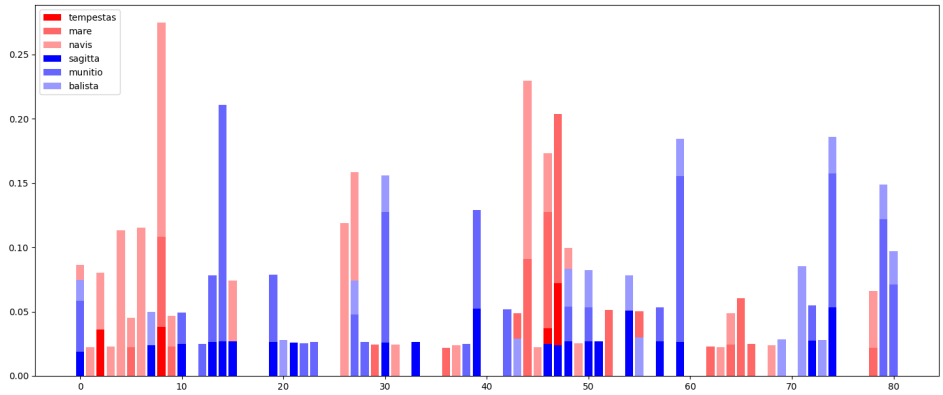


Figure 1. Comparative frequency of some words from the navigation topic (red) and from the missiles topic (blue), across consecutive two-page segments.

These groups of words, which we call topics, can be constructed automatically, without sophisticated mathematical tools, by clustering words based on their relative frequencies. Clustering is a generic technique used in data analysis to group similar data points together based on shared characteristics or patterns. In other words, we group together words that exhibit similar patterns of occurrence across the pages. There are numerous clustering techniques, such as *k*-means and agglomerative clustering, each of which has its own advantages and disadvantages depending on the structure of the data.[13]

Here, we apply a *k*-means algorithm, which groups together words with similar patterns of occurrence. For instance, according to Table 2, *balista* and *sagitta* are likely to be assigned to the same cluster, while *baptizo* will belong to a different one. For the sake of readability, we display raw counts rather than relative frequencies. However, in the computation we use term frequency–inverse document frequency (TF-IDF). TF-IDF is computed by multiplying a word's term frequency in a document by its inverse document frequency, $log(\frac{N}{df})$, where N is the total number of documents and df is the number of documents containing the word. As a result, common words receive lower weights, while distinctive words receive greater importance.

13 The scikit-learn documentation provides a useful overview of these various algorithms: https://scikit-learn.org/stable/modules/clustering.html (accessed 5/10/2026).

Table 2. Raw occurrence counts of selected words across text blocks 49–64.

block	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
balista	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
sagitta	0	1	1	0	0	2	0	0	1	0	1	0	0	0	0	0
baptizo	0	1	0	0	0	0	0	3	0	1	0	0	3	3	7	2

Note that choosing a clustering method based on an arbitrary division of the text means that one occurrence of the first word in block 55 versus two occurrences of the second word in block 54 may contribute to an increase in the distance between the two words, even though they refer to the same event. In fact, the two occurrences of *balista* are in the same paragraph:

> *De obsidione castri Wendorum. Descendebant quoque sagittarii fratrum milicie de castro suo, intrabant ad Wendos et* ***balistis*** *suis multos Ruthenorum interficiebant et quam plurimos vulnerabant. (...) Mane quoque facto rex Nogardie videns multorum nobilium suorum lesionem et aliorum interfectionem castrumque ipsum Wendorum se capere non posse considerans, cum sit tamen minus castellum, quod habet Lyvonia, pacifice loquutus est fratribus milicie. Qui pacem talem non curantes* ***balistis*** *eos a se repulerunt.* (HCL XXII.5, pp. 151–152)
>
> The archers of the Brothers of the Militia came down from their fort, got through to the Wends, and with their ballistas killed many Russians and wounded a great many of them (...). When it was morning the king of Novgorod, seeing that many of his nobles were wounded and others were killed, and considering that he could not take the fort of the Wends, though, indeed, it was one of Livonia's lesser castles, spoke peacefully to the Brothers of the Militia. They cared nothing for such a peace and drove the Russians away with ballistas.

In this case, a straightforward solution would be to use Henry's textual structure instead of page numbers, since the former is closer to a semantic division than the latter. However, this would not solve another problem, which is that successive chapters would not be considered closer to each other than any other pair. A more effective approach is to use a progressive distance between word occurrences. Instead of defining $d(x, y) = 0$ if x and y belong to the same entity (page, paragraph, etc.) and 1 otherwise, we can use something like $d(x, y) = \log(1 + n)$, where n is the number of words, sentences, paragraphs, or pages between the two occurrences. The choice of a logarithmic function reflects the idea that when two words

are very far apart, the exact distance matters little, whereas variations in close proximity are more significant.

In Table 3, we display the first six words of selected topics identified by the algorithm. We observe a clear proximity among words within each topic. Topic 8, for instance, is strongly centred around the baptism process, while topic 7 is clearly more concerned with military action. A large topic, topic 6, contains the vocabulary that is distributed fairly regularly throughout Henry's chronicle. This is an artefact of the method: since a word cannot belong to several topics at once, a very common word like *exercitus* or *castrum* cannot be attached to less common words from the same lexical field that follow a more irregular pattern. For the same reason, all topics are marginally present on every page, because some of their more common words might appear here and there, which makes the results quite noisy.

Table 3. Examples of topics identified by k-means over TF-IDF (first words, n=20). Topic 6 has been omitted since it contained more than half of the vocabulary and was the 'default' topic, while topic 2 only contained three very common words (omnis, qui, sum).

1	acceptus	captivans	circuiverunt	consequor	consolatio	fraternitas
3	apostolicus	audeo	benedicens	benigne	celestia	concilium
4	capio	congrego	duco	evado	fluvius	fuga
5	abhorrens	abicio	adaptatus	affigo	affirmo	antiquus
7	adoro	advocacias	affligens	affligo	andreas	animadverto
8	accipio	baptismum	beatus	captivus	christus	dominus

Table 4. The beginning of Henry's chronicle, with each word mapped to a topic. Some words (for example *ydolatrie*) were not lemmatised because they are absent from the dictionary and different forms of the same word may therefore be assigned to different topics. A deeper study of the chronicle could include pre-lemmatisation orthographic normalisation and/or adding text-specific entries to the dictionary.

(divina, 6) (providencia, 6) (memor, 5) (raab, 6) (babilonis, 6) (videlicet, 9) (confusus, 15) (gentilitas, 6) (nostris, 6) (modernus, 6) (tempus, 10) (volens, 9) (modernus, 6) (livones, 6) (ydolatras, 6) (ydolatrie, 14) (peccatum, 6) (sompno, 6) (taliter, 1) (ignis, 6) (amor, 6) (excitavit, 6) . . .

This method tends to place extra emphasis on specificity, which may introduce artefacts. Typically, if a group of words appears several times in one location and nowhere else, it is likely to form a cluster. For instance,

as shown in Figure 2 (upper graph), topics 1 and 3 consist almost entirely of the words found on just a couple of pages (respectively HCL XX.6, p. 138 and HCL XXIX.8, p. 214).

Even these poorly clustered topics can still offer useful clues. For instance, topic 16 (*archiepiscopus*, *aspergo*, *dacie*, *dania*, *imperium*, *osiliam*, *vironie*) maps onto specific parts of Henry's chronicle: mainly chapters XXIII and XXIV, which are indeed concerned with the conflicts between the kingdom of Denmark and the Swordbrothers over the conquest of Estonia (see Figure 2, lower graph).

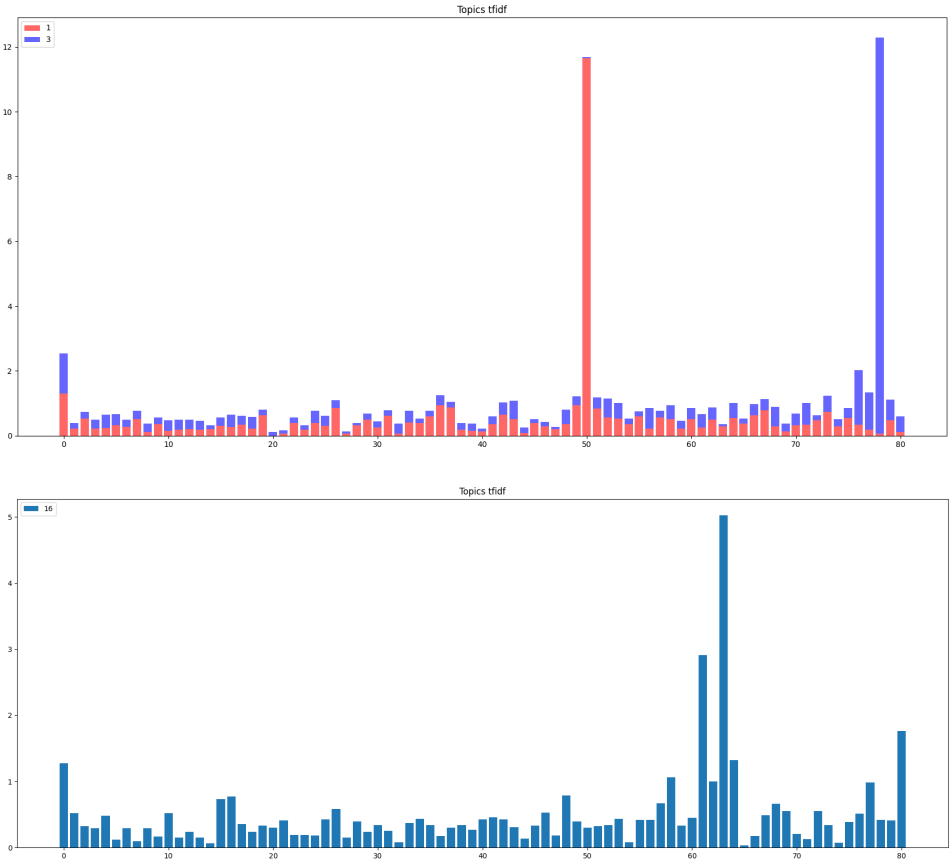


Figure 2. Comparative frequency of some TF-IDF topics across consecutive two-page segments.

CROSS-TEXT VOCABULARY

Lexicometry does not readily allow for cross-language analysis either. Certainly, in most cases, *frater* and *paganus* have the same meaning in Henry's chronicle as *brûdere* and *heide* in the *Reimchronik* – except when *frater* is used in the familial sense, for instance. In any case, direct lexicometric comparison is possible through word-for-word translation, which is often misleading. Moreover, while someone who reads both Latin and German may recognise that the following passages tell the same story, they frame it so differently that a word-count approach would not detect the correspondence.

> mit urloube daz geschach,
> bie der Dune ûf einen berc
> dâ bûweten sie ein êrlich werc,
> eine burc sô veste,
> daz die selben geste
> mit vride wol dar ûffe bliben
> und iren kouf lange triben.
> Ickesculle wart iz genant (...)
> Nû was als ich hân vernomen
> ein wîser man mit in komen,
> der in sanc und las,
> wen er ein reiner prîster was;
> der hêre hiez Meynhart. (RC, ll. 216–223, 229–233, p. 6)
>
> With native permission they built a worthy place on a mountain by the Dvina, a castle so large that these traders could remain there in peace and conduct trade for protracted periods of time. It was named Uexküll (...). Now, as I have heard it told, there had come with them a man who read and chanted the Mass for them. He was a holy priest, a man named Meinhard.

> *Fuit vir vite venerabilis et venerande caniciei, sacerdos ex ordine beati Augustini, in cenobio Sigebergensi (...) Accepta itaque licencia prefatus sacerdos a rege Woldemaro de Ploceke, cui Lyvones adhuc pagani tributa solvebant, simul et ab eo muneribus receptis, audacter divinum opus aggreditur, Lyvonibus predicando et ecclesiam in villa Ykescola construendo.* (HCL I.2–3, p. 2)
>
> In the monastery of Segeberg there was a man of worthy life, and with venerable gray hair, Meinhard by name, a priest of the order of Saint Augustine (...). After receiving, therefore, the permission of King Vladimir of Polozk, to whom the Livonians, while still pagan, paid tribute, and, at the same time, after receiving gifts from him, this priest boldly set out upon the divine work, preaching to the Livonians and building a church in the village of Uexküll.

Even texts written in the same language pose difficulties when they come from different social contexts or are written by different authors.

Clustering based on frequency is generally more effective at classifying authors than at classifying the events being described. In the next sections, we introduce two methods that can partially address these issues. Latent Dirichlet Allocation is particularly useful for identifying thematic structures across texts that share a language or sufficiently comparable vocabularies. Embeddings, by contrast, represent words or passages through their distributional contexts rather than through lexical identity alone. This makes them potentially useful for cross-linguistic comparison, but only when the representations from different languages are explicitly aligned or generated within a shared multilingual model.

LATENT DIRICHLET ALLOCATION

We have seen that the process of creating topics through clustering words based on TF-IDF variables has several shortcomings. One of them is that, by definition, clustering assigns a given word systematically to a single topic, even though it may belong to different lexical fields depending on the context (think again of *frater* and *dominus*). Another limitation is that this approach tends to favour monolithic blocks of text and struggles to account for the fact that several topics can coexist within the same statistical unit, such as a page or paragraph, depending on how the variables are constructed.

The Latent Dirichlet Allocation (LDA) topic model was first introduced in 2003.[14] It is a probabilistic generative model that represents each document as a mixture of latent topics, and each topic as a probability distribution over words. Here, a 'topic' denotes a latent statistical pattern inferred from the distribution of words across a corpus, rather than a predefined category or necessarily a theme consciously recognised by the author or reader.[15] In simplified terms, the model treats an observed text as if it had been generated by repeatedly selecting a topic from the document's topic distribution and then selecting a word from that topic's word distribution. A passage describing an expedition against the Lithuanians may therefore receive high probabilities for topics characterised by vocabulary relating to warfare and diplomacy, while receiving lower probabilities for topics dominated by seafaring

14 D. Blei, A. Ng, M. Jordan. Latent Dirichlet Allocation. – Journal of Machine Learning Research, 2003, 3, 993–1022. https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf.

15 This should come as no surprise to historians, who are accustomed to developing concepts that make sense for studying a given period, even though the people living at the time would certainly never have used them themselves – feudalism, for instance.

or ecclesiastical terminology. The analytical task is therefore to infer, from the observed distribution of words, both the word distributions associated with the latent topics and the mixture of topics associated with each textual unit. Since exact inference is computationally intractable, these distributions are estimated using approximate inference methods; here we use the variational approach introduced by Blei et al.

Latent Dirichlet Allocation has been widely used in many areas of the humanities and social sciences, as it is a very powerful tool for extracting information from a large corpus in an unsupervised context, that is, when classes are not defined a priori. In practice, topic models tend to produce more readily interpretable results when the corpus contains a sufficiently large number of textual units, when individual units concentrate on a limited range of topics, when topics recur across multiple units, and when there is enough shared vocabulary across the corpus for recurring patterns to emerge. Newspaper collections are a common example of such a corpus. Some historical corpora, including certain cartularies or collections of texts produced within a relatively coherent institutional or linguistic context, may also display these characteristics, although this depends strongly on their internal heterogeneity and on how the texts are segmented. In Table 5, we present an excerpt from the topics generated through LDA on Henry's chronicle.

Table 5. Examples of topics identified by LDA (most significant words, number of topics = 10).

1	mare	ecclesia	mater	magnus	dominus	christus	episcopus
2	exercitus	paganus	peregrinus	fides	frater	lyvones	episcopus
3	seculum	fides	verbum	abbas	legatus	christus	rego
4	expeditio	exercitus	villa	terra	equus	vir	castrum

Once the topics are generated, they can be used as new variables for the study of vocabulary. This drastically reduces the size of the variable space from more than 10,000 word forms in this case to only ten topics, thus making visualisations possible, albeit with a significant loss of information. We can, for example, study the intensity of topics over time or analyse the correlations between them.

Let us now consider the distribution of our first four topics across Henry's chronicle (Figure 3). In contrast to the topics produced through clustering, we observe much less noise: a small number of

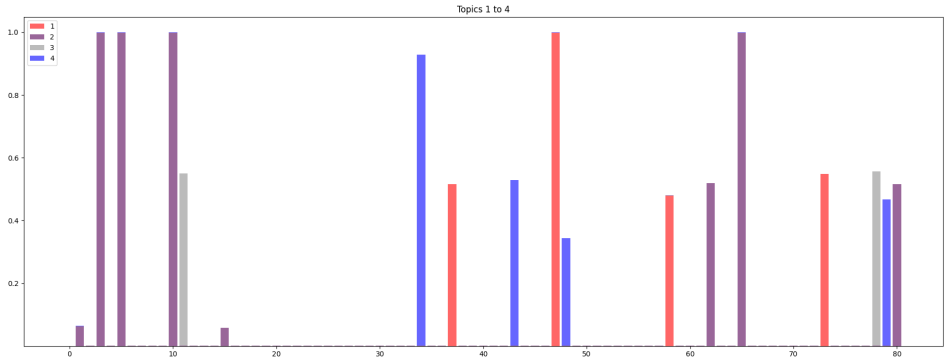


Figure 3. Comparative frequency of some LDA topics across consecutive two-page segments.

topics predominate within each textual unit, and their distribution is more clearly localised. Topic 2, for instance, which contains primarily vocabulary relating to crusading armies, is especially prominent during the initial expedition (HCL II–X) and during the conquest of Estonia (HCL XXIII and XXIX). By contrast, topic 1 appears mainly in sermons in the second part of the chronicle, while topic 3 concerns two specific episodes in which negotiators are sent between rival powers (HCL IX and XXIX).

EMBEDDINGS

One way of moving beyond direct lexical comparison is to represent words according to the contexts in which they occur. Embedding methods provide such representations, although they do not in themselves eliminate differences of language, genre, authorship, or style.

An embedding is a numerical representation of a linguistic unit (such as a word, sub-word, sentence, or passage) as a vector in a multidimensional space. Different model families construct such representations in different ways. Static word-embedding models, such as Word2Vec and GloVe, assign each word type a single vector learned from patterns of co-occurrence in a corpus. Contextual language models, including transformer-based models, instead generate representations that depend on the surrounding text, so that the same word may receive different vectors in different contexts. Encoder–decoder models constitute a further architecture in which such internal representations are used to transform one sequence into another, for example in machine translation.

Instead of treating each word[16] as an isolated symbol, embeddings place words in a continuous vector space where their positions reflect semantic and syntactic relationships. In the following examples we use a static word-embedding approach. Word2Vec learns vector representations by exploiting patterns of word co-occurrence in a corpus, so that words occurring in similar linguistic contexts tend to occupy nearby regions of the resulting vector space.[17] In Figures 4 and 5, we show examples of such embeddings applied to the vocabulary of Henry's chronicle and the *Livonian Rhymed Chronicle*, respectively.

A latent space is the abstract, multidimensional space in which these vectors reside. The term 'latent' refers to the fact that the dimensions of this space are not directly interpretable; they are learned features that capture underlying patterns in the data rather than explicit, human-defined properties. Each dimension may encode a mixture of linguistic traits, such as formality, plurality, or semantic category, although not in a way that can be clearly labelled. What matters is not the meaning of any single dimension, but the geometric relationships between points. Distance and direction in this space correspond to similarity and relational structure.

t-SNE (t-distributed stochastic neighbour embedding) is a widely used method for visualising high-dimensional data in two dimensions. Its principal strength lies in preserving local neighbourhood relationships: points that are close in the original high-dimensional space will often remain close in the two-dimensional projection. Unlike PCA (principal component analysis), t-SNE is non-linear and is therefore capable of representing complex local structures that may be difficult to visualise using linear projection. This advantage comes with an important limitation, however: distances between widely separated points, the relative positions of clusters, and even the apparent size of clusters in a t-SNE plot should not be interpreted as faithful representations of the original space. The visualisations in Figures 4 and 5 should therefore be

16 Whereas words can be directly embedded in a latent space, modern language models generally operate on tokens rather than words. Tokenisation differs fundamentally from lemmatisation: whereas lemmatisation reduces each word to a common lexical form while leaving the syntactic structure of the sentence unchanged, tokenisation represents a sentence as a sequence of tokens. A token is an abstract unit that typically corresponds to part of a word, although it may sometimes correspond to an entire word or, in some cases, to several words. Crucially, the division of a word into one, two, or even several tokens is not guided by grammatical or linguistic considerations, but is determined primarily by the statistical distribution of linguistic units in the corpus on which the model is trained.

17 L. van der Maaten, G. Hinton. Visualizing Data Using t-SNE. – Journal of Machine Learning Research, 2008, 9, 2579–2605. https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf; T. Mikolov, K. Chen, G. S. Corrado, J. Dean. Efficient Estimation of Word Representations in Vector Space. – International Conference on Learning Representations, 2013. https://doi.org/10.48550/arXiv.1301.3781.

read primarily as exploratory representations of local proximity rather than as maps of the global structure of the embedding space.

During training, the model processes text and attempts to predict relationships between words. Each word begins as a random vector, but through repeated exposure to text and iterative optimisation, the

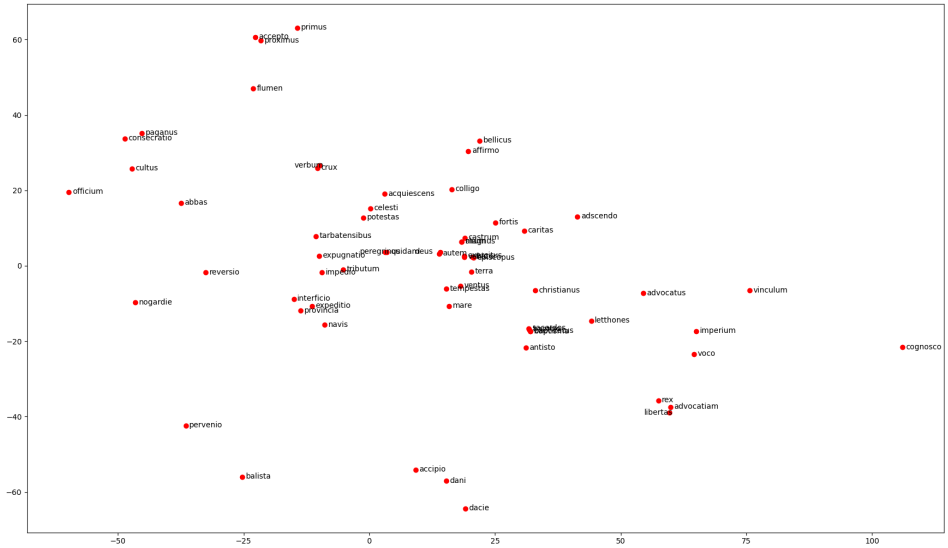


Figure 4. t-SNE projection of selected words from HCL. Local proximity may indicate similarity in distributional context; global distances in the two-dimensional projection should not be interpreted directly.

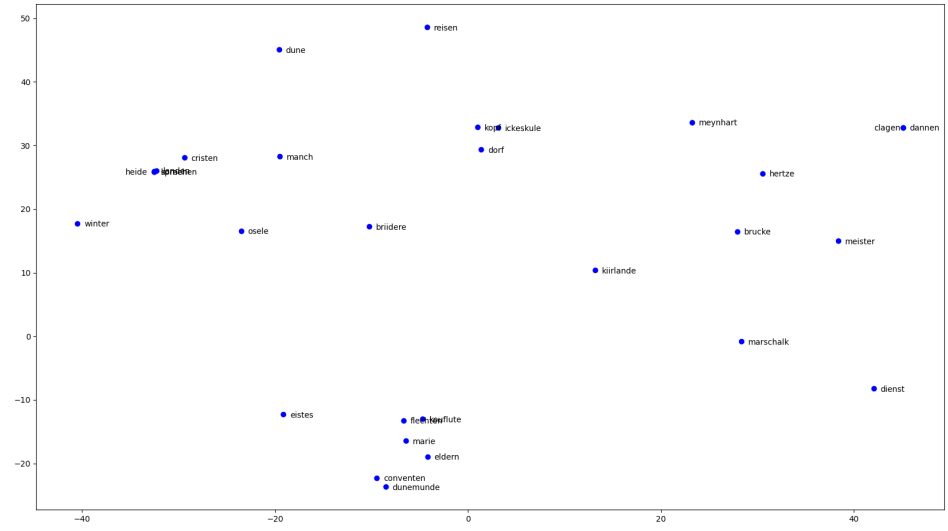


Figure 5. t-SNE projection of selected words from RC.

vectors gradually adjust to capture meaningful patterns. Over time, coherent structures emerge: words such as *conventen* and *dunemunde* end up close to each other in the space, and relationships such as gender or tense distinctions can be encoded as consistent vector offsets. If we had sufficient material for more extensive training and time to explore the results in greater depth, we might even hope to discover arithmetic relationships similar to those observed in English word embeddings, where subtracting the vector for 'man' from that for 'king' and adding the vector for 'woman' yields a vector close to that for 'queen'. In a similar spirit, one could imagine relations such as *rex* + *ecclesia* − *miles* ~ *episcopus*, although this does not currently emerge in our data.

It is indeed important to keep in mind that, while it is possible to train embeddings on a small corpus for the purpose of clustering, training a model capable of more complex tasks such as question answering requires significantly more resources and training data. Pretrained language models capable of processing Latin are increasingly available, although their performance varies according to period, genre and task. Comparable resources for Middle High German remain more limited, particularly for the kinds of context-sensitive tasks considered here.

SIMILARITIES BETWEEN CHRONICLES

The problem becomes more difficult when we attempt to compare passages across the two chronicles. A direct lexicometric comparison is strongly affected by language and vocabulary. Even when both chronicles are analysed through their English translations,[18] TF-IDF tends largely to separate the two texts rather than identify passages describing comparable events (Figure 6). The following experiment therefore adopts a different strategy. Rather than directly comparing independently trained Latin and Middle High German embeddings, we construct thematic representations separately for the two chronicles and then identify approximate correspondences between topics on the basis of their most significant words. The resulting comparison is thus partly model-based and partly researcher-mediated: the models identify patterns within each corpus, while the correspondence between thematic

18 The Chronicle of Henry of Livonia. Trans. by J. A. Brundage; The Livonian Rhymed Chronicle. Trans. by J. C. Smith, W. L. Urban. Indiana University, Bloomington, 1977.

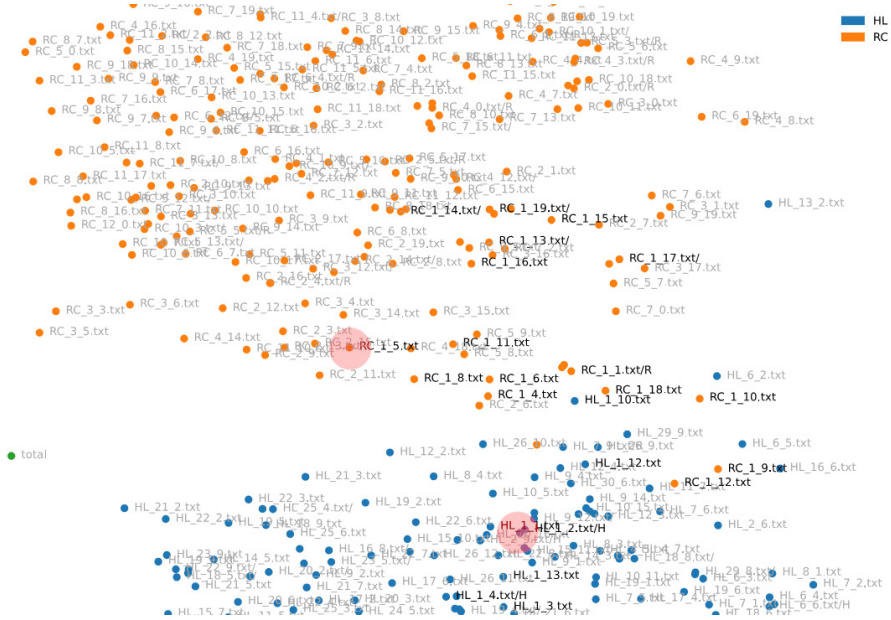


Figure 6. t-SNE projection of chapters from Henry's chronicle and the Livonian Rhymed Chronicle based on word-level TF-IDF in the English translations. The pink circles show the position of two texts describing the same event (the arrival of Meinhard).

structures across the two languages is established through historical and linguistic interpretation.

Once these approximate thematic correspondences have been established, passages from the two chronicles can be compared in terms of their topic composition. Most passages remain associated primarily with one chronicle, but several cases display comparable thematic profiles across the two texts. These similarities should not be understood as having been discovered through a language-neutral embedding space; rather, they emerge from a combination of statistical modelling within each corpus and scholarly alignment between the resulting topics.

The passages represented by neighbouring red and blue points in Figure 7 deal with warfare against Russian princes, as illustrated by HCL XXII.2 (pp. 148–149).

If we look closely at the topic distribution in this chapter, we see that topic 4 accounts for more than 90 per cent of the estimated topic mixture in this passage, which deals with warfare, the description of armies, and Russian powers. Similar topics are omnipresent in the *Rhymed Chronicle*, hence the association.

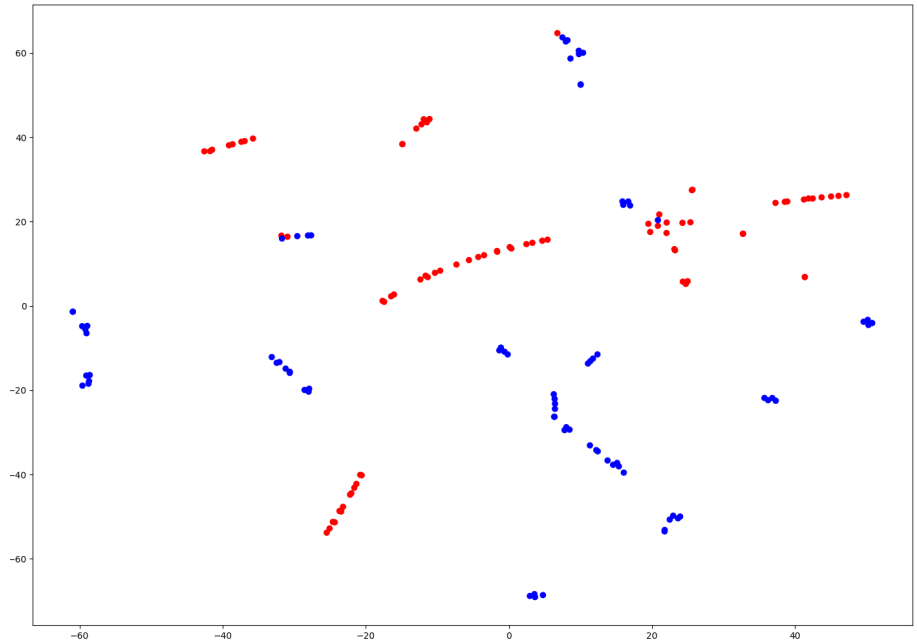


Figure 7. t-SNE projection of pages from Henry's chronicle (red) and the Livonian Rhymed Chronicle (blue) based on thematically aligned LDA topic representations. Topic models were generated separately for the two corpora; cross-language correspondences between topics were established manually.

VERIFICATION OF RESULTS

Evaluating the results produced by a statistical tool is a problem as old as statistics itself. An important part of the data scientist's work consists in designing and applying control procedures throughout the analytical process, both to assess and improve the results and, ultimately, to provide an estimate of the validity of the final outcome. This process rests essentially on two pillars: metrics and control samples.

Data-science textbooks provide well-established principles for designing robust evaluation metrics. A simple accuracy score, for instance, can be misleading when the classes are unbalanced. A rate of 95 per cent correct classifications may not be particularly impressive when distinguishing sentences that do or do not refer to King Philip of Swabia in Henry's chronicle (just answer 'no' to all and you have it), whereas it would be a perfectly respectable result for identifying paragraphs containing a reference to Holy Scripture.

It is often easier to identify false positives (is the selected paragraph actually about baptism in Vironia?) than false negatives (can I be sure

that no passage dealing with siegecraft has been discarded?). The former requires only a quick examination of the selected passage, whereas the latter may require an exhaustive inspection of the material that was not selected. Since both false positives and false negatives need to be assessed with equal care, the construction of control samples is essential. Such samples must be large enough to be representative, yet small enough that a reliable ground truth can be established with a reasonable investment of effort.

In data analysis, many evaluation metrics are automated for the sake of efficiency. In historical research, however, scholars will often rely on what might be called 'manual' metrics, for instance, their own assessment of the quality of a translation.[19] Although manual evaluation is more time-consuming, it does not fundamentally alter the principles governing the construction of samples and metrics.

Once an evaluation framework has been established, the algorithm can be modified and re-run iteratively until its performance on the control sample reaches a satisfactory level. If substantial changes are made to the model or its design during this process, however, it is safer to revise the control sample periodically, in order to reduce the risk of overfitting to it. Finally, a separate test sample, which has not been used at any stage of the development process, can be used to compute the final evaluation metrics and provide an independent assessment of the results.

STATISTICAL MODELS AS PARATEXTUAL TOOLS

Throughout the modern period, archivists and scholars have sought to develop tools such as printed editions of manuscripts, compilations, marginal notes, indexes and regesta. All were designed to spare future researchers the need to begin the study of manuscripts from scratch and to facilitate the gradual construction of a collective understanding of historical sources.

Emerging progressively from medieval glosses and early modern scholarly conventions, these tools became increasingly standardised between the seventeenth and nineteenth centuries, enabling historians to organise vast amounts of information, reference sources, and construct

19 More complex architectures can blur this distinction, particularly when language models are used to evaluate the outputs of other language models. Such approaches may raise difficult questions even for machine-learning specialists, and most historians are unlikely to need to venture into this particularly treacherous territory.

arguments with greater transparency and rigour. Indexes allowed readers to access themes non-linearly, while footnotes and references created a traceable network of verifiable evidence, provided that the original materials were preserved. In many ways, these paratextual systems externalised the historian's method, making visible the processes of selection, citation, and interpretation.

This has helped historians work with increasingly large corpora while grounding their claims in systematic surveys of the sources rather than impressions derived from reading a limited number of texts. However, it was not without drawbacks. Researchers have long emphasised the importance of maintaining direct contact with original materials: studying palaeography in order to read manuscripts directly, treating regesta as second-hand material, systematically verifying comments and references, and publishing regularly updated editions intended to correct errors and incorporate advances in scholarly knowledge. Nevertheless, very few historians today would argue for abandoning the tools of critical editions altogether.

The large-scale digitisation of textual materials and the rise of the internet can be understood as a contemporary parallel to the tradition of critical editions of sources, extending their core aims while transforming their scale and accessibility. Digital archives further the purpose of earlier critical editions by providing searchable transcriptions and layered metadata that allow users to trace provenance, compare versions, and examine editorial decisions. Online platforms also make it possible to juxtapose multiple editions or manuscripts instantly. As with earlier paratextual tools, the promise of accessibility and critical transparency depends on maintaining rigorous standards, as well as on developing digital literacy within the scholarly community, in order to avoid drawing unfounded conclusions from partially understood materials.

In this context, statistical methods – and especially more recent approaches based on large language models – occupy a particular position, since it is virtually impossible for the vast majority of historians to acquire the highly specialised skills required to fully understand and critically assess the way their outputs are produced. While palaeography, Latin, archival classification systems, XML encoding, and even Python scripting can be integrated into a historian's initial training to the point of making the use of digital materials both safe and fruitful, statistical methods can only be mastered up to a certain point. Even specialists working in the field of machine learning must constantly update their knowledge, navigating a landscape in which the dissemination of research is often

hindered by economic competition and by the difficulty of maintaining an accessible and up-to-date body of knowledge within the community.

However, indexes and marginal notes were certainly very useful to historians who could not access original manuscripts directly. As long as they were familiar with primary sources in general and understood the process of critical editing, there was little need for them to decipher the particularly difficult handwriting of a given fifteenth-century lawmaker when working from a printed or digital transcription. In a similar way, we argue that a solid understanding of basic statistical tools may be sufficient for scholars to make effective use of enhanced textual editions that incorporate more advanced statistical methods, provided these tools are designed so that the additional information remains easy to interpret and verify.

ML-ENRICHED METADATA

Many readers are probably already familiar, if not with the specific conventions of TEI-XML, then at least more generally with the use of tags to integrate metadata into the digital edition of a historical source. This practice is a direct evolution of the footnotes and marginal comments of earlier printed editions, designed to be revealed or hidden on demand, and to be easily searchable using software tools. These tags may be used to indicate information about the original manuscript's appearance, its current location, or its successive versions. They can also provide references to secondary materials or help categorise texts within a broader corpus, among many other uses.

```
< chapter type="chapter" n="1.4" >
< topics allocation="lda8tfidf42" >
  < topic id="1" p="0.7"/ >
  < topic id="5" p="0.3"/ >
< /topics >
< head > (4) < /head >
< p > And in the same village < fn > Ükskülll, nowadays Ikšķile < /fn > Ylo, the father of Kulewene,
and Viezo, the father of Alo, were the first to be baptized, while the others followed in their turn. < p >
...
< allocation id="lda8tfidf42" method="LDA" n="8" >
  < topic id="1" >
    < word text="baptized" p="0.632"/ >
    < word text="holy" p="0.611"/ >
    < word text="baptism" p="0.598"/ >
  < /topic >
  < !-- Additional topics omitted -- >
< /allocation >
```

Figure 8. Example of a metadata-enriched edition using custom XML markup to link a chapter to topic-modelling results.

These metadata-enriched editions may represent the most effective integration of statistical tools into a historian's engagement with sources. Consider, for instance, a scenario in which a data scientist, in collaboration with a historian specialising in the relevant field, produces a topic model of a given corpus. The results of this modelling can then be integrated into a digital edition in the form of metadata (Figure 8).

In this example, we use the tag "topics" to indicate the existence of such an analysis, with the attribute "allocation" referring to a methodological section where the details are clearly specified. In this section, the method and its parameters are documented in order to ensure the process is reproducible. In addition, the full results are presented within the same methodological framework: lists of words for each topic, along with the model's confidence in the assignment of each item.

Thus, a scholar interested in identifying all texts in a corpus dealing with, for instance, medieval piracy or the ordination of canons could use these topic-based metadata through a search engine in much the same way as they would use a traditional index. Conversely, another scholar with deeper mathematical literacy and a specific interest in a topic not captured by the initial metadata could adjust the parameters until a topic relevant to their research emerges from the process.

COLLABORATIVE GENERATION OF METADATA WITH LLMS

When it comes to generating metadata for historians, topics – whether produced through frequency-based clustering or Latent Dirichlet Allocation – represent only one form of computationally generated metadata. For most uses, they may already be outdated or superseded by more recent methods. Beyond the validity of their theoretical framework, their main limitation is that their output is difficult to adapt to the specific questions a researcher may wish to ask, much like traditional indexes, which were often restricted to personal names, place names, or core concepts. By contrast, large language models (LLMs) offer historians a powerful means of generating and organising metadata at scale.

Large language models can automatically extract key attributes such as dates, locations, personal names, topics, and genres, even when the original documents are unstructured or written in archaic languages. Their ability to perform such tasks derives from large-scale language modelling, through which they learn contextual patterns and

associations that can be mobilised through natural-language instructions. In addition, the use of natural language allows researchers to be far more directive in shaping the form and scope of the output. For instance, it is relatively straightforward to ask a model to produce a table of all sieges mentioned in a chronicle, with precise references to the relevant passages. The accuracy and verifiability of the resulting tables, however, remain separate questions. As a simple illustration, we asked a locally run Llama 3 (8B) model to extract keywords, a task at which such models perform reasonably well, since it primarily requires an understanding of local context, and to append them as tags at the end of paragraphs in Henry's chronicle:

> Anno domini 1198 venerabilis Albertus, Bremensis canonicus, in episcopum consecratur. Post consecrationem estate proxima Gothlandiam vadit et ibidem circa quingentos viros signo crucis ad eundum in Lyvoniam signat. Inde per Daciam transiens munera regis Canuti et ducis Woldemari et Absolonis archiepiscopi recepit.
>
> <keywords>1198, consecration of bishop, Albert of Bremen, crusading recruitment, Gotland, Livonia mission, Denmark, King Canute, Duke Valdemar, Archbishop Absalon, medieval church networks</keywords>
>
> Reversus in Theuthoniam in natali domini Magdeburch plures signat. Ubi rex Philippus cum uxore coronatur. Et coram eodem rege in sentencia queritur, si bona in Lyvoniam peregrinancium sub tuicione pape ponantur, sicut eorum, qui Ierosolimam vadunt. Responsum vero est ea sub protectione apostolici comprehendi, qui peregrinacionem Lyvonie in plenariam peccaminum remissionem iniungens, vie coequavit Ierosolimitane.
>
> <keywords> 1199, church assembly, recruitment of crusaders, coronation of king, papal protection, legal status of pilgrims, Livonian crusade, comparison to Jerusalem pilgrimage, indulgences, ecclesiastical authority </keywords>
>
> Anno secundo sui episcopatus cum comite Conrado de Tremonia et Harberto de Yborch cum multis peregrinis in Lyvoniam vadit, habens secum in comitatu 23 naves.
>
> <keywords> 1200, second episcopal year, expedition to Livonia, Count Conrad of Dortmund, Harbert of Iburch, crusading fleet, maritime campaign, medieval colonisation, missionary expansion </keywords>

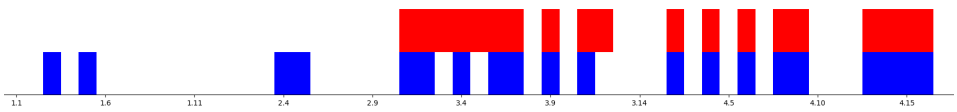


Figure 9. Chapters in Henry's chronicle assigned the keyword 'seafaring' in two runs of an LLM using slightly different prompts.

Although they perform reasonably well overall, models may introduce inaccuracies, such as confusion over dates, or overly confident

interpretations, for instance, assigning "colonisation" as a keyword in the last paragraph, which could be debated but is certainly not supported by that specific passage. More subtly, they can produce outputs that appear broadly appropriate but contain irrelevant or weakly informative items; for example, "second episcopal year" would probably not constitute a meaningful keyword. Figure 9 illustrates how the chapters assigned a particular keyword can vary with the prompt.

While topic models, for instance, provide probabilities of classification that can be useful for identifying errors, raw token probabilities generated by large language models do not provide straightforward or well-calibrated confidence estimates for the factual correctness of generated metadata. Moreover, proprietary systems may not expose these probabilities to users at all. As a result, such probabilities cannot by themselves be used as reliable indicators of whether generated metadata is historically accurate, and the outputs therefore require independent validation. Conversely, these models can be highly effective when used to generate metadata within a well-defined and controlled process.

While a single historian or data scientist could easily be misled by an LLM's output, it is important to recognise that the construction of trustworthy metadata is a collective and incremental process. The *Liv-, Esth- und Curländisches Urkundenbuch*, initiated by Friedrich Georg von Bunge, was corrected, curated, and expanded by many scholars over the following century. Leonid Arbusow and Friedrich Benninghoven, in particular, systematically revised its coverage of the period up to 1300,[20] but many additional pieces of information have since been disseminated through the scholarly literature. Before the rise of language models, digitisation and collaborative editions already helped accelerate this process, enabling proofreading, the inclusion of new materials, and the addition of references to occur much more rapidly under the supervision of the scholarly community, although the mechanisms that make such supervision reliable remain, in themselves, the subject of extensive debate.

Large language models amplify both the possibilities of these tools and their associated challenges. The time required for generation – even including prompt design and preliminary review of small samples – is generally much shorter than the time required for validation. Depending on the specificity of the language, the size of the corpus, the need to

20 Liv-, est- und kurländische Urkundenregesten bis zum Jahre 1300. Hrsg. von F. G. v. Bunge, mit Ergänzungen von L. Arbusow jun., neu herausgegeben von F. Benninghoven, Geleitwort von P. Johansen. Hamburg, 1959.

incorporate external information or the possibility of framing the question in precise terms, the use of LLMs for metadata generation may or may not be justified.

Moreover, language models, as well as more conventional software and statistical tools, can be useful in the incremental improvement of such metadata. For the specific task of generating keywords alone, a non-exhaustive list of applications includes standardising keyword sets, clustering similar names, removing hapaxes, and compiling references. The metadata generated can thus be processed using a range of established digital analysis tools.

At a later stage in the workflow, scholars can use an LLM with a large data repository more effectively if the underlying texts have been enriched with metadata. In 2024, retrieval-augmented generation (RAG) was the cornerstone of such techniques, although many architectures combining the model itself with function calling and various search engines can now be used, particularly with models capable of handling large input contexts. In any case, these systems generally benefit from curated metadata, much as classical search tools do. Thus, alongside traditional software, LLMs can be useful at both ends of the process, provided that strong mechanisms of control are maintained in each case.

Before concluding, it is worth emphasising that this article has explored only the tip of the iceberg when it comes to LLM-assisted research in history, which is a very rapidly expanding topic. A wide range of applications are currently being developed and critically examined, many of which the author has not tried himself at the time of writing. Readers interested in exploring this rapidly evolving field may wish to begin with Andres Karjus's review of some of the most recent and promising experiments in the area.[21]

CONCLUSION

This article has examined what different levels of computational text analysis can contribute to the study of medieval historical texts, while also emphasising the methodological limitations that accompany them. Applied to the *Chronicle of Henry of Livonia* and the *Livonian Rhymed Chronicle*, the methods discussed here show that computational analysis

21 A. Karjus. Machine-assisted quantitizing designs: augmenting humanities and social sciences with artificial intelligence. – Humanities and Social Sciences Communications, 2025, 12, 277. https://doi.org/10.1057/s41599-025-04503-w.

can identify recurring patterns, reveal relationships between passages, support carefully designed comparisons between texts written in different languages, and assist in the production of searchable metadata. At the same time, the experiments demonstrate that the evidential value of these results depends heavily on corpus size and quality, textual segmentation, linguistic variation, model design, and the procedures used to validate the output.

These dependencies become more important as the methods themselves grow more complex. Relatively transparent techniques such as frequency analysis and clustering allow researchers to inspect fairly directly how results are produced, even if they remain sensitive to lexical ambiguity and arbitrary textual divisions. Topic models, embeddings, and transformer-based language models can capture more complex patterns, but they also introduce additional layers of modelling and interpretation. Their outputs therefore require correspondingly stronger forms of methodological documentation, validation, and scholarly mediation. Increasing computational sophistication should not be confused with increasing evidential autonomy.

The emergence of transformers and large language models further expands these possibilities by enabling, among other applications, the automatic extraction of metadata, semantic tagging, and complex textual queries. Nevertheless, these systems also introduce new risks. Their outputs may appear coherent and authoritative while containing subtle inaccuracies, anachronisms, or unsupported interpretations. Because the internal logic of such models is often opaque even to specialists, the historian's task cannot simply become one of accepting algorithmic results. Instead, we must approach these tools critically, understanding that their value lies not in producing definitive interpretations but in assisting the construction, organisation, and exploration of historical corpora.

For this reason, we argue that the most productive role of machine learning in historical research is not as a substitute for scholarship, but as part of a collaborative and transparent paratextual infrastructure. Just as indexes, critical editions, footnotes, and regesta transformed historical practice in earlier centuries, ML-generated metadata and digitally enriched editions can facilitate the collective accumulation, organisation, and revision of knowledge today. Their usefulness depends on rigorous documentation, reproducibility, clear provenance, and continuous scholarly verification. Statistical methods are most useful

when integrated into workflows in which historians, linguists, and data scientists jointly design, evaluate, and refine the tools being used.

Ultimately, the responsible integration of computational methods into medieval studies depends on balancing technological possibilities with methodological control. Historians do not need to become machine-learning specialists, but they do need sufficient statistical literacy to understand the assumptions and limitations of the systems they use. Likewise, developers of computational tools must prioritise interpretability, accessibility, and scholarly transparency over technical sophistication alone. Used critically and collaboratively, such methods can greatly enrich the study of medieval texts, while contributing to more accessible, interconnected, and verifiable forms of historical knowledge.

Acknowledgements

The author used ChatGPT to proofread the English throughout the text and to generate initial drafts of the abstract and summary. In the section 'Collaborative Generation of Metadata with LLMs', he used a local copy of Llama 3 (8B). All other analyses and data processing were carried out using the scikit-learn library.

AJALOOLANE JA TÄNAPÄEVASED STATISTILISED MEETODID: RAKENDUSI 13. SAJANDI LIIVIMAA KROONIKATE UURIMISEL

Nicolas Bourgeois

Artiklis uuritakse, kuidas saab tänapäevaseid statistilisi keeletöötlusmeetodeid rakendada keskaegsete ajalootekstide uurimisel, keskendudes eeskätt 13. sajandi Liivimaa kahele peamisele kroonikale: ladinakeelsele Henriku Liivimaa kroonikale ja keskkõrgsaksakeelsele Liivimaa vanemale riimkroonikale. Autori eesmärk ei ole piirduda tehniliste meetodite tutvustamisega, vaid arutleda ka praktiliste ja tunnetuslike probleemide üle, millega ajaloolased tehisintellekti tööriistu kasutades kokku puutuvad.

Sissejuhatuses osutab autor, et tehisintellekti, eriti loomuliku keele töötluse (NLP), kiire areng on tekitanud märgatava lõhe selle vahel, mida need süsteemid suudavad teha, ja selle vahel, kuivõrd teadlased

nende süsteemide toimimisest tegelikult aru saavad. Ajaloolased võivad neid tööriistu keerukuse tõttu kas umbusaldada või, vastupidi, usaldada liigselt tulemusi, mille kujunemist ja usaldusväärsust nad piisavalt hinnata ei oska. Seetõttu peab autor vajalikuks kasutada eeskätt selliseid tööriistu, mille toimimine on võimalikult läbipaistev ja tulemused tõlgendatavad ning mida saab lõimida koostöisele uurimistööle tuginevatesse teaduspraktikatesse.

Esimeses, metodoloogilises osas käsitletakse leksikomeetriat. Lihtne sõnasageduste analüüs võib olla eksploratiivses uurimistöös kasulik, kuid selle võimalused on piiratud. Sama sõna eri vorme võidakse käsitada eri sõnadena, samal ajal kui üks ja sama sõna võib eri kontekstides kanda erinevat tähendust. Autor näitab, kuidas sageli koos esinevate sõnade klasterdamise abil on võimalik tuvastada näiteks sõjapidamise, religiooni või meresõiduga seotud teemad. Sellised meetodid võivad aga anda ka juhuslikke või eksitavaid tulemusi, sest sõnad paigutatakse jäigalt ühte kategooriasse ning teksti meelevaldne jaotamine väiksemateks osadeks võib moonutada sõnadevahelisi semantilisi seoseid.

Seejärel tutvustatakse artiklis latentse Dirichlet' jaotuse meetodit (*Latent Dirichlet Allocation*, LDA), mida kasutatakse tõenäosuslikus teemamudeldamises. Erinevalt lihtsast klasterdamisest lähtub LDA eeldusest, et iga tekst sisaldab korraga mitut varjatud teemat ning iga teemat iseloomustab omakorda sõnade tõenäosusjaotus. Nii võib üks ja sama sõna kuuluda sõltuvalt kontekstist erinevatesse semantilistesse väljadesse. Henriku kroonikale rakendatuna võimaldab LDA eristada sidusamaid ja tekstis paremini lokaliseeritavaid teemasid ning jälgida seeläbi näiteks ristisõdade, diplomaatia või jutlustamisega seotud ainest kogu kroonika ulatuses. Samas toimib see meetod kõige paremini suurte ja suhteliselt ühetaoliste tekstikorpuste puhul, mida keskaja-uuringutes sageli kasutada ei ole.

Järgmises osas käsitletakse sõnavektoreid ja latentseid ruume. Sellised mudelid nagu Word2Vec esitavad sõnu vektoritena mitmemõõtmelises ruumis, kus tähenduslik sarnasus väljendub sõnade geomeetrilises läheduses. Näiteks võimaldab t-SNE neid seoseid ka visuaalselt esitada. Eri keeltes tekstide võrdlemisel eeldab sõnavektorite kasutamine eri keelte vektorruumide joondamist või ühist mitmekeelset mudelit. Artiklis esitatud kahe kroonika võrdlus põhineb eraldi loodud teemamudelite tulemuste käsitsi joondamisel. Nii tuvastatakse sarnase teemakoostisega tekstikohti, näiteks sõdu vene vürstide vastu käsitlevaid lõike.

Artiklis eristatakse tokeniseerimist ja lemmatiseerimist. Tokeniseerimine jagab teksti mudeli töödeldavateks üksusteks, näiteks sõnadeks, sõnaosadeks või märkideks; lemmatiseerimine koondab sama sõna eri vormid ühise algvormi alla. Keskaegsete keelte puhul võivad nii kirjaviis kui ka sõnavormid suuresti varieeruda, mis võib raskendada mõlemat töötlusetappi ning mõjutada analüüsi tulemusi.

Artikli lõpuosas vaadeldakse tehisintellekti ja statistilisi mudeleid eeskätt paratekstiliste tööriistadena, mida võib võrrelda registrite, glosside ja kriitiliste väljaannetega, millele ajaloolased on oma töös traditsiooniliselt toetunud. Autori hinnangul ei pea ajaloolane tehisintellektisüsteemidest kasusaamiseks täielikult valdama nende matemaatilist ülesehitust, nii nagu trükitud allikaväljaande kasutamine ei eelda tingimata paleograafia täielikku valdamist. Praktiliseks vahevõimaluseks peetakse seetõttu hoolikalt koostatud digitaalseid väljaandeid, mida on rikastatud läbipaistvalt esitatud metaandmetega.

Artiklis tehakse ettepanek lõimida masinõppe abil loodud metaandmed teaduslikesse väljaannetesse läbipaistvate märgendussüsteemide abil. Teemamudelite ja suurte keelemudelite (LLM) abil saab automaatselt luua märksõnu, liigitada tekstilõike ning tuvastada selliseid olemeid nagu isikud, kohad ja sündmused. Autor rõhutab siiski, et tulemusi ei saa käsitleda lõplikena ja neid tuleb alati teaduslikult kontrollida. Suured keelemudelid võivad teha raskesti märgatavaid vigu, pakkuda liiga kaugele minevaid tõlgendusi või anda eksitavaid liigitusi. Selliseid eksimusi võib olla keeruline avastada, sest mudelite sisemine toimimine ei ole kasutajale läbipaistev.

Kokkuvõttes rõhutab artikkel, et tehisintellekti tööriistad ei peaks asendama kriitilist ajaloouurimist, vaid toetama ajaloolise materjali indekseerimist, annoteerimist ja korrastamist kui koostöist teaduspraktikat. Tehisintellekti tööriistade suurim väärtus ei seisne autoriteetsete ja iseseisvate tõlgenduste loomises, vaid kollektiivsete, läbipaistvate ja pidevalt täiendatavate teadustaristute kujundamise toetamises.

Acta Historica Tallinnensia 32/2 https://doi.org/10.3176/hist.2026.2.03

From Manuscript to Visualisation: A Digital Approach to Medieval Livonian Craft Ordinances

Rūta Bruževica

Tallinn University, School of Humanities, Institute of History, Archaeology and Art History, Narva mnt 25, 10120 Tallinn, Estonia; ruta.bruzevica@gmail.com

Abstract. Medieval Livonian craft ordinances have often been reproduced and examined by researchers, but they have not previously been compiled into a single edition. The digitisation of these sources is not just a matter of access and preservation; it also offers an opportunity for further scholarly analysis by transforming dispersed manuscripts into a structured corpus of data. This corpus enables new forms of comparison, corpus analysis, and visualisation, while also exposing the limits of interpretation in data-driven approaches to the normative sources of medieval towns.

Keywords: Medieval Livonia, Riga, Tallinn, craft guilds, Transkribus, digital history, Nodegoat

INTRODUCTION

Medieval Livonian craft ordinances have attracted legal and scholarly interest since at least the seventeenth century, when medieval documents were copied into manuscripts reflecting the writing practices of the time.[1] Printed editions subsequently made these texts accessible to a

1 For example, the 1459 craft ordinances of Tallinn masons and sculptors survive in a seventeenth-century copy: Kivi- ja kujuraidurid / Stein- und Bildhauer, Schragen, 1459, 17th-century copy. Tallinna Linnaarhiiv (Tallinn City Archives, TLA), Tallinn, 190-2-651.

© 2026 Author. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International License CC BY 4.0 (http://creativecommons.org/licenses/by/4.0).

wider scholarly readership. In a way, the digitisation initiatives of the twenty-first century continue this process of transmission, transforming medieval documents into structures and formats usable by today's researcher. Printed editions are now increasingly supplemented, and in some cases replaced, by digital source editions containing images of original sources and digital transcriptions of those texts.

The digitisation of medieval Livonian craft ordinances, as part of the global digitisation trends, first improves access to scattered sources. However, the process also allows the transformation of manuscripts, later copies and printed editions into a reusable corpus that enables new forms of scholarly comparison, corpus analysis, and visualisation. This article introduces methodological opportunities through a case study in digitising and analysing digitised medieval sources.

The central research question of the article is: How does the digitisation of medieval Livonian craft ordinances change the possibilities for analysing this source group, and what methodological limits must be considered when normative texts are transformed into searchable and visualisable data? To answer this question, the article examines the source group and its digital transformation within the Digital Livonia project[2] and presents case studies of applying digital analytical approaches: corpus analysis and visualisation.

A digital source edition makes it possible to compile manuscripts and their fragments that are defined not only by their location in specific archives and collections, but also by new types of collections defined by their relevance to particular topics that reach across several centuries and modern national borders. This idea underlies the Digital Livonia collection "Craft guild statutes of medieval Livonia" that contains digitised and transcribed craft ordinances from Riga and Tallinn, dating from the early fourteenth century to the year 1550, thereby compiling these sources in a single collection for the first time.[3] Such a source edition, including materials preserved at the Tallinn City Archives (TLA) and the Latvian State Historical Archive (LVVA), makes it possible to treat the many individual medieval craft ordinance documents as a joint corpus, which supports access for traditional history research, while allowing new approaches through data analysis.

The craft ordinances are especially suitable for this kind of methodological reflection because they contain several levels of

2 Digital Livonia. About the Project. https://dl.tlu.ee/en/project/about (accessed 20/09/2026).

3 R. Bruževica. Craft Guild Statutes of Medieval Livonia, 2025. https://dl.tlu.ee/en/source-editions/gildide-skraad (accessed 20/09/2026).

information, including legal, economic, religious and social regulations for urban and craft life. The artisan ordinances often regulated not only economic activity and work, but also the social life of the organisation, relations between members, and religious life and practice, which was an important aspect in the life of medieval townspeople. These ordinances, which primarily hold information about the craft organisations and communities that compiled and observed them, also include reflections of customs observed in medieval towns, including traditions that stretched across the landscape of the city, and social customs that were advisable or inappropriate during social interaction.

The application of digital history tools to the corpus of craft ordinances will enhance the study of the craft community as a place of sociability in the Livonian towns, allowing comparison between towns, crafts, and centuries, based on concordances, connections and networks in the digitised manuscript collection. During research work with the craft ordinance corpus, a traditional approach based on qualitative source reading and analysis has been combined with experiments and case studies performed with digital humanities methods. The opportunities and limitations encountered through this process form the basis of the article's methodological argument: digital tools can widen the scope of analysis and help locate patterns, but they do not produce historical interpretation automatically.

To address this topic, this article includes a historiographical overview and an analysis of the specific character of medieval craft ordinances in the Livonian manuscript material, followed by a description of the concepts, opportunities and the process of the transformative and reflective approaches imposed on the manuscript collection by digitisation. Based on the main outcomes from the initial digitisation process and opportunities of applying digital history tools, two methods are examined in detail to evaluate their contribution to the analysis of medieval sources. These examples of corpus analysis and visualisation through network analysis introduce the accessible tools for a historian of medieval sources and highlight the strengths, challenges, and limits of application. Finally, the discussion summarises the contribution of digitised medieval craft ordinances to research and the role of source-critical analysis as a critical feature paired with future analysis opportunities at the present stage of digital medieval history research on Livonia.

CRAFT ORDINANCES OF MEDIEVAL LIVONIA

In my ongoing research about crafts as spaces of sociability in the towns of medieval Livonia,[4] craft ordinances from Riga and Tallinn have been explicitly chosen as the main source group. The focus on source material from these two towns is largely connected to the survival of craft sources from these towns in reasonable quantity, whereas comparable sources from Tartu have not survived to the present. In many smaller Livonian towns, the population, and therefore the number of artisans, was too limited to sustain independent craft guilds. For this reason, the digitisation process and further analysis concentrate on Tallinn and Riga, where the documentary record and institutional craft life were more fully developed and remain accessible today, although the state and survival of sources from these two towns is diverse.

With regard to the available source material from Tallinn, the Tallinn City Archives hold 33 craft ordinances, comprising both original documents and contemporary or later copies. An overview of the Tallinn ordinances, including archival references and earlier editions, has been compiled by Maija Ojala in Table 3.1 of her study of craft trade culture in the Baltic Sea region.[5] For the present project, 30 of these manuscripts constitute the digitised corpus. The remaining items were omitted because suitable digital images were unavailable or their survival as later copies made them less suitable for inclusion in the present digital edition. Although several statutes had been digitised previously and published via the Saaga[6] platform, given the rapidly changing possibilities of digitisation, their quality did not meet the goals of digitisation in the Digital Livonia project. Digital photographs of Tallinn craft ordinances were made by photographer Stanislav Stepashko in 2010 and were made available to the author for digitisation. Several transcriptions of the ordinances exist; however, they remain dispersed across books and other publications and have not been compiled into a single collection.

For the study of Riga's craft guilds, scholars have mainly relied for more than a century on a single foundational edition: the transcribed compilation of the statutes of Riga's craft and merchant guilds until

4 PhD thesis in preparation at Tallinn University, provisionally titled "Craft Guild as a Place of Sociability in Medieval Livonian Towns: Religious and Everyday Practices".

5 M. Ojala. Protection, Continuity and Gender: Craft Trade Culture in the Baltic Sea Region (14th–16th Centuries). Tampereen Yliopisto, Tampere, 2014. 319–320, Table 3.1.

6 Saaga: Digiteeritud arhiiviallikad. https://www.ra.ee/dgs/explorer.php (accessed 16/09/2026).

Table 1. Overview of the medieval Livonian craft ordinance corpus. The table summarises the source material by craft, survival in Riga and/or Tallinn, number of ordinances, date scope, available archival originals, and printed-edition basis. It shows the mixed character of the corpus, which combines archival manuscripts, later copies, and nineteenth-century printed editions. Table prepared by the author.

Craft	Survival (R/T)	No. of ordinances	Date scope	Archival originals available	Printed edition
Bakers	R, T	5	1235–1527	Bäcker Schragen 1438 TLA.190.2.281, 1r–7v; Bäcker Schragen 1527 TLA.230.1.Ac5, 98r–103r	Stieda, Mettig 235; Stieda, Mettig 239; Stieda, Mettig 242
Barbers	R	1	1494	—	Stieda, Mettig 248
Beer carters	R, T	2	1456–1466	Bierträger Schragen 1456 TLA.190.2.680, 1r–3r	Stieda, Mettig 252
Blacksmiths	R, T	6	1382–1459	Schmiede Schragen 1382 LNA_LVVA, F4038_2_1117, 41r; Schmiede Schragen 1382 LNA_LVVA, F8_1_15, 2r–4r; Schmiede Gesellen Schragen ~1399 LNA_LVVA, F8_1_15, 17r–18r; Schmiede Schragen 1415–1437 TLA.190.2.554, 1r–4v; Schmiede Schragen 1459 TLA.190.2.555, 1r–7v	Stieda, Mettig 463
Boatmen	T	2	1506–1531	Mündriche Schragen 1506 TLA.190.2.663, 1r–3v; Mündriche Schragen 1531 TLA.190.2.664, 1r–11r	—
Brewers	T	2	1498–1505	Brauer Schragen 1498 TLA.190.2.666, 1r–3r; Brauer Schragen 1505 TLA.190.2.667, 1r–2v	—
Butchers	T	4	1394–1528	Fleischer Schragen 1394 TLA.190.2.118, 1r–11r; Fleischer Schragen 1509 TLA.190.2.121, 1r–8r; Knochenhauer Schragen 1509 TLA.230.1.Ac5, 45r–48r; Knochenhauer Schragen 1528 TLA.230.1.Ac5, 61r–63v	—
Carpenters	T	1	1508	Zimmerleute Schragen 1508 TLA.190.2.668, 1r–7r	—
Carvers	R	2	1436/1536–1541	Schnitzer Schragen 1436 or 1536 LNA_LVVA, F4038_2_1117, 127r–136v	Stieda, Mettig 515
Coopers	R, T	4	1375–1556	(Bodeker) Bottcher Schragen 1375 LNA_LVVA, F8_1_15, 12r–14r; Bottcher Schragen 1435 TLA.190.2.472, 1r–4r; Bottcher Schragen 1515 TLA.190.2.473, 1r–7r; Böttcher Schragen 1556 TLA.190.2.474, 1r–7r	—
Dockers	R	1	1463	—	Stieda, Mettig 407
Fabric cutters	R	1	1383	Lakenschrerere Schragen 1383 LNA_LVVA, F4038_2_1117, 40r	—
Fishers	R	1	1403	—	Stieda, Mettig 275
Furriers	R, T	3	1397–1553	Kürschner Schragen 1397 LNA_LVVA, F8_1_15, 6v–8v; Kürschnere Schragen 1553 TLA.230.1.Ac5, 68r–70v	Stieda, Mettig 383
Girdler and Belt makers	R	1	1512	—	Stieda, Mettig 338

Continued on the next page

Table 1. Continued

Craft	Survival (R/T)	No. of ordinances	Date scope	Archival originals available	Printed edition
Glaziers	R	2	1541–1542	Glaser Schragen 1541 LNA_LVVA, F224_1_2472, 1r–4v	Stieda, Mettig 289
Goldsmiths	R, T	5	1360–1542	Goldschmiede Schragen 1360 LNA_LVVA, F8_1_15, 4v–6r; Goldschmiede Schragen 1393 TLA.190.2.069; Goldschmiede Schragen 1453 TLA.190.2.070, 1r–5r; Goldschmiede Schragen 1537 TLA.230.1.Ac5, 64r–67v; Goldschmiede Schragen 1542 LNA_LVVA, F4038_2_1117, 1r–15r	—
Hemp spinners	R, T	2	1436–1462	Hanfspinner Schragen 1462 TLA.190.2.642, 1r–6v	Stieda, Mettig 346
Linen weavers	R	1	1458	Leineweber Schragen 1458 LNA_ LVVA, F4038_2_1117, 45r	—
Masons	R	2	1390–1546	Maurer Schragen 1390–1459 LNA_ LVVA, F4038_2_1117, 118r–122r	Stieda, Mettig 428
Masons, Sculptors	T	1	1459	Maurer, Steinhauer Schragen 1459, 17th century copy TLA 190-2-651, 1r–7r	—
Painters (wood carvers, glaziers, box-makers)	T	1	1513	Maler, Glaser, Schnitker und Kontor-macher Schragen 1513, TLA.230.1.Ac5, 5r–11v	—
Porters	R, T	2	1450-1529	Träger Schragen 1529 TLA.190.2.641, 1r–3r	Stieda, Mettig 414
Saddle makers	T	1	1459–1543	Sattel- und Gürtelschmiede Schragen 1459 & 1543 TLA.190.2.549, 1r–8r	—
Sculptors	T	1	1402	Steinhauer Schragen 1402 TLA.190.2.650, 1r–5r	—
Shoemakers	R, T	4	1399–1536	Schuhmacher Schragen ~1399 LNA_ LVVA, F8_1_15, 9r–11v; Schuhmacher Schragen 1416–1477 TLA.190.2.023, 1r–9v; Schuhmacher Schragen 1481–1536 TLA.190.2.024, 1r–12r	Stieda, Mettig 532
Tailors	R, T	4	1399–1519	(Schrodere) Schneider Schragen ~1399 LNA_LVVA, F8_1_15, 14v–16r; Schneider Schragen 1413–1594 TLA.190.2.511, 1r–8r; Schneider Gesellen Schragen 1519 TLA.190.2.522, 22r–24v	Stieda, Mettig 482
Town servants	R	1	1415	—	Stieda, Mettig 655
Wagoners	T	2	1435–1545	Führleute Schragen 1435 TLA.190.2.678, 1r–4v; Führleute Schragen 1545 TLA.190.2.679, 1r–2v	—
Workers, Porters	R	1	1535	—	Stieda, Mettig 233

1621, edited by Wilhelm Stieda and Constantin Mettig and published in 1896.[7] An overview of the Riga ordinances, including archival references

7 Schragen der Gilden und Aemter der Stadt Riga bis 1621. Bearb. von W. Stieda, C. Mettig. Alexander Stieda's Buchhandlung, Riga, 1896.

and editions, has also been compiled by Maija Ojala in Table 2.1 of her study of craft trade culture in the Baltic Sea region.[8] The present research sought to return, wherever possible, to the surviving originals and integrate them into a digital edition. Fourteen ordinances were located in the fonds of the Latvian State Historical Archive (LVVA), digitised, and provided with revised semi-diplomatic transcriptions; Stieda and Mettig's transcriptions were used as a starting point where available. The transcription preserves manuscript u/v distinctions and capitalisation, represents punctuation with modern Unicode characters, and expands most abbreviations. The nineteen other statutes, whose manuscripts are no longer available, are based on the nineteenth-century edition and consequently retain more of its editorial characteristics. This mixed-source basis makes the available Riga ordinances accessible as a large corpus but also means that transcription and editorial differences must be considered in digital comparison and term-count analysis (see Table 1).

CRAFT ORDINANCES AS A SOURCE

The craft ordinances from Riga and Tallinn belong to a normative and legally descriptive source type with a relatively stable external structure: first identifying the craft concerned and the date of confirmation, then referring to approval by the town council, and often mentioning the type of organisation and its dedication to a patron saint. Through these formulae, the opening and closing paragraphs of the documents often present the structure of formally authorised texts, with the main body of the statute consisting of a sequence of listed regulations. They regulate admission to apprenticeship and mastership, workshop practice, production standards and prices, relations between members, expected conduct during meetings and feasts, participation in religious ceremonies, and the obligations of members in cases of sickness, death, or poverty. Often, the regulations are accompanied by a sanction, usually a fine, although other forms of punishment, such as exclusion, may also occur. Most of the texts included in the corpus follow a similar structural and thematic convention, which, together with the use of penalties, suggests the normative character of these sources.

The normativity of the craft ordinances is one of the main methodological challenges in using them as historical sources. Although

8 M. Ojala. Protection, Continuity and Gender, 315–316, Table 2.1.

it would be misleading to read the ordinances as direct reflections of the functioning of these communities, the normative character of these documents does not make them detached from the social reality. The repeated confirmation of ordinances, the presence of different hands, additions inserted into the text, corrections, cancellations, marginal notes, and later supplements all indicate that these documents remained active within the institutional life of the craft community. Regulations could establish a formal framework that was approved by the town council and the community, while actual practice could develop in more varied ways.

Research in combination with other sources concerning artisans has suggested the flexibility of craft ordinances, which does not necessarily contradict the surviving regulations. This can be seen, for example, in recent research on cases of widows' rights within craft communities.[9] Ordinances might allow a widow to continue the workshop for a limited time period, but individual cases could extend this right beyond what the regulatory content of the statute suggested. Such cases suggest that statutory restrictions could be relaxed in practice while remaining grounded in written regulations.

Other types of evidence, such as fine and account books, also suggest that statutory regulations were not merely theoretical. The late fifteenth-century fine and account book of the Riga Porters' guild records offences and links them to specific individuals, showing that at least some regulations were actively monitored and enforced within craft communities.[10] Even when such parallel evidence survives only exceptionally, it is important as it demonstrates the practical administrative context in which statutory rules could operate and offences could be addressed within the late medieval town.

A further indication of the connection between statutes and lived practice is the specificity of some regulations. Many additions or later clauses address very particular situations: forms of inappropriate conduct, disputes between members, obligations during specific rituals or gatherings, or procedures for circumstances not covered by earlier regulations. Such clauses are often too precise to be understood only as abstract legal formulae and their specificity suggests that the craft ordinances could respond to concrete problems that had arisen within the community and that written regulation was a tool to adjust, change, or redefine existing practices.

9 M. Ojala. Protection, Continuity and Gender.

10 G. Strenga. Donations, Discipline and Commemoration: Creating Group Identity in the Transport Workers Guilds of Mid Fifteenth-Century Riga. – Journal of Medieval History, 2022, 48, 1, 103–128. https://doi.org/10.1080/03044181.2021.2012502.

The craft ordinances of medieval Riga and Tallinn are best approached as legal texts that hold formal frameworks of institutional practice but were influenced by the practices of the craft communities and urban traditions. The value of craft ordinances as sources of medieval urban life lies in the duality between their normativity and practicality that reveals what craft communities and town councils considered necessary to regulate, what kinds of behaviour required sanction, which relationships needed formal definition, and how written rules were revised over time. These features of the ordinances make them suitable material for digitisation, since recurring regulations and terms can be studied with comparative analysis tools, while prioritising qualitative interpretation for specific regulations that reveal particularities of the crafts under study.

FROM MANUSCRIPT TO DATA

One of the aims of the Digital Livonia project is to create a comprehensive online portal that brings together written sources related to medieval Livonia.[11] The craft ordinances of Riga and Tallinn are among the source editions, combining the diverse formats that are today found in different archives and collections across Latvia and Estonia, both as original manuscripts and as physical and digitised printed editions.[12] This source edition brings together sources from different towns, centuries and organisations, regardless of their belonging to separate or common collections in the archives.[13] The work with this collection was planned according to modern standards and the practices of the most recent theoretical developments in the field of cultural heritage and manuscript digitisation.[14] To achieve the most diverse possibilities of the digital edition, it should include a high-quality photograph of the object and a layer of data, including a transcription that lets the

11 See the article by Marek Tamm and Hembo Pagi in this issue.
12 R. Bruževica. Craft Guild Statutes of Medieval Livonia.
13 For example, the Riga craft statutes – six were a part of the compiled book of "Olde Renthebock" (Olde Renthebock, 1382. Latvian National Archives, Latvian State Historical Archive (LNA_LVVA), Riga, 8-1-15); several others were found in the records of compiled documents from the corresponding crafts, (Glaser 1541 LNA_LVVA 224-1-2472) several were found sewn into a collection of many different late medieval and early modern pieces of handwritten documents of different conditions and materials, (LNA_LVVA 4038-2-1117) not to mention all of these sources being part of different archival fonds.
14 M. Terras, J. Nockels, S. Ames, P. Gooding, A. Stauder, G. Mühlberger. On Automating Editions: The Affordances of Handwritten Text Recognition Platforms for Scholarly Editing. – Scholarly Editing, 2024, 41. https://doi.org/10.55520/5V92VK9A.

researcher use the prepared transcription, while being able to observe the state of the original.

The transformation of the manuscripts into digital objects and data can be viewed as part of the tradition of reproducing text-based documents. In reality, this process produces new versions of sources and is never purely technical. As highlighted by Lars Björk[15], the reproduction of a document is adapted by the goals of use and standards of preservation of the producer – the historian or the institutional staff, meaning that the producer of the digital version also takes part in shaping the source as a digital object. This is also to be considered in the source editions of craft ordinances: the digitised transcription of the document is created through work with original manuscripts, but includes editorial decisions concerning layout, uncertain readings, strikethroughs and additions, and the degree of character normalisation for the digital format. These decisions also shape the conditions under which the digitised source material can be accessed, compared and interpreted. As a result, the digital format of the source represents the original object through data that preserves or suppresses certain features, while enabling different possibilities for scholarly work through digital humanities methods.

The digitised versions of craft ordinances were created by using Transkribus,[16] a research infrastructure for producing line-based transcriptions and automated recognition of handwritten text.[17] The Transkribus platform allows manuscript images to be divided into text regions and lines, which are then linked to corresponding transcription lines. These image-text pairs can also be used as "ground truth" for training Handwritten Text Recognition (HTR) models, which produce automatic transcriptions of visually and textually similar material.[18]

In the case of the Livonian craft ordinances, Transkribus was used as the main environment for transforming original manuscript images into structured digital text. Images of the ordinances were uploaded to the platform, segmented into text regions and lines, and then transcribed, corrected, or, if earlier transcriptions existed, aligned with

15 L. Björk. How Reproductive Is a Reproduction? Digital Transmission of Textbased Documents. University of Borås, Borås, 2015, 36.

16 Transkribus. READ-COOP. https://readcoop.eu/transkribus/ (accessed 20/09/2026).

17 J. Nockels, K. Gooding, M. Terras, M. Berlinska. Understanding the Application of Handwritten Text Recognition Technology in Heritage Contexts: A Systematic Review of Transkribus in Published Research. – Archival Science, 2022, 22, 3, 367–392, here 368. https://doi.org/10.1007/s10502-022-09397-0.

18 G. Muehlberger, L. Seaward, M. Terras, S. Ares Oliveira, V. Bosch, M. Bryan, S. Colutto, et al. Transforming Scholarship in the Archives through Handwritten Text Recognition: Transkribus as a Case Study. – Journal of Documentation, 2019, 75, 5, 954–976. https://doi.org/10.1108/JD-07-2018-0114.

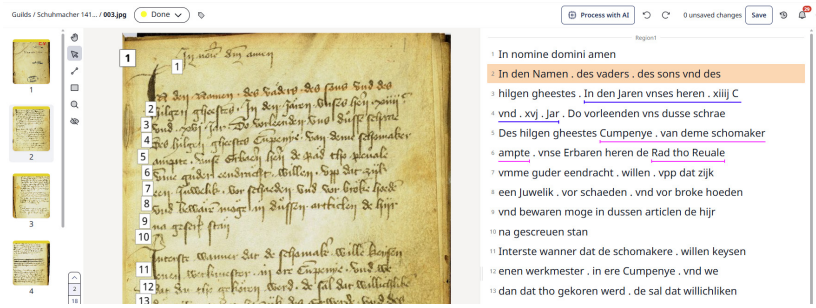


Figure 1. Manuscript image, layout segmentation, and line-based transcription of Shoemaker ordinances in Transkribus. The screenshot shows how the digitisation environment connects the manuscript layout with the corresponding transcription line, creating a layered digital object. Screenshot by the author; manuscript image from Tallinn City Archives (TLA), Shoemakers' ordinances [Schuhmacher Schragen] 1416–1477, 190.2.023, fols. 1r–9v.

the relevant manuscript images. The use of Transkribus as a digitisation environment has made it possible to connect the manuscript layout with the corresponding line in the transcription. The result is not only a readable text, but a layered digital object in which image, layout, and transcription remain linked (see Figure 1). For the study of Livonian craft ordinances, this changes the working conditions of the source: the transcription can be searched and exported for analysis, while the manuscript image remains available for verification and correction.

Creating the digital version of craft ordinances with Transkribus keeps the original source within close reach for any researcher who would re-approach it through the digital edition. If a reading is uncertain or open to correction, the image remains directly accessible, does not require a physical visit to the archive, and if the platform is maintained, corrections can also be added later. This accessibility of the source material can be used by researchers who want to examine the statutes of a specific craft, century or city, as well as approach craft guild history more broadly through search engines or data analysis. The material can also be used in teaching palaeography and reading Middle Low German in university curricula, especially for students planning to advance in medieval studies.

When a digitised version of the craft ordinances moves beyond the simple online display of manuscript images and offers a critically produced, usable, and documented representation of the sources, it can be understood as a digital scholarly edition. As highlighted by Melissa Terras and others,[19] images and machine-readable transcriptions can already form the basis of online scholarly editions, especially when they

19 M. Terras, J. Nockels, S. Ames, P. Gooding, A. Stauder, G. Mühlberger. On Automating Editions.

allow browsing, reading, searching, and further analysis; at the same time, such editions gain scholarly value through editorial intervention, quality control, metadata, statements on provenance, language, temporality, editorial practice, technical method, and links to related data. In the case of the Livonian craft ordinances, the Digital Livonia version fulfils the basic conditions of such an edition because manuscript image, line-based transcription, archival context, and structured metadata are brought together in one environment.

Work with the sources and their transcriptions on the Transkribus platform also created an opportunity to train HTR models on local source material. Since late medieval Livonian handwriting has not been the focus of many existing models, the project drew on comparable work with fifteenth-century German-language material, including models based on Hanseatic diet protocols[20] and Viennese property registers[21] from the fifteenth century. Early in the digitisation process the craft ordinance data was also included in the creation of a broader HTR Generic Model for fifteenth and sixteenth-century German.[22] The model was based on late medieval German manuscripts from Austria and the Hanseatic Baltic region and included both Middle Low German and Early New High German texts and was later used as a base model[23] for the creation of the now public Medieval Livonia DL[24] HTR model, which included data from transcriptions of the craft guild ordinances, civic ordinances from Livonian towns and probate inventories from medieval Tallinn.

The application of digitisation and HTR models to the craft ordinances resulted in digital source versions that enabled export in XML formats and implementation of the digital editions in the Digital Livonia platform (see Figure 2). Yet not all data produced during the digitisation process can be included in the public edition due to technical restrictions and the risk of sharing database or descriptive data without the context

20 I. Lange. Hanserezesse 15th Century v.2.0. – Transkribus / READ-COOP SCE, 2024. https://www.transkribus.org/models/hanserezesse-15th-century-v20 (accessed 20/09/2026).

21 J. Helmchen. Die Digitalisierung der spätmittelalterlichen Grundbücher Wiens mit Transkribus. – Mittelalter, 12 June 2025. https://doi.org/10.58079/143DO.

22 J. Helmchen. Generic Model 15th–16th Century German (prototype). Transkribus / READ-COOP SCE, 2022. https://www.transkribus.org/models/generic-model-15th-16th-century-german-prototype (accessed 20/09/2026).

23 H. Prokop. Should I Use a Base Model When Training a Custom Model? Transkribus, 7 February 2025. https://www.transkribus.org/blog/using-base-model-when-training-a-custom-model (accessed 20/09/2026).

24 Rūta Bruževica, 'Medieval Livonia DL', Software, Transkribus, Transkribus / READ-COOP SCE, 18 March 2024, https://www.transkribus.org/models/medieval-livonia-dl (accessed 20/09/2026).

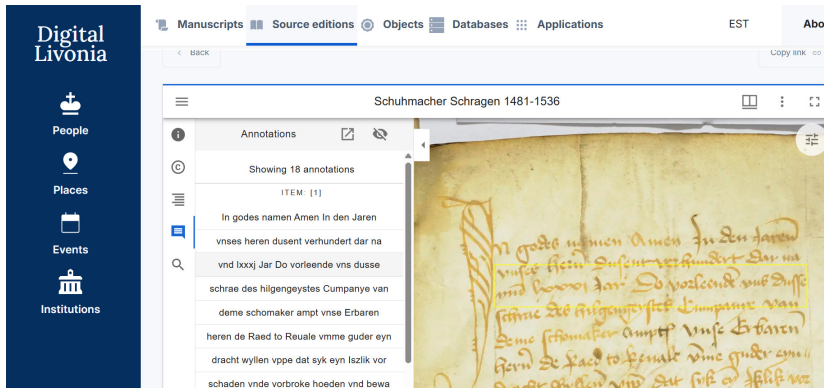


Figure 2. Digital edition of Shoemaker ordinances in the Digital Livonia platform. The screenshot shows the implementation of the digitised source as a source edition, where the manuscript image, transcription, and annotation environment are connected within the same interface. Screenshot by the author; manuscript image from Tallinn City Archives (TLA), Shoemakers' ordinances [Schuhmacher Schragen] 1481–1536, 190.2.024, fols. 1r–12r.

in which the researcher created it.[25] Such data, in the form of XML files, text files, databases, and encodings remains in the private administration of the researcher, and can be applied in other analysis tools for further insights based on specific, data-driven research questions that span the corpus of the now digitised manuscripts.

TEXT CORPUS AS A DATA ANALYSIS SOURCE

Apart from the digital edition and its display on the Digital Livonia website, another valuable result of the digitisation of the craft ordinances is the textual corpus of the source collection. This body of textual data on a common topic can be a starting point for several research directions that are possible with the tools of the digital humanities field and allow research in other disciplines beyond history. The text files, exported from transcriptions of each ordinance separately, hold information about the period, town and craft that they correspond to. The corpus, which consists of more than 400 transcribed pages and the additional texts from the previously published edition,[26] forms the basis for the following case studies.

25 M. Terras, J. Nockels, S. Ames, P. Gooding, A. Stauder, G. Mühlberger. On Automating Editions, 6–8.
26 Schragen der Gilden und Aemter der Stadt Riga bis 1621.

Corpus analysis, a method originally developed in linguistics, has recently been adapted for historical research,[27] especially in work with large bodies of similar texts from the same period, such as historical newspapers. For the present corpus, two especially useful methods, borrowed from corpus linguistics, are concordance and term-count analysis. Concordance allows the researcher to trace how a selected word appears in different textual contexts and alongside which other terms, making it possible to study recurring associations and to identify the meanings a term acquires within particular regulations.[28] Another simple method, term-count analysis, measures how often specific words or phrases occur across the corpus and can therefore indicate which themes, roles, or forms of regulation are especially prominent.[29] In the case of medieval craft ordinances, both methods are valuable as exploratory tools: concordance helps to interpret usage in context, while term counts help to detect broader textual patterns.

Corpus analysis tools can be used to reveal patterns of specific topics in the texts, and to visualise the results of these quantitative approaches. Although these methods could be applied to the whole corpus of craft ordinances, the clearest comparisons can be produced by choosing a smaller set of documents for a focused case study. To demonstrate the possibilities of corpus analysis, this section presents visual results based on four Shoemaker ordinances from Riga and Tallinn, with attention to terms referring to women in the ordinance texts. The free text analysis software CATMA was used to analyse and visualise both term distribution and concordance.[30]

In medieval ordinances, one of the challenges of applying corpus analysis is the inconsistency of words, not only in the variety of the terms used to denote the concepts mentioned, but also different versions of writing the same word or its inflectional forms. The application of Natural Language Processing (NLP) and lemmatisation of text content can reduce this problem for modern texts in many languages, where NLP pipelines are available for easy use, although Middle Low German is

27 L. Gauthier. Digital Humanities and the Notion of Corpus in Ancient History. – Digital Studies / Le Champ Numérique, ahead of print, 9 July 2025, 4, https://doi.org/10.16995/dscn.18169.

28 S. Hunston. Methods in Corpus Linguistics: Interpreting Concordance Lines. – Corpora in Applied Linguistics, Cambridge Applied Linguistics. Cambridge University Press, Cambridge, 2002, 38–66. https://doi.org/10.1017/CBO9781139524773.004.

29 C. Tumbe. Corpus Linguistics, Newspaper Archives and Historical Research Methods. – Journal of Management History, 2019, 25, 4, 533–549, here 536. https://doi.org/10.1108/JMH-01-2018-0009.

30 E. Gius, J. C. Meister, M. Meister, M. Petris, D. Gerstorfer, M. Akazawa, S. Messner. CATMA. – Zenodo, 10 March 2025. https://doi.org/10.5281/zenodo.15000951.

not among these languages. However, recent developments of Artificial Intelligence tools and specifically Large Language Models (LLMs) have advanced the accessibility of lemmatisation for historical languages.[31] Researchers can use LLMs to assist this process by supplying historical context, dictionaries, and text examples.

To enhance the quality of corpus analysis and reduce variation in the text, the four ordinance transcriptions were processed with an LLM. As UTF-8 text files, they were automatically tokenised, lightly normalised, and lemmatised using Middle Low German (MLG) dictionary-style lemma forms. The original spellings were preserved in the XML attributes, while separate CATMA-oriented outputs were created with either the original token or the lemma as the running text. The lemma-based version was used for the CATMA analysis presented in this article. Each token was enriched with basic linguistic metadata, including normalised form, lemma, part of speech, and a confidence marker distinguishing straightforward forms from linguistically or contextually uncertain ones requiring verification. After automated processing, the terms retrieved for the CATMA queries and their immediate KWIC contexts were checked manually against the transcriptions. This checking verified the retrieved occurrences rather than constituting a systematic search for forms potentially missed during automated lemmatisation. The lemmatised files should therefore be understood as a controlled working layer for exploratory corpus analysis, not as a complete linguistic annotation or critical normalisation of the ordinances. The precise prompt and lexical resources used during this experimental LLM-assisted processing were not preserved as part of the research documentation and cannot be reconstructed reliably; this limits the reproducibility of this stage and is one reason why the resulting data is not used for independent philological conclusions. The lemmatisation and XML structuring were produced with GPT-5.5 Thinking.[32]

After lemmatisation, the four Shoemaker ordinances from Riga[33] and Tallinn[34] were uploaded to CATMA and analysed with two tools:

31 C. Vidal-Gorène, B. Kindt, F. Cafiero. Under-Resourced Studies of under-Resourced Languages: Lemmatization and POS-Tagging with LLM Annotators for Historical Armenian, Georgian, Greek and Syriac. – arXiv, 17 February 2026. https://doi.org/10.48550/arXiv.2602.15753.

32 OpenAI. ChatGPT-5.5 Thinking. ChatGPT, May 2026. https://chatgpt.com/.

33 Schuhmacher Schragen, ca. 1399, Riga. Latvian National Archives, Latvian State Historical Archive (LNA_LVVA), Riga, F8-1-15, fols. 9r–11v; Schuhmacher Gesellen Riga 1414–1480. – Schragen der Gilden und Aemter der Stadt Riga bis 1621, 532–537.

34 Schuhmacher Schragen 1416–1477, Tallinn. TLA, 190.2.023, fols. 1r–9v. https://dl.tlu.ee/en/source-editions/gildide-skraad (10/05/2026); Schuhmacher Schragen 1481–1536, Tallinn. TLA, 190.2.024, fols. 1r–12r. https://dl.tlu.ee/en/source-editions/gildide-skraad (accessed 20/09/2026).

Distribution to visualise term occurrence, and Keyword in Context (KWIC) for concordance analysis. Lemmatisation substantially reduced the spelling and inflectional variation visible in the original transcriptions, but did not eliminate it completely. For example, a form such as *vrowenn* could be reduced to *vrowen*, while related forms could remain separate in the lemma-based text. This residual variation reflects both the difficulties of automatically lemmatising historical MLG and the mixed textual basis of the corpus, which combines semi-diplomatic manuscript transcriptions with an edition-derived text. The CATMA searches therefore used nine known forms referring to women: *junckvrowen*, *suster*, *susterne*, *vrouwe*, *vrowen*, *wif*, *frouw*, *frowen*, and *frouwenn*. These represent four broader lexical groups – *junckvrowe*, *suster*, *vrowe*, and *wif* – while retaining residual spelling and inflectional variants. The forms were queried separately, using their known word beginnings where necessary to capture remaining variants, and their occurrence counts were not pooled into a single category. This distinction was retained because the lexical groups can refer to different social positions or relationships, including unmarried women, wives, and women described as members or 'sisters' of the community.

Figure 3 demonstrates how corpus analysis can be used to compare the occurrence and position of selected terms across a small group of related ordinances. Each panel represents one ordinance separately, while its horizontal axis divides the document into ten equal sections, each corresponding to 10% of its character length. CATMA calculates positions through character offsets, including spaces and punctuation, but displays them here as relative positions from 0 to 100, which makes texts of different lengths comparable. The vertical axis shows the number of occurrences of a queried form within each 10% section rather than a cumulative count for the whole corpus. Each colour represents a separate queried form; points of the same colour are connected across sections for visual clarity, but the connecting lines do not indicate a relationship between individual occurrences.

The distribution graph does not in itself explain the role of women in the craft community, nor can the number of occurrences be treated as a direct measure of social importance. It does, however, make it possible to compare where different terms occur within the four ordinances and whether their occurrences are dispersed or concentrated in sections. For example, forms of *suster* are concentrated in particular parts of the ordinances, whereas forms belonging to the *vrowe* lexical group occur at several positions. Such differences provide a reason to return to

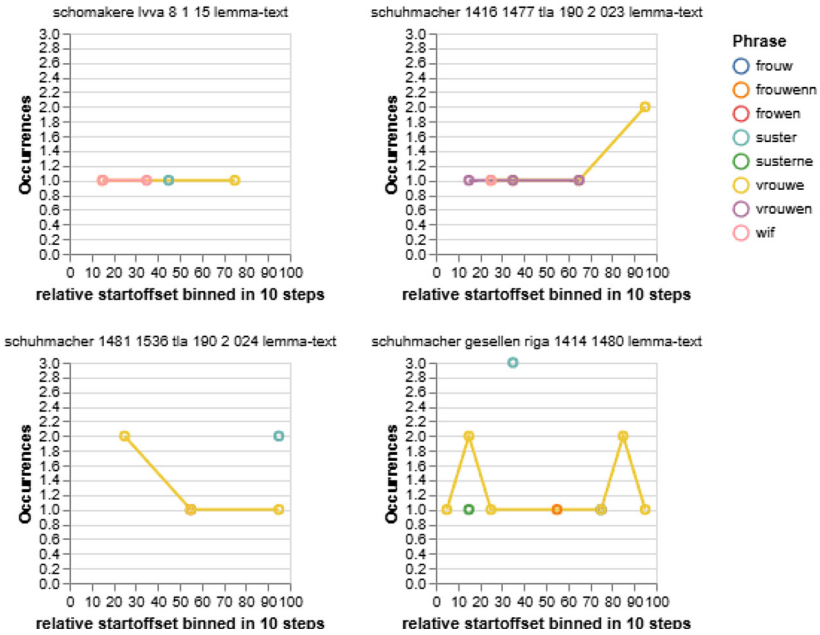


Figure 3. Distribution of selected Middle Low German terms referring to women in four sampled Shoemaker ordinances. Each panel represents one ordinance divided into ten sections of equal relative character length; the vertical axis shows the number of occurrences within each section. Colours represent individual queried forms, while connecting lines link occurrences of the same form across sections for visual clarity. The queried terms were *junckvrowen*, *suster*, *susterne*, *vrouwe*, *vrowen*, *wif*, *frouw*, *frowen*, and *frouwenn*. The graph was generated with the CATMA Distribution tool after lemmatisation of the ordinance texts. Graph prepared by the author.

the relevant passages: the concentration of *suster*, for example, can be examined in relation to clauses addressing the community as brothers and sisters, while other terms may occur in different regulatory contexts. The distribution visualisation therefore functions primarily as a means of locating textual patterns and identifying passages for subsequent qualitative reading.

While the Distribution tool visualises where selected terms occur within the sampled ordinances, the Keyword in Context tool provides a closer view of their immediate textual surroundings. The KWIC table (see Figure 4) lists each occurrence by document name, queried keyword, and a brief surrounding context, making it possible to examine how terms referring to women appear within specific clauses. This is important because concordance analysis moves corpus work closer to qualitative reading: the searched word is no longer treated only as a countable item, but as part of a regulation whose meaning depends on its legal and textual context. Using KWIC analysis, it is possible to follow the use of specific terms and compare their positions in text qualitatively. The term *suster*,

Document Name	Keyword	KeyWord In Context
schomakere lvva 8 1 15	junckvrowen	welk man unse werk ein ***junckvrowen*** nympth de in unse werk
schomakere lvva 8 1 15	suster	mogen item welk broder edder ***suster*** de olderman ofte de bisitter
schomakere lvva 8 1 15	vrouwe	vd denen item wesen ein ***vrouwe*** in de werk de er
schomakere lvva 8 1 15	vrouwe	wesen willen al ein bederv ***vrouwe*** . vortmer se scholen he
schomakere lvva 8 1 15	vrowen	werk al ein par knopeder ***vrowen*** scho . ein par wyve
schomakere lvva 8 1 15	vrowen	ein kost beyde man unde ***vrowen*** unde he scholen geven ein
schomakere lvva 8 1 15	wif	komen van buten to de ***wif*** unde kind mit sik brengen
schomakere lvva 8 1 15	wif	jemant he wesen man edder ***wif*** ofte kind se scholen de
schuhmacher 1416 1477 tla 190 2 023	vrouwe	ampt sterven . dat si ***vrouwe*** edder man se scholen man
schuhmacher 1416 1477 tla 190 2 023	vrouwe	de vesper de gelik de ***vrouwe*** item vart up de dag
schuhmacher 1416 1477 tla 190 2 023	vrouwe	. he edder sin wif ***vrouwe*** offt ymant anders van er
schuhmacher 1416 1477 tla 190 2 023	vrouwe	vortmer se ein scholen ne ***vrouwe*** in de bod stan .
schuhmacher 1416 1477 tla 190 2 023	vrouwe	ge wesen dat mans unde ***vrouwe*** er besten ledder scholen an
schuhmacher 1416 1477 tla 190 2 023	vrouwen	edder bu Ķ¦ ten sunder ***vrouwen*** schol vortmer wenick werkckmester to
schuhmacher 1416 1477 tla 190 2 023	vrouwen	solde de man vor de ***vrouwen*** betalen ein ber ber j
schuhmacher 1416 1477 tla 190 2 023	vrouwen	man se scholen man unde ***vrouwen*** . van unse ampt dar
schuhmacher 1416 1477 tla 190 2 023	wif	wesen . he edder sin ***wif*** vrouwe offt ymant anders van
schuhmacher 1481 1536 tla 190 2 024	frowen	se scholen de man unde ***frowen*** van unse ampt dar wesen
schuhmacher 1481 1536 tla 190 2 024	suster	dat ein jewelik broder unde ***suster*** de amptes scholen weszen to
schuhmacher 1481 1536 tla 190 2 024	suster	crispinian ein jewelik broder unde ***suster*** scholen antheen sin besten kledere
schuhmacher 1481 1536 tla 190 2 024	vrouwe	unse ampt sterven de si ***vrouwe*** ofte man se scholen de
schuhmacher 1481 1536 tla 190 2 024	vrouwe	dar to offerende id si ***vrouwe*** edder man item de dar
schuhmacher 1481 1536 tla 190 2 024	vrouwe	solden de man vor de ***vrouwe*** betalen j mark vortmer wesen
schuhmacher 1481 1536 tla 190 2 024	vrouwe	brode id he edder sin ***vrouwe*** edder yemant anders van er
schuhmacher gesellen riga 1414 1480	vrouwe	in de ehre unse leven ***vrouwe*** sante maria to holden .
schuhmacher gesellen riga 1414 1480	frouw	kompanie , de ein echt ***frouw*** hebben , unde de thovornn
schuhmacher gesellen riga 1414 1480	frouwenn	de pelendeı ferding unse levenn ***frouwenn*** elendigheidt allmiszen wesen .
schuhmacher gesellen riga 1414 1480	suster	, se scholen wi unnse ***suster*** unde broder laten beghann ,
schuhmacher gesellen riga 1414 1480	suster	scholen de gemeinen broder unde ***suster*** to komen to d .
schuhmacher gesellen riga 1414 1480	suster	) item welk broder edder ***suster*** sterven van unse elendigheidt ,
schuhmacher gesellen riga 1414 1480	susterne	scholen ne vrouwe nemen to ***susterne*** , se wesen de uthem
schuhmacher gesellen riga 1414 1480	vrouwe	brödern de werde , de ***vrouwe*** edder er gesinde vordret edder
schuhmacher gesellen riga 1414 1480	vrouwe	werdenn , de ein echt ***vrouwe*** hebben . 27) item
schuhmacher gesellen riga 1414 1480	vrouwe	bröderschop de elendigheit unse levenn ***vrouwe*** van wesen to wesen ,
schuhmacher gesellen riga 1414 1480	vrouwe	) item men scholen ne ***vrouwe*** nemen to susterne , se
schuhmacher gesellen riga 1414 1480	vrouwe	geselle de compenye unse leven ***vrouwe*** elendigheidt willen avergevenn mit hogmode
schuhmacher gesellen riga 1414 1480	vrouwe	unnsen geselle ginge to hovescher ***vrouwe*** hus , unde dar jenig
schuhmacher gesellen riga 1414 1480	vrouwe	unde schuldig bleve unse leven ***vrouwe*** elendigheidt unde darna wedder to

Figure 4. Keyword in Context concordance of selected Middle Low German terms referring to women in four sampled Shoemaker ordinances. The table was generated with the CATMA Keyword in Context tool and contains the columns "Document Name", "Keyword", and "Keyword In Context", showing the immediate textual context around each queried term. The queried terms correspond to those used in Figure 3: *junckvrowen*, *suster*, *susterne*, *vrouwe*, *vrowen*, *wif*, *frouw*, *frowen*, and *frouwenn.* Table prepared by the author.

for example, is paired with the masculine *broder*, defining the addressed community as brothers and sisters. In the Tallinn Shoemaker ordinance of 1481–1536, a regulation associated with the feast of St Crispinian states that "ein jewelik broder unde suster scholen antheen sin besten kledere" (every brother and sister shall put on their best clothes).[35] In other contexts, by contrast, *vrouwe* is the principal term used to refer to women, demonstrating why differences between these terms require contextual rather than purely quantitative analysis.

Even though it could be argued that the same analysis could be done through qualitative reading, using corpus analysis tools for visualisation and concordance analysis allows one to grasp a wider picture of the selected manuscripts through the approached question. The value of such analysis is not that the data becomes a research result automatically, or that the number of occurrences can be treated as a direct measure of historical importance. Rather, it helps the historian move across a larger body of sources and identify where closer reading should begin. The present comparison uses absolute occurrences rather than frequencies normalised according to document length. Differences in the number of hits between the four ordinances should therefore not be interpreted as direct quantitative evidence of the relative prominence of women in one text or craft community.

In the case of the Shoemaker ordinances, the searches reveal differences between the sampled texts: in absolute terms, references to women occur more often in the Journeymen's ordinance and in the later statutes, while the Distribution visualisation additionally shows whether these occurrences are concentrated in particular parts of a text or dispersed across several regulations. This does not yet prove the role of women in the craft, but it does lead to a more precise hypothesis: in the Shoemaker material, women appear in contexts that should be examined further in relation to apprenticeship, craft membership, and the regulation of the craft community. The KWIC table makes it possible to return directly to the relevant clauses and examine the legal and textual context in which these terms occur. The examples discussed above show the advantage of corpus analysis in a qualitative approach, revealing where terms may function differently depending on their location and direct context in the clause. The method therefore works best as a bridge between data and interpretation. Lemmatisation can reduce some problems caused by Middle Low German spelling variation, but it does

35 Schuhmacher Schragen 1481–1536, Tallinn. TLA, 190.2.024, fol. 10r. https://dl.tlu.ee/en/source-editions/gildide-skraad (accessed 20/09/2026).

not remove the difficulties of synonyms, semantic nuance, context, or the normative character of the statutes. Corpus analysis is thus most useful not as a substitute for qualitative reading, but as a way of locating patterns, passages, and comparisons that would be difficult to keep in view across the whole source group.

FROM DATA TO VISUALISATION

The transformation of the craft ordinances into searchable text creates the conditions for another form of analysis: the visualisation of source content through tagging and relational modelling. While corpus analysis works through selected terms in their original contexts, visualisation requires the source material to be structured in a specific way. People, organisations, events and topics found among regulations must first be identified and encoded before they can be compared or represented as networks. This means that the visualisation is not produced directly from the manuscript or even from the transcription alone. It is produced from a second analytical layer created on the basis of the transcribed text.

The tool used for visualising craft ordinance data is Nodegoat, a research environment that allows the creation of relational databases for humanities data.[36] In this case, the database includes source references, original images, transcriptions, and manually encoded entities or attributes, such as craft organisations, people, events, and thematic factors. The advantage of such a platform is that the database can be modelled according to specific research questions, not only according to archival or bibliographical structure. Through tagging, a regulation can be linked to a person, office, organisation, event, or thematic category, allowing relations to be traced that are not always visible through concordance analysis alone. At the same time, the resulting data is not extracted neutrally from the digital edition: it is shaped by the researcher's reading of the ordinance, the chosen categories, and the questions being asked. Nodegoat then makes it possible to visualise these encoded relations within the same environment, for example as links between sources, people, crafts, institutions, or thematic categories. The same data can also be exported for further analysis in Gephi, R, or Python, but for this article Nodegoat is sufficient as a tool for preparing, modelling, and visualising data without requiring advanced programming skills.

36 P. van Bree, G. Kessels. Nodegoat: Enabling Explorative Research. Paper presented at DH2017, 2017.

Database building for historical documents is most valuable when prosopographical information is available.[37] The information concerning specific people can be structured in objects that hold personal information, including family and given names, gender and ethnicity, occupations, ties to organisations and other available data. These, then, through common categories, can be visualised, highlighting their connections through organisations, time frames, or appearance in the same documents.

The craft ordinances of late medieval Riga and Tallinn contain a relatively small amount of prosopographical data and only a few ordinances list personal names. The named individuals are usually town council members who were included in the composition and approval of the statute, and the acting aldermen or assessors of the craft. For example, the Riga councillor Johan van Lynden appears in the ordinances of several crafts, including the Furriers, Coopers, and Smiths, providing a named example of an individual connecting otherwise separate craft records.[38] Some ordinances also mention other names of members, with the mention usually being connected to their position in the craft, a milestone in professional life or donations made to the guild. However, some identifications remain uncertain, and visualising the prosopographical networks can help to resolve them. Comparing variant name forms and recorded affiliations helped identify the roles of some of the people mentioned, usually highlighting town councillors and craft members with long-term positions who were therefore repeatedly mentioned by the same craft. Figure 5 visualises people mentioned in the medieval craft ordinances of Riga and Tallinn, connecting them to the relevant statutes and, where possible, to their affiliation with a craft, town council, or other organisation.

The visualisation shows that most craft ordinances contain no, or only very few, personal names. When names do appear, they are often connected with the approval, confirmation, or administration of the statutes. In the Riga material especially, town councillors form an important part of the network, since they are frequently mentioned in connection with the formal approval of craft ordinances. The density of this part of the network suggests that some council members appeared repeatedly in relation to different crafts and over several years. This may

37 C. Raffensperger. Visualizing Prosopography through Digital Humanities. – Medieval Prosopography, 2019, 34, 207–220.

38 Kürschner Schragen 1397. LNA_LVVA, F8-1-15, fols. 6v–8v; Böttcher Schragen 1375. LNA_LVVA, F8-1-15, fols. 12r–14r; Schmiede Riga 1426–1530. – Schragen der Gilden und Aemter der Stadt Riga bis 1621, 463.

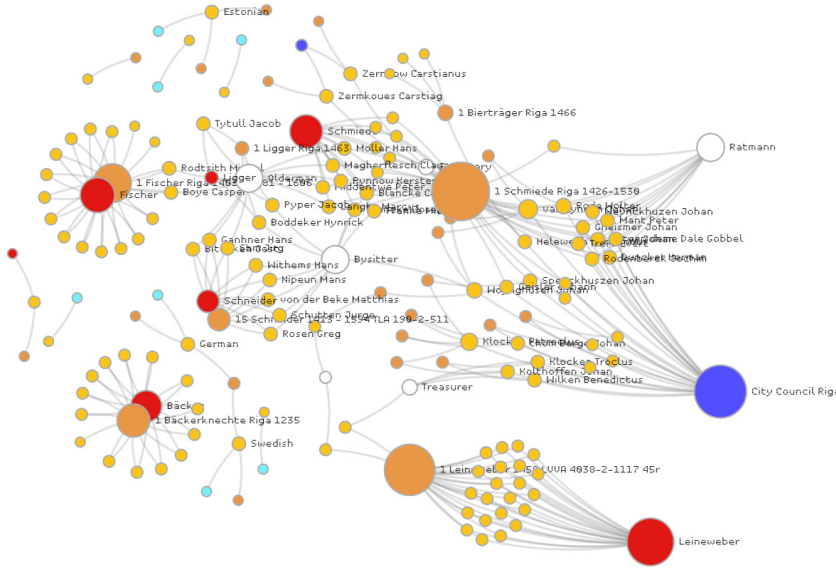


Figure 5. Network visualisation of people mentioned in medieval Livonian craft ordinances. The graph connects named individuals (yellow nodes) to the ordinances in which they appear (orange nodes) and, where possible, to their craft, position (white nodes), town council (blue nodes), or other organisational affiliation. The visualisation highlights the limited but significant presence of prosopographical data in the ordinance material, especially the role of craft office holders and town council representatives in the approval and regulation of craft statutes. Visualisation created by the author in Nodegoat.

indicate a stronger administrative connection with the craft communities, although such a conclusion requires further research. Other council representatives appear only in a few ordinances or alongside more frequently mentioned figures, which may point to a more limited or temporary role in the regulation of the crafts.

The same visualisation also makes visible the difference between ordinances that mention only external authorities and those that preserve more information about members of the craft community itself. The figure does not prove personal relationships in a strict sense, but it indicates where repeated interactions between crafters, aldermen, and councillors may be studied more closely.

The second visualisation example in Figure 6 concerns thematic factors encoded in the ordinance texts that are chosen for analysis. These factors include responsibility, behavioural, work, legal, religious, economic, personal, and social regulations. “Responsibility” was used for clauses concerning obligations of members and duties of members elected in positions; “behavioural” for rules regulating acceptable or unacceptable conduct; “work” for regulations connected with craft practice, production, tools, competition, or labour; “legal” for clauses

concerning confirmation, procedure, or formal regulation; "religious" for references to worship, saints, masses, processions, candles, and church as a space; "economic" for fixation of prices, payments, wages, and provision of material; "personal" for references to individual status, to personal achievements, family relations, illness, death; "social" for feasts, gatherings, communal meals, and participation in the collective life of the craft.

Unlike named people, these categories are not entities found directly in the text in a simple form. They have been created by the researcher through manual interpretation of individual regulations. In the Nodegoat data model, each ordinance forms an object containing individual manuscript pages as sub-objects, while transcription lines within these pages are encoded with relevant categories, including the thematic factors used here. A passage may therefore receive more than one factor: for example, a regulation concerning attendance at a feast may be encoded as religious, social, and responsibility related, while a rule describing craft production standards may be linked to work, legal, and economic factors. Figure 6 does not count individual clauses as mutually exclusive units. Its visual unit is the manuscript page: a factor is connected to a page when the transcription on that page contains material encoded with that category, and the page is in turn connected to its parent ordinance. Multiple coding therefore allows the same page to be connected to several factors.

Figure 6 visualises the presence of these manually encoded factors across the Shoemaker ordinances. The categories are analytical constructs created to compare the thematic orientation of the statutes rather than objective divisions inherent in the sources. The size of a factor node reflects how many pages in the selected material are linked to that factor: a larger node indicates that the category occurs across more encoded pages, not a greater number or proportion of individual regulations. The figure should therefore be read as an exploratory overview of the distribution and co-presence of thematic factors rather than as an exact quantitative measurement of ordinance content. It can reveal which factors recur across more pages and how these pages connect to the four ordinances, directing attention towards differences that can subsequently be examined through close reading. It should not, however, be read as a direct measure of what mattered most in the craft community, but as a visualisation of how the written regulations appear and were classified within the database.

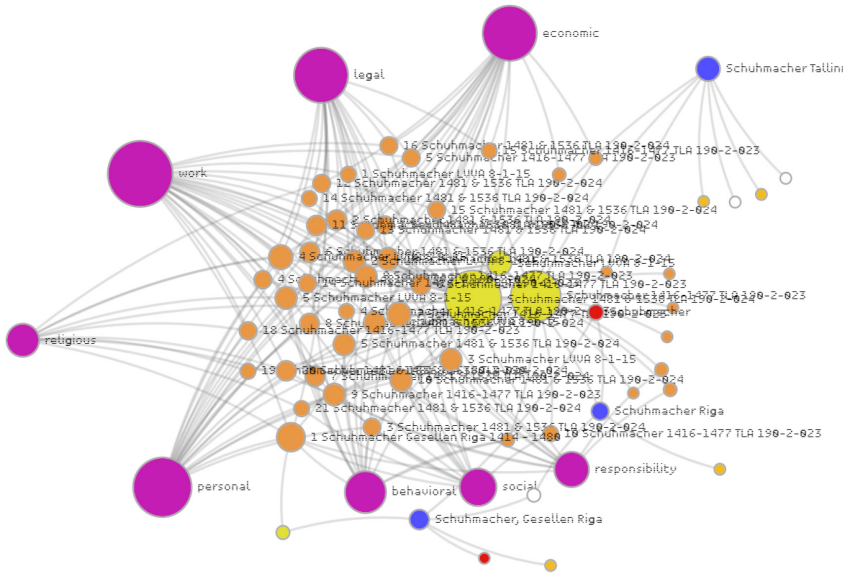


Figure 6. Visualisation of manually encoded thematic factors in four selected Shoemaker ordinances from Riga and Tallinn. The factors (pink nodes) include responsibility, behavioural, work, legal, religious, economic, personal, and social regulations. The visualisation shows how often regulations connected to these factors appear in the selected ordinance texts (orange nodes) and allows cautious comparison between these factors in the sampled ordinances. Visualisation created by the author in Nodegoat.

This visualisation of the thematic factors in Shoemaker ordinances suggests that among all four documents from both Riga and Tallinn, the most common themes in all regulations combined are connected to work, personal, economic and legal factors. The behavioural, religious, social and responsibility factors are represented less frequently, but hold a similar frequency among each other. If this visualisation is used for quantitative analysis, it can suggest that the medieval Livonian shoemakers, based on their normative documents, devoted more clauses to work and economic life than to religious and social events and responsibilities. However, these results must be read cautiously, as the data reflects both the source material and the author's decisions about classification. This observation can be formulated as a hypothesis for closer reading, as the figure shows the relative frequency of encoded factors, not practical importance and so must be combined with close reading of the context.

Apart from analysing the results of factor-based encoding in several ordinances of the same craft, the same visualisation strategy can be used for other encoded ordinances. Comparing the thematic coding results with those for other crafts might reveal which crafts had

a stronger emphasis on clauses related to social or religious factors in their ordinances. Such results do not reveal the full reality of craft life, but they do reveal what different craft communities chose, or needed, to formalise in writing. The factors revealed in the visualisation can be paired with close reading and corpus analysis tools not only to approach these factors, but also to question the type of the craft guild and its main goals. Depending on terms used to describe the craft, its members and the visualisation of revealed factors, there can be an argument as to whether some of the craft communities placed greater written emphasis on work-related issues, or managed the craft as a guild with a strong religious and social component. Such combined approaches can be used as case studies when approaching specific craft guilds, their events or overall traditions in the addressed town or region.

The usefulness of this form of visualisation lies in its ability to make a large and internally varied corpus more navigable. This helps the researcher to approach the data visually to identify clusters, absences, and contrasts between ordinances. For example, one craft may appear to place greater emphasis on religious and social obligations, while another may be more strongly represented through work-related or economic regulations. Such patterns do not become conclusions in themselves. They are indications that direct the researcher back to the relevant clauses and to the specific legal and social context of each ordinance.

Visualisation is the next level of working with extracted and digitised data. The network visualisations discussed here depend on an additional layer of information encoded by the researcher, whereas the corpus analysis visualisations represent patterns in the transcribed text. This layer consists of encoded objects and topics selected in relation to the research questions. The strength of such visualisations, based on network analysis in the created database, lies in showing visual patterns, clusters, and topical or personal relations and makes these easier to perceive across a larger body of material. Their main value is exploratory as they help locate concentrations, repeated relations, and unexpected connections that can then be tested through renewed close reading. A limiting argument is the researcher's bias in the collected database and therefore the visualisation, since the output reflects the author's categories, interests and interpretation of the source. However, historians in their work are always subject to choices, and the digital condition of sources shapes the research and interpretation practices.[39] The critical

39 S. Schwandt. Opening the Black Box of Interpretation: Digital History Practices as Models of Knowledge. – History and Theory, 2022, 61, 4, 77–85, here 80. https://doi.org/10.1111/hith.12281.

question is not whether the visualisation is fully objective, but whether its categories, links, and limits are made clear enough for the reader to understand how the pattern was produced. Digital visualisation of such data should therefore be treated not as a final representation of social reality in craft communities, but as a structured way of identifying patterns and showing relations that are traceable across many manuscripts of similar source material.

DISCUSSION: PRODUCING HISTORICAL KNOWLEDGE THROUGH DIGITISATION?

For a scattered source group preserved across different archives, manuscript units, digital photographs, printed editions, and previous transcription practices, digitisation creates conditions for approaching the ordinances as a whole. The Digital Livonia platform, and craft guild ordinances as part of it, does not aim to replace the archive but to create an accessible, research-oriented digital version of the source, in which manuscripts can be approached through a common framework. The craft ordinance, as a normative document with similarities across the region and several centuries, becomes a possible object of comparison when digitised as part of a larger corpus. Such a corpus makes it possible to compare regulations across crafts, towns, and periods, and to ask whether some occupational communities produced more detailed normative texts than others. Research into these questions depends on access to several digital layers of the source, such as the manuscript image, transcription, metadata, archival reference, and information about earlier editions. The transcribed ordinances, structured by layout and transcription, form a corpus that allows the historian to explore the material in ways that are difficult through linear reading alone. Terms can be traced across ordinances, their distribution can be visualised, and their immediate context can be examined. The corpus can also be exported into other research environments, where it may be tagged, visualised, and compared with other source groups.

The examples of visualised data suggest that digitisation itself involves interpretive decisions by the researcher or author of the digital editions. As discussed by Silke Schwandt with reference to Michel

de Certeau, history writing can be understood as an interpretative practice. From this perspective, the digital corpus likewise does not simply reproduce the past,[40] it makes explicit how the sources have been selected, arranged, transcribed, and prepared for analysis. The corpus is therefore both a representation of medieval documents and a product of present scholarly practice. However, the data and questions selected for visualisation should also be carefully chosen by the researcher. Bruno Latour warns that images, diagrams, lists, archives, files, collections, and other forms of inscription can explain "almost everything or almost nothing", depending on how they are brought into focus. Their importance does not lie simply in being visual, written, or printed.[41] They become analytically powerful when they allow resources to be mobilised, compared, recombined, and brought into new relations. When applied to medieval ordinances, this idea suggests that the visualisations based on quantitative data cannot replace traditional historical research and qualitative reading, but can introduce new questions and suggest patterns that have to be explored with combined methods.

Quantitative approaches to medieval ordinances introduce specific technical challenges and amplify some difficulties also encountered in qualitative research. Among these are the terminological inconsistencies, imperfect textual data and metadata, software limitations and the level of proficiency required to use digital tools. However, the increasing availability of LLMs can assist humanities researchers in organising data and carrying out case studies with the source corpus. As a digital body of data, the medieval text corpus also creates possibilities for research beyond the interests of the author and, when shared among researchers, can assist studies in legal and economic history, historical linguistics and other fields. The digital editions can also be approached in connection with other source groups, such as account books, town council records, or other available urban sources from medieval Riga and Tallinn.

Future developments in digital humanities and tools designed for humanities researchers could further improve the computational analysis opportunities of such sources. In addition to the development of HTR for source transcription and NLP for lemmatisation, opportunities with Named Entity Recognition[42] tools could make historical language

40 Ibid., 79.

41 B. Latour. Visualization and Cognition: Thinking with Eyes and Hands. – Knowledge and Society, 1986, 6, 1–40, here 4.

42 T. Hiltmann, S. Dumont, L. Körfer, A. Monroy, S. Thode. NER4all or Context Is All You Need: Using LLMs for Low-Effort, High-Performance NER on Historical Texts: A Humanities Informed Approach. – arXiv, 4 February 2025. https://doi.org/10.48550/arXiv.2502.04351.

more accessible for digital analysis, with automated identification of people, places, and other named entities in medieval documents. Even so, automatic encoding may become accessible in the future, but manual classification still has the advantage of keeping the researcher close to the text. The requirement to read the text and assign categories through interpretation makes the digital version of the source serve both as a basis for advanced quantitative and visual analysis, and as a corpus that represents the medieval document and allows traditional approaches to the study of medieval history.

CONCLUSION

From accessibility to visualisation, the digitisation of medieval manuscripts produces and reintroduces several layers of historical knowledge. This article has asked how the digitisation of medieval Livonian craft ordinances changes the possibilities for analysing this source group, and what methodological limits must be considered when normative texts are transformed into searchable and visualisable data. Based on the source work and case studies, it can be concluded that digitisation improves access, enables comparison, and supports exploratory analysis, but only when it remains connected to source-critical reading. Quantitative methods and visualisation can suggest patterns and raise new questions for the researcher who has access to the corpus and the skills to pair the data with qualitative interpretation of the original sources. As the digitised corpus of ordinances becomes an established digital source, it can continue to support the exploration of specific questions of medieval craft practices, with visualisations that reflect patterns and suggest associations in support of historical analysis.

Acknowledgements

This research has received funding from the Estonian Research Council grant PRG1276. All figures included in this article are either screenshots or graphics created by the author. The reproduced manuscript images are based on digitised material from the Latvian State Historical Archive (LVVA) and the Tallinn City Archives (TLA).

KÄSIKIRJAST VISUALISEERINGUNI: DIGITAALNE LÄHENEMINE KESKAJA LIIVIMAA KÄSITÖÖLISTE SKRAADELE

Rūta Bruževica

Artikkel käsitleb keskaja Liivimaa käsitööliste skraade digiteerimist ja uurib võimalusi, mida digiteerimine avab selle allikarühma analüüsimiseks, arutledes ühtlasi metodoloogiliste piirangute üle, mis kaasnevad normatiivsete tekstide muutmisel otsitavaks ja visualiseeritavaks andmestikuks. Artikkel tugineb keskaja Liivimaa tsunftide skraade allikakollektsioonile, mis on osa „Digitaalse Liivimaa" projektist ning kuhu on esmakordselt ühtseks korpuseks koondatud Riia ja Tallinna digiteeritud ja transkribeeritud käsitööliste skraad 14. sajandi algusest kuni 1550. aastani.

Käsitööliste skraad on normatiivsed dokumendid, mis reguleerisid keskaegses linnas nii majanduslikke ja õiguslikke tsunftiküsimusi kui ka tsunftikogukonna sotsiaalset ja religioosset elu. Skraad sisaldavad teavet käitumisnormide kohta koosviibimistel ja pidustustel, usutalitustel osalemise kohta ning kohustustest kogukonnaliikmete ees haiguse, surma või vaesuse korral. Kuigi skraad on ametlikult kinnitatud tekstid, näitab nende korduv kinnitamine, hilisemad täiendused ja üksikjuhtumeid käsitlevad erisätted, et need ei olnud kogukonnaelust lahutatud jäigad õigustekstid, vaid kasutusel olevad dokumendid, mida kujundas ja muutis tsunftikogukonna igapäevane elu.

Tsunftide käitumispraktika uurimine tugineb digitaalsele tekstikorpusele, mis on koostatud Tallinna Linnaarhiivis ja Läti Riiklikus Ajalooarhiivis asuvate originaalkäsikirjade digiteerimise käigus. Lisaks on kaasatud korpusesse skraade transkriptsioonid, mis on kättesaadavad üksnes Wilhelm Stieda ja Constantin Mettigi 1896. aastal ilmunud trükiväljaandes, kuna originaalkäsikirjad ei ole säilinud. Arhiivides asuvate käsikirjade digitaalseid fotosid töödeldi platvormil Transkribus, kus need jaotati tekstiväljadeks ja ridadeks ning seoti vastavate transkriptsiooniridadega, luues sel moel mitmekihilise digitaalse väljaande, kus käsikirja kujutis ja transkribeeritud tekst jäävad omavahel püsivalt seotuks. Töö Transkribusega Liivimaa käsitööliste skraade ning sama piirkonna ja perioodi teiste allikatega toetas keskaja Liivimaa alamsaksakeelsete allikate käsikirjatuvastusmudeli (HTR-mudeli) väljatöötamist. Valminud digitaalne editsioon ühendab käsikirja kujutise, transkriptsiooni ja

metaandmed ühtses keskkonnas, vastates seega tänapäevase teadusliku digiväljaande põhitingimustele.

Artiklis esitatud juhtumiuuringud illustreerivad ajalooliste dokumentide transkribeeritud tekstikorpuseks muutmise interdistsiplinaarseid võimalusi. Korpuslingvistika meetodite rakendamist käsitööliste skraade analüüsimisel tutvustatakse tööriistaga CATMA loodud visualisatsioonide kaudu. Selline analüüs andis terminijaotuse visualisatsiooni, millest ilmneb näiteks, kuidas naisi tähistavate terminite kobardumisi konkreetsetes skraades saab nende esinemissageduse alusel ära tunda ja esile tõsta. Teine CATMA analüüsitööriist, mis koostab märksõnade konteksti loendi (KWIC), võimaldas terminid koos nende vahetu ümbritseva kontekstiga eraldada, et neid seejärel kvalitatiivselt analüüsida.

Käsitööliste skraade sisu visualiseerimise juhtumiuuringud on koostatud andmebaasitööriista Nodegoat abil. Nende juhtumiuuringute tulemusena valmis kaks visualisatsiooni: esiteks skraades nimeliselt mainitud isikute võrgustik, mis peegeldab olemasolevat, kuigi piiratud prosopograafilist teavet tsunftiliikmete ja raehärrade kohta; ning teiseks valitud skraade rühma temaatiliste tegurite ülevaade, mis hõlmab selliseid kategooriaid nagu töö, majandus, õigus, religioon, sotsiaalne elu, käitumine, vastutus ja isiklikud asjaolud. Mõlemad visualisatsioonid sõltuvad uurija määratud kategooriatest ning on esitatud uurimist suunavate vahenditena: need ei kinnita järeldusi, vaid aitavad uurijal tuvastada mustreid, kobardumisi ja lünki, mille põhjal saab sõnastada edasisi uurimisküsimusi nii kvantitatiivse kui ka kvalitatiivse analüüsi jaoks.

Artiklis jõutakse järeldusele, et keskaja Liivimaa dokumentide digiteerimine parandab hajutatud allikmaterjali kättesaadavust uurijatele ning võimaldab seda käsitleda tervikliku korpusena. Juhtumiuuringute põhjal võib digiteerimine toetada võrdlusi eri linnade, tsunftide ja ajastute vahel ning võimaldada uusi analüüse. Samas ei saa kvantitatiivne analüüs ega visualiseerimistehnikad kunagi asendada allikakriitilist lugemist ega originaalallikate kvalitatiivset tõlgendamist. Kuna ajalooliste tekstide transkribeerimise, lemmatiseerimise ja nimeüksuste tuvastamise tööriistad arenevad edasi ning võivad aidata püstitada uusi uurimisküsimusi, jääb teksti lähilugemine jätkuvalt uue ajaloolise teadmusloome keskmesse.

Acta Historica Tallinnensia 32/2 https://doi.org/10.3176/hist.2026.2.04

Digital Methods and the Study of Fragmentary Medieval Grave Slabs

Anu Mänd

University of Tartu, Institute of History and Archaeology, Ülikooli 18, 50090 Tartu; anu.mand@ut.ee

Abstract. The study of medieval grave slabs is often hampered by wear, fragmentation, reuse, and the poor legibility of inscriptions. This article examines the potential of digital methods for the study of grave slabs from medieval Livonia, focusing on Reflectance Transformation Imaging (RTI) and digital photomontage. Case studies demonstrate how RTI can enhance the legibility of worn inscriptions and carved decoration, while digital photomontage enables the virtual reassembly of dispersed fragments that can no longer be physically reunited. The article also explores the digital integration of surviving fragments with historical drawings, showing how such reconstructions can facilitate their interpretation and presentation in churches and museums. These methods can generate new art-historical and epigraphic knowledge while also enhancing the accessibility and public understanding of medieval grave monuments.

Keywords: medieval grave slabs; medieval Livonia; Reflectance Transformation Imaging; epigraphy; inscriptions

Grave slabs are invaluable sources for the study of medieval social history, visual culture, epigraphy, and heraldry. Large-scale, artistically sophisticated, and aesthetically striking memorials were costly commissions that could generally be afforded only by rulers, members of the nobility, and high-ranking clergy. From the late Middle Ages onwards, grave slabs also survive for prominent urban elites, such as town councillors and wealthy merchants. These monuments were decorated with human figures, religious symbols, inscriptions, and coats of arms or

© 2026 Author. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International License CC BY 4.0 (http://creativecommons.org/licenses/by/4.0).

merchants' marks that both identified and commemorated the deceased through visual and textual means.[1]

The size, decoration, and location of grave slabs within the church interior served as markers of social status. The study of their material, style, and ornamentation provides new insights into medieval art history and the working methods of medieval stonemasons. Their inscriptions yield valuable information about the deceased, including the date of death and family relationships, while also shedding light on broader aspects of medieval mentality, religious belief, and literate culture. Examining grave monuments together with their spatial and liturgical context offers new perspectives on the culture of commemoration in the Middle Ages.[2]

The study of medieval grave slabs, and especially their inscriptions, presents exceptional challenges for several reasons. As individuals of high social status were generally buried inside churches (and, in monastic complexes, also in other spaces such as chapter houses and cloisters[3]), the floor slabs marking their graves have been subjected to centuries of wear and are often now almost illegible. In general, the older the slab, the poorer its state of preservation. Preservation also depends on the material used in a particular region: marble and brass, for example, tend to survive better than limestone or sandstone. The carving technique is another important factor. High-relief figures and inscriptions are usually better preserved than low-relief decoration or incised slabs.

Deciphering and transcribing inscriptions is particularly difficult because parts of the text may have been worn away entirely. The task is further complicated by the frequent use of abbreviations, missing letters, and spelling errors.[4] Such errors may have been introduced not only by medieval stonemasons but also by nineteenth- and twentieth-century restorers, who were not always able to read Gothic minuscule script correctly and inadvertently altered the original text.[5] Moreover, grave slabs

1 See, e.g., N. Saul. English Church Monuments in the Middle Ages: History and Representation. Oxford University Press, Oxford, 2009, 92–93, 108–111, 114–119; K. Krüger. Corpus der mittelalterlichen Grabdenkmäler in Lübeck, Schleswig, Holstein und Lauenburg (1100–1600). Thorbecke, Stuttgart, 1999, 173–190.

2 N. Saul. English Church Monuments, 120–142; A. Mänd. Keskaegne kunstiteos kui kirjalik allikas. – Vana Tallinn, 2023, 32 (36), 68, 74–76. For commemoration practices in medieval Livonia, see G. Strenga. Remembering the Dead: Collective Memory and Commemoration in Late Medieval Livonia. Brepols, Turnhout, 2023.

3 T. Coomans. Life inside the Cloister. Understanding Monastic Architecture: Tradition, Reformation, Adaptive Reuse. Leuven University Press, Leuven, 2018, 104–107.

4 See, e.g., V. Brandis, J. L. de Jong, R. Seidel. Introduction. – Funerary Inscriptions in Early Modern Europe: Shaping Identities to Remember. Ed. by V. Brandis, J. L. de Jong, R. Seidel. Brill, Leiden, Boston, 2025, 4.

5 A. Mänd. Visuaalne mälestamine: Liivimaa ordumeistrite ja käsknike hauaplaadid (14.–16. sajand). – Kunstiteaduslikke Uurimusi / Studies on Art and Architecture, 2019, 28, 3–4, 83;

were often reused, both during the Middle Ages and the early modern period. As a result, the original inscriptions were frequently damaged by the addition of new texts, coats of arms, and other commemorative symbols.

The surviving medieval grave monuments from Livonia (roughly corresponding to present-day Estonia and Latvia), together with those known from historical drawings, date from around 1200 to the mid-sixteenth century.[6] In Estonia alone, more than two hundred inscribed slabs are thought to survive, although their exact number cannot be established because no comprehensive catalogue has yet been compiled and the National Register of Cultural Monuments includes only protected monuments.[7] Most grave slabs were carved from local varieties of limestone, abundant particularly in northern and western Estonia. Written sources also mention a small number of brasses,[8] but none have survived.[9]

Relatively few grave slabs remain intact and in good condition. Most survive only as fragments, either as single pieces or as several fragments that can no longer be reassembled into a complete monument. Grave slabs have been damaged or destroyed by fires and warfare,[10] but many were also deliberately broken or relocated. This occurred, for example, when church chancels or naves were rebuilt, when space was needed for new memorials in the early modern period, or when church floors were repaired or replaced in later centuries.

A. Mänd. Grave Slabs of Ecclesiastical and Secular Dignitaries in Livonia, *c.1470–c.1560.* – Funerary Inscriptions in Early Modern Europe: Shaping Identities to Remember. Ed. by V. Brandis, J. L. de Jong, R. Seidel. Brill, Leiden, Boston, 2025, 55–56.

6 H. Loeffler. Die Grabsteine, Grabmäler und Epitaphien in den Kirchen Alt-Livlands vom 13.–18. Jahrhundert. G. Löffler, Riga, 1929; K. Markus. Esimesed kristlikud hauatähised. – Eesti kunsti ajalugu 1, 1100–1520. Koost. K. Markus. Eesti Kunstiakadeemia, Tallinn, 2023, 387–390; A. Mänd. Hauaplaadid 14. sajandist 1550. aastateni. – Eesti kunsti ajalugu 1, 1100–1520. Koost. K. Markus. Eesti Kunstiakadeemia, Tallinn, 2023, 391–395; E. Grosmane. Stone Sculpture. – Art History of Latvia I, 9th Millennium BC – 1562. Ed. by E. Grosmane. Art Academy of Latvia, Riga, 2024, 256–261.

7 A. Mänd. Arutlusi hauatähistest ja muinsuskaitsest: Keskaegsed hauaplaadid Tallinna toomi käärkambris. – Aja lugu: Muinsuskaitse ja restaureerimise ajaloost. Toim. A. Randla. (Eesti kunstiakadeemia toimetised, 21). Eesti Kunstiakadeemia, Tallinn, 2016, 10–14.

8 I. Põltsam-Jürjo. Liivimaa väikelinn Uus-Pärnu 16. sajandi esimesel poolel. Argo, Tallinn, 2009, 406; A. Mänd. Church art, commemoration of the dead and the saints' cult: Constructing individual and corporate *memoria* in late medieval Tallinn. – Acta Historica Tallinnensia, 2011, 16, 24.

9 The only known exception is a bronze roundel bearing the symbol of St Mark the Evangelist, which once adorned the corner of a limestone grave slab found in the Dominican church in Tallinn. M. Kurisoo. *Ars moriendi* – suremise kunst. Näitus Niguliste muuseumis 02.11.2012–02.06.2013. Eesti Kunstimuuseum, Tallinn, 2013, 60.

10 See, e.g., M. Kurisoo. Patching Together a Thousand Pieces. Tombstones of Tallinn's St Nicholas' Church: History, Restoration and Layout in the Exposition of the Niguliste Museum in the 1970s. – Kunstiteaduslikke Uurimusi / Studies on Art and Architecture, 2013, 22, 3–4, 153–176.

Wear, damage, fragmentation, and limited accessibility are the principal reasons why the study of medieval grave slabs is so time-consuming and challenging. Moreover, art historians are generally not trained to read inscriptions in Latin and Middle Low German. The monochrome character of Estonian and Latvian grave slabs, together with the relative rarity of figural decoration, has probably also contributed to their limited appeal for art historians.

Although digital methods for the study of grave monuments – including digital epigraphy – have advanced rapidly across Europe in recent years,[11] they have so far seen only very limited application in Estonia and Latvia. Rather than discussing the full range of digital methods that might prove promising for future research, this article adopts a user-oriented perspective based on my own experience. It examines the methods I have tested and assesses their effectiveness in reading inscriptions and identifying the deceased. In particular, I focus on Reflectance Transformation Imaging (RTI) and digital image processing, including photomontage. Through a series of case studies, I demonstrate how and to what extent these methods facilitate the analysis and visual reconstruction of fragmentary grave slabs.

FRAGMENTS AND REFLECTANCE TRANSFORMATION IMAGING (RTI)

The condition and degree of preservation of medieval grave slab fragments vary considerably depending on their age, location, and other environmental factors. Within churches, fragments may be embedded in the floor, mounted vertically on walls, or stored loose against a wall, in the sacristy, beneath the tower, or in the porch. In some cases, they have been removed from view altogether and deposited in the attic space above the vaults. Because dressed stone was a valuable building material, grave slabs were frequently reused, and fragments are therefore often found in a wide range of secondary contexts. They commonly served as construction material: for example, they were built into the exterior or interior walls of churches and bell towers (as at Valjala and Kaarma

11 See, e.g., Epigraphy in the Digital Age: Opportunities and Challenges in the Recording, Analysis and Dissemination of Inscriptions. Ed. by I. Velázquez Soriano, D. Espinosa Espinosa. Archaeopress, Oxford, 2021; B. Luci. Unveiling the Hidden Stories: The Impact of Digital Technologies in Medieval Latin Epigraphic Field Documentation. – Digital Applications in Archaeology and Cultural Heritage, 2026, 40, https://doi.org/10.1016/j.daach.2025.e00488.

churches on the island of Saaremaa) or reused as lintels (for example in the window of the north porch of Koeru Church). Grave slab fragments also occur as stair treads, both inside churches (as on the staircase leading to the vaults in the north-east corner of Haapsalu Cathedral) and outside them (as on the sacristy steps of Umurga Church in Latvia[12]).

The preservation of grave slabs and their fragments is strongly influenced by their location, i.e. whether they are exposed outdoors or kept indoors and, in the latter case, whether they remain embedded in the floor or have been mounted on walls for display. Unsurprisingly, monuments exposed to the elements suffer the greatest weathering, especially those lying at ground level or fixed vertically to exterior walls. Indoors, the most vulnerable are slabs set into the floors of heavily trafficked areas of churches (or of former churches now functioning as museums). Among the most damaging forms of reuse is the conversion of a medieval grave slab into a church threshold. In such cases, the slab is subjected both to constant foot traffic and, when located at an exterior entrance, to the effects of weathering. As a result, its decoration often becomes almost completely obliterated, making it impossible to determine its original owner, date, or decorative programme.

Two parish churches in Estonia provide particularly striking examples. While the threshold slab of Haljala Church is located just inside the south portal, the corresponding slab at Jüri Church lies directly outside the entrance (Fig. 1). In both cases, centuries of foot traffic combined with weathering have worn the surfaces so severely that



Figure 1. Worn late medieval grave slab serving as the threshold of Jüri Church, Estonia. Photo by Stanislav Stepaško

12 A. Caune, I. Ose. Latvijas viduslaiku mūra baznīcas 12. gs. beigas – 16. gs. sākums: Enciklopēdija. Latvijas vēstures institūta apgāds, Rīga, 2010, 439.

only faint traces of the Gothic minuscule inscription and fragments of the corner ornaments survive along the edge nearest the wall.

One digital method that has proven particularly useful for the study of worn grave slabs is Reflectance Transformation Imaging (RTI), a technique that combines multiple photographs taken under light from different angles. In recent decades, RTI has become an established method for the digital documentation and analysis of archaeological and cultural heritage surfaces. RTI enhances the visibility of shallow engravings and inscriptions and allows researchers to relight the surface interactively, making it possible to read inscriptions and identify motifs that are invisible or barely visible to the naked eye.[13] In Estonia, RTI has been used, for instance, in the study of daguerreotypes[14] and medieval seal impressions[15], but has so far been less frequently applied to large surfaces such as grave slabs.

At St Olaf's Church in Tallinn, one of the city's two medieval parish churches, fragments of medieval and early modern grave slabs have been relocated to the porch at the western end of the nave. Most of the fragments are stored loose along the south wall, while two larger medieval fragments have been mounted on the west wall of the porch, one on each side of the portal. There are no labels providing information about these slabs. The northern slab is somewhat better preserved, whereas the slab to the south of the entrance, which bears a fragmentary merchant's mark incorporating an unidentified coat of arms at its centre, is extremely worn.[16] Previous researchers have been unable to decipher a single word of its inscription.[17] Although the inscription is indeed illegible to the naked eye, digitally processed photographs (Fig. 2) nevertheless made it possible to decipher individual letters and even a few words from the inscription on the left-hand side of the slab.[18]

13 For the use and potential of this method, see, e.g., G. Earl, K. Martinez, H. Pagi. Reflectance Transformation Imaging (RTI) System for Ancient Documentary Artefacts. – VAST 2010: The 11th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, 116–119; G. Earl, K. Martinez, T. Malzbender. Archaeological Applications of Polynomial Texture Mapping: Analysis, Conservation and Representation. – Journal of Archaeological Science, 2010, 37 (8), 2040–2050; H. Mytum, J. R. Peterson. The Application of Reflectance Transformation Imaging (RTI) in Historical Archaeology. – Historical Archaeology, 2018, 52 (2), 489–503.

14 H. Pagi, J. Miles, A. Uueni, S. Hogarth, K. Sikka. Reflectance Transformation Imaging in Daguerreotype Investigation. – Archiving 2017 Final Program and Proceedings, 2017, 116–121.

15 See the contribution by Juhan Kreem in this volume.

16 Monument no. 2148. https://register.muinas.ee/public.php?menuID=monument&action=view&id=2148 (accessed 25/06/2026).

17 M. Loit. Keskaegsest surmakultuurist ja hauatähistest reformatsioonieelse Tallinna kirikutes ja kloostrites. – Vana Tallinn, 2006, XVII (XXI), 64.

18 A. Mänd. Oleviste kiriku keskaegsest sisustusest ja annetajate ringist. – Acta Historica Tallinnensia, 2014, 20, 41–43, figs. 4–5.



Figure 2. Fragment of a burgomaster's grave slab in St Olaf's Church, Tallinn. Digitally enhanced image. Photo by Stanislav Stepaško

At my request, the professional photographer Stanislav Stepaško produced RTI images of the two medieval grave slab fragments mounted on the west wall of the porch. The resulting images made it possible to read both the inscriptions and the fragmentary coats of arms carved in the centre of each slab more clearly. RTI proved particularly valuable for the study of the slab south of the entrance, whose surface is heavily worn and damaged. Its inscription, carved in Gothic minuscule, is further complicated by unusually elongated letterforms, making the text especially difficult to decipher. In addition to the words that had already been identified during my earlier examination, RTI revealed the deceased's office on the left-hand side of the inscription: *de bo[r]gemest[e]r* (the burgomaster). It also showed that I had previously misread one of the words. Thanks to RTI, the text along the left edge of the fragment can now be read in its entirety as *de bo[r]gemest[e]r de got gnade.* The final word, gnade, did not fit within the inscription band and was therefore carved above it in smaller letters.

Unfortunately, RTI did not improve the legibility of the inscription along the upper edge and right-hand side of the fragment, where the surface is even more severely worn than on the left. There, I was still able to identify only a few isolated letters. It is possible that future investigation of this and other heavily damaged fragments could benefit from 3D scanning and photogrammetry. This grave slab fragment from St Olaf's Church will be discussed again later in the article.

DIGITAL PHOTOMONTAGE OF FRAGMENTS

In most cases, fragments belonging to the same grave slab are kept in the same location, whether in a church or a museum. However, this does not necessarily mean that they are situated next to one another. In the early modern period, when space was needed for new grave monuments, and during later repairs to church floors, medieval grave slabs were sometimes cut into two or more pieces and re-laid at different locations within the floor. As a result, visitors – including art historians – may fail to recognise that separate fragments once formed part of the same monument. Even when the connection is noticed, it is usually impossible to lift the fragments from the floor and physically reassemble them.

A notable exception is provided by the fragments of the grave slab of the vassal Woldemar Sorsever and his wife, dating from the 1370s, which were removed from the floor of the nave of Tallinn Cathedral in 2016. The slab, surviving in at least thirteen fragments, was subsequently conserved and has since been displayed on the wall of the cathedral's Fersen Chapel.[19] The principal arguments for rescuing the monument were its early date, its status as a listed monument, and the imminent risk of further deterioration, as the fragments were located in one of the busiest parts of the church.[20] Clearly, however, it is neither feasible nor desirable to remove all medieval grave slabs from church floors, and in most cases there is no suitable space available for their display.

Where fragments cannot be physically removed and reassembled, they can instead be digitally reunited by means of photomontage, using

19 T. Sakermaa. Woldemar Sorseveri hauaplaadi konserveerimine. – Muinsuskaitse aastaraamat 2017. Muinsuskaitseamet, Tallinn, 2018, 76.

20 See M. Klammer, T. Sakermaa. Tallinna toomkirik. Woldemar Sorseveri hauaplaat: seisukord ja ettepanekud. Ettepanekud hauaplaatide säilitamiseks. Tallinn, 2013. Manuscript in the Archives of Tallinn Urban Planning Department, TLPA MO, n. 12, s. 504.

software such as Adobe Photoshop. This approach not only enables the virtual reconstruction of a monument's original appearance and facilitates the identification of the deceased, but also enhances the scholarly value of small, isolated fragments that might otherwise pass almost unnoticed when embedded in the church floor. The following case studies illustrate the potential of this method.

Three worn fragments of medieval grave slabs are embedded in the chancel floor of St Maurice's Church in Haljala (historical Viru County, northern Estonia). The largest fragment is situated in the western part of the chancel, on the boundary between the nave and the chancel; the second lies along the north wall, and the third, partially covered by the altar rail, is located south of the altar. Thanks to a drawing made by Danish illustrator Søren Abildgaard in 1754 (Fig. 3), it is possible to establish that the first two fragments belonged to the grave slab of the nobleman Hinrick Hastfer, who died in 1523. Admittedly, Abildgaard recorded only the year of death: he did not identify the deceased, and on the right-hand inscription band he left a blank space where the deceased's first name should have appeared.[21] It is possible that in the mid-eighteenth century an object – perhaps a lectern – stood on top of the slab and prevented Abildgaard from seeing Hinrick's name, which remains clearly legible on the larger surviving fragment today.

A digital photomontage combining the larger fragment with the smaller strip (Fig. 4) makes it possible to reconstruct much of the slab's original appearance. Only the fragment from the lower right corner is missing. The two coats of arms in the centre of the slab, belonging to Hinrick Hastfer and his wife, a member of the Taube family, have suffered considerably more wear than is evident in the mid-eighteenth-century drawing. In contrast, the inscription is relatively well preserved. Combining Abildgaard's drawing with the surviving fragments makes it possible to reconstruct the entire Middle Low German inscription: *Int iar m v^{c} / xxiii starf hinrick / hastuer va[n] / katte[n]tack de got* (In the year 1523 died Hinrick Hastfer of Kattentack.[22] May God [have mercy on his soul]). The inscription records the year of death, but not the exact day or month. In written sources, Hinrick Hastfer (Hastever)

21 P. Grinder-Hansen. Søren Abildgaard (1718–1791): Fortiden på tegnebrættet. Nationalmuseet, Copenhagen, 2010, 650, no. 842. Poul Grinder-Hansen was unaware that the fragmentary grave slab has survived in the church. Abildgaard's drawing is also available at: https://samlinger.natmus.dk/DMR/asset/56149 (accessed 25/06/2026).

22 Kattentack refers to Aaspere manor in Haljala parish, in the historical Viru County.

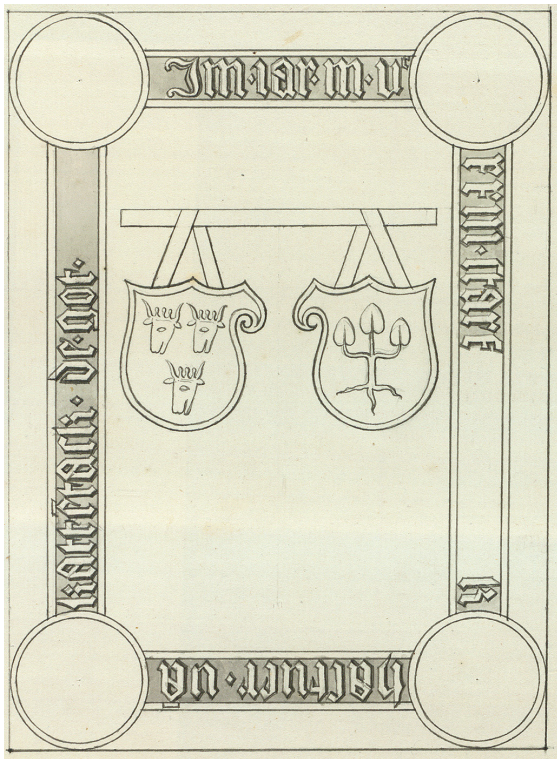


Figure 3. Grave slab of Hinrick Hastfer (†1523) and his wife. Haljala Church, Estonia. Drawing by Søren Abildgaard, 1754. National Museum of Denmark



Figure 4. Grave slab of Hinrick Hastfer (†1523) and his wife. Haljala Church, Estonia. Photos and photomontage by Stanislav Stepaško

is last mentioned on 19 April 1523, when he was serving as district judge (*Mannrichter)* of Viru County.[23]

A comparison between Abildgaard's drawing and the surviving fragments also reveals that the drawing is not entirely accurate, either in its proportions or in certain details. The original slab was somewhat shorter in relation to its width than Abildgaard depicted. Moreover, he left the corner roundels blank, whereas they originally contained the symbols of the Evangelists. The upper left roundel can still be identified as the symbol of St Luke, while the lower left roundel preserves the symbol of St Matthew.

A second example comes from St Nicholas' Church in Tallinn (since 1984 the Niguliste Museum), where the floor of St Anthony's Chapel contains, among other monuments, a fragment of the grave slab of the city's journeyman minters (Ger. *Münzgesellen*). The slab bears a Gothic minuscule inscription along its border and, in the centre, a coat of arms with the date 1533 in Arabic numerals. The grave slab was already fragmentary by the late nineteenth century, when it was first described by Eugen von Nottbeck.[24] However, Nottbeck failed to recognise that a second fragment of the same slab had also survived, a strip approximately 40 cm wide and 160 cm long, preserving the upper corner of the shield enclosed within a roundel and part of the marginal inscription. It is possible that before the Second World War this strip was not yet embedded in the floor of St Anthony's Chapel, since, according to the art historian Mai Lumiste (1932–1985), who directed the restoration of St Nicholas' Church and its grave slabs after the devastating fire of 1944, the fragment was discovered beneath the floor of the nave.[25] Lumiste did not realise that it belonged to the larger fragment of the journeyman minters' slab.[26] The connection likewise escaped Mari Loit (1977–2016), who compiled the catalogue of the Niguliste Museum's grave slabs in 2002.[27]

23 Est- und Livländische Brieflade, T. 1, Bd. 1. Hrsg. von F. G. von Bunge, R. von Toll. Kluge und Stroem, Reval, 1856, 484, no. 902.

24 E. von Nottbeck, W. Neumann. Geschichte und Kunstdenkmäler der Stadt Reval. Bd. 2, Die Kunstdenkmäler der Stadt. 2nd edn., Kluge, Reval, 1904, 169–170, no. 84. In Nottbeck's description, only the first three digits of the date (153) are recorded. The fourth digit may have been concealed beneath a pew or another heavy object and was therefore not visible.

25 M. Lumiste. Niguliste kirik. Antoniuse kabeli hauaplaatide paigaldamine. Manuscript in the Archives of Tallinn Urban Planning Department, n. 5, s. 250C. Tallinn, 1973, 8, no. 30. See also the floor plan at the beginning of her manuscript.

26 Ibid., 6, no. 23.

27 M. Loit. Tallinna Niguliste kiriku hauaplaatide kataloog. Manuscript in the Art Museum of Estonia. Tallinn, 2002, 18, no. 29, cf. 16, no. 23.



Figure 5. Grave slab of journeyman minters, 1533. Niguliste Museum, Tallinn. Photos and photomontage by Stanislav Stepaško

Because the floor of St Anthony's Chapel is completely paved with grave slabs, it is impossible to remove and physically reassemble the two fragments. Here again, digital photomontage provides a solution by allowing the fragments to be reunited virtually (Fig. 5).

The partially preserved inscription along three sides of the journeyman minters' grave slab remains reasonably legible: *<v>nde dritteyn vort / hebben hir dissen sten g[...] / [...]st de*[28] *muntergsellen (sic!) to ev*[29] *[...]r weg*[30] (... and thirteen have here this stone ... the journeyman minters at [Reval]). If the word *dritteyn* (thirteen) refers to the year 1513, an obvious question arises: why does the coat of arms in the centre bear the date 1533? Both the inscription and the central motif were recut during the restoration campaign of the 1970s, as was the case with several other grave slabs in St Anthony's Chapel.[31]

28 The letters *d and e form a ligature.*
29 Nottbeck read *r[evel]* (Geschichte und Kunstdenkmäler, 170), whereas the first letter now appears to be *e*. The error may have been introduced when the inscription was recut, since the Gothic minuscule forms of *r* and *e* are relatively similar.
30 The letters *w* and *e* form a ligature.
31 M. Kurisoo. Patching Together a Thousand Pieces, 158–161. See also M. Lumiste, R. Kangropool. Niguliste kirik. Kunst, Tallinn, 1990, 60.

While the fragments of the journeyman minters' slab remain in the same room, more complicated cases also occur, in which fragments of a single monument have ended up in different buildings because their relationship has gone unrecognised. One such example is the grave slab of the burgomaster discussed above, whose largest surviving fragment is mounted on the west wall of the porch of St Olaf's Church (see Fig. 2). Another, smaller fragment from the same slab, originating from the lower right corner, also survives in the porch but has been placed along the south wall among numerous other grave slab fragments. This corner piece preserves a double roundel containing an Evangelist's symbol and the fragmentary Middle Low German inscription *starf her ioha[n]* (died Master Johan).[32]

More remarkably still, one or perhaps two[33] additional fragments of the same monument are embedded in the floor of the south aisle of St Nicholas' Church (now the Niguliste Museum). Their connection with the fragments in St Olaf's Church has hitherto gone unnoticed. By means of digital photomontage, all fragments can be reunited virtually (Fig. 6). In the resulting reconstruction, the Niguliste fragments occupy the lower part of the left-hand side of the slab.

How the fragments of the St Olaf's grave slab came to be incorporated into the floor of St Nicholas' Church remains unclear. One possible explanation is that, during the post-war decades when St Nicholas' Church was being restored and converted into a museum, some of its architectural stonework was temporarily stored in St Olaf's Church, and when the floor of the nave at St Nicholas' was laid in the late 1970s or the early 1980s,[34] some grave slab fragments discovered at St Olaf's may inadvertently have been transferred along with the material destined for reinstallation. At present, however, this remains only a hypothesis, as I have found no documentary evidence confirming that stonework from St Nicholas' Church was ever stored at St Olaf's.[35]

Another comparable case dates from the transitional period between the Middle Ages and the early modern era. In the westernmost bay (the space beneath the tower) of St Mary's Church at Pöide on the island of Saaremaa, several fragments of medieval grave slabs are

32 A. Mänd. Oleviste kiriku keskaegsest sisustusest, 43–44, fig. 6.

33 As the adjoining fragments are partially obscured by cement, it is uncertain whether they represent one fragment or two.

34 M. Lumiste. Ettepanek hauaplaatide *pro* fragmentide paigaldamiseks pikihoone, koori ja Väikese kabeli põrandatesse ning maa-aluse vestibüüli seintele. Tallinn, 1979. Manuscript in the Archives of Tallinn Urban Planning Department, n. 5, s. 250B.

35 Rein Kaur, a restorer of the grave slabs, now aged 86, does not recall any of the carved stones having been kept in St Olaf's Church or transferred from there to St Nicholas' Church. Telephone conversation with the author, August 2026.



Figure 6. Burgomaster's slab. Two larger fragments in St Olaf's Church, Tallinn (cf. Fig. 2); two smaller fragments in the floor of the Niguliste Museum. Photos and photomontage by Stanislav Stepaško

preserved. Among them are three fragments belonging to the right-hand side of the grave slab of Johann Möller (Müller), lord of Kingli (Ger. Möllershof) Manor, and his wife Anna Treiden. The corner roundels are decorated with stylised double roses, while one fragment preserves part of the Treiden family coat of arms. The fragments also retain portions of inscriptions from two different periods, the earlier of which includes the date 1552.

Two additional fragments of the same grave slab survive some 50 kilometres away, in the collection of carved stone monuments at the Saaremaa Museum in Kuressaare Episcopal Castle. The fragment forming the upper left corner had previously been recognised by museum staff as belonging to the Möller–Treiden slab, as it preserves the Möller family coat of arms and part of the marginal inscription containing the family name *MVLLER*.[36] The museum also houses the lower left corner of the slab, but this fragment has not previously been associated with the monument.[37] Although it is somewhat more worn than the other pieces, the double rose in the corner roundel, together with the inscription – whose script, content, and layout correspond closely to those of the other fragments – clearly demonstrates that they belong to the same grave slab.

By digitally combining all five fragments (Fig. 7), it is possible to reconstruct most of the slab's original appearance, although some losses remain in both the central field and along the margins. The inscription running around the border reads: *ANNO 1552 / DEN* [blank left for the day of the month] *IS DER EDEL ERENFEST / VNT MANHAF/DIGE IOHAN MVLLER IN GOT / ENTSLAPEN* (In the year 1552, on the [day left blank], the noble, honourable, and valiant Johann Möller fell asleep in God). A later, four-line inscription was subsequently added beneath the coats of arms in the centre of the slab: *ANNO 1586 IS IOHAN / MVLL[ERS SIN]E ELICHE HVS/FRUWE ANNA TREIDEN / IN GOT ENTSLAPEN* (In the year 1586, Johann Möller's lawful wife Anna Treiden fell asleep in God).

Because the inscription has previously been misread, reference works have given Johann Möller's year of death as 1568.[38] The grave slab, however, clearly demonstrates that he died considerably earlier, in 1552.

Evidence concerning when the right-hand portion of the slab was discovered and when the left-hand fragments were transferred to Kuressaare Castle is incomplete. When Martin Körber (1817–1893) described the artistic heritage of Pöide Church in the late nineteenth century, he mentioned a broken grave slab lying in front of the church entrance, from which only the date 1586 and the family name Möller –

36 Inv. no. SM 10438:23, available at: https://www.muis.ee/museaalview/304136 (accessed 25/06/2026).

37 Inv. no. SM 10438:22, available at: https://www.muis.ee/museaalview/304135 (accessed 25/06/2026).

38 Genealogisches Handbuch der Baltischen Ritterschaften. Teil: Oesel. Bearb. von N. von Essen. Saaremaa Üldkasuliku Ühingu kirjastus, Tartu, 1935–1939, 186–187, no. VI.1. The source of the dates of death for Johann Möller and Anna Treiden given in the inscription on p. 187 is unknown.



Figure 7. Grave slab of Johann Möller (†1552) and Anna Treiden (†1586). The fragments on the left are in the Saaremaa Museum (Kuressaare Castle); those on the right are in Pöide Church. Photos and photomontage by Stanislav Stepaško

from which the name of Möllershof (Estonian: Kingli) Manor was derived – could still be read.[39] It is therefore possible that, in Körber's time, only the left half of the slab, or perhaps only its upper left fragment, remained at Pöide Church. A drawing in a manuscript by the graphic artist Johann Naha, dated 1939, depicts both fragments from the left-hand side of the slab, indicating that they had already been transferred to Kuressaare Castle by that date.[40]

39 M. Körber. Oesel einst und jetzt. Bd. 3. Arensburg, 1915, 40.

40 Johan Naha käsikirjad. Kuressaare. Tallinn, 1939, 35, fig. [4]. Manuscript in the Archives of the National Heritage Board, ERA.5025.2.12265, available at: https://register.muinas.ee/ftp/DIGI_2013/pdf/era5025_002_0012265.pdf (accessed 25/06/2026).

According to the Estonian Museums Public Portal, the lower left corner of the slab was discovered in the former cemetery of Kuressaare, located on the site of the present-day town park.[41] The reliability of this information, however, is open to question.

The fragments from the right-hand side of the slab, now preserved beneath the tower of Pöide Church, are in considerably better condition than those from the left. As these fragments were not documented in the church before the Second World War, they probably came to light later, most likely during archaeological investigations. It has not been possible to determine exactly when this occurred. It must, however, have taken place during the second half of the twentieth century, since the investigations carried out at Pöide Church in the twenty-first century are exceptionally well documented.[42]

Because most of the Möller–Treiden grave slab can now be reconstructed, it would be highly desirable to physically reunite the fragments that are currently dispersed between different locations. After conservation, the reassembled monument should ideally be returned to its original location in Pöide Church.

COMBINING HISTORICAL DRAWINGS AND FRAGMENTS DIGITALLY

Whereas the previous section examined grave slabs of which the greater part survives despite their fragmentary condition, this section focuses on monuments whose surviving fragments provide little indication of their original appearance. How can the broader public be made aware that even such isolated fragments retain considerable value as historical sources?

Many grave slabs that have now been wholly or partially destroyed are documented in eighteenth- or nineteenth-century drawings. One way of giving surviving fragments a new life is therefore to place them digitally onto these historical drawings, thereby visualising which part of the original monument each fragment represents. This approach is feasible, however, only if the historical drawing is reasonably accurate in its proportions, something that cannot always be expected, particularly

41 https://www.muis.ee/museaalview/304135 (accessed 25/06/2026).

42 For a brief history of archaeological studies, see: V. Kadakas. Arheoloogilised uuringud Pöide kirikus. Aruanne. Tallinn, 2023, 5–6. Manuscript in the Archives of the National Heritage Board, available at: https://register.muinas.ee/ftp/Arheoloogiliste%20uuringute%20aruanded/L-9102_Aeelu_2012_KadakasV_kirik-21058_Saaremaa-Pöide-Maarja.pdf (accessed 25/06/2026).

of eighteenth-century artists. For example, the surviving fragments of the Hastfer grave slab in Haljala Church discussed above cannot be combined with Søren Abildgaard's 1754 drawing (see Figs. 3 and 4) without substantially distorting the proportions of the drawing.

An instructive example of how surviving fragments can be identified and visualised with the aid of a nineteenth-century drawing is provided by the fifteenth-century grave slab of Brun Warendorp from St Nicholas' Church in Tallinn. Like many other grave slabs in the church, it was shattered during the Soviet air raid of March 1944 and the subsequent collapse of the nave vaults a few years later. The subsequent history of this monument, however, differs from that of the slabs that were restored during the 1970s and reinstalled in the floor of the church, which was then being converted into a museum.

In 2002, Tiina Mägi, director of the Kodulinna House in Tallinn, discovered several hundred fragments of limestone in the former studio of the sculptor and conservator Paul Horma (1940–1988). Some of the fragments preserved parts of inscriptions and carved decoration. Mägi transferred the material to the Estonian Academy of Arts for safekeeping. The stone conservator Isabel Aaso-Zahradnikova, together with her students, began to study the finds and established that they represented the remains of two grave slabs. Their provenance, however, remained unknown. The mystery began to unravel in 2016, when an interview with Rein Kaur, formerly the stone conservator at the Niguliste Museum, revealed that two grave slabs had disappeared during the restoration of the church in 1981. One of them, represented by 206 surviving fragments, was identified by Merike Kurisoo as the monument of Casper Goldberg (†1635).[43]

The second slab, whose Gothic minuscule script indicates a late medieval date, survives in only fifteen fragments. Most preserve no more than two or three letters, although one fragment contains the word *geboren* (born), and another bears the later inventory number *No. 37*. On the basis of these details, together with Eugen von Nottbeck's late nineteenth-century description and drawing,[44] I was able to identify the monument as the grave slab of Brun Warendorp, a merchant from Münster who became a member of the Table Guild of Tallinn's Great Guild in 1428.[45] By the late nineteenth century, the upper edge of the

43 I. Aaso-Zahradnikova, M. Kurisoo, A. Mänd. Killustikust kunstiteoseks: Niguliste kaks taasleitud hauaplaati. – Muinsuskaitse aastaraamat 2017. Muinsuskaitseamet, Tallinn, 2018, 77.

44 E. von Nottbeck, W. Neumann. Geschichte und Kunstdenkmäler, 162, no. 58, fig. 119.

45 Tallinn City Archives, TLA.191.2.1, fol. 10v.

slab, which originally bore the year of Warendorp's death, had already been lost. The inscription surviving on the remaining three sides read: *fridages na svnte mathaeus dach do starf / brvn warendorp va[n] / mvnster geboren biddet ghode vm[m]e de sele* (On the Friday after St Matthew's Day died Brun Warendorp, born in Münster. Pray to God for his soul).

A comparison of the surviving fragments with Nottbeck's description and drawing reveals that, although the drawing is accurate in its proportions, some of its details are incorrect. The corner ornaments are depicted incorrectly: instead of simple roundels, they originally consisted of a combination of square and circular frames. In addition, Nottbeck omitted the later inscription *Johannes Andreä, med. Doct.,* which had been added after the slab was reused by the physician Johannes Andreä in 1631.[46]

From 1 to 5 November 2017, the Niguliste Museum hosted a small exhibition displaying the fragments of both grave slabs. The more visually striking exhibit was the Goldberg monument, whose 206 surviving fragments made it possible to reconstruct the slab's outline with reasonable accuracy.[47] In contrast, the fifteen fragments of the Warendorp slab, simply arranged side by side against the wall (Fig. 8), attracted considerably less attention.

Only later did it occur to me that the larger fragments, whose original positions on the slab could be identified, might be digitally superimposed on Nottbeck's nineteenth-century drawing (Fig. 9).[48] This approach not only enhances the interpretative value of the surviving fragments but also demonstrates that the corner ornaments were in fact of a different design from those shown in Nottbeck's drawing.



Figure 8. Fifteen fragments of the grave slab of Brun Warendorp on display in the Niguliste Museum in 2017. Photo by Stanislav Stepaško

46 E. von Nottbeck, W. Neumann. Geschichte und Kunstdenkmäler, 162, no. 58.
47 I. Aaso-Zahradnikova, M. Kurisoo, A. Mänd. Killustikust kunstiteoseks, 77, fig. 1.
48 Ibid., fig. 2.

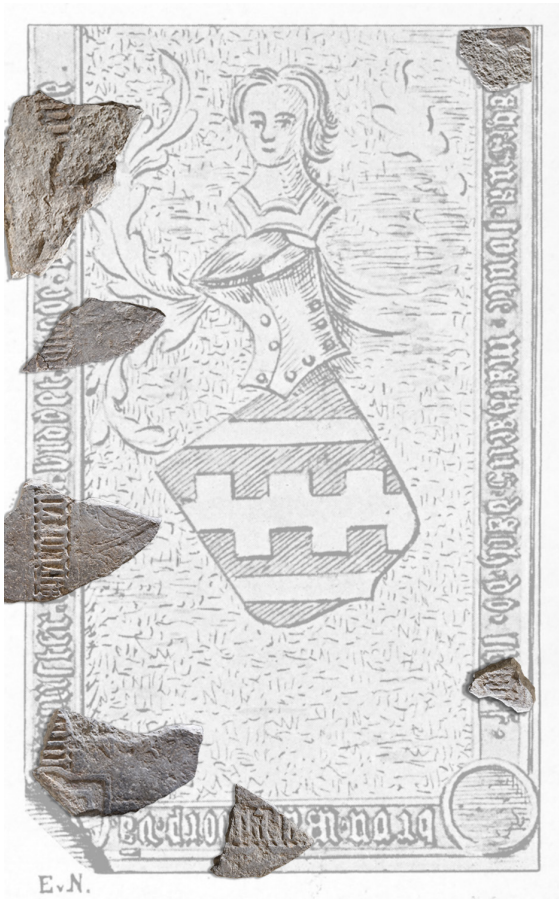


Figure 9. Grave slab of Brun Warendorp. First half of the 15th century. Niguliste Museum, Tallinn. Drawing by Eugen von Nottbeck, 1890s. Photos by Stanislav Stepaško; photomontage by Villu Plink

Following the exhibition, the fragments of both grave slabs were transferred to the Art Museum of Estonia, with the hope that a suitable location for their permanent display might eventually be found in the Niguliste Museum. To date, however, no such space has become available, and the fragments remain in museum storage, inaccessible to the public. The principal reason is presumably the limited exhibition space available in the museum, where visually more appealing works of art – such as paintings and wooden sculptures – are generally given priority over fragments of grey limestone grave slabs.

The method described above has also proved applicable to several other medieval grave slabs. One example is the monument of Heidenricus Savijerwe (†1460), a vassal and the brother of Bartholomäus Savijerwe, Bishop of Tartu. Numerous fragments of his grave slab survive in the collections of the Estonian National Museum in Tartu, but they are too incomplete to form either a coherent reconstruction or even half of the original monument.[49] Fortunately, the slab was recorded in a historical drawing by Johann Christoph Brotze (1742–1823), who most likely copied a drawing by Eduard Philipp Körber (1770–1850).[50] Two of the larger surviving fragments, originating from the upper left and lower

49 A. Mänd. Art in Medieval Tartu. – Baltic Journal of Art History, 2024, 27, 105–107; A. Mänd. Kunstist keskaegses Tartus. – Õpetatud Eesti Seltsi aastaraamat 2024 / Annales Litterarum Societatis Esthonicae. ÕES, Tartu, 2026, 91–92.

50 Johann Christoph Brotze *Estonica*. Ed. by A. Hein, I. Leimus, R. Pullat, A. Viires. Estopol, Tallinn, 2006, 258. On Brotze's copying of several drawings by Körber, including that of this particular slab, see ibid., VIII.

right corners of the slab, can therefore be digitally superimposed onto the historical drawing.[51] Through this method, the surviving pieces are transformed from what might otherwise appear to be mere fragments of stone into recognisable parts of the funerary monument of an influential medieval vassal.

A similar case is provided by the grave slab of Wolter von Plettenberg (†1535), Master of the Livonian Branch of the Teutonic Order, which originally lay in front of the high altar of St John's Church in Cēsis. A drawing of the monument by Johann Christoph Brotze, dating from 1771, has survived. Although several smaller pieces of the slab were already missing at that time, the centrally placed effigy of the deceased and the marginal inscription were still preserved in their entirety and fully legible. Today, only three fragments of the slab survive, all originating from its lower part. These preserve only Plettenberg's knees and parts of his sword and shield.[52] In Latvia, the fragments are regarded as important cultural monuments and are protected by the state.[53] Because they are visually rather unremarkable, however, historians usually reproduce Brotze's drawing and simply note that the grave slab has been destroyed.[54]

One may, of course, question when a grave slab should truly be considered destroyed: when it has been broken into pieces, when only a few fragments remain, or only when no fragments survive at all? Both the complete slab and, later, its surviving fragments have been relocated several times within St John's Church.[55] Today, the fragments are mounted on the wall of the south aisle, accompanied by a label reproducing Brotze's drawing.

Two of the three surviving fragments can be digitally superimposed onto Brotze's drawing, which proves to be remarkably accurate in its details (Fig. 10). The third fragment, however (the leftmost of the three), lacks the decoration that should occupy this area according to the drawing, while the surviving portion of its inscription band differs in design from the inscription preserved on the right-hand side of the

51 A. Mänd. Art in Medieval Tartu, 106, fig. 5; A. Mänd. Kunstist keskaegses Tartus, 92, fig. 3.

52 A. Mänd. Visuaalne mälestamine, 74, fig. 9; A. Mänd. Grave slabs of Ecclesiastical, 59, fig. 3.3.

53 No. 3094, Kapa plāksnes fragments V. Pletenbergam, https://likumi.lv/ta/id/343726-valsts-nozimes-makslas-pieminekla-izmantosanas-un-saglabasanas-noradijumi (accessed 25/06/2026).

54 E.g., J. Kreem. Ordu sügis: Saksa ordu 16. sajandi Liivimaal. Tallinna Linnaarhiiv, Tallinn, 2022, 97.

55 F. von Klocke. Zu den Bildnissen und zum Grabstein Wolter von Plettenbergs. – Westfalen. Hefte für Geschichte, Kunst und Volkskunde, 1941, 26/1, 34; G. Kalniņš. Cēsu Svētā Jāņa baznīca. Cēsu Kultūras un tūrisma centrs, Cēsis, 2015, 14 (fig. 22), 44, 50 (fig. 62); A. Mänd. Visuaalne mälestamine, 85; A. Mänd. Grave slabs of Ecclesiastical, 60.

slab. This raises the question of whether it belongs to Plettenberg's monument at all.

In any case, digital photomontage conveys more effectively that the grave monument of the Livonian Master has not been completely lost. Even the surviving fragments remain valuable sources: they allow the dimensions of the inscription band and the Gothic minuscule lettering to be measured, the provenance of the limestone to be investigated, and the carving techniques employed by the stonemason to be studied.



Figure 10. Grave slab of Wolter von Plettenberg (†1535). St John's Church, Cēsis, Latvia. Drawing by Johann Christoph Brotze, 1771. Photos and photomontage by Stanislav Stepaško

CONCLUSIONS AND FUTURE PERSPECTIVES

The widespread adoption of digital photography during the twenty-first century, together with increasingly sophisticated image processing techniques, has created new opportunities for the study and digital documentation of medieval grave slabs. Standard digital photographs can be enhanced by adjusting contrast, lighting, and other image parameters to improve the visibility of worn surfaces. Even greater potential is offered by Reflectance Transformation Imaging (RTI), which has proved particularly effective in deciphering weathered inscriptions and symbols on grave slabs.

Digital photomontage provides another valuable tool by enabling the virtual reassembly of fragments that are physically separated, for example, when they are embedded in different parts of a church floor. Since there is little prospect of dismantling church or museum floors simply to reunite dispersed fragments, photomontage offers a practical means of visualising their original relationship.

Some grave slabs have become even more widely dispersed over time, with fragments ending up in different buildings. The Möller–Treiden monument discussed in this article is one such example, its fragments now divided between Pöide Church and Kuressaare Castle. In this case, physical reassembly, conservation, and redisplay of the monument would in fact be feasible. It is hoped that the digital photomontage presented here will help draw attention to this possibility.

For several grave slabs documented in eighteenth- or nineteenth-century drawings, only a few fragments survive today. These fragments are often visually unremarkable and tend to be overlooked by visitors to churches and museums. One way to increase their interpretative value would be to display reproductions of the historical drawings alongside the surviving fragments, as has already been done in St John's Church in Cēsis. An even more informative approach would be to combine photographs of the surviving fragments with historical drawings digitally, thereby demonstrating precisely which part of the original monument each fragment represents. Such digital montages would make grave slab fragments more meaningful and engaging both for specialists (including art historians, museum professionals, and heritage conservators) and for the wider public, while at the same time communicating the significance of medieval grave slabs as an important category of cultural heritage. However, this method can be successfully applied only when the

proportions of the historical drawings are sufficiently accurate, which is not always the case.

Although this article has focused primarily on fragmentary monuments, digital methods are equally valuable for the study of complete or largely intact grave slabs. One approach that has not yet been implemented in either Estonia or Latvia would be to provide interpretative panels accompanying reused grave slabs, with the decorative and epigraphic layers from different periods distinguished by colour. Such panels could also include a digital reconstruction of the slab's original appearance. A single example will suffice. Embedded in the floor of the former south aisle of St Nicholas' Church, now the Niguliste Museum, is the grave slab of Hinrick Lussenberch (†1512), alderman of Tallinn's Great Guild. In addition to Lussenberch's merchant's mark or house mark, the central field of the slab contains the shields with house marks of two later owners from different periods, together with a later inventory number.[56] To the untrained eye, these successive historical layers are difficult to distinguish. An interpretative panel – whether in printed or digital form – would therefore greatly facilitate the understanding of the monument. A virtual reconstruction could also illustrate the slab's original appearance and, where known, its original location within the church interior.

The study of heavily worn grave slabs and their inscriptions could also benefit significantly from 3D scanning and photogrammetry, techniques that have not yet been applied systematically to the investigation of medieval grave monuments in Estonia or Latvia. Their limited use is due partly to a lack of institutional interest and partly to financial constraints. At the same time, the potential benefits of such time-consuming and often costly digital methods should always be weighed against the expected results, particularly in cases where a grave slab is so severely worn that successful documentation or interpretation remains uncertain.

Finally, it should be emphasised that, despite the wide range of digital tools now available to twenty-first-century researchers, the direct examination of monuments in situ remains indispensable. The trained eye of an experienced specialist can often identify details, establish connections, and propose interpretations through close examination of the monuments and comparison with other historical sources. In

56 A. Mänd. Hauaplaadid 14. sajandist 1550. aastateni, 393, fig. [3].

this sense, digital methods should be regarded not as a replacement for traditional scholarship but as a powerful complement to it.

Acknowledgements

This research was funded by Estonian Research Council grant PRG1276 Digital Livonia: For a Digitally Enhanced Study of Medieval Livonia (c. 1200–1550), and by the University of Tartu grant PHVAJARENG22 Illusions of Plenty.

DIGITAALSED MEETODID FRAGMENTAARSETE KESKAEGSETE HAUAPLAATIDE UURIMISEL

Anu Mänd

Keskaegsed hauaplaadid on väärtuslikud allikad sotsiaalajaloo, kunstiajaloo, epigraafika ja heraldika uurimisel. Plaatidele raiutud inimfiguurid, religioossed sümbolid, vapid, peremärgid ja tekstid võimaldasid surnut identifitseerida ning jäädvustasid tema sotsiaalse staatuse. Hauaplaatide materjali, kujunduse, raidtehnika ja ruumilise konteksti uurimine võimaldab seetõttu saada teavet nii keskaegse kunsti kui ka mälestamispraktikate, religioossete tõekspidamiste ja kirjakultuuri kohta.

Keskaegsete hauaplaatide uurimist raskendab nende halb säilivusseisund. Jõukamad ja kõrgema staatusega inimesed maeti enamasti kirikutesse, aga põrandas paiknevad hauaplaadid on sajandite jooksul väga kulunud. Eriti halvasti on säilinud uurendatud ja madalreljeefsed kujutised ja pealiskirjad. Lisaks on hauaplaate hilisematel sajanditel korduvalt taaskasutatud ja uute lahkunute tarvis ümber kujundatud. Enamasti on keskaegsetest hauatähistest alles vaid üksikud fragmendid, mis võivad paikneda kiriku põrandas või mõnes sekundaarses asukohas, näiteks müürituna trepiastmeks või aknasilluseks või paigutatuna lävekiviks (ill 1).

Selliste (fragmentaarsete) hauatähiste uurimiseks pakuvad uusi võimalusi digitaalsed meetodid. Artiklis lähtutakse kasutajakesksest vaatenurgast, keskendudes meetoditele, mida autor ise on katsetanud. Põhitähelepanu all on pinnamuutuste pildistamine (*Reflectance Transformation Imaging*, RTI) ning digitaalne pilditöötlus ja fotomontaaž. Juhtumiuuringute kaudu hinnatakse, mil määral aitavad need meetodid

lugeda kulunud raidkirju, identifitseerida hauaplaatide omanikke, ühendada füüsiliselt lahus olevaid fragmente ja rekonstrueerida monumentide kunagist välimust.

Matemaatilisel algoritmil põhinev RTI on osutunud eriti kasulikuks tugevasti kulunud raidkirjade ja kujutiste uurimisel. Meetod põhineb ühe ja sama pinna pildistamisel eri nurkade alt langeva valgusega ning võimaldab tuua nähtavale palja silmaga raskesti märgatavaid detaile. Tallinna Oleviste kiriku läänеportaali kõrval asuva tugevalt kulunud keskaegse hauaplaadi fragmendi (ill 2) puhul võimaldas RTI lugeda gooti minusklites pealiskirja osi, mida varem polnud suudetud dešifreerida. Muu hulgas õnnestus kindlaks teha surnu amet – bürgermeister (*borgemester*). Samas näitas see katse ka meetodi piiratust: pealiskirja kõige tugevamalt kulunud osades ei aidanud RTI teksti loetavust oluliselt parandada. Sarnastel juhtudel võiks tulevikus katsetada 3D-skaneerimist ja fotogramm-meetriat.

Teiseks oluliseks meetodiks on digitaalne fotomontaaž. Varauusajal ja hilisemate kirikuremontide käigus on hauaplaadid sageli mitmeks osaks lõigatud ning fragmendid paigutatud põrandas eri kohtadesse. Neid ei ole enamasti võimalik ega otstarbekas põrandast eemaldada ja füüsiliselt taasühendada. Digitaalne montaaž võimaldab aga eraldi paiknevad fragmendid virtuaalselt kokku panna ning seeläbi rekonstrueerida hauaplaadi algse välimuse. Meetodi kasutamist demonstreeritakse artiklis mitme Eesti hauaplaadi näitel.

Haljala kiriku kooriruumis paiknevad kaks fragmenti kuuluvad 1523. aastal surnud Hinrick Hastferi ja tema naise hauaplaadile. Nende fragmentide omavahelise seose kindlakstegemisel on oluline allikas Søren Abildgaardi 1754. aastal tehtud joonis (ill 3). Fragmentide digitaalne ühendamine võimaldab taastada suure osa hauaplaadi kunagisest välimusest (ill 4), ajaloolise joonise ja säilinud osade võrdlemine aga näitab ühtlasi, et Abildgaardi kujutis ei ole õiges proportsioonis ega detailitäpne. Näiteks on ta tühjaks jätnud plaadi nurkades asuvad sõõrid, kuigi säilinud fragmentidel võib seal näha evangelistide sümboleid.

Tallinna Niguliste kirikus paikneva mündisellide hauaplaadi puhul võimaldas digitaalne fotomontaaž ühendada kaks fragmenti, mille omavaheline seos oli jäänud varasematele uurijatele märkamatuks (ill 5). Veelgi tähelepanuväärsem on Oleviste kiriku bürgermeistri hauaplaadi juhtum. Selle fragmendid paiknevad tänapäeval kahes eri hoones: kaks suuremat tükki Oleviste kiriku eeskojas, üks või kaks fragmenti aga Niguliste muuseumi põrandas. Digitaalne montaaž võimaldab kõik säilinud osad taas

üheks tervikuks ühendada (ill 6). Kuidas ühe hauaplaadi osad kahte kirikusse sattusid, ei ole õnnestunud kindlaks teha.

Eriti kõnekas on Johann Mölleri ja tema abikaasa Anna Treideni hauaplaadi näide. Kolm fragmenti asuvad Pöide kirikus, kaks ülejäänut aga Kuressaare piiskopilinnuses Saaremaa Muuseumi raidkivihoidlas. Raidkirja, dekoori ja kujunduse võrdlus näitab, et kõik viis fragmenti pärinevad samast hauaplaadist. Digitaalselt ühendatuna võimaldavad need taastada suurema osa hauaplaadi algsest välimusest (ill 7) ja lugeda eri aegadel tahutud pealiskirju. Selgub, et Johann Möller suri 1552., mitte 1568. aastal, nagu varasemas kirjanduses ekslikult väidetud. Kuna märkimisväärne osa hauaplaadist on säilinud, oleks võimalik ja soovitav fragmendid tulevikus ka füüsiliselt ühendada, konserveerida ja Pöide kirikus eksponeerida.

Digitaalne fotomontaaž pakub olulisi võimalusi ka juhul, kui hauaplaadist endast on säilinud ainult väikesed fragmendid, kuid selle terviklikku välimust tuntakse ajaloolis(t)e joonis(t)e kaudu. Sellisel juhul saab fragmentide fotod asetada digitaalselt ajaloolisele kujutisele, näidates täpselt, millisest monumendi osast need pärinevad. Meetod toimib siiski üksnes sel juhul, kui ajalooline joonis on proportsioonidelt piisavalt täpne.

Hea näite pakub Tallinna Niguliste kirikust pärinev 15. sajandi kaupmehe Brun Warendorpi hauaplaat, mis purunes 1944. aasta märtsipommitamise ja mõni aasta hiljem toimunud võlvivaringute tagajärjel. Hauaplaadist on säilinud vaid viisteist väikest fragmenti, millest enamikul leidub kõigest paar raidkirja tähte (ill 8). Eugen von Nottbecki 19. sajandi lõpu kirjelduse ja joonise abil õnnestus fragmendid siiski identifitseerida ja osa neist joonisel õigesse kohta paigutada (ill 9). Digitaalne rekonstruktsioon näitab ühtlasi, et Nottbecki joonis on küll proportsioonidelt täpne, kuid eksib mõnes detailis, näiteks nurgakaunistuste kujutamisel. Digitaalne montaaž muudab visuaalselt tagasihoidlikud kivifragmendid vaatajale mõistetavaks, sest näitab nende seost kunagise tervikliku hauamonumendiga.

Sama põhimõtet sai rakendada Eesti Rahva Muuseumi kogus hoitavatele Heidenricus Savijerwe (†1460) hauaplaadi fragmentidele. Samuti on võimalik Johann Christoph Brotze 1771. aasta joonisega ühendada kaks ordumeister Wolter von Plettenbergi (†1535) hauaplaadi fragmenti Cēsise Jaani kirikus (ill 10).

Digitaalsetel meetoditel on seega kaks omavahel seotud funktsiooni. Esiteks aitavad need käibesse tuua uut teaduslikku teavet: muuta kulunud raidkirju loetavamaks, tuvastada eri paikades asuvate

fragmentide omavahelisi seoseid, korrigeerida varasemaid atribueeringuid ja rekonstrueerida hauatähiste algset kujundust. Teiseks on neil oluline roll kultuuripärandi vahendamisel. Üksikud hallid paekivifragmendid ei pruugi muuseumis või kirikus vaatajale palju öelda. Kui need aga paigutada digitaalselt oma algsesse konteksti või ühendada ajaloolise joonisega, muutub arusaadavaks, millise suurema ja sageli rikkalikult kujundatud mälestusmärgi osad need kunagi olid.

Seetõttu võiks muuseumides ja kirikutes säilinud fragmentide eksponeerimisel kasutada senisest enam ajaloolisi jooniseid ja digitaalseid rekonstruktsioone. Samuti oleks võimalik visuaalselt eristada hauaplaatide ajaloolisi kihistusi, näiteks keskajal loodud kujundust ning hilisema taaskasutamise käigus lisatud raidkirju, vappe või peremärke.

Tulevikus võiks Eesti ja Läti keskaegsete hauaplaatide uurimisel senisest süsteemsemalt rakendada ka 3D-skaneerimist ja fotogrammmeetriat. Need meetodid võivad osutuda eriti kasulikuks tugevasti kulunud pindade ja raidkirjade dokumenteerimisel, kuid nende rakendamise ajakulu ja maksumust tuleb alati võrrelda loodetava teadusliku tulemusega. Kõige tugevamalt kahjustatud raidkivide puhul ei pruugi ka tehniliselt keerukad meetodid anda piisavalt uut informatsiooni.

Seega ei asenda digitaalsed meetodid hauaplaatide vahetut uurimist. Kohapealne vaatlus ja uurija kogemusel põhinev võime märgata detaile, võrrelda raidkirju ja kujundust ning luua seoseid eri fragmentide vahel jäävad asendamatuks. RTI, digitaalne pilditöötlus, fotomontaaž, 3D-skaneerimine ja fotogramm-meetria toimivad kõige tulemuslikumalt traditsiooniliste kunstiajalooliste uurimismeetodite täiendusena. Nende suurim väärtus seisneb võimaluses muuta halvasti säilinud või fragmentaarne materiaalne pärand taas loetavaks, analüüsitavaks ja visuaalselt mõistetavaks.

Acta Historica Tallinnensia 32/2 https://doi.org/10.3176/hist.2026.2.05

Imaging Medieval Livonian Seals: Techniques and Technologies in the History of Sigillography

Juhan Kreem

Tallinn City Archives, Tolli 6, Tallinn 10133; Juhan.Kreem@tallinnlv.ee

Abstract. This article reviews different techniques for rendering images of medieval Livonian seals from the eighteenth century to the present. The first Baltic collectors made freehand drawings and took moulds of originals to cast copies of seals. In the 1860s Baron Robert von Toll introduced an engraving machine into Baltic sigillography to transfer seal reliefs to printing plates. Photography and the digital turn have boosted the number of images, and large databases have emerged, but the creation of metadata has not kept pace. While some older imaging technologies were quickly abandoned, others continued to be used alongside newer ones.

Keywords: sigillography, historiography, relief copying, engraving technology, photography

INTRODUCTION

Seals represent not only a fascinating form of medieval miniature art, used to authenticate or close documents, but also the attributes and insignia of authority. Seals also carry meanings and communicate the self-perception and self-representation of their owners. Sealing is a universal cultural practice, and thus sigillography is an important auxiliary science in general. The extensive use of sealing in medieval European legal

© 2026 Author. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International License CC BY 4.0 (http://creativecommons.org/licenses/by/4.0).

contexts makes sigillography especially relevant to medieval studies.[1] In Livonia, from the thirteenth century onwards, sealing was among the standard procedures of clerical and secular overlords and institutions, as well as towns; over the course of the following centuries, numerous nobles and town burghers likewise adopted private seals.[2]

Seal impressions are the result of mechanical reproduction, made by applying a harder matrix to a softer material, mainly wax.[3] Impressions vary in quality, and although the overall number of preserved specimens is immense, they remain very fragile and pose a continuing challenge for conservators. Some seal matrices, especially those of institutions, were in use for a very long time; others, most notably personal seals, were used for a shorter time and were usually destroyed after the death of their owner. In general, matrices rarely survive, and researchers have to use impressions to reconstruct the matrix, the very heart of the *hic et nunc*[4] of sealing.

As with every type of visual source, adequate reproduction of preserved specimens is also crucial in sigillography.[5] Images in publications provide an important tool for historians in various fields. They not only enable the identification and study of seals and sealing practices but also contribute to their preservation. The aim of this article is to give an overview of imaging techniques used in the history of Livonian sigillography. Recent research has pointed out that scientific images participate in discourses of power and help to produce, assert and sustain power relations.[6] As seals themselves are, among other things, attributes of power, the effort put into their study is strongly motivated by the wish to sustain this power, to trace its origins or, in the event of its loss, to recover it.[7] But in discussing technical solutions for imaging, authenticity and precision re-emerge as central endeavours, which are essential to every kind of scientific imaging (drawings, photographs, etc.).

1 A. Stieldorf. Siegelkunde. Hahnsche Buchhandlung, Hannover, 2004, 11.
2 Est- und Livländische Brieflade. Vierter Theil. Siegel und Münzen der weltlichen und geistlichen Gebietiger über Liv-, Est- und Curland bis zum Jahre 1561 nebst Siegeln einheimischer Geschlechter. Aus dem Nachlasse von Baron Robert von Toll hrsg. von J. Sachssendahl. Kluge & Ström, Reval, 1887.
3 For multilingual terminology see: Vocabulaire international de la sigillographie. Conseil International des Archives, comité de sigillographie, Roma, 1990.
4 W. Benjamin. The Work of Art in the Age of Mechanical Reproduction. Transl. by J. A.Underwood. Penguin, London, 2008.
5 T. Diederich. Sphragistik. – Historische Hilfswissenschaften. Stand und Perspektiven der Forschung. Hrsg. von T. Diederich, J. Oepen. Böhlau, Köln, 2005, 35–59, here 55–58.
6 The Technical Image. A History of Styles in Scientific Imagery. Ed. by H. Bredekamp, V. Dünkel, B. Schneider. University of Chicago Press, Chicago, 2015; Kunst või teadus = Art or science. Ed. by L. Kaljundi, K. Polli, K. Tiideberg. Tartu Ülikooli Kirjastus, Tartu, 2022.
7 E. Hagström Molin. Colonisation, Decline and Loss: Making Livonian Time in Nineteenth-Century Russia. – Time & Society 2026, 35, 2, 236–255.

Rather than constituting a linear progression towards ever greater fidelity, the history of seal imaging reveals a recurrent tension between accuracy, legibility, reproducibility and the preservation of the original object.

FIRST SOLUTION: FREEHAND DRAWING

A basic way to represent a seal's appearance is through verbal description, but a drawing is in many ways better and was used very early on.[8] The quality of drawings could of course vary from crude sketches to impeccable facsimiles. Richard von der Hardt (d. 1714), a German-born book auctioneer and bibliographer active in Stockholm and Uppsala, is known to have made several pictures of Livonian coins and seals, some of which were published by Johann Michael Heineccius (d. 1722) in his ground-breaking scholarly treatment of sigillography in 1709.[9] When Johann Gottfried Arndt (1713–1767) was compiling a set of seal descriptions in Livonia for the appendix to the second part of his Chronicle, he could rely on these rather clumsy images.[10]

In the late eighteenth century, freehand drawing was the main method used by the great recorders of Baltic heritage, such as Joachim Christoph Brotze (1742–1823) and Eduard Philipp Körber (1770–1850). Both scholars collected drawings and watercolours of antiquities and exchanged images, including drawings of seals.[11] The results are a series of manuscript volumes preserved, in Brotze's case, in the University of Latvia Library in Riga and, in Körber's case, in the Estonian Literary Museum in Tartu. Brotze's seal images are above all to be found in his *Sylloge Diplomatum Livoniam illustrantium*,[12] and Körber's in the second volume of his *Vaterländische Merkwürdigkeiten* (see Fig. 1).[13]

8 For example O. D. Harris. Fragments of the Past: The Early Antiquarian Perception and Study of Seals in England. – Companion to Seals in the Middle Ages. Ed. by L. J. Whatley. Brill, Leiden, Boston, 2019, 129–154, here 148.

9 J. M. Heineccius. Syntagma historicum de veteribus Germanorum aliarumque nationum sigillis. Sumtibus Nicolai Foersteri, Frankfurt, Leipzig, 1709, Tab. XI 6–10, Tab. XIII 9, Tab. XV 13.

10 J. G. Arndt. Der Livländischen Chronik anderer Theil von Liefland unter seinen Herren Meistern welche die alte Geschichte des Ordens und der benachbarten Völker erleutert. Johann Justinus Gebauer, Halle, 1753, 303, Tabelle II. Von den Sigillen der Ordens- und anderer geistlichen Personen nebst den ältern und neuern Sigillen und Wappen der Städte.

11 A. Taimiņa. Johana Kristofa Broces (Brotze, 1742–1823) dzīvesgājums un viņa rokrakstu kolekcija LU Akadēmiskajā bibliotēkā, Diss. Riga, 2013, 115–127, 251–256.

12 Latvijas Universitātes Akadēmiskā bibliotēka (LUAB), Johana Kristofa Broces (Brotze) rokrakstu kolekcija, inv. Nr. R4979, R4980: Brotze, Johann Christoph, Sylloge diplomatum Livoniam illustrantium. Tomus primus, Tomus secundus. https://dspace.lu.lv/handle/7/61386 (accessed 28/08/2026).

13 Eesti Kultuurilooline Arhiiv (EKLA), f 192, m 57:1(M.B): Eduard Philipp Körber.

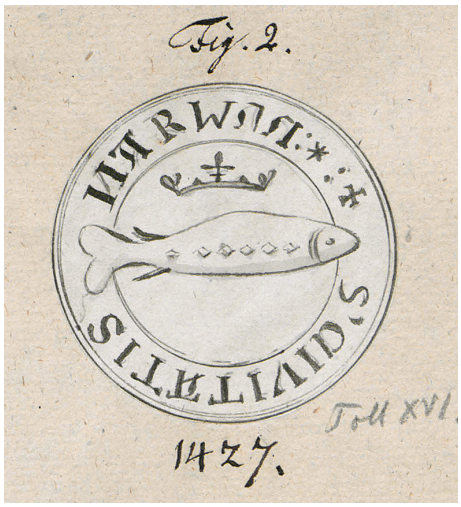


Figure 1. Narva seal by Eduard Philipp Körber. Eesti Kultuurilooline Arhiiv (Tartu), ÕES, MB 57, fol. 276



Figure 2. Narva seal by Johann Naha. – A. Süvalep. Narva ajalugu. Narva Ajaloo Selts, Narva, 1936, p. 268

These collections, although they remained in manuscript, became the basis for a major publication on Baltic seals, discussed below.[14]

Freehand drawing depends on the attentive eye and skill of the illustrator but offers no guarantee of exact resemblance. For coins, Brotze suggested a method for achieving greater precision, which was also used by Körber: wet paper had to be pressed over the coin and rubbed with an ivory stick, and the imprint then drawn over and painted.[15] Not all coins could be copied in this way, especially very thin bracteates. Applying this method to seals was out of the question because of the softness of the wax impressions.

Freehand drawing remained competitive with other methods of imaging throughout the nineteenth and twentieth centuries. For example, for the history of Narva published in 1936, seal images were drawn directly from originals in the Tallinn City Archives, with the scale also indicated, and printed as woodcuts by the Estonian artist Johann Naha (1902–1982) (Fig. 2).[16] Freehand drawings of seals continue to appear in publications in the twenty-first century.[17]

Vaterländische Merkwürdigkeiten II. Beyträge zur Kenntnisz der Liefländischen Diplomatie mit 260 Original Siegel Abbildungen… https://kivike.kirmus.ee/index.php?pid=EKLA-11278-43212-77047&oid=0&t=1&module=400&op=3#EKLA-11278-43282-07977 (accessed 28/08/2026).

14 J. Sachssendahl. Vorwort. – Est- und Livländische Brieflade. Vierter Theil, p. IX–XVII, here p. XII–XIII.

15 V. Dāboliņš. New footprints of Eduard Philipp Körber's (1770–1850) numismatic research. – From Hoard to Archive. Numismatic Discoveries from the Baltic Rim and Beyond. Ed. by E. Russow, V. Dāboliņš, V. Lang, Tartu UP, Tartu, 2023, 17–36, here p. 23. On Körber see: I. Leimus. Zu den Anfängen der Numismatik in Estland: Eduard Philipp Körber (1770–1850). – Geisteswissenschaften und Publizistik im Baltikum des 19. und frühen 20. Jahrhunderts. Hrsg. von N. Angermann, W. Lenz, K. Maier. LIT, Berlin, 2011, 43–63.

16 A. Süvalep. Narva ajalugu. Narva Ajaloo Selts, Narva, 1936, 35, 38, 268, 269, 327.

17 Eesti ajalugu II. Eesti keskaeg. Koost. ja toim. A. Selart. Tartu Ülikool, Tartu, 1012, 30, 55; Tallinna ajalugu I. 1561. aastani. Koost. ja peatoim. T. Kala, toim. T. Tamla. Tallinna Linnaarhiiv, Tallinn, 2019, 79–80.

MECHANICAL COPIES

Different materials can be used to mechanically copy seals, but in principle the technique remains the same: one has to take a mould from the original to cast the image in positive.[18] Brotze had already collected copies of seals in sealing wax made in this way,[19] but the greatest contribution in this field was made by Robert von Toll (1802–1876), a former soldier who in 1851 dedicated himself to the study of history and the collection of medieval documents and seals. An important milestone in his career was a journey to Stockholm in 1860 with Professor Carl Schirren (1826–1910), where they searched for the remains of the archives of the Teutonic Order.[20] While Carl Schirren worked on the first register of Livonian sources in Sweden,[21] Toll was busy making moulds of seals in gutta-percha, a high-quality latex sourced from the Malaysian tree *Palaquium gutta*. This material is rigid, naturally biologically inert, resilient, electrically non-conductive and thermoplastic, and was used extensively instead of rubber until synthetic materials became available.

The result of Toll's work was a collection known as the Kukruse (Ger. Kuckers) collection, after the manor in Estonia where it was originally located. In 1910 the Toll family deposited it in the Archives of the Estonian Noble Corporation in Tallinn.[22] When the Corporation was dissolved in 1920, the Kukruse collection was relocated by its owner to the Library of the Estonian Literary Society (*Estländische Literärische Gesellschaft*). Only in 1935 did the Republic of Estonia purchase it for 5,000 crowns, making it part of the collection of the State Central Archives in Tartu.[23] Toll's seal collection was inventoried in the archives

18 The practice was widespread and has not been entirely abandoned even today. See e.g. M. Libert. The seal casts collection and the digitisation of the sigillographic collections of the national archives of Belgium. – Vestnik of Saint Petersburg University. History, 2017, 62: 2, 255–266, here 256–258 https://doi.org/10.21638/11701/spbu02.2017.203 (accessed 09/09/2026).

19 LUAB, Inv. nr. R4104, Brotze, J. Chr. Siegel in Siegellack, adliche.

20 J. Götz. Das Archiv des livländischen Deutschordenszweiges. Eine archivgeschichtliche Untersuchung. – Die Kirche im mittelalterlichen Livland. Hrsg. von R. Biskup, J. Götz, A. Radzimiński. UMK, Toruń, 2019, 9–77, here 10–11.

21 C. Schirren. Verzeichniss livländischer Geschichts-Quellen in schwedischen Archiven und Bibliotheken. Dorpat, 1861.

22 P. Osten-Sacken. Bericht über das Estländische Ritterschaftsarchiv. – Jahrbuch für Genealogie, Heraldik und Sphragistik 1909 und 1910, 311; P. Osten-Sacken. Bericht über das Estländische Ritterschaftsarchiv fürs das Triennium 1911–1913. – Jahrbuch für Genealogie, Heraldik und Sphragistik 1914, 39; See also K. Neitmann. Ständisches Archivwesen: Aufstieg und Fall der deutsch-baltischen Ritterschaftsarchive (ca. 1880–1920). – Estonian Archives in the European Context. Papers of the Conference in the Tallinn City Archives on September 15–16, 2005. Ed. by L. Kõiv and P. Pillak. Estonian Society of Archivists, Tallinn, 2012, 15–73.

23 O. Liiv. The Central State Archives 1 IV. 1932–1. IV 1937. – Riigi keskarhiiv 1932–1937 ja arhiivinduslikke eriküsimusi = The Central State Archives 1932–1937 & Special Archival Questions. Eesti Riigi Keskarhiiv, Tartu 1937, 43–54. Estonian National Archives (RA),



Figure 3. Seal of Christoph I, King of Denmark, sealing wax copy, Eugen von Nottbeck collection. RA, EAA.2071.1.10

after World War II in three phases. The first inventory was made in 1971 and contains 599 items from the special cabinet commissioned by Toll for storing his seal copies.[24] Two other inventories, compiled in 2003 (478 items) and 2006 (462 items), cover objects that Toll himself had already separated from the main collection.[25]

Apart from the Kukruse collection, the Estonian National Archives holds the collection of seals and seal copies assembled by the lawyer and historian Eugen von Nottbeck (1842–1900).[26] Nottbeck criticised the use of gutta-percha and gypsum in making moulds because of the potential damage to the seal. He used a mixture of bread and talc to make the mould, which he then used to press a new image into sealing wax (Fig. 3).[27] Nottbeck's widow donated the collection to the Archives of the Estonian Noble Corporation,[28] from where it passed directly into the custody of the Estonian state and is now held in the National Archives. Finally, there is a seal collection created in the National Archives in 1980 by bringing together seals and matrices of various provenances, including

EAA.661.1.217, fol. 148 (contract); EAA.R-271.1.1287.13 (photo of the collection).

24 The cabinet is currently in the custody of the Estonian Government Office.

25 RA, EAA.2069.3, 4, 5.

26 RA, EAA.2071.

27 E. von Nottbeck. Siegel aus dem Revaler Rathsarchiv nebst Sammlung von Wappen der Revaler Rathsfamilien. H. G. Rahtgens, Lübeck, 1880, 7.

28 I. Jürjo. Die Reformen im Estländischen Ritterschaftsarchiv am Vorabend des Ersten Weltkrieges. – Estonian Archives in the European Context, 74–111, here 96.

those from the collection of the Estonian Noble Corporation.[29] All these collections in the National Archives of Estonia, which contain matrices as well as seal copies, comprise 7,658 units. They have been photographed and have been digitally available online since 2012.[30]

WAGNER'S ENGRAVING MACHINE

In publishing images of seals, the main problem was transferring three-dimensional images to two-dimensional paper. A solution was actively sought in the 1830s.[31] In 1830 the French industrial designer Achille Collas (1795–1859) invented a mechanical engraving machine, a pantograph, which enabled printing plates to be engraved with mathematical exactness, and he presented his prints at the Paris Salon in 1833.[32] This so-called numismatic engraving was first used to reproduce coins,[33] but the new method was quickly introduced for the reproduction of seals as well.[34] Although the method has since been widely associated with the name of Collas, other inventors were also working in the field, for example Christian Gobrecht (1785–1844) and Joseph Saxton (1790–1873) in America and John Bate in England (who called his machine an anaglyptograph).[35]

In the German-speaking world, a *Relief-maschine* for a similar purpose was designed and described in a publication in 1836 by Karl Karmarsch (1803–1879), director of the Polytechnic School in Hanover.[36] In Berlin, however, the academic artist Friedrich Gottlieb Wagner Jr.

29 RA, EAA.2077.

30 https://www.ra.ee/pitserid/ (accessed 28/08/2026). Unfortunately, only about one quarter of the images are described in the database (1,852 units). The remaining images have only brief references from the inventory.

31 T. Fawcett. Plane Surfaces and Solid Bodies: Reproducing Three-Dimensional Art in the Nineteenth Century. – Art History through the Camera's Lens. Ed. by H. E. Roberts. Gordon and Breach, Amsterdam, 1995, 59–85.

32 See also: V. Nolte. Memorial of Facts Connected with the History of Medallic Engraving and the Process of M. Collas. – Museum of Foreign Literature, Science, and Art, 1838, 32, 1, 113–119.

33 Trésor de numismatique et de glyptique, ou Recueil général de médailles, monnaies, pierres gravées, bas-reliefs tant anciens que modernes... gravés par les procédés de M. Achille Collas. Paris, 1834–1850.

34 The Great Seals of England from the Time of Edward the Confessor to the Reign of His Majesty William the Fourth: Illustrated by Thirty-eight Plates, Containing Exact Fac-similes of the Seals in their Present Condition, Engraved by the Process of Achilles Collas; with Historical and Descriptive Notices. Henry Hering, London, 1837.

35 Handbuch der Schriftgiesserei und der verwandten Nebenzweige. Hrsg. von A. Henze. Voigt, Weimar, 1844, 276–278; Handbuch der Litographie. Nach dem Gegenwartigen Stande der Technik. Hrsg. von G. Fritz. W.Knapp, Halle, 1901, 282–286.

36 K. Karmarsch. Beschreibung einer Relief-maschine zur getreuen bildlichen Darstellung von Münzen, Medaillen und anderen Reliefs, auf ganz mechanischem Wege. Helwingsche Hof- Buchhandlung, Hannover, 1836.

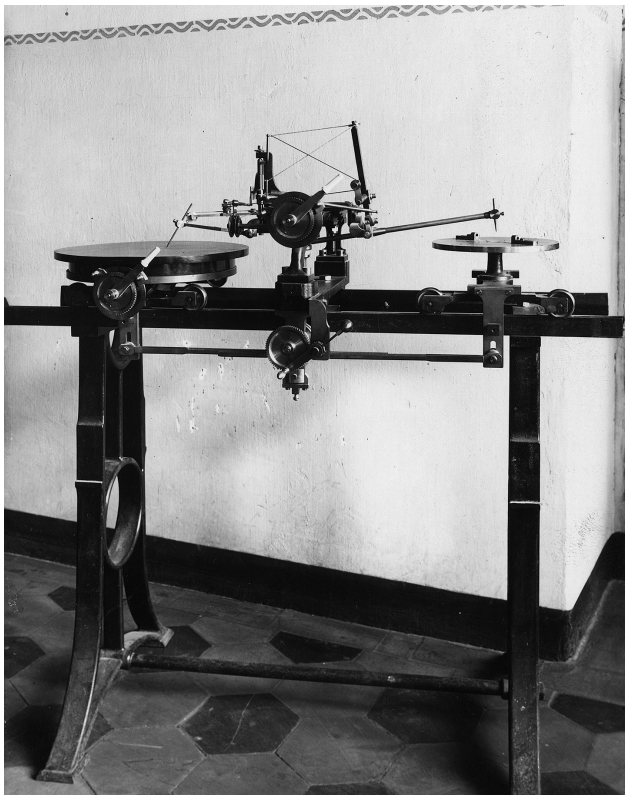


Figure 4. Wagner's engraving machine in Estonian State Central Archives in the 1930s. RA, EAA.R-271.1.1287.30

(1803–1875) developed his own version (Fig. 4).[37] Very little is known of Wagner's life. He seems to have made his first adaptations of Collas's method in 1834.[38] In 1836 he published his first successful prints[39], for which he was awarded a gold medal by the crown prince of Prussia.[40] The next year he was granted the title of *Hofmechanicus*.[41] A catalogue of the Berlin industrial exhibition of 1844 claims that he had already been developing the machine for 15 years and had by then exported 150 machines, large and small, to St Petersburg, Moscow, Kyiv, Warsaw, Prague, Stuttgart, Switzerland and Sweden; the prices of Wagner's machines depended on their size and ranged from 50 to 500 *Reichsthaler*, with the 150-*Thaler* model being the most sought after.[42] Wagner's prints were also praised by the Polytechnic Society in Berlin.[43] In 1860 the Swedish Royal Archivist Emil Hildebrand (1806–1884)[44] introduced Robert von Toll to Wagner's

37 Neues allgemeines Künstler-Lexicon. Bearb. v. G.K.Nagler. Bd 21. E.A. Fleischmann, München, 1851, 54.

38 A. Raczynski. Geschichte der neueren Deutschen Kunst. Aus dem Französischen übersetzt von F. H. von der Hagen. Bd III. Berlin, 1841, 418.

39 Abbildungen geschnittener Steine und Medaillen, ausgeführt durch F. G. Wagner d. jüngern. Steger, Berlin 1836.

40 Morgenblatt für gebildete Stände / Kunst-Blatt, Nr. 52, 30. Juni 1836, 220 (news from 25 May).

41 Allgemeine Preußische Staats-Zeitung, Nr. 96, 7. April 1837, 385: *Se. Majestät der König haben dem akademischen Künstler Friedrich Gottlieb Wagner jun. das Prädikat eines Hof-Mechanikus und Optikus beizulegen geehrt.*

42 Amtlicher Bericht über die allgemeine Deutsche Gewerbe-Ausstellung zu Berlin im Jahre 1844, zweiten Theils erste Abtheilung. Carl Reimarus, Berlin, 1846, 531–532.
A commemorative medal from this exhibition is preserved in the Kukruse collection: EAA.2069.3.598.

43 Bericht über die Verhältnisse und die Wirksamkeit der Polytechnischen Gesellschaft in Berlin, zweites Heft. (vom März 1841 bis Februar 1843). Moeser und Kühn, Berlin, 1846, 69.

44 On him see: E. Norberg. Emil Hildebrand. Langenskiöld, Stockholm, 2018.

engraving machine, and Toll subsequently travelled to Berlin to acquire one.[45]

The problem with this kind of image transfer was that the relatively soft wax seals were not suitable for the pointer of the engraving machine. Thus, a copy in a harder material was needed first. For this purpose, Toll prepared galvanoplastic (electrotype) positive images from his gutta-percha moulds (Fig. 5). In very many cases, both versions are still preserved in the Kukruse collection and can be matched through close study.[46] An example of a copper printing plate[47] and a test print have also been preserved.[48] Furthermore, evidence of experimentation with different materials, as well as failed attempts, can be found in the Kukruse collection.[49] Although the electrotype copies are hard, their use in the engraving machine has nevertheless left thin lines on their surfaces, which can be seen under magnification.[50]

From 1866 Toll suffered from deteriorating eyesight, and in 1876 he handed his work over to Johannes Sachssendahl (1851–1903),



Figure 5. Seal of the Master of the Teutonic Order, gutta-percha mould and galvanoplastic cast, Kukruse collection. RA, EAA.2069.3.13, 12



Figure 6. Seal of the Master of the Teutonic Order, printing plate engraved with Wagner's machine. Brieflade, Tafel F (fragment)

45 Sachssendahl, Vorwort, X.
46 For example RA, EAA.2069.3.13 gutta-percha, 12 galvanoplastic, 11 galvanoplastic on a metal base (Master of the Teutonic Order in Livonia); EAA.2069.3.116, 115 (Vogt of Karkus); EAA.2069.3.124, 123 (Vogt of Oberpalen); EAA.2069.3.126, 125 (Land Marschal).
47 RA, EAA.2069.3.34.
48 RA, EAA.2069.3.4.
49 RA, EAA.2069.3.4; EAA.2069.3.5, EAA.2069.3.32.
50 For example RA, EAA.2069.3.30.

who published the seals together with extensive descriptions in 1887 as the fourth volume of *Est- und Livländische Brieflade*, which remains the central publication on medieval Livonian seals to date. In the publication, however, only plates D, E and F, with 22 seal images in total, were produced using Wagner's machine (Fig. 6). The overall number of electrotype copies in the Kukruse collection is much larger; they simply have not been used in the engraving machine. The reason for abandoning this method was probably that the results remained less than satisfactory. Although the printed image was mathematically exact, it was difficult to read, and not all the details of such small objects were rendered visible. Wagner's machine seems to have seen little use in sigillography; nevertheless, it was widely employed, for example, in engraving banknote printing plates.[51]

The overwhelming majority of the plates in *Brieflade* once again rely on freehand drawing. Toll had started ordering lithographic images from the Tallinn lithographer F. W. Macdonald as early as 1857,[52] using both original seals and freehand drawings by Brotze, Körber and others.[53] On his journey to Berlin, Toll commissioned the seal carver Carl Julius Milde (1803–1875) in Lübeck to make him three plates (A, B, C), which were printed in Lübeck by H. G. Rahtgens.[54] In the course of editing the whole book, Sachssendahl based his descriptions of the seals, wherever possible, on the originals in various Tallinn archives and copies in the Kukruse collection. He was fully aware of the limitations of freehand drawings and regarded them as only general sketches; he appreciated the results of Wagner's machine but above all stressed the quality of Milde's work.[55]

Wagner's machine was still in the House of the Knighthood when the building was inventoried during the dissolution of the Noble Corporation in 1920.[56] When the collection was purchased for the State

51 A. Heinonen. Vom Zettel zum Geldschein: Deutsche Einflüsse auf das Design und den Druck von Banknoten im Großfürstentum Finnland. Suomen Pankki, Helsinki, 2024, 65.

52 He had his workshop in Tallinn between 1838 and 1861, see: H. Treumann. Kivitrükikodade ajaloost Eestis 1818–1958. – H. Treumann. Vanemast raamatukultuuriloost. Eesti Raamat, Tallinn, 1977, 92–102, here 94.

53 Sachssendahl, Vorwort, XII–XIII.

54 Sachssendahl, Vorwort, X.

55 Sachssendahl, Vorwort, XVI: *Die Siegelabbildungen müssen als gute Skizzen angesehen werden, da sie aus technischen Schwierigkeiten nicht mit der Treue angefertigt werden konnten, wie es an sich wünschenswerth gewesen wäre. Es ist daher in denselben eine genaue Wiedergabe der kleinen Trennungszeichen, Buchstabenformen, Perlenkreise und dergleichen mehr nicht zu erwarten*.

56 RA, EAA.854.1.475, on the ground floor in the dark room between the Archives and Library: *graweerimise masin, Hermann Tolli omandus* (Engraving machine, property of Hermann Toll). The inventory is not dated, but as it is in Estonian, it must have been compiled by the authorities of the newly established Republic of Estonia.

Archives in 1935, it also included the engraving machine.[57] Although hopes were expressed that the machine could be put into use,[58] these seem not to have been realised.[59] At some point during or shortly after World War II, the machine disappeared from the archives, most probably having been scrapped. In contrast, photographs of the galvanoplastic copies made by Toll remain popular as book illustrations. For example, in a recent treatment of medieval Livonian seals, five of the 16 illustrations depict galvanoplastic copies rather than originals.[60]

PHOTOGRAPHY AND BEYOND

At the time when Toll was putting much effort into making galvanoplastic copies and using them for engraving, another invention with a much brighter future began to be implemented: photography.[61] It took some time for photography, and especially the printing of photographs, to become sufficiently developed for use with small objects.[62] In Estonia, there is evidence from 1875 of the first photographs of archaeological objects in the Learned Estonian Society (*Gelehrte Estnische Gesellschaft*).[63] Although the collection of Eugen von Nottbeck in the National Archives contains early photographs of seals,[64] he does not appear to have used them in his publication on Tallinn seals in 1880. The images in his publication were drawn from the sealing-wax copies by the lithographer Hermann Viegelmann in Lübeck and printed by H. G. Rahtgens, just like some of Toll's plates.[65] Nevertheless, while textual descriptions remained paramount, by the end of the nineteenth century the use of photographs to present seals in catalogues had become established.[66]

57 It is not mentioned in the contract, but in the receipt (*akt*) RA, EAA.661.1.217, p. 149. A photograph was also taken at that time: EAA.R-271.1.1287.30. See also: Haruldasim Eesti eraarhiiv riigi keskarhiiwile. – Päevaleht. 30. Oct. 1935, 4.

58 RA, EAA.661.1.217, 154.

59 J. Lehtaru. Arhivaalide säilitamine Rahvusarhiivis läbi sajandi I. – Tuna 2020, 4, 60–81, here p. 72–73.

60 I. Leimus. Pitserid. – Eesti kunsti ajalugu 1 (1100–1520). Koost. toim. K. Markus. EKA, Tallinn, 2023, 447–455.

61 G. Batchen. Desiring Production. – Each Wild Idea: Writing, Photography, History. MIT Press, Cambridge, Mass., London, 2001, 3–24.

62 Fawcett, Plane Surfaces and Solid Bodies, 67–71.

63 T. Liibek. Fotograafiakultuur Eestis 1839–1895. Diss. Tallinna Ülikool, Tallinn, 2010, 231.

64 RA, EAA.2071.1.1889.

65 Nottbeck, Siegel aus dem Revaler Ratsarchiv, 7.

66 Katalog der Ausstellung zum X. Archäologischen Kongress in Riga 1896. W. F. Häcker, Riga, 1896, Tafel 32; Geschichte des Siegels der Stadt Dorpat. Aus dem Nachlass des weil. Stadtarchivar H. Lichtenstein; bearb. und hrsg. von A. Feuereisen. C. Mattiesen, Dorpat, 1907, Tafel 1–5.

Although much has happened since the invention of photography, most notably the shift from chemical to digital recording, the basic elements of the imaging chain (the subject, capturing device and display) have remained the same. Lighting conditions are crucial when capturing small three-dimensional seals made of wax whose colour is indistinct. It is the shadows that make the relief of the seal visible. As examples of high-quality images, the photographs of seals on parchment charters from the Tallinn City Archives taken in the early 1980s at the *Geheimes Staatsarchiv* in Berlin must be singled out (Fig. 7).[67] These charters were lent to Berlin by the Federal Archives of Germany (Koblenz), where part of the City Archives was deposited at that time.[68] A number of high-quality reproductions of the Teutonic Order's seals from this collection were also published in the exhibition catalogue commemorating the 800th anniversary of the Order in 1990.[69] The negatives made in Berlin and the catalogue compiled during this project by Ursula Benninghoven were given to the Tallinn City Archives in 2021. Another project, carried out by the Estonian National Archives and Tallinn City Archives at the Kanut conservation and digitisation centre between 2010 and 2012, aimed to digitise parchment charters.[70] As the focus was on the text of the charters and the light was directed more or less straight at the parchment, the images of seals came out not as clear as in the older Berlin photographs.[71]



Figure 7. Seal of the Master of the Teutonic Order, photograph from the 1980s. TLA.230.1-I.178

67 Available on the platform Digital Livonia. https://dl.tlu.ee/en/objects/Pitserid (accessed 28/08/2026).

68 L. Biewer. Pergamenturkunden des Stadtarchivs von Reval/Tallinn. Erinnerungen an eine unvollendete archivarische Arbeit und einige biographische Bemerkungen. – Archivalische Zeitschrift, 2022, 98, 331–351, here p. 346–347.

69 800 Jahre Deutscher Orden. Ausstellung des Germanischen Nationalmuseums Nürnberg in Zusammenarbeit mit der Internationalen Historischen Kommission zur Erforschung des Deutschen Ordens. Hrsg. von G. Bott, U. Arnold. Bertelsmann, Gütersloh, 1990, 380–383; 800 Jahre Deutscher Orden. Ausstellung des Germanischen Nationalmuseums 1990. Ergänzungen und Korrekturen. Zusammengestellt von I. von Andrian-Werbung. – Anzeiger des Germanischen Nationalmuseums 1992, 7–50, here p. 43.

70 https://www.ra.ee/apps/pargamendid (accessed 28/08/2026).

71 See also: Pitsereid Tallinna Linnaarhiivist = Siegel aus dem Tallinner Stadtarchiv = Seals from Tallinn City Archives. Koost. M. Habicht, J. Kreem, S. Feldberg. Tallinna Linnaarhiiv, Tallinn 2013.

One of the limitations of photographic images is that light can come from only one direction at a time to create shadows in an image. RTI (Reflectance Transformation Imaging) allows the user to change the apparent light direction as needed.[72] In preparation for the Digital Livonia project, RTI photography was undertaken in the Tallinn City Archives in 2020.[73] The effort put into this method of imaging pays off especially in the case of seals pressed through paper, which have traditionally been difficult to read or photograph. The quality of the original nevertheless remains crucial. In the case of clearly readable originals, RTI does not add much extra value; in the case of poor impressions, RTI may reveal details that the naked eye cannot easily register. There remain, however, cases where even RTI cannot reveal much new information. Photogrammetry and other forms of three-dimensional surface recording offer a complementary strategy by capturing the geometry of the object itself and allowing the relief to be examined independently of a single photographic viewpoint or lighting condition. Their implementation in Estonia lies in the future.

Internationally the digital turn in sigillography extends beyond the production of better images.[74] Recent projects combine computational imaging with structured data, standardised vocabularies and searchable digital corpora, enabling researchers not only to inspect individual seals but also to compare impressions across collections and repositories. The emerging field of digital sigillography thus shifts attention from the reproduction of single objects towards the integration of images, descriptions and contextual information.[75] This development also makes the quality of metadata increasingly important: large collections of digital images remain only partly useful for research if the objects cannot be reliably identified, described and linked to their documentary contexts.

72 For an overview of the RTI workflow, see: M. Schaeben, M. T. Catalano. Byzantine Sigillography and RTI: Insights from the DigiByzSeal Project in Cologne. – Digital Medievalist, 2024, 17. http://journals.openedition.org/digitalmedievalist/418 (accessed 28/08/2026).

73 RTI imaging was carried out by Hembo Pagi (Archeovision). https://archaeovision.eu/en (accessed 28/08/2026).

74 See e.g. M. Filosa, C. Neuefeind, C. Sode. Digital Approaches to Medieval Sigillography: An Introduction. – Digital Medievalist, 2025, 18.1, 1–16. https://doi.org/10.16995/dm.18559.

75 See e.g. J. McEwan. New Approaches to Old Questions: Digital Technology, Sigillography, and digisig. – Digital Medieval Studies: Practice and Preservation. Ed. by L. K. Morreale. Arc Humanities Press, Amsterdam, 2022, 33–48. For the DIGISIG project, see: https://www.digisig.org/ (accessed 14/09/2026).

CONCLUSION

The imaging of seals has been a continuous quest for authenticity and precision of representation. Throughout history, different technologies have been applied, but while some have been quickly abandoned, others have remained in use alongside newer ones. Dissatisfied with the arbitrariness of freehand drawing, Robert von Toll introduced Wagner's engraving machine into Livonian sigillography in the 1860s. The results were not entirely satisfactory, and freehand drawing retained its position. Photography has enabled better documentation, but freehand drawing has continued to play a role in seal depictions because the material of seals is difficult to capture and reproduce in a clearly readable image. Digital imaging techniques such as RTI have increased the possibilities for examining and representing poorly preserved seal impressions, while searchable databases and structured metadata have begun to transform the ways in which larger bodies of sigillographic evidence can be compared and analysed.

Sigillography has not been a very active research field within the study of medieval Livonia. The Toll–Sachssendahl *Brieflade* volume remains the central reference work, despite the fact that most of its plates feature freehand drawings. Although the number and quality of available images have increased tremendously in the meantime, the development of metadata for these images has lagged behind. Catching up with metadata would be a major challenge for the future.

Acknowledgements

This research was funded by Estonian Research Council grant PRG1276: Digital Livonia: For a Digitally Enhanced Study of Medieval Livonia (c. 1200–1550). For help and good advice, I am indebted to Tõnis Türna (Estonian National Archives), Ingrid Sahk (University of Tartu Museum) and Marek Tamm (Tallinn University).

KESKAEGSE LIIVIMAA PITSERITE KUJUTAMINE: TEHNIKA JA TEHNOLOOGIA SIGILLOGRAAFIA AJALOOS

Juhan Kreem

Pitserite sõnalise kirjeldamise kõrval on sigillograafias alati olnud tähtis ka pitserite pildiline edasiandmine. Vaba käega tehtud pitserjoonistusi leidub hulgaliselt juba esimeste suurte Balti antikvaaride, Joachim Christoph Brotze ja Eduard Philipp Körberi kogudes. Ülesjoonistamise kõrval valmistati pitseritest koopiaid ka originaalist vormi võtmise teel. Suurimad teened selles on parun Robert von Tollil, Kukruse isandal, kelle *gutta-percha* vormide kogu on tänapäeval hoiul Rahvusarhiivis. Teine suur pitserite koguja, tallinlane Eugen von Nottbeck, kasutas vormi võtmiseks peenleiva ja talgi segu ning ei soovitanud *gutta-percha*'t ega kipsi, sest need kahjustavad liialt originaale. Võetud vormi järgi sai pitserist teha kas kipsivalandi või pressida kujutis pitserilaki sisse.

Et reljeefsete pitserite kujutisi tasasele paberile trükkida, võttis Robert von Toll 1860. aastatel kasutusele Wagneri graveerimismasina. Kuna graveerimismasin nõudis tugevast materjalist originaali, valmistas Toll *gutta-percha* negatiivis vormidest galvaanoplastilised positiivid. Rahvusarhiivi pitserite lähemal vaatlemisel on võimalik ka tänapäeval kokku viia samale pitserile tagasiviidav vorm ja galvanoplastika. Kukruse kogus on säilinud ka mõned prooviplaadid ja mõned praakeksemplarid galvanoplastikast. „Est- und Livländische Brieflade" neljandas köites ilmunud tahvlitest on siiski vaid kolm valmistatud Wagneri koopiamasina abil. Tulemus annab pitseri suuruse ja kujutise edasi küll täpselt, kuid ebapiisava detailsusega, mistõttu on arusaadav, et masina kasutamisest edaspidi loobuti.

Enamik Tolli avaldatud pitseritest on vabakäejoonistused, mida ei tule võtta enama kui skemaatilise kujutisena. Vaid Lübeckis Carl Julius Milde poolt tehtud kujutised kolmel pilditahvlil vastavad originaalile paremini. Ümberjoonistamine jäi uute tehnoloogiate kõrval konkurentsivõimeliseks pitseri kujutamise viisiks ka 20. sajandil, mõnikord kohtab selliseid kujutisi ka päris viimase aja publikatsioonides. Joonistuste põhiprobleemiks jääb, et miski ei garanteeri nende sarnasust originaaliga.

19. sajandi lõpul võimaldas fotograafia areng teha pitseritest juba päris kvaliteetseid ülesvõtteid ja pitserite fotod hakkasid ilmuma

kataloogides, kuigi pitseri sõnaline kirjeldus püsis endiselt kesksena. Kvaliteetse pitseriülesvõtte saamisel on otsustav valgus, mille oskuslikust kasutamisest on hea näide Berliinist, kus 1980. aastate alguses pildistati üles Tallinna Linnaarhiivi pitserid. Tulemus on kohati kvaliteetsem kui 2010–2012 pärgamentürikute digiteerimise käigus Tallinnas tehtud ülesvõtted. Digiteerimise tulemusel on sündinud ulatuslikud pildipangad, milles metaandmed on jäänud kahjuks vaeslapse ossa. Liivimaa sigillograafia keskseks käsiraamatuks jääb ka täna Robert von Tolli ja Johannes Sachssendahli publikatsioon 1887. aastast.

Acta Historica Tallinnensia 32/2 https://doi.org/10.3176/hist.2026.2.06

The Tallinn *Ex quo* Vocabulary: Textual Tradition and Historical Context

Tiina Kala

Tallinn City Archives, Tolli 6, 10133 Tallinn, Estonia;
tiina.kala@tallinnlv.ee

Kaspar Kolk

Tallinn University, School of Humanities, Narva mnt 25, 10120 Tallinn, Estonia; kaspar.kolk@tlu.ee; Tallinn University Academic Library, Rävala pst 10, 15042 Tallinn, Estonia; University of Tartu Library, W. Struve 1, 50091 Tartu, Estonia; kaspar.kolk@ut.ee

Abstract. The article examines the textual contents, composition and historical background of a Latin–Low German *Ex quo* vocabulary manuscript preserved in the Tallinn City Archives. It aims to reconstruct the process of its compilation, identify its distinctive textual features, and assess its function within the religious and educational landscape of late medieval Estonia. Previous studies concerning the place of the Tallinn manuscript among the various textual versions of the *Ex quo* vocabularies have not been unanimous. Although it can be confirmed that the Tallinn text belongs to the most widespread redaction of the *Ex quo* manuscripts, its immediate master copy may have combined features characteristic of different sub-redactions. The binding and ownership inscription illuminate the manuscript's connections with the Tallinn Dominican friary, while the identity and institutional affiliation of its scribe remain uncertain.

Keywords: Manuscript studies, Low German vocabularies, Dominican Order, liturgy, medieval Livonia

© 2026 Authors. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International License CC BY 4.0 (http://creativecommons.org/licenses/by/4.0).

INTRODUCTION

From the beginning of the written transmission of the German language in the eighth century, numerous Latin–vernacular glossaries and vocabularies circulated in German-speaking regions.[1] In the late Middle Ages one of the most widespread among them was the anonymous Latin–German vocabulary *Ex quo* which takes its name from the opening words of its introduction.[2] The vocabulary survives in almost 300 manuscripts (some of them incomplete) from approximately 1410–1502 and in more than 40 printed editions from 1467–1505,[3] representing different regional varieties of German. The differences between individual copies and groups of copies are considerable, reflecting a complex process of textual development and the emergence of different redactions retaining, however, a number of substantial common features in both compilation principles and layout.

The present article is devoted to an *Ex quo* manuscript preserved in the Tallinn City Archives (Cm 16).[4] It aims to investigate the ways in which general features observed in *Ex quo* vocabularies are reflected in the Tallinn manuscript, to discuss the circumstances of its production, and to situate it within the framework of local literary culture. After the introductory sections on *Ex quo* manuscripts in general and the codicological features of the Tallinn manuscript, the article proceeds to a text-critical examination of the *Ex quo* text, followed by a discussion of the position of *Termini iuristarum* annexed to *Ex quo* in the Tallinn manuscript. The final section will examine the manuscript's origins and the potential background of its scribe.

The Tallinn manuscript was first brought to scholarly attention during the cataloguing of the former Tallinn Magistracy archive in the 1890s.[5] In the early twentieth century it was briefly touched upon by the

1 See K. Grubmüller. Vocabularius Ex quo. Untersuchungen zu lateinisch-deutschen Vokabularen des Spätmittelalters. C. H. Beck, München, 1967, 45–64; S. Stricker. Old High German Glossaries. – Medieval Glossaries from North-Western Europe. Tradition and Innovation. Ed. by A. Seiler, C. Benati, S. M. Pons-Sanz. Brepols, Turnhout, 2023, 375–377, and other related essays in the quoted volume.

2 B. Schnell. *Vocabularius Ex quo*. – Medieval Glossaries from North-Western Europe. Tradition and Innovation. Ed. by A. Seiler, C. Benati, S. M. Pons-Sanz. Brepols, Turnhout, 2023, 504.

3 See 'Vocabularius Ex quo'. Überlieferungsgeschichtliche Ausgabe. Bd. I. Einleitung. Hrsg. von B. Schnell, K. Grubmüller et al. Max Niemeyer, Tübingen, 1988, 4. B. Schnell. *Vocabularius Ex quo*, 504–506, 510, and https://www.handschriftencensus.de. The first printed edition appeared in Eltville in 1467.

4 Full shelf mark: Tallinn City Archives (Tallinna Linnaarhiiv), TLA.230.1.Cm16 (hereafter Cm 16). Digital images of the manuscript are available through Saaga, the online platform of the National Archives of Estonia (https://www.ra.ee/dgs/explorer.php, reference code TLA.230.1.Cm16).

5 Katalog des Revaler Stadtarchivs. Hrsg. von G. v. Hansen. Reval, 1896, 4, no. 28:

Baltic German historian and archivist Hermann von Bruiningk but not identified as an *Ex quo* vocabulary.[6] In the 1940s, the Finnish philologist Pekka Katara, whose interest lay primarily in the philological study of the manuscript, edited select parts of the manuscript and presented a short general description.[7] In the 1960s, the Estonian cultural historian Hans Treumann devoted an article to Cm 16 that focused on the codicological characteristics of the manuscript and its historical background. Since its first publication in 1968, this article has remained the principal source of information about the manuscript for Estonian scholars.[8]

In Germany, research on *Ex quo* vocabularies has been more intense since the 1960s. In addition to collecting and classifying the lexicographical material, researchers have examined how the *Ex quo* manuscripts reflect the character and change of language use and vocabulary in late medieval literary texts as well as the language proficiency of an averagely educated person.[9] For decades the Tallinn manuscript Cm 16 had only a marginal role in these studies. As the manuscript was not accessible to Western scholars after the Second World War, it was long considered to be lost, and only small portions of the text on microfilm were at the disposal of German researchers.[10] In 1996, Sabine Jordan brought the survival of the Tallinn manuscript to the attention of German scholarship,[11] and it has now (re)entered the standard lists of the *Ex quo* manuscripts.[12]

"Lateinisch-niederdeutsches Wörterbuch, scriptum in Turgel 1454 (geschrieben von einem ungenannten katholischen Priester an jener Kirche in Jerwen)".

6 H. von Bruiningk. Messe und kanonisches Stundengebet nach dem Brauche der Rigaschen Kirche im späteren Mittelalter. Nicolai Kymmel, Riga, 1904, 16.

7 P. Katara. Ein Lat.-nd. Vokabular des Stadtarchivs zu Reval. – Niederdeutsche Mitteilungen 1946, 2, 35–55; P. Katara. Versifizierte Glossen im Vokabular des Stadtarchivs zu Reval. – Neuphilologische Mitteilungen 1948, 49, 1–17. Katara's attention to the manuscript was drawn by Finnish history professor Arvi Korhonen in 1928. The manuscript was sent to Katara to Helsinki for inspection. (P. Katara. Ein Lat.-nd. Vokabular, 35)

8 H. Treumann. Seigake inkunaablisajandist. Poole tuhande aastane Türi koodeks "Ex quo". – Keel ja Kirjandus 1968, 1, 21–24, reprinted in H. Treumann. Vanemast raamatukultuuriloost. Eesti Raamat, Tallinn, 1977, 14–17 (hereafter quoted according to the 1977 edition).

9 'Vocabularius Ex quo', 4–5.

10 Ibid., 88, 213. The microfilm contained the text part with the initial letter *A*. In addition, text parts published by Katara (see P. Katara. Ein Lat.-nd. Vokabular, 45–55) could have been used. The facsimile mentioned by Schnell et al. contained in another Katara's article (P. Katara. Versifizierte Glossen) presents *Termini iuristarum* (Cm 16, fols 232v–234r); see note 64.

11 S. Jordan. Der Revaler 'Ex quo' ist nicht verschollen! – Korrespondenzblatt des Vereins für niederdeutsche Sprachforschung 1996, 103, 20–22.

12 See Handschriftencensus: https://handschriftencensus.de/5591; C. Benati. Middle Low German Glossaries. – Medieval Glossaries from North-Western Europe. Tradition and Innovation. Ed. by A. Seiler, C. Benati, S. M. Pons-Sanz. Brepols, Turnhout, 2023, 542–551 (the Tallinn manuscript: p. 543).

PURPOSES AND CHARACTERISTICS OF *EX QUO* MANUSCRIPTS

Initially, one of the main purposes of the *Ex quo* vocabularies was to facilitate understanding of the literal meaning of the Bible and other Latin texts.[13] At the same time, the vocabularies contain a large number of general words, differing in this respect from other types of medieval glossary devoted to some specific field of knowledge.

Strictly speaking, the *Ex quo* cannot be described as a Latin–German vocabulary, but rather as a Latin vocabulary intended for German-speaking users with some knowledge of Latin, or for those familiar with only a limited range of Latin texts. The main sources of the vocabulary are listed in the introduction: the *Magnae Derivationes* of Huguccio of Pisa, the *Catholicon* of Johannes Balbus of Genoa,[14] and an anonymous *Brevilogus*.[15] As indicated already in the wording of the introduction, this list is not complete.[16] Other vocabularies or schoolbooks could have been used for the various redactions of the *Ex quo*. Some text parts originated from primary sources, others from secondary works.[17] The purpose of listing famous authors and works in the introduction was probably not primarily to refer to the actual sources of the vocabulary, but rather to lend the work greater authority and credibility.

According to the introduction, the *Ex quo* was intended for "poor students" (*pauperes scholares*) and many of the known scribes or owners of *Ex quo* manuscripts were indeed students or teachers. Judging from its contents, however, the vocabulary could also have served as a practical reference work for preachers. Given its unsystematic character, it was not the best tool to be used in academic learning. The question is whether the characteristic features of individual sub-redactions or manuscripts

13 'Vocabularius Ex quo', 5; this aim has been set already in the introduction of the *Ex quo*.

14 See G. Powitz. Zum "Catholicon" des Johannes de Janua. Das Autorexemplar und die Tradition der Exemplare des Franciscus de Agaciis. – Archivum Fratrum Praedicatorum, 1983, 53, 203–218.

15 For the works normally listed in the introduction of the *Ex quo*, see B. Schnell. *Vocabularius Ex quo*, 504.

16 See the introduction of the Tallinn vocabulary in Cm 16, fol. 2r: *Ex quo vocabularij varij auctentici videlicet / hvgvcio katholicon breuilogus alijque codices / sunt in comparacione preciosi in collectione prolixi / intellectione obscuri et in numero multi Ita quod pauperes / scholares eosdem de facili et pro precio competenti ratione / eorum paupertatis habere ac sibi procurare non valebunt / vt eo tamen ~~sal~~ facilius sacram scripturam literaliter intelli / gere potuerint ymmo quodlibet scriptum latinicale pro vtilitate / et necessitate collectus est presens vocabularius / secundum ordinem alphabeti conscriptus Ita quod latinum / precedat et teutunicum subiungitur Post hoc cuius / generis et declinacionis seu cuius partis orationis quelibet / dictio fuerit est annexum.*

17 'Vocabularius Ex quo', 10–12.

can provide insight into developments in education or into the specific needs of users in particular fields, regions or periods.[18]

As for the lexicography, the focus of the *Ex quo* vocabularies is somewhat diffuse: it is not always clear according to what principles the lemmas were included in them, particularly when the words were in broad and common usage. The average number of vocabulary articles per manuscript is relatively limited (ca. 12,000–15,000). Alphabetical order is usually followed up to the fourth letter of the word. In individual manuscripts, however, this order may be disturbed by interpolations or by addition of entire blocks of lemmas. Occasionally, inflectional endings of lemmas have been included, although these and other grammatical details are often presented inconsistently. The same applies to the explanations or equivalents of the lemmas: sometimes they consist of merely a Latin synonym,[19] sometimes of a more extensive explanation in Latin,[20] sometimes of a German equivalent, or a German explanation.[21] These different presentations may also be combined within a single entry.[22] Biblical names are usually translated;[23] in some cases mnemotechnic verses are employed.[24] Most of these features can also be observed in the Tallinn manuscript Cm 16.

Although many *Ex quo* manuscripts are similar in layout and decoration, groups of manuscripts can be distinguished not only on the basis of their contents but also according to their codicological features. Most *Ex quo* manuscripts have been written in bastard or cursive script, suggesting that they were copied for practical use, often by the scribes themselves, although copies in neat *textualis* also exist. Rubricated initials serve not merely as decoration but also as finding aids. Manuscripts and printed editions can be further grouped according to the number of columns (one or two), the version of their introduction (short or long, including various subversions), and the position of abbreviations indicating the grammatical categories of lemmas (before or after the entry).

18 Ibid., 15–16.
19 E.g. Cm 16, fol. 30r: *Cianga dicitur caliga f p*.
20 E.g. Cm 16, fol. 20r: *Bymatus est spacium duorum Annorum m s*; fol. 50r: *Dialogus est sermo duorum uel est liber gregorij In quo habetur vita sancti benedicti m s d*; fol. 157v: *Prora est anterior pars nauis uel capitur pro tota naui*.
21 E.g. Cm 16, fol. 4v: *Adamas eyn sten m t*; fol. 18v: *Beare salichmaken a p.*
22 E.g. Cm 16, fol. 4v: *Adaminia sunt quidam heretici ketter f p*; fol. 32v: *Clades eyne plaghe des dodes scilicet quando homines fame moriuntur ex defectu frugum m t*; fol. 83r: *Grates danck namicheit et habet solum grate in ablatiuo singulari et In plurali habet grates*.
23 E.g. Cm 16, fol. 101r: *Laban Interpretatur de albacio et fuit frater rebecce*.
24 'Vocabularius Ex quo', 6–10.

THE CODICOLOGICAL FEATURES OF THE TALLINN MANUSCRIPT

The Tallinn manuscript Cm 16 consists of 240 paper leaves measuring 20.0 × 13.7 cm with approximately 30 to 40 lines per page. It is a single-column manuscript mainly written in relatively neat bastarda. The vocabulary occupies fols 2r–232v and contains approximately 13,000 lemmas. It is directly followed by a versified text known under the title *Termini iuristarum* (fols 232v–234r) and a colophon (fol. 234r), all written in a single hand.[25] Minor variations in handwriting and letter forms (in some cases, the scribe has used different forms for one and the same letter), pen and ink, occurring primarily in the initial part of the manuscript up to fol. 40r, might indicate either frequent changes in the writing conditions or the copying of the text over a longer period of time.[26] The somewhat careless execution of the manuscript – uneven numbers of lines per page, fluctuation between strictly cursive and partly textualis letter forms, and interruptions in writing – suggests that it was written by a presumably clerical person for his own use, not by a professional scribe.

The first page of the Tallinn manuscript, containing the introduction and the first seventeen lemmas with the initial letter *A*, follows the layout typical of the opening page of single-column *Ex quo* manuscripts, also adopted in some early printed editions (see Fig. 1).[27] A similar opening page layout can be found in some other late medieval vocabularies.[28] The manuscript follows the traditional structure of the *Ex quo* vocabularies: each Latin lemma is followed by its equivalent or explanation in Low German, in Latin, or both, and abbreviations

25 For codicological descriptions, see P. Katara. Ein Lat.-nd. Vokabular, 35–36; H. Treumann. Seigake inkunaablisajandist, 14–15; T. Kala. Mittelalterliche Handschriften in den Sammlungen des Stadtarchivs Tallinn und des Estnischen Historischen Museums. Katalog. Tallinna Linnaarhiiv, Tallinn, 2007, 185–189. K. Kolk. Ladina-alamsaksa sõnaraamat "Vocabularius Ex quo". – Keskaja Liivimaa allikad. Tekstid, kujutised, esemed. Koost ja toim M. Tamm. Tallinna Ülikooli Kirjastus, Tallinn, 2026, 833–839. Katara uses his own foliation starting from fol. 1r as the first page of the vocabulary (see P. Katara. Ein Lat.-nd. Vokabular, 36). Subsequent descriptions and references use the foliation inserted by an employee of Tallinn City Archives in May 1967, in connection with Hans Treumann's research of the manuscript. This foliation also includes parchment leaves used as pastedowns of the volume. The vocabulary starts on fol. 2r (see H. Treumann. Seigake inkunaablisajandist, 17 notes 2 and 4).

26 Changes always appear with a new page (fols 12v, 14r, 17v, 18v, 20v, 21r, 21v, 22v, 23r, 24r, 27r, 29v, 30r, 30v, 32v, 40r).

27 For example, in Eltville: [Nicolaus Bechtermünze] 1472; [Speyer]: [Peter Drach der Mittlere] [ca. 1485]. But not for example in [Heidelberg]: [Heinrich Knoblochtzer] [ca. 1489], [Ulm]: Johann Zainer der Ältere, 1480 and others.

28 See e.g. Johannes Melber. Vocabularius Jodoci doctoris et predicatoris sacre scripture. [Straßburg]: [Martin Schott] [nicht nach 1484]. GW M22731: https://gesamtkatalogderwiegendrucke.de/docs/M22731.htm.



Fig. 1. The opening page of the Tallinn *Ex quo* manuscript containing the introduction and the first seventeen lemmas of the vocabulary (right); the parchment fragment used as the front pastedown originating from a breviary and containing texts for the office of St Andrew (left). TLA.230.1.Cm16, fols 1v–2r (before conservation).

indicating its grammatical categories. In some cases, however, grammatical categories have been omitted. In Cm 16, the abbreviations are placed after the entry, whereas in most manuscripts and printed editions they are located before the Latin lemmas. The position of grammatical categories is usually specified in the introduction of the *Ex quo*. This is also the case in Cm 16.[29] There are deviations from the alphabetical order of the lemmas, and some lemmas occur twice.[30]

In many *Ex quo* vocabularies, large decorated initials mark either the beginnings of the groups of lemmas with a new initial letter or a change in the second letter of the word, thus serving as a finding aid. This pattern has also been followed in Cm 16. In addition, the first letters of lemmas have been marked in red throughout the manuscript. There are, however, numerous cases in Cm 16 in which the change in the second letter of the word has not been marked with a decorated initial, or, conversely, there are red initials where no such change occurs.[31] Since the large initials have been executed after the main text, in spaces left for them by the scribe, this inconsistency could have characterised already the master copy. On fols 215v and 216r, the large initials are absent, perhaps due to turning two leaves instead of one, while the first letters of lemmas are marked in red according to the same pattern as elsewhere in the manuscript. Thus, the various elements of the rubrication were added at different stages of the production of the manuscript. Text beginnings of the sections with the initial letters *I* (fol. 86r), *P* (fol. 136r), *Q* (fol. 177r), *S* (fol. 188r), *T* (fol. 212v), *V* (fol. 222v), *X* (fol. 231v) and *Z* (fol. 232r) have been marked as a separate chapter (*capitulum*) each, while the text sections with other initial letters are not.

According to the colophon, the manuscript was copied (or at least completed) in 1454 "in Turgel", which has been identified as Türi,

29 See the text of the introduction of Cm 16 in note 16. There are cases in which the location of the categories mentioned in the introduction and their location in the vocabulary proper do not coincide, suggesting that the master texts of the introduction and the vocabulary might have originally been different. For example: Národní knihovna České republiky, XXIII.F.107 (https://www.manuscriptorium.com/hub/catalog/default/detail/single/manuscriptorium%7CAIPDIG-NKCR__XXIII_F_107_3S9PY9C-cs?lang=cs); Biblioteca Apostolica Vaticana, Pal. lat. 1779 (https://digi.ub.uni-heidelberg.de/diglit/bav_pal_lat_1779) and Pal. lat. 1790 (https://digi.ub.uni-heidelberg.de/diglit/bav_pal_lat_1790). These manuscripts contain the shortest version of the introduction, in which the categories are said to be found after the lemmas. In the following text of the vocabulary, however, they are to be found before the lemmas.

30 For example, Cm 16, fol. 145v: *pestiferus sukich m s*; fol. 160r: *pestiferus sukich m s*; fol. 145v: *Phalanx est idem quod turba f t d*; fol. 160v: *Phalanx turba*; fol. 151v: *Practica .i. manualis sciencia / Practicare ouen*; fol. 167v: *Practica est manualis scientia uel piscis quidem f p.*

31 For example, Cm 16, fols 41v, 44v, 54r, 61v, 62v, 63v, 84v etc.

a rural parish centre in central Estonia. The wording of the colophon may reflect the influence of theological treatises or liturgical formulae.[32]

At the end of the volume, on an originally blank page (fol. 236v), there are four short, partly illegible additions: two Low German medical recipes against fever, short excerpts from liturgical and biblical texts in Latin, and an ownership inscription from the Tallinn Dominican friary (*Iste liber pertinet conuentui Reualiensi*). These texts have been written by three different hands, none of which seems to be identical with the hand of the principal text.

The codex has thick wooden boards covered in tanned leather tooled in blind with two fillets and three small stamps: a rose, a six-pointed star, and tiny circles forming triangles. Only the metal catch-plate has survived of the fastening of hook-clasp type on the front cover. Four small holes in the back cover, one of which still retains a nail, indicate that the manuscript was once fitted with a chain and belonged to the chained library of the Dominican friary. The binding, too, was most likely executed in the friary. According to Hans Treumann there was also a strip of parchment (now lost) with the title *Ex quo* detached from the front cover.[33] Similar parchment strips bearing book titles have survived on covers of other volumes which could have once belonged to the friary.

According to dendrochronological analysis, the oak used for the back cover was felled between 1454 and 1466, indicating that the codex was bound a few years later, after an appropriate seasoning period.[34] Art-historical analysis of the cover design suggests that the binding was most probably carried out between the 1470s and the 1490s.[35]

Fragments of two parchment manuscripts have been used as pastedowns. The rear pastedown is a leaf from a thirteenth- or early fourteenth-century copy of the Constitutions of the Friars Preachers[36] in neat *textualis* with secondary marginal additions by at least two

32 The colophon has been edited by Katara (P. Katara. Ein Lat.-nd. Vokabular, 55). The wording of the original differs slightly from his edition. Cm 16, fol. 234r: *Et sic finitus est uocabularius iste anno domini milesimo / quadringentesimo quinquagesimo 4°* [Katara: quanto] *feria sexta / ante annunciationem beate marie virginis hora completorij uel / quasi pro quo fine integra diuinitas* [Katara: deitas] *totaque deitas nec / non virginalis integritas sit benedicta in seculorum secula a/men Scriptum in turgel in die festo et alijs quibus / supra*.

33 During Treumann's research, this parchment strip was found between the pages of the manuscript. See H. Treumann. Seigake inkunaablisajandist, 14, 17 note 4.

34 A. Läänelaid, K. Sohar, A. Daly, A. Van Ham-Meert, P. Paiste, K. Kolk, K. Paloveer, R. Suni. New evidence of re-use of an oak panel in Estonia: Covers of the fifteenth century Codex of Türi. – Journal of Cultural Heritage, 2025, 73, 52–61.

35 K. Kolk. Late Gothic Bookbinding in Tallinn. – Baltic Journal of Art History, 2024, 27, 41–88, esp. 81.

36 For the fragment of the Constitutions, see T. Kala. Mittelalterliche Handschriften, 188–189 and the entry in the virtual collection Fragmenta membranea Estoniae (http://hdl.handle.net/10062/54965).

different hands. The copy of the Constitutions was most likely discarded after becoming outdated by the late fifteenth century. This fragment is one of the principal pieces of evidence linking the bookbinding workshop responsible for the binding to the Dominican friary. The front pastedown originates from a breviary and contains the opening of the sanctorale, namely the vigil of St Andrew and part of matins from his feast day. Remarkably, the breviary has probably been written by the same hand as *Ex quo* and *Termini iuristarum* and rubricated in the same style. We shall return to the breviary fragment in the section on the historical background of Cm 16.

THE TALLINN VOCABULARY AND THE CLASSIFICATION OF *EX QUO* VOCABULARIES

As closest to the Tallinn *Ex quo* vocabulary, both with regard to the choice of lemmas and their equivalents, Pekka Katara identified a manuscript in the Royal Library in Copenhagen (Cod. Thott. 111.4), copied in 1448.[37] Using the vocabulary of Lorenz Diefenbach[38] as a basis for comparison, Katara was additionally able to identify two manuscripts from the Mainz City library (Mainz, Stadtbibl., Hs. I 594 and Hs. I 595) whose German equivalents or explanations are close to those in the Tallinn manuscript, whereas the overlap in Latin lemmas is considerably less extensive. Both Mainz manuscripts belong to the first half of the fifteenth century.[39]

Apart from similarities between all four manuscripts there are also significant differences. For example, in cases where German equivalents of Latin lemmas in the Tallinn manuscript coincide with German equivalents in some of the three other manuscripts, the orthography of the Tallinn manuscript is often different. These differences do not, however, appear to reflect distinct linguistic varieties: Katara admits that he was not able to identify any local (Baltic) features in the Tallinn

37 P. Katara. Ein Lat.-nd. Vokabular, 40–42. About this manuscript, see C. Borchling. Mittelniederdeutsche Handschriften in Skandinavien, Schleswig-Holstein, Mecklenburg und Vorpommern. Zweiter Reisebericht. – Nachrichten von der Königl. Gesellschaft der Wissenschaften zu Göttingen. Philologisch-historische Klasse. Beiheft. Lüder Horstmann, Göttingen, 1900, 28–29; 'Vocabularius Ex quo', 71 (here Kh2).

38 L. Diefenbach. Glossarium latino-germanicum mediae et infimae aetatis. J. Baer, Frankfurt am Main, 1857.

39 P. Katara. Ein Lat.-nd. Vokabular, 37. About these manuscripts, see 'Vocabularius Ex quo', 74 (here Ma3 and Ma4). On Hs. I 595 see also K. Grubmüller. Vocabularius Ex quo, 322–329.

manuscript. All four manuscripts represent generally used forms of Low German, specifically northern *ostelbisch* and eastern *nordalbingisch*.[40]

In their fundamental study on *Ex quo* vocabularies, Bernhard Schnell et al. distinguished four principal and two additional redactions of the *Ex quo* vocabularies. The original redaction S was, judging from the oldest surviving manuscripts, probably compiled in the Low German-speaking area around 1410,[41] and redactions K, M and P derived from S. These four principal redactions have been in turn divided into a number of sub-redactions.[42] The redactions have been distinguished on the basis of the selection of original Latin lemmas, the addition or deletion of individual lemmas or groups of lemmas, the expansion or abbreviation of Latin explanations as well as the addition, abbreviation, or deletion of German equivalents, the use of alternative German equivalents and regional word forms. German equivalents constitute the most variable component of the *Ex quo* manuscripts.[43] Different redactions reflect not only regional preferences or influences of particular sources, but also the development of the vocabulary over time, whereas individual manuscripts can reveal the personal preferences of their scribes.[44]

Schnell et al. classified the Tallinn manuscript as belonging to redaction P, described as the most complex of the principal redactions. Approximately 2,300 lemmas have been added to the ca. 12,700 original lemmas of redaction S in redaction P. These additions typically consist of a Latin lemma and its German equivalent, but do not include longer explanations. The sources of these additions have not been established. Redaction P is represented in more than 50 manuscripts. Their number begins to decline after the middle of the fifteenth century; the latest dated manuscripts of redaction P belong to the 1470s.[45]

The Copenhagen manuscript (Cod. Thott. 111.4 = Kh2) has been classified by Schnell et al. as belonging to branch Y of redaction P, while the Tallinn manuscript and one of the two Mainz manuscripts inspected by Katara (Hs. I 595 = Ma4) were assigned to branch X of redaction P.[46] The other Mainz manuscript (Hs. I 594 = Ma3) has been classified

40 P. Katara. Ein Lat.-nd. Vokabular, 42–43.
41 On this redaction, see 'Vocabularius Ex quo', 19, 113, 119–123.
42 'Vocabularius Ex quo', 12–15; B. Schnell. *Vocabularius Ex quo*, 506 (B. Schnell et al. also describe a redaction Me in which Latin interpretamenta are significantly expanded and its direct connection to other redactions cannot be established). See the stemma of the redactions in 'Vocabularius Ex quo', 116. The additional redactions I and W derive from M.
43 'Vocabularius Ex quo', 12–15.
44 See for example B. Merrilees. "Copiste et compilateur": Transmission and Individuality in Medieval Glossaries. – "Li premerains vers". Essays in Honor of Keith Busby. Ed. by C. M. Jones. Brill, Boston, 2011, 281–290.
45 B. Schnell. *Vocabularius Ex quo*, 507; 'Vocabularius Ex quo', 206–210.
46 'Vocabularius Ex quo', 208.

as belonging to redaction S.[47] This classification differs from Katara's assessment, but it should be kept in mind that Schnell et al. were not able to examine the entire Tallinn manuscript,[48] whereas Katara had access to both the Tallinn and the Copenhagen manuscripts in person and over a longer period of time.[49] As for the language of the Tallinn manuscript, Schnell et al. designated it as *baltisch*,[50] which differs from Katara's judgement described above.

Schnell et al. placed the Tallinn manuscript Cm 16 (Re1)[51] in a subgroup of five manuscripts together with Mainz, Stadtbibl., Hs. I 595 (Ma4),[52] Wolfenbüttel, Herzog August Bibliothek, Cod. Guelf. 369 Helmst. (Wo2),[53] Berlin, SB Preuß. Kulturbesitz, Ms. Lat. fol. 868 (B15),[54] and Basel, UB, Cod. F IV 46 (Bs6).[55] Of these, B15, Re1 and Bs6 do not fully follow the redaction P (B15 and Bs6 comply with the features of this redaction only in the beginning of the alphabet).[56]

At least two of the manuscripts from this group of five are single-column (TLA, Cm 16; Basel, UB, Cod. F IV 46) and two are two-column manuscripts (Mainz, Hs. I 595; Wolfenbüttel, Cod. Guelf. 369 Helmst.). The short version of the introduction can be found at least in the Tallinn manuscript Cm 16, the Mainz manuscript Hs. I 595,[57] and the Wolfenbüttel manuscript Cod. Guelf. 369 Helmst.; in those three manuscripts the abbreviations indicating grammatical categories are placed after the entries.[58]

As regards the content of the vocabularies, the classification of the above-mentioned five manuscripts as one distinct group by Schnell et al. is disputable. On the basis of the comparison of Cm 16 (Re1) and Cod. Guelf. 369 Helmst. (Wo2) and partial comparison of Cm 16 with Hs. I 595 (Ma4), it can be said that the Tallinn manuscript does not fully follow either the selection or the order of lemmas in other manuscripts

47 Ibid., 121–122.
48 See p. 281 and note 10.
49 See P. Katara. Ein Lat.-nd. Vokabular, 35, 40.
50 'Vocabularius Ex quo', 209.
51 Ibid., 88 (here the manuscript's old shelf mark Cod. 28 has been used and the manuscript has been marked as lost).
52 Ibid., 74; Handschriftencensus: https://www.handschriftencensus.de/3376.
53 Ibid., 98; Handschriftencensus: https://www.handschriftencensus.de/6707.
54 Ibid., 57; Handschriftencensus: https://handschriftencensus.de/4619.
55 Ibid., 54; Handschriftencensus: https://www.handschriftencensus.de/7378.
56 Ibid., 208, 213.
57 See the edition of this introduction according to the Mainz manuscript (Hs. I 595) in K. Grubmüller. Vocabularius Ex quo, 333. For a detailed description of the manuscript, see Christopher Winterer (2018): https://handschriftenportal.de/workspace, Wissenschaftliche Stadtbibliothek Mainz, Hs. I 595.
58 No detailed description of B15 (Berlin, SB Preuß. Kulturbesitz, Ms. Lat. fol. 868) has been available to the authors of this article.

of this group. In many cases German explanations of lemmas are also differently worded.[59] In some sections of Cm 16 additional lemmas have been inserted on the same line as a preceding entry (e.g. fols 33r, 33v, 34r, 34v, 36v), although no change in handwriting or writing conditions can be observed. For example, eleven additional lemmas occur on fol. 33v, while in corresponding text parts in Hs. I 595, fol. 29r and Cod. Guelf. 369 Helmst., fol. 17r all lemmas, including those added in Cm 16 on the same line as the preceding entry, are written on a separate line each. This suggests that the master copy of Cm 16 probably contained secondary additions corresponding to textual versions represented in Hs. I 595 and Cod. Guelf. 369 Helmst., while the scribe of Cm 16 simply followed the layout of his master copy. This copy, in turn, probably originally belonged to some earlier redactional stage than, for example, Hs. I 595 or Cod. Guelf. 369 Helmst.

Several *Ex quo* manuscripts, including the Tallinn manuscript, contain errors caused by misinterpretation of the master copy. The Tallinn manuscript, however, does not necessarily reproduce the errors observed in manuscripts previously considered closely related to Cm 16; rather, it contains a different set of errors. For example:

Mainz, Hs. I 595, fol. 178v: *Zonifragium est zonarum fraccio n s*[60]
Tallinn, Cm 16, fol. 232v: *Zonifragium est sonifraccio n s*

Wolfenbüttel, Cod. Guelf. 369 Helmst., fol. 2ra: *Adangere to denken a s q / Addere to voghen uel to don a t q*
fol. 4rb: *Amen id est vere pariter t sit sine ~~fide~~ fine*
fol. 22ra: *Curis est dos uel curia plebis* (!) *f t*

Tallinn, Cm 16, fol. 4v: *Adangere to oken* (!) *a s / Addere to vogen to doen a t*
fol. 8v: *Amen idem vere pariter t sit sine sine* (!)
fol. 44r: *Curis .i. des* (!) *uel est curia plebani f T*

Schnell et al. distinguished four sub-versions – X, X1, X3, and X4 – situated between the group of five manuscripts described above and the initial redaction P,[61] with Cm 16 sharing features of sub-versions X3 and X4.[62] A closer comparison of Cm 16 with the characteristic features of X3 and X4, however, indicates that the Tallinn manuscript

59 For example, Hs. I 595, fol. 29r: *Cofinus .i. cophinus en draghe korf m s*; Cm 16, fol. 40v: *Cophinus eyn kype uel eyn brotkorff m s*.
60 K. Grubmüller. Vocabularius Ex quo, 323.
61 'Vocabularius Ex quo', 208.
62 Ibid., 212–213.

does not conform precisely to either of these sub-redactions.[63] At the same time, Cm 16 displays certain similarities to deviations from the original redaction P observed in X1, although with its own features, including incorrect readings.

Additional entry in X1:

> *f p Cimera est monstrum uel aliquid fictum, quod tamen in rei veritate nihil est* (›Vocabularius Ex quo‹, 211)
> Cm 16, fol. 30v: *Cimera est idem cum ~~cymera~~ chymora f p*

Expanded interpretation of a lemma in X1:

> *n s Burellum ein duch das von line vnd von wollen gemacht ist + est pannus de lino et lana stamine contextus* (›Vocabularius Ex quo‹, 211)
> Cm 16, fol. 22r: *Burellum est pannus de lino uel lano en sleger n t / stucke uel en beyer wantt*

Incorrect readings of interpretations of lemmas in X1:

> *Rachel quis* (!) *interpretatur* instead of *m t Rachel ouis interpretatur* (›Vocabularius Ex quo‹, 211)
> *Regan eyn groddans* (!) *uel parens interpretatur* instead of *m n Regan egrotans uel parens interpretatur* (›Vocabularius Ex quo‹, 211)
>
> Cm 16, fol. 180v: *Rachel q. ouis interpretatur*
> Cm 16, fol. 182v: *Reganen* (!) *grote* (!) *aut uel parens Interpretatur*

Although Cm 16 does not strictly correspond to any of the subdivisions discussed above (X1, X3, X4), it still tends to be closer to X1 rather than to X3 and X4. At the same time, it is not always possible to determine which peculiarities of the Tallinn manuscript reflect one or other redactional stage of the *Ex quo* vocabularies and which should instead be regarded as individual features or scribal errors of Cm 16.

63 In X3 due to skipping one line the entries *f p Ramata est ciuitas / f in Ramasse est ciuitas in finibus Egypti* appear as follows: *f p Ramata est ciuitas quedam in finibus Egypti* ('Vocabularius Ex quo', 212). In Cm 16, fol. 180v the text is correct: *Ramatha est ciuitas quedam f p / Ramassa est ciuitas In finibus egipti f p.* As one of the characteristic features of X4, Schnell et al. mention additional interpretations of the following lemmas: *n s Analogium eyn pulte + super quod ponuntur libri*; *m s Bucellus ein stog, eyn snyde uel eyn bißen + .i. morcellus* ('Vocabularius Ex quo', 213). In Cm 16, fol. 9v the same entries are not following this pattern: *Analogium en prekek* (!) *stol f p*; fol. 21v: *Bucellus en snede en bete en stucke m s* (here *stucke* is an additional German, not Latin interpretation).

TERMINI IURISTARUM IN THE TALLINN MANUSCRIPT

After more than 461 pages of *Ex quo*, Cm 16 concludes with another, considerably shorter lexicographical text, commonly known as *Termini iuristarum*. It extends over just three and a half pages and consists of 110 verses. Although far less widely dispersed and less universal than *Ex quo*, the mnemonic verse vocabulary *Termini iuristarum* nevertheless belongs to the widespread category of short late medieval Latin–German lexica. Originating in the late thirteenth century, its earliest version presumably consisted exclusively of legal glosses, but by the mid-fifteenth century it had been expanded to include various other topics. Following the conventional opening with legal vocabulary, the Tallinn version contains brief sections devoted to the names of tools, animals, plants, and so forth. Pekka Katara published a transcription of the Tallinn text in 1948 and compared its contents with those of other published versions of *Termini*.[64]

The textual transmission of *Termini* does not display a uniform pattern: verses have been added and omitted on a large scale, their order varies considerably, and half-verses or even smaller metrical units are combined in different ways. Like *Ex quo* and many other Latin–German vocabularies, *Termini* also survives in both High and Low German and in their various regional forms. Since the equivalents of the Latin lemmas in these dialects could differ greatly and often varied in length, the exact wording likewise had to be rearranged, at least where an attempt was made to preserve the metrical structure of the hexameter. The *Handschriftencensus* of medieval German-language manuscripts currently lists 51 codices containing the text of *Termini* in one form or another, ranging from fragments comprising only a few verses to interpolated versions exceeding one hundred lines.[65] At 110 verses, the Tallinn text belongs among the longest, together with the versions in Mainz (Stadtbibl., Hs. I 160, 118 verses), Wrocław (Universitätsbibl., Cod. IV F 86, approximately 110 verses written as prose), and a lost manuscript from Strasbourg (Stadtbibl., Cod. C 107, 111 verses).[66] Most surviving versions are considerably shorter. The longest texts

64 P. Katara. Versifizierte Glossen, transcription 5–17; cf. W. Wackernagel. Deutschlateinische Hexameter. – Zeitschrift für deutsches Alterthum, 1845, 5, 413–416; J. Zacher. Die nomina volucrum und die termini juristarum. – Zeitschrift für deutsche Philologie, 1880, 11, 299–324.

65 https://handschriftencensus.de/werke/2901 (accessed 04/07/2026).

66 Ibid.; https://bibliotekacyfrowa.pl/publication/32955 (accessed 04/07/2026); W. Wackernagel. Deutschlateinische Hexameter, 413–416; P. Katara. Versifizierte Glossen, 3.

are essentially the most heavily interpolated or supplemented variants. These expanded versions do not constitute a coherent group but instead combine material from different sources.

Within the manuscript tradition, such a short text occupying only a few pages invariably forms a small proportion of the entire codex. It never appears independently but always in combination with one or more texts of diverse character. The principal contents of the codices vary considerably and include liturgical, hagiographical, devotional, medical, and historical works, either in Latin or German, or in both languages. Nevertheless, *Termini* is quite frequently appended to other vocabularies. In twenty-three of the fifty-one recorded witnesses, the main part of the manuscript consists of another, more extensive Latin–German vocabulary, usually extending over several hundred leaves. According to the data provided by the *Handschriftencensus*, *Termini* most frequently appears – on eleven occasions – as an addition to *Ex quo*.

Does the coexistence of these two texts in the Tallinn codex reveal anything about the manuscript's history and origins? It seems reasonable to assume that *Ex quo* and *Termini* already occurred together in the exemplar from which our manuscript was copied. Since the colophon states that the manuscript was completed in a small rural settlement, it is unlikely that several vocabulary manuscripts would have been available there simultaneously. Furthermore, after completing the text of *Ex quo*, the scribe appears to have continued writing without interruption and without any noticeable change in ductus or scribal practice. Indeed, there are neither concluding words marking the end of *Ex quo* nor introductory phrases to *Termini*, although the latter are present in many other manuscripts. No change of script, enlarged initial, or blank line marks the beginning of a new text; instead, the writing continues as though *Termini* were simply the next section of the preceding work. This contrasts sharply with the numerous earlier points within *Ex quo* where the scribe's handwriting and even the page layout change noticeably, presumably owing to interruptions in the copying process. The initial *E* marking the beginning of *Termini* closely resembles an ordinary initial *E* elsewhere in the codex marking the beginning of the subsection containing words beginning with the letters *E* and *L* (see Fig. 2).[67] The three and a half pages of *Termini* were evidently copied without interruption, with no discernible change in ductus. Maintaining the same appearance throughout, the text concludes with the colophon

67 Cm 16, fol. 57v.

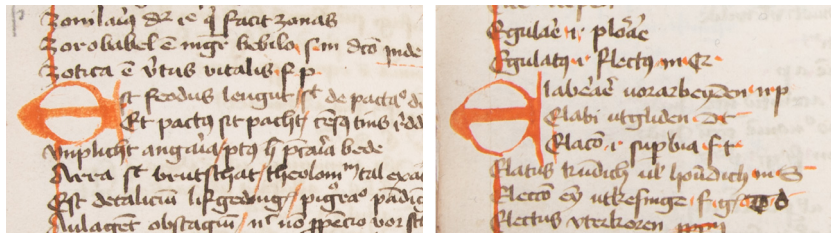


Fig. 2. The beginning of *Termini iuristarum* compared with the section containing words beginning with the letters *E* and *L* (TLA.230.1.Cm16, fols 232v and 57v).

of the entire manuscript, where the work is referred to as *vocabularius*, just as in the incipit of the codex.

Such a layout is by no means typical of other manuscripts containing *Termini*. The beginning of *Termini*, usually following another lexicographical work, is invariably marked clearly. An interruption in copying, the scribe's deliberate decision to indicate the beginning of a new text, or both are immediately apparent. There may be blank space between the texts[68] or even an entirely blank page, the explicit of the preceding text may be rubricated,[69] or the new text may begin with a large decorative initial.[70] Since nothing comparable occurs in the Tallinn manuscript, it seems reasonable to assume that *Ex quo* and *Termini* were already closely connected in the exemplar, with little or no perceptible break between them, leading the scribe to regard them as a single work.

Let us now turn to the corpus of manuscripts containing both *Ex quo* and *Termini* in order to examine whether the close connection between the two texts characteristic of the Tallinn manuscript also occurs elsewhere in the textual tradition. The following table presents the present locations of the manuscripts, the positions of both texts within them, and the approximate or exact extent of the respective versions of *Termini*.

Table 1. The manuscripts containing both *Ex quo* and *Termini*.

	Ex quo	*Termini*
Altenburg (NÖ), Stiftsbibl., Cod. AB 13 B 2	1r–86v (Ab2)	119rv
Altenburg (NÖ), Stiftsbibl., Cod. AB 15 B 4	1r–93v (Ab1)	132v

Continued on the next page

68 Wrocław, Biblioteka Uniwersytecka, Cod. IV F 86, fol. 10v.
69 Wolfenbüttel, Herzog August Bibliothek, Cod. Guelf. 576 Helmst., fol. 113r.
70 Basel, Universitätsbibl., Cod. F V 41, fol. 385r.

Table 1. Continued

	Ex quo	*Termini*
Basel, Universitätsbibl., Cod. F V 41	1r–381r (Bs9)	385r–386v (65 verses)
Klagenfurt, Universitätsbibl., Pap.-Hs. 16	1r–365v (Kf3)	366v–367r (38 verses)
Munich, Staatsbibl., Cgm 668	1r–189v (M16)	191r
Vatican, Bibl. Apostolica Vaticana, Cod. Pal. lat. 1781	6r–143r (R2)	191rv (51 verses)
Solothurn, Zentralbibl., Cod. S I 212	1r–130r (So1)	243rv (68 verses)
Stuttgart, Landesbibl., Cod. HB VIII 8	2r–160r (S6)	1r
Tallinn, Linnaarhiiv, Cm 16	2r–232v (Re1)	232v–234r (110 verses)
Vienna, Österr. Nationalbibl., Cod. 2951	1r–304r (W4)	II
Wolfenbüttel, Herzog August Bibl., Cod. Guelf. 576 Helmst.	125r–204r (W03)	113r (31 verses)

It is immediately apparent from the table that the overwhelming majority of these manuscripts are now preserved in southern Germany, Austria, Switzerland, and the Vatican. Their places of origin largely coincide with their present locations: almost all originated in Bavaria, Austria, or the Upper Rhine region and contain High German alongside Latin. Even the only manuscript outside the southern German area apart from the Tallinn codex, Wolfenbüttel, Cod. Guelf. 576 Helmst., the only witness of Lower Saxon origin (Göttingen), displays Middle German linguistic features.[71] It contains only thirty-one verses of *Termini*, is confined to legal vocabulary, and differs from the Tallinn manuscript in both composition and order. The Tallinn version of *Termini* is by far the longest among these eleven manuscripts; the next longest, the Solothurn codex, contains only sixty-eight verses. The positions of *Ex quo* and *Termini* within the manuscripts likewise vary considerably. The only other manuscript in which *Termini* immediately follows *Ex quo* is the Klagenfurt codex, where the two texts are separated merely by a blank page. Nevertheless, the versions of both texts differ from those

71 B. Lesser. Wolfenbüttel, Herzog August Bibliothek, Cod. Guelf. 576 Helmst., 2017 (https://handschriftenportal.de/workspace, HSP-ID: HSP-70b09d77-ffec-3c14-b613-014ebf939709, accessed 07/07/2026).

preserved in the Tallinn manuscript. In all the remaining manuscripts, the two texts are separated by other works or, in a few cases, *Termini* precedes *Ex quo*. Finally, only two further manuscripts containing both *Ex quo* and *Termini* – those of Solothurn and Stuttgart – belong to the same redaction (P) of *Ex quo* as the Tallinn manuscript, although they represent distant branches within that redaction.[72] It may therefore be concluded that the surviving manuscripts containing both *Ex quo* and *Termini* display no particularly close relationship. The decision to include both texts in a single codex appears generally to have been an independent choice on the part of the scribe and could therefore occur in geographically distant regions. Although Katara characterises the Tallinn text of *Termini* as confused,[73] the manuscript itself presents a neat and orderly transcript, suggesting that the text was not compiled from several sources but copied straight from one model, which in turn may have been responsible for the composition of the text. That exemplar may well have been another student's or scholar's manuscript compiled at a university or another educational institution in the Low German-speaking area and then incidentally transcribed by a presumably clerical person, at least partly in an Estonian rural area.

THE HISTORICAL BACKGROUND OF THE TALLINN MANUSCRIPT

While the owner's inscription in Cm 16 and specific binding features indicate that the manuscript belonged to the Tallinn Dominican friary, the questions of who copied the texts and what the connection of the rural settlement mentioned in the colophon might have been to the Tallinn friary are not so easy to answer.

Given the educational aspirations of the mendicants in general and the Friars Preachers' multilevel education system in particular, the manuscript complies well with the needs of a Dominican friary.[74]

72 'Vocabularius Ex quo', 208.

73 P. Katara. Versifizierte Glossen, 4: *zeigt … einen bunten Wirrwarr*.

74 For the general features of the mendicant education system, see for example M. M. Mulchahey. "First the Bow is Bent in Study": Dominican Education before 1350. Pontifical Institute of Mediaeval Studies, Toronto, 1998; A. Rüther. Educational Communities in German Convents of the Franciscan and Dominican Provinces before 1350. – Medieval Education. Ed. by R. B. Begley, J. W. Kotersky. Fordham University Press, New York, 2005, 123–132. On Dominican biblical studies, see for example W. A. Hinnebusch. The History of the Dominican Order. Intellectual and Cultural Life to 1500. Vol. 2. Alba House, New York, 1973, 99–115. On Dominican schools and libraries in

Although there are only eleven surviving handwritten books and a few dozen manuscript fragments which can be connected to the Tallinn Dominican friary with greater or lesser certainty,[75] they still include texts which refer to studies of the arts, such as fourteenth-century fragments from verse grammars of Alexander de Villa Dei and Eberhardus Bethuniensis[76] as well as notes on orthography in a *Sammelband* from the end of the fifteenth century.[77]

It is not known when the vocabulary entered the friary's library. In January 1525, during the Reformation, the Tallinn Magistracy dissolved the friary and gradually took over the rest of its possessions, which probably included the meagre remains of its library.

Under the introduction of the *Ex quo* in the Tallinn manuscript, the scribe of the vocabulary has inserted the name *her marten* (Sir Marten) in red ink. Partly relying on the description in the archival catalogue from 1896, Hermann von Bruiningk suggested that it was the name of the scribe who was probably also the parish priest of Türi.[78] This assumption was generally accepted in subsequent historiography,[79] although no reliable sources concerning parish priests of Türi survive from the fifteenth century. Hans Treumann tentatively added the alternative possibility that the scribe might have been a Dominican friar.[80] In the 1470s a friar named Merten, the son of Wolder Kann from Tallinn, occurs among the Tallinn Dominicans. He was probably identical with the Merten Wolteri mentioned in 1482 as one of the leading brethren and with brother Merten who died before 1495.[81] However,

the northern Baltic Sea region (including Tallinn), see J. G. G. Jakobsen. The Dominican Order and Urban Literacy in Medieval Scandinavia. – Urban Literacy in the Nordic Middle Ages. Ed. by K. H. Andersen, J. B. Netterstrøm, L. M. Imer, B. Poulsen, R. Steenholt Olesen. Brepols, Turnhout, 2021, 375–410.

75 About the remains of this library, see K. Kolk. Trükitud raamatu esimene aastasada Eestis: raamatud, raamatukogud ja lugejad 15. sajandi teisest poolest 16. sajandi keskpaigani. Tartu Ülikooli Kirjastus, Tartu, 2025, 61–68, here particularly pp. 64–65.

76 Fragments of both verse grammars survive as binding material of manuscript books (TLA.230.1.Cm4 and TLA.230.1.Cm13 respectively).

77 TLA.230.1.Inc13a, fols 100r–102r.

78 H. von Bruiningk. Messe und kanonisches Stundengebet, 16 note 4. According to Bruiningk, the name was added by a different hand. This interpretation might have been caused by the fact that the scribe of Cm 16 often uses different forms of one and the same letter, *a* among them. Treumann was convinced that the name and the principal text were written by one and the same person (see H. Treumann. Seigake inkunaablisajandist, 14, 17 note 1).

79 Relying on the study by Hermann von Bruiningk, Leonid Arbusow has included Marten in the list of parish priests of Türi (L. Arbusow. Livlands Geistlichkeit vom Ende des 12. bis ins 16. Jahrhundert. Dritter Nachtrag. Steffenhagen, Mitau, 1913, 133, 316).

80 H. Treumann. Seigake inkunaablisajandist, 14.

81 T. Kala. Jutlustajad ja hingede päästjad. Dominiiklaste ordu ja Tallinna Püha Katariina konvent. Tallinna Ülikooli Kirjastus, Tallinn, 2013, 127, 167–168; H. Treumann. Seigake inkunaablisajandist, 14; Liv-, est- und kurländisches Urkundenbuch. Zweite Abteilung, Bd. 1. Deubner, Riga, Moskau, 1900, no. 242.

whether this friar (or these friars) might have had anything to do with the name under the introduction of the vocabulary cannot be said.[82]

In the medieval period Türi (80 km from Tallinn in a straight line) belonged to the Teutonic Order's *Komturei* Viljandi (Fellin) and the diocese of Tartu (Dorpat). The parish church was in existence by 1300[83] and stood on the frequently used road from Tallinn to Viljandi.[84] Thus it would not have been difficult for a Tallinn friar to visit the place. However, no records of a direct link between the Tallinn friary or an individual friar and Türi have been preserved. No sources that would mention specific place names survive from the travels of the Tallinn friars in the countryside in the middle of the fifteenth century. Such records survive only from the 1520s. In 1524 friar Petrus Romer travelled from Tallinn to Paide (Weissenstein), the centre of the Teutonic Order's *Vogtei* Järva (Jerwen).[85] The distance between Türi and Paide is approximately 12 km. We cannot exclude the possibility that the Tallinn friary had a kind of stopping place in Türi, for example for brethren who were collecting alms, and some contacts with this region may have existed in the 1450s when the vocabulary was copied.

As mentioned above, the breviary leaf containing the beginning of the office of St Andrew and used as the front pastedown of the codex was probably written by the same hand as the main part of the manuscript, the vocabulary (see Fig. 1). The fragment's textual contents may therefore help to determine the scribe's background. To investigate the liturgical affiliation of the breviary, its text was collated with other late medieval breviaries. Since the breviaries of Tartu, the 'home' diocese of Türi, and of the neighbouring bishopric of Tallinn have not survived, diocesan breviaries from the Baltic Sea region that were presumably close to the uses of Tartu and Tallinn were used instead. The Diocese of Tartu was a suffragan of Riga and that of Tallinn of Lund,[86] although the liturgies

82 The parish church of Türi was dedicated to St Martin, but regarding the form of the phrase in question it should be kept in mind that invocations of this kind addressed to patron saints are not known to have been common in manuscripts.

83 Baltisches historisches Ortslexikon. Teil I. Estland. Hrsg. von H. von zur Mühlen, bearb. von G. Westermann. Böhlau, Köln, Wien, 1985, 610; P. Johansen. Die Estlandliste des Liber Census Daniae. H. Hagerup, Kopenhagen; F. Wassermann, Reval, 1933, 106, 173–174 (in 1560 Türi was transferred to the Teutonic Order's *Vogtei* Jerwen).

84 P. Johansen. Die Estlandliste, 241.

85 T. Kala. Jutlustajad ja hingede päästjad, 330. In winter 1524 the Tallinn friars received grain from Järva (ibid., 151). Whether the journey of the friar Petrus was somehow connected to these grain deliveries is not known; the accounts of the friary do not specify the aim of his journey.

86 For the subordination of the bishopric of Tallinn, see M. Maasing. Die Metropolitanverbindung Rigas mit den preußischen Bistümern zur Zeit des Erzbischofs Wilhelm von Brandenburg. – Die baltischen Länder und Europa in der Frühen Neuzeit. Hrsg. v. N. Angermann, K. Brüggemann, I. Põltsam-Jürjo. Böhlau, Köln, Weimar, Wien, 2015, 165–187, here 168–169.

of both may also have been influenced by contacts with other North German traditions, particularly that of Lübeck. Due to the Dominican ownership of the vocabulary manuscript and the potential connection of the scribe with this order, the standard text of the late medieval Dominican breviary was included in the comparison. The structure of the office from the beginning of the vigil of St Andrew to the second lesson of matins of his feast day is most conveniently presented in the table annexed as an appendix to the present study.

Considerable variation between different liturgical uses occurs in almost every element of the office of St Andrew. Only the Gospel reading, the collect *Quaesumus omnipotens Deus* (at vespers), the invitatory antiphon *Adoremus victoriosissimum regem*, the responsory and versicle following the first lesson of matins of the feast day, and a few other chants remain invariable throughout the corpus. The readings for the matins of the vigil in all versions derive from Bede's sixteenth homily on the Gospels,[87] but the selection and sequence of the passages vary considerably. The rigid structure of the concise Dominican lessons, quoted almost verbatim from Bede but arranged in a modified sequence, also appears in the Kamień breviary and, to some extent, in that of Uppsala. All the remaining breviaries, including the Tallinn fragment, preserve the original order of the paragraphs (with omissions) and begin at the opening of the homily, although the subsequent extracts differ in every case. Most of the diocesan breviaries supplement the readings with short introductory phrases. While most diocesan breviaries display only minor divergences in the subsequent office of none, both the Tallinn fragment and the Lund breviary omit it altogether, as does the Dominican breviary. The vespers reflect the differences between the German tradition (including the Tallinn fragment) and the Scandinavian tradition (Lund and Uppsala), whereas the Dominican use retains its more concise form. The same dichotomy between the German (including the Tallinn fragment) and Scandinavian traditions continues in the sung portions of the first nocturn, with certain peculiarities in the Riga breviary; here, however, the usual 'German' version coincides with the Dominican tradition. The lessons of matins of the feast day are drawn from hagiographical tradition in all the breviaries, but whereas most diocesan liturgies (including the Tallinn fragment), with the sole exception of Kamień, preserve a uniform sequence of passages from the

87 Beda Venerabilis. Homeliarum evangelii libri II. Ed. D. Hurst. (Beda Venerabilis. Opera homiletica. Opera rhythmica. Corpus Christianorum, Series Latina, 122.) Brepols, Turnhout, 1955, lib. 1, hom. 16.

Passion of St Andrew, the Dominican version instead introduces a series of anecdotal, exemplum-like miracle stories.

The Cm 16 fragment consistently follows one or other diocesan liturgical tradition. It coincides with Dominican practice only where the respective diocesan use also overlaps with it, and it is clearly distinct from the otherwise highly standardised composition of the Dominican breviary. Most notably, the fragment does not contain the Dominican miracle stories incorporated into the lessons of matins, which appear to have been particularly well suited to preaching. In omitting the office of none on the vigil of St Andrew, it resembles the Lund breviary, but overall it is closest to the Lübeck use. The combination of North German liturgical traditions with certain Scandinavian features suggests that the fragment may represent the liturgical use of Tallinn, although the surviving comparative material does not allow a secure attribution. Such Scandinavian connections would be considerably less expected in the case of Tartu, where affinity with the Riga use would be more likely; yet the fragment displays no distinctive features of the Riga liturgy.

This leads to the conclusion that the clergyman who copied the breviary, presumably for his own use, was at that time serving as a secular priest. Whether this took place in Türi or elsewhere remains uncertain; indeed, it is not even known whether the connection between the vocabulary manuscript and Türi extended beyond a merely ephemeral association. The unbound vocabulary and the disused breviary may have entered the bookbinder's workshop together. The vocabulary was intended to become part of the Dominican library and was therefore provided with a durable binding and a chain. A diocesan breviary may have been of less value to the friary, which could help explain its reuse as binder's waste, although the reason for its disposal remains uncertain.[88] Whether the scribe of both the breviary and vocabulary was initially one of the friars who may have temporarily served secular parishes, or whether the friary obtained the unbound manuscript from a secular priest, himself perhaps a future friar, remains unclear.

88 Discarding often depended on the physical condition of the manuscript. Perhaps a Dominican breviary would have continued to serve the needs of the friars at a time when printed breviaries had not yet become common: the very first breviary was printed around 1473 for the Use of Constance (Breviarium Constantiense. [Konstanz?/Basel?/Augsburg?], not after 1473, ISTC ib01150975, GW 5315), while the first Dominican breviary appeared in 1476 (Breviarium Fratrum Praedicatorum. Milan: Christoph Valdarfer, 1476, ISTC ib01138500, GW 5214). However, Dominican liturgical books could occasionally be used as binding material for other Dominican books; cf., e.g., the rear pastedown of New Haven, Beinecke, Music Deposit 60 (we owe thanks for this reference to the anonymous reviewer of this paper).

CONCLUSION

In studies on late medieval German glossaries the Tallinn manuscript has received uneven attention. Among the sub-redactions of the *Ex quo* vocabularies, it has been somewhat misplaced, partly apparently since it was temporarily not accessible to German researchers working on these vocabularies.

With its short introduction and location of grammatical categories behind the entries, the Tallinn manuscript seems to belong to a relatively early version of the *Ex quo* vocabularies that originated in the first half of the fifteenth century. Thus, in the 1450s, when the Tallinn manuscript was copied, it did not represent the latest textual tradition of these vocabularies. The comparison of the Tallinn vocabulary with manuscripts which have previously been described as standing textually close to it demonstrates that a meticulous division between individual sub-redactions might not always be justified, while the text of many manuscripts, the Tallinn *Ex quo* among them, could have been strongly influenced by random circumstances or scribal preferences. The Tallinn version of *Termini iuristarum* shows no close textual relationship to the other versions occurring in manuscripts also containing *Ex quo*. This leads to the conclusion that placing these texts side by side was usually the scribe's choice in each case, not an established textual convention. However, in the case of the Tallinn manuscript, both texts were apparently present in the codex that served as the model for Cm 16, since the scribe of Cm 16 regarded them as a single text, as evidenced by textual layout.

Like many other *Ex quo* manuscripts, the Tallinn vocabulary was meant for everyday use and was accessible to different people, as indicated by secondary additions on the last page of the volume. At some point the manuscript entered the library of the Tallinn Dominican friary. The breviary fragment, apparently copied by the same scribe as the vocabulary and later reused in its binding, follows a diocesan rather than a Dominican liturgical tradition. The principal scribe, even if he was not a friar himself, might still have had contact with the friary, an institution where this kind of manuscript could have been needed and a master copy easily obtained.

Acknowledgements

This research was funded by Estonian Research Council grant PRG1276 Digital Livonia: For a Digitally Enhanced Study of Medieval Livonia (c. 1200–1550).

APPENDIX. TEXTUAL CONTENTS AND LITURGICAL AFFILIATION OF THE BREVIARY FRAGMENT IN CM 16

The verso side of the fragment is easily readable; the text on the glued recto, particularly in the first column, is more poorly preserved and partly illegible. Nevertheless, the structure of the office, including the textual passages serving as the readings for the matins and the chants for the vespers of the vigil of St Andrew (contained on the recto side), can be reconstructed with a high degree of confidence from the surviving text, with the exception of the introductory phrases to the readings.

The fragment's text was collated with late medieval Dominican breviaries and several diocesan breviaries from the Baltic Sea region. The Dominican breviaries included in the comparison comprised three early printed editions[89] presenting virtually identical texts and three late medieval manuscripts,[90] which exhibit only minor divergences. The textual composition of the Dominican breviaries, both in terms of the selection of readings, chants, and prayers and of their precise wording, is so uniform that all Dominican breviaries may be treated as a single unit of comparison.[91] The diocesan breviaries included in the comparison were the printed editions of Riga, Lübeck, and Kamień Pomorski, together with a manuscript copy from Schwerin in the Holy Roman Empire, as well as a manuscript breviary of Lund and a printed edition of Uppsala from

89 Breviarium Fratrum Praedicatorum. Venice: Franciscus Renner de Heilbronn and Petrus de Bartua, [about 1477–78] (ISTC ib01139000, GW 5215); Breviarium Fratrum Praedicatorum. Nuremberg: Anton Koberger, 1485 (ISTC ib01141300, GW 5219); Breviarium Fratrum Praedicatorum. Basel: Jacobus Wolff de Pforzheim, for Jacobus de Kirchen, 1492 (ISTC ib01142000, GW 5224).

90 Aarau, Aargauer Kantonsbibliothek, MsMurQ 5 (c. 1300); Karlsruhe, Badische Landesbibliothek, Cod. St. Peter perg. 32 (2nd quarter of the 15th century, provenance: Strasbourg, St. Nicolaus in Undis); Helsinki, Kansalliskirjasto, C.IV.10 (mid-13th century, England?).

91 For general characteristics of the Dominican liturgy, see W. R. Bonniwell. A history of the Dominican liturgy. J. F. Wagner, New York, 1944; for the Dominican office of the saints, A.-É. Urfels-Capot. Le sanctoral du lectionnaire de l'office. – Aux origines de la liturgie dominicaine: le manuscrit Santa Sabina XIV L 1. Sous la direction de L. E. Boyle et al. Institut de Recherche et d'Histoire des Textes (IRHT), Aubervilliers, 2004, 319–353.

Scandinavia.[92] These diocesan breviaries display substantial differences from one another and must therefore be considered individually.

The titles in the table represent both full texts and headings in breviaries; the liturgical cross-references are given in brackets. The titles are normally standardised, except for the recto of Cm 16 where reconstructed transcriptions of the headings are provided, where possible; where the titles have not survived (the readings of the matins of the vigil), the text passage has been reconstructed from surviving sections. For Cm 16 and the Dominican breviary, all items are listed; for other breviaries, all explicitly cited titles are likewise included, but cross-references to other liturgical material appearing in only one or two breviaries are omitted. Medieval terminology, which appears in the source material, is employed for liturgical occasions and genres.

92 Breviarium Rigense. [Paris], s.n., voor Willem Corver (Amsterdam), 1513; Breviarium Lubicense (Lübeck). Nuremberg: Georg Stuchs, 18 Aug. 1490 (ISTC ib01164300, GW 5375); Breviarium Caminense. [Bamberg/Nuremberg?: Printer of the 'Breviarium Caminense', about 1480?] (ISTC ib01150550, GW 5298); Breviarium Suerinense, Pars hiemalis. Kirchenbibliothek Barth, 4° D 54 (Diözese Schwerin, about 1430–1440); Breviarium Lundense. Lund, Universitetsbiblioteket, Medeltidshandskrift 34 (Lund, c. 1500); Breviarium Upsalense. Stockholm: Johannes and Anna Fabri (Smedh), 30 Sept. 1496 (ISTC ib01187000, GW 5499).

Table 2. The structure of the office of St Andrew from the matins of the vigil to the second lesson of matins of his feast day in the breviary fragment in Cm 16 and a selection of late medieval breviaries. (Ant. – Antiphona, Ev. – Evangelium, Cap. – Capitulum, Coll. – Collecta, Hymn. – Hymnus, Invit. – Invitatorium, Lect. – Lectio, Magn. – Magnificat, Or. – Oratio, Ps. – Psalmus, Resp. – Responsorium, Vers. – Vers(icul)us; the equals sign indicates agreement with the entry in the previous column.)

	Riga	Lübeck	Kamień Pomorski	Tallinn, Cm 16	Schwerin	Lund	Uppsala	Dominican
Vigilia S. Andreae								
Ev.	Stabat Iohannes et ex discipulis eius duo (Io 1:35)	=	=	=	=	=	=	=
Lect. I (Bedae)	Tanta ac talis ... Ecce enim narrante Evangelio ... stabat Iohannes et ex discipulis (l. 4)	Ecce narrante Evangelio stabat Iohannes et ex discipulis (l. 4)	Stabat Iohannes et respexit Iesum (l. 23)	... [l. 4/7]	Narrante evangelio ... stabat Iohannes et ex discipulis (l. 4)	Huius verba lectionis ... Ecce narrante Evangelio ... stabat Iohannes et ex discipulis (l. 4)	Stabat Iohannes et respexit Iesum (l. 23)	Stabat Iohannes et respexit Iesum (l. 23)
Lect. II	Statimque cognovimus (l. 7)	Ceteraque huiusmodi (l. 21)	Stabant cum illo et discipuli (l. 27)	... [l. 27]	Quid ... Stabat Iohannes et respexit Iesum (l. 23)	Sed non est putandum (l. 16)	Stabant cum illo et discipuli (l. 27)	Stabant cum illo et discipuli (l. 27)
Lect. III	Hoc est testimonium (l. 14)	Ambulatio Iesu dispensationem (l. 30)	Ecce agnus dei. Cognovimus quare (l. 7)	Sequitur ... [l. 45]	Conversus autem Iesus et videns (l. 54)	Stabat etiam quia illam virtutum (l. 26)	Ambulatio Iesu dispensationem (l. 30)	Ecce agnus dei (l. 7)

Continued on the next page

Table 2. Continued

	Riga	Lübeck	Kamień Pomorski	Tallinn, Cm 16	Schwerin	Lund	Uppsala	Dominican
Nona								
Hymn.	Benedictus Deus	(Hymnus ferialis)	-	-	(Hymnus ferialis)	-	-	-
Ant.	Qui persequebantur	Qui persequebantur	(Ant. Absterget Deus; Vers. Iustorum animae)	-	Andreas Christi famulus	-	-	-
Cap.	Scimus quoniam diligentibus	Benedictio Domini super caput iusti	Benedictio Domini super caput iusti	-	Benedictio Domini super caput iusti	-	Benedictus Deus et pater Domini Iesu Christi	-
Resp.	Nimis honorati	(Require in communi)	Nimis honorati	-	Nimis honorati sunt	-	Nimis honorati	-
Vers.	Annuntiaverunt opera	Nimis honorati	Dedisti hereditatem	-	-	-	Dilexit Andream Dominus	-
Coll.	Quaesumus omnipotens Deus	Quaesumus omnipotens Deus	Quaesumus omnipotens Deus	-	Quaesumus omnipotens Deus	-	Quaesumus omnipotens Deus	-

Continued on the next page

Table 2. Continued

	Riga	Lübeck	Kamień Pomorski	Tallinn, Cm 16	Schwerin	Lund	Uppsala	Dominican
Vesperae								
Ant.	Ambulans Iesus iuxta mare Galilaeae	Ambulans Iesus iuxta mare Galilaeae	Ambulans Iesus iuxta mare Galilaeae	Ambulan[s Iesus iuxta] m[are] g[ali]ee	Ambulans Iesus iuxta mare Galilaeae	Ambulans Iesus; Unus ex duobus; Videns Andreas crucem; Beatus Andreas orabat; Dilexit Andream Dominus	Ambulans Iesus; Unus ex duobus; Videns Andreas crucem; Beatus Andreas orabat; Dilexit Andream Dominus	Unus ex duobus qui secuti sunt
Cap.	-	Dominus mihi astitit et confortavit me	(ut in nona)	B[enedictio Dni super caput ...] (?)	Dominus mihi astitit et confortavit me	Fratres Corde creditur ad iustitiam	Fratres Corde creditur ad iustitiam	-
Resp.	Homo Dei	Homo Dei	Homo Dei	[Homo] dei (?)	Homo Dei	Homo Dei	Homo Dei	Homo Dei
Hymn.	Exultet caelum	Exultet caelum	Exultet caelum	[Exult]et (?)	Exultet caelum	Andrea pie	Andrea pie	-
Vers.	Dilexit Andream Dominus	In omnem terram	In omnem terram	[In] o[mnem] terram (?)	In omnem terram	In omnem terram	Dilexit Andream Dominus	-
Magn. Ant.	Mox ut vocem Domini	=	=	=	=	=	=	Ambulans Iesus iuxta mare Galilaeae
Coll./Or.*	-	(ut in nona)	(ut in nona)	Quaesumus omnipotens Deus	Quaesumus omnipotens Deus	Quaesumus omnipotens Deus	Quaesumus omnipotens Deus	Quaesumus omnipotens Deus
Si transferatur Or.							Maiestatem tuam Domine	Maiestatem tuam Domine

Continued on the next page

Table 2. Continued

	Riga	Lübeck	Kamień Pomorski	Tallinn, Cm 16	Schwerin	Lund	Uppsala	Dominican
(Festum S. Andreae, Matutinum)								
Invit.	Adoremus victoriosissimum regem	=	=	=	=	=	=	=
Ps.	Venite	Venite	=	=	(de apostolis)	(de apostolis)	=	=
Hymn.	Aeterna Christi munera	-	-	-	-	-	Aeterna Christi munera	-
Primum nocturnum								
Ant.	Christus me misit ad istam provinciam	Vidit Dominus Petrum et Andream	Vidit Dominus Petrum et Andream	Vidit Dominus Petrum et Andream	Vidit Dominus Petrum et Andream	Andreas apostolus dixit ad Aegeam	Andreas apostolus dixit ad Aegeam	Vidit Dominus Petrum et Andream
Ps.	(de apostolis)	Caeli enarrant	Caeli enarrant	Caeli enarrant	(de apostolis)	(de apostolis)	(de apostolis)	Caeli enarrant
Ant.	Ego si crucis patibulum	Venite post me	Venite post me	Venite post me	Venite post me	Ego crucis Christi servus	Ego crucis Christi servus	Venite post me
Ps.	(de apostolis)	Benedicam	Benedicam	Benedicam	(de apostolis)	(de apostolis)	(de apostolis)	Benedicam

Continued on the next page

Table 2. Continued

	Riga	Lübeck	Kamień Pomorski	Tallinn, Cm 16	Schwerin	Lund	Uppsala	Dominican
Ant.	Andreas apostolus Domini magnum operatus	Relictis retibus secuti sunt Dominum	Relictis retibus secuti sunt Dominum	Relictis retibus secuti sunt Dominum	Relictis retibus secuti sunt Dominum	Cumque carnifices apostolorum	Cumque carnifices apostolorum	Relictis retibus secuti sunt Dominum
Ps.	(de apostolis)	Eructavit	Eructavit	Eructavit	(de apostolis)	(de apostolis)	(de apostolis)	Eructavit
Vers.	In omnem terram	In omnem terram	In omnem terram	In omnem terram	In omnem terram	(de apostolis)	(de apostolis)	In omnem terram
Lect. I	Proconsul Aegeas Patras civitatem ingressus	Proconsul Aegeas Patras civitatem ingressus	Cum doceret et praedicaret verbum Dei	Proconsul Aegeas Patras civitatem ingressus	Proconsul Aegeas Patras civitatem ingressus	Proconsul Aegeas Patras civitatem ingressus	Proconsul Aegeas Patras civitatem ingressus	Ex gestis. Beatus Andreas post passionem Domini
Resp.	Dum deambularet Dominus iuxta mare	=	=	=	=	Cum perambularet Dominus iuxta mare	Dum deambularet Dominus iuxta mare	Cum perambularet Dominus iuxta mare
Vers.	Erant enim piscatores et ait	=	=	=	=	=	=	=
Lect. II	Cui Aegeas dixit Tu es Andreas qui destruis	Crastina vero die de carcere eductus	Andreas respondit Ego sum	Cui Aegeas dixit Tu es Andreas et destruis	Andreas respondit Romani principes	Cui Aegeas dixit Tu es Andreas qui destruis	Cui Aegeas dixit Tu es Andreas qui destruis	Cum iuvenis quidam Christianus

*Diocesan breviaries and Cm 16 use *Coll(ecta)*, while the Dominican breviary has *Or(atio)*.

„EX QUO“ SÕNASTIKU TALLINNA KÄSIKIRI: TEKSTITRADITSIOON JA AJALOOLINE TAUST

Tiina Kala ja Kaspar Kolk

Artikkel käsitleb Tallinna Linnaarhiivis säilitatava käsikirjalise ladina-alamsaksa sõnastiku „Ex quo“ ülesehitust, sisu, koostamisprotsessi ning ajaloolist tausta ja funktsiooni hiliskeskaegse Eesti intellektuaalses ja hariduslikus olustikus. Umbes 13 000 märksõna sisaldavale sõnastikule järgneb pealkirja „Termini iuristarum“ all tuntud värsivormis lühisõnastik, mis lõpeb kolofooniga. Kõik need tekstiosad on kirja pannud üks ja seesama kirjutaja, lisaks leidub käsikirja lõpulehekülgel lühimärkusi (sh omanikumärge) veel kolmelt kirjutajalt. Kolofooni järgi valmis käsikiri 1454. aastal Türil, omanikumärke ja köitetunnuste põhjal otsustades kuulus raamat Tallinna dominiiklaste konvendile. Varasematest uurijatest on käsikirja põgusalt käsitlenud Hermann von Bruiningk (1904), selle köitetunnuseid ja ajaloolist tausta kirjeldanud Hans Treumann (1968), keelelistel ja tekstoloogilistel aspektidel peatunud Pekka Katara (1946, 1948), Klaus Grubmüller (1967, 1988) ja Bernhard Schnell (1988).

„Ex quo“ sõnastiku algne otstarve oli hõlbustada piibli ja teiste ladina tekstide mõistmist. Sõnastiku eksemplarid olid hiliskeskajal saksa keeleruumis laialt levinud, neis on esindatud paljud keelevariandid nii ülem- kui alamsaksa alalt. Säilinud on ligi 300 käsikirja (osa neist mittetäielikud) aastatest umbes 1410 kuni 1502 ja enam kui 40 trükiväljaannet aastatest 1467–1505. Kuigi nende koostamise aluspõhimõtted ja tihti ka märkimisväärne osa tekstist kattuvad, esineb käsikirjade või käsikirjarühmade vahel ka erinevusi, nagu üksikute märksõnade või tervete märksõnaplokkide lisamine või ärajätmine, lahknevused märksõnaselgituste sõnastuses ja keeles (ladina või saksa), samuti sissejuhatuse sõnastuses. Schnell ja Grubmüller on välja toonud neli peamist „Ex quo“ redaktsiooni, jagades need omakorda arvukateks alamredaktsioonideks. Selles, millisesse neist paigutada Tallinna käsikiri, ei ole olnud täit üksmeelt. Suuresti seletab seda tõik, et 1960.–1980. aastail, mil saksa teadlased „Ex quo“ sõnastikuga kõige intensiivsemalt tegelesid, ei olnud Tallinna käsikirja terviktekst neile kättesaadav.

Schnelli ja Grubmülleri väitel kuulub Tallinna käsikiri 1410. aasta paiku koostatud ja umbes 12 700 märksõna sisaldavast algredaktsioonist (S) välja kasvanud redaktsiooni (P), mis sisaldab arvukalt täiendavaid märksõnu. Selle redaktsiooni leviku kõrgaeg hakkas 15. sajandi keskpaiku läbi saama; niisiis esindab Tallinna käsikiri „Ex quo“ sõnastiku selleks

ajaks juba vanemat versiooni. Sellele osutavad ka vormistuslikud eripärad, mis on pigem iseloomulikud 15. sajandi algupoole käsikirjadele. Tallinna käsikirja võrdlus redaktsiooni P alamredaktsioonides (X, X1, X3 ja X4) täheldatud kõrvalekalletega näitab, et Tallinna tekst ei ole lähedane mitte alamredaktsioonidele X3 ja X4, nagu seni väidetud, vaid pigem alamredaktsioonile X1, kuigi sedagi oma erijoontega. Käsikirju võivad üksteisest eristada unikaalsed, näiteks kirjutaja isiklikest eelistustest tulenevad tunnused, mistõttu iga üksiku käsikirja paigutamine ühte või teise uurijate piiritletud alamredaktsiooni ei tarvitse alati olla võimalik ega otstarbekas.

Lühike „Termini iuristarum" esineb käsikirjades tavaliselt lisandusena mõnele teisele leksikograafilisele tekstile, kõige sagedamini (üheteistkümnes säilinud käsikirjas) just „Ex quole". Sellest võiks oletada, et tekstide koosesinemisest kujunes omaette käsikirjaline traditsioon. Analüüs näitab, et nii see siiski ei olnud: eri käsikirjades paiknevad koos kõige erinevamad nii „Termini iuristarumi" kui „Ex quo" tekstivariandid, sageli teises järjekorras ja eraldatuna teistest tekstidest. Seega on tekstide valik käsikirjades tulenenud üldjuhul pigem iga üksiku kirjutaja otsustest. Tallinna käsikiri tugineb tõenäoliselt siiski juba sellele eeskujuks olnud, ent praeguseks tundmatu käsikirja kompositsioonile: mõlemad tekstid ja lisaks kolofoon on vormistuslikult täiesti kokku sulanud, millest võib järeldada, et ümberkirjutaja ei pruukinud neid eri tekstidena tajudagi.

Tallinna käsikirja võimaliku kirjutaja isik on tekitanud eri arvamusi. Teksti sissejuhatuse lõppu lisatud nime *her Marten* (härra Marten) pidas Bruiningk kirjutaja nimeks ja oletas, et tegu oli Türi kogudusepreestriga. Treumann püüdis kirjutajat samastada mõne Marteni või Merteni nime kandnud Tallinna dominiiklaste konvendi vennaga, ent 15. sajandi viimase kolmandiku allikad, kust sellist nime võib leida, ei osuta kuidagi Tallinna konvendi Merteni(te) seosele „Ex quo" käsikirja või Türiga. Paar 1520. aastatest pärit teadet viitab küll Tallinna dominiiklaste sidemetele Järvamaaga, kuid milles need täpsemalt seisnesid ja kas nende põhjal võiks eeldada konvendi või mõne selle venna kokkupuuteid Türiga juba 1450. aastatel, ei selgu. Lõpuks ei pruugi *her Marten* viidatagi üldse kirjutajale. Käsikirja köitmisel kaanelehena kasutatud makuleeritud breviaari fragment, mille on tõenäoliselt samuti kirjutanud „Ex quo" kirjutaja, esindab piiskopkondlikku (oletatavasti Tallinna piiskopkonna), mitte dominiiklaste liturgiat. Sellest võib järelduda, et kirjutaja oli breviaari kirjutamise ajal seotud ilmikkoguduse teenimisega.

Arvestades dominiiklaste suurt tähelepanu haridusele, sealhulgas piibliõpingutele, ning hästi välja arendatud koolisüsteemi, samuti asjaolu,

et Tallinna konvendist säilinud kümmekonna käsikirjalise raamatu ja mõnekümne käsikirjafragmendi seas leidub kolm grammatikaõpingutele viitavat tekstikatket, sobitub „Ex quo“ sõnastik siiski hästi hiliskeskaegse Tallinna konvendi eeldatavate vajadustega.

Acta Historica Tallinnensia 32/2 https://doi.org/10.3176/hist.2026.2.07

"We Are from the North": Livonica in the Vatican Archives and the Special Case of Aleksander Kurtna[1]

Kristjan Toomaspoeg

Università del Salento, Department of Human and Social Sciences, viale San Nicola, 73100 Lecce, Italy; kristjan.toomaspoeg@unisalento.it

Abstract. This paper examines the historical development of research concerning medieval and early modern Livonica within the Vatican Archives from the nineteenth century to the present. This survey is inextricably linked to the singular case of Aleksander Kurtna, one of the most brilliant Estonian intellectuals of the twentieth century. While Kurtna spent years conducting high-level research in the Vatican, he was simultaneously one of the most prominent intelligence agents in the Eternal City, blending archival scholarship with clandestine espionage.

Keywords: medieval Livonia, medieval history, Vatican Archives, Baltic historiography, espionage

This article focuses on two distinct yet closely intertwined themes that hold considerable significance for the cultural history of the Baltic States. First, it offers an overview of the state of research on medieval and early modern Livonian sources preserved in the Vatican Apostolic Archive (formerly the Vatican Secret Archive). Second, it demonstrates that this enquiry is inseparable from the figure of Aleksander Kurtna,

1 "We are from the North", the coded opening of a message sent by Soviet secret services to their agents in Rome (see below), has already been used as a chapter title in D. Alvarez and R. A. Graham. Nothing Sacred: Nazi Espionage Against the Vatican, 1939–1945. Frank Cass, London, 1997, 114–139.

© 2026 Author. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International License CC BY 4.0 (http://creativecommons.org/licenses/by/4.0).

who spent more time working in the Vatican Archive than any other Baltic scholar. Kurtna's case is exceptional: his intricate biographical circumstances deeply shaped his archival endeavours, meaning that his scholarly contribution can only be fully assessed through a reconstruction of his life. Moreover, existing historical and popular literature lacks a sufficiently thorough and rigorous biographical account.

Accordingly, this study adopts a dual methodology, complementing a historiographical review with a biographical analysis. The latter – unfolding as a true Second World War espionage narrative grounded in contemporary historical methodology – allows for multiple interpretations and leaves several questions unresolved. While the article may consequently seem dual-natured or even fragmented, both thematic strands ultimately converge on a single issue: the persona of the historian within the archive who can never disentangle private existence from professional life.

INTRODUCTION: FROM TURGENEV TO KARTTUNEN

Established in its modern form in 1612 by Pope Paul V and further organised under Urban VIII, the Archivio Apostolico Vaticano houses over 600 archival fonds across 85 kilometres of shelving, containing millions of documents.[2] The Vatican Archives serve as one of the primary repositories for the history of medieval Livonia – encompassing present-day Estonia and Latvia – from the twelfth to the sixteenth centuries. While a significant body of these documents is already known to Estonian and Latvian scholars, they represent only a fraction of the relevant material preserved in the archives.

Investigating this field is a formidable task, primarily due to the vast scale of the Vatican's holdings. A further challenge stems from the fragmented nature of Estonian and Latvian archival research in the Vatican, which remains marked by significant gaps and methodological shortcomings.[3]

2 See L. E. Boyle, O. P. A Survey of the Vatican Archives and of Its Medieval Holdings. Pontifical Institute of Mediaeval Studies, Toronto, 1972 (revised edition Pontifical Institute of Mediaeval Studies, Toronto, 2001); F. X. Blouin Jr. Vatican Archives: An Inventory and Guide to Historical Documents of the Holy See. Oxford University Press, New York, 1998. These works provide the essential foundational framework for any researcher approaching the central archives of the Holy See.

3 See K. Toomaspoeg. Ühest Brugge kirstust, Riia peapiiskoppidest Avignonis ja arhiiviuuringutest Vatikanis. – Vana Tallinn, 2017, 28 (32), 202–224. This article discusses several of the methodological issues and specific archival challenges addressed in the present study.

Systematic archival research in the Vatican is generally traced back to 1881, when Pope Leo XIII opened the archives – then known as the Archivio Segreto Vaticano – to all qualified researchers.[4] This pivotal event coincided with the establishment of several foreign historical research institutes in Rome, including the Austrian Historical Institute in 1881 and the Prussian Historical Institute in 1888.[5] Prior to this opening, official access was not standard practice, though it remained possible under exceptional circumstances.

Regarding Livonia, two scholars gained access to these collections prior to 1881: Alexander Turgenev in the 1840s for his *Historiae Russiae Monumenta*, a work of remarkable quality for its time[6], and Augustin Theiner for his *Vetera Monumenta Poloniae et Lithuaniae* in the 1860s.[7] Turgenev's research in the Vatican covered the entire territory of the then-Russian Empire, with a particular focus on Livonia. His success was facilitated by his status as a prominent Russian aristocrat with close ties to the Holy See, as well as by his personal reputation for erudition. Theiner, a native of Breslau (Wrocław), conducted his research from a position of authority, having served as Prefect of the Vatican Archives since 1855.

Hermann Hildebrand, the city archivist of Riga, was the first Baltic historian to enter the Vatican Archives following the granting of free access in 1881; he conducted his research there from October 1885 to April 1886. His publications – *Livonica, vornämlich aus dem 13. Jahrhundert, im Vatikanischen Archiv*, as well as volumes 7, 8, and 9 of the *Liv-, Est- und Kurländisches Urkundenbuch* (LEKUB)[8] – contain several sources from the Vatican. Furthermore, he was the first Livonian scholar to consult not only the papal registers but also other fonds, such as the *Dataria* – a collection that remains relatively obscure even today.[9] As Hildebrand noted, despite being warmly welcomed, researchers still faced significant practical hurdles, such as limited opening hours and the absence of archival inventories; consequently, scholars could

4 See, for example, K. J. Walsh. The Opening of the Vatican Archives (1880–1881) and Irish Historical Research. – Archivium Hibernicum, 1981, 35, 34–43.

5 Das Deutsche Historische Institut in Rom. 1888–1988. Bearb. von R. Elze, A. Esch. Niemeyer, Tübingen, 1990.

6 A. Turgenew. Historiae Russiae Monumenta. Eduard Pratz, St. Petersburg, 1841–1848 (reprint Herculaneum Press, West Vancouver, 1973).

7 A. Theiner. Vetera Monumenta Poloniae et Lithuaniae 1–4. Typis Vaticanis, Roma, 1860–1864 (reprint Zeller, Osnabrück, 1969).

8 H. Hildebrand. Livonica, vornämlich aus dem 13. Jahrhundert, im Vatikanischen Archiv. Deubner, Riga, 1887; Liv-, Est- und Kurländisches Urkundenbuch (LEKUB). Bd. 7. 1423 Mai–1429 Mai. Bearb. von H. Hildebrand. Deubner, Riga, Moskau, 1881; LEKUB, Bd. 8, 1429 Mai–1435. Bearb. von H. Hildebrand. Deubner, Riga, Moskau, 1884; LEKUB, Bd. 9, 1436–1443. Bearb. von H. Hildebrand. Deubner, Riga, Moskau, 1889.

9 H. Hildebrand. Livonica, 9, see an example in ibid., 70, no. 53.

only request items they had already identified through references in existing literature.[10] Although his objective was to systematically record and publish a comprehensive collection of available sources, his work remained a selection of documents, proving more exhaustive for the fifteenth century than for earlier periods.

Research on Livonica in the Vatican has frequently been integrated into broader international projects. A notable example is the Finnish initiative conducted between 1907 and 1915 and directed by Henry Biaudet,[11] during which Liisi Karttunen transcribed a series of documents concerning Livonia from the papal registers and *miscellanea* fonds. This work took place within the framework of the so-called Finnish Historical Expedition to Rome, for which Karttunen served as secretary; during this period, she also prepared her doctoral thesis on the sixteenth-century papal diplomat Antonio Possevino.[12] The National Archives of Estonia (RA) in Tartu preserve five dossiers containing Karttunen's transcriptions – primarily from the papal registers – covering documents dated from 1219 to 1571.[13]

This archival series has a curious history: as Karttunen noted in her own hand, the transcriptions were prepared in April 1919 in Rome for the "Estonian Embassy in Italy", and the dossiers themselves consist of paper from the *Cartoleria Ricci* in Rome.[14] However, since an official Estonian Embassy did not yet exist in Rome in 1919, Karttunen's work should be linked to the activities of the Estonian diplomatic delegation, particularly to Kaarel Robert Pusta, who visited Italy several times between 1918 and 1919.[15] Having ceased her own Vatican research after 1915 to serve as a

10 H. Hildebrand. Livonica, 6.

11 E. Garritzen. Lähteiden lumoamat: Henry Biaudet, Liisi Karttunen ja suomalainen historiantutkimus Roomassa 1900-luvun alussa. Suomalaisen Kirjallisuuden Seura, Helsinki, 2011.

12 L. Karttunen. Antonio Possevino. Un diplomate pontifical au XVIe siècle. Pache-Varidel & Bron, Lausanne, 1908. See K. Salonen. The Vatican Archives and the Nordic Countries. – Friedensnobelpreis und historische Grundlagenforschung: Ludwig Quidde und die Erschließung der kurialen Registerüberlieferung. Bearb. v. M. Matheus. De Gruyter, Berlin, 2012, 552–561, here 554–555; F. Zuliani. L'expédition historique finlandaise à Rome (1909–1915) e la storiografia italiana novecentesca. – Bridges to Scandinavia. Ed. by A. Meregalli, C. Storskog. Ledizioni, Milan, 2016, 157–169.

13 Rahvusarhiiv (National Archives of Estonia, RA), EAA.854.2.560, Urkunden aus dem Vatikanischen Archiv aus Jahren 1219–1389. Bestätigte Kopien und Dubletten. Lateinisch. Bd. I.; RA, EAA.854.2.561, Urkunden aus dem Vatikanischen Archiv aus Jahren 1220–1478. Bestätigte Kopien und Dubletten. Lateinisch. Bd. II.; RA, EAA.854.2.562, Urkunden aus dem Vatikanischen Archiv aus Jahren 1228–1324. Bestätigte Kopien und Dubletten. Lateinisch. Bd. III.; RA, EAA.854.2.563, Urkunden aus dem Vatikanischen Archiv aus Jahren 1404–1533. Bestätigte Kopien und Dubletten. Lateinisch. Bd. IV; RA, EAA.854.2.564, Urkunden aus dem Vatikanischen Archiv aus Jahren 1432–1571. Bestätigte Kopien und Dubletten. Lateinisch. Bd. V. All these materials date from 1919.

14 See RA, EAA.854.2.560, after every document there is the date of transcription: 11 April 1919.

15 T. Mulla, P. Pillak, J. Trei, H. Vilbiks. Eesti välisteenistus. Biograafiline leksikon 1918–1991. Välisministeerium, Tallinn, 2006, 94–95.

diplomat, Karttunen probably compiled these transcriptions from her previous work. These copies were later transferred to the State Archives and inserted – somewhat illogically – among the records of the Estonian Knighthood (*Ritterschaft*), the territorial self-government of the German nobility abolished in 1920. A note in the first dossier indicates a revision dated 29 February 1924. While some of Karttunen's papers had been previously published, most remained unedited; nonetheless, her transcriptions are not critical editions, as they consist of rapid copies of papal letters without diplomatic or palaeographic apparatus.

Shortly before Karttunen, between 1904 and 1905, the Prussian historian Arthur Motzki – while examining sources related to the Teutonic Order during the Avignon Papacy – copied and subsequently published a valuable series of acts concerning Livonia from the registers of papal supplications (*Registra Supplicationum*).[16] These supplications consist primarily of requests for the concession of church benefices; additionally, petitioners sought the confirmation of indulgences, dispensations from marriage impediments, and occasionally absolution from various offences. As such, they possess immense historical value. Although the collection comprises more than 7,300 volumes, Motzki consulted only the first 43, covering the period from 11 October 1342 to 9 May 1366, from which he collected 152 documents.

The young Baltic States successfully combined classical academic tradition with the evolving demands of medieval studies, largely thanks to figures such as Leonid Arbusow Jr. (1882–1951). The first state-organised research project in the Vatican took place in the late 1920s and early 1930s, when Arbusow Jr. conducted four expeditions to Rome supported by the Latvian government (specifically from November 1926 to March 1927, the first half of 1928, the summer of 1929, and from November 1930 to February 1931). His third expedition was carried out alongside the historian and nationalist politician Augusts Tentelis. While the results of Arbusow's work were extensively published,[17] many of his unedited notes and transcriptions are preserved in the Latvian State Historical

16 A. Motzki. Livonica aus den Supplikenregistern von Avignon (1342 Okt. 11–1366 Mai 9). – Mitteilungen aus der Livländischen Geschichte, 1921, 21, 2, 101–172.

17 L. Arbusow. Römischer Arbeitsbericht. I. Mit einem Anhang von Papstregesten des 13. und 14. Jahrhunderts und von Urkunden. – Latvijas Universitātes Raksti. Acta Universitatis Latviensis, 1928, 17, 285–423; L. Arbusow. II. Römischer Arbeitsbericht. Nebst Papstregesten und -urkunden des 14. Jahrhunderts. – Latvijas Universitātes Raksti. Acta Universitatis Latviensis, 1929, 20, 475–657; L. Arbusow. III. Römischer Arbeitsbericht. – Latvijas Universitātes Raksti. Acta Universitatis Latviensis, Filoloģijas un Filosofijas Fakultātes serija 1929–1931, 1, 3, 65–160; L. Arbusow. IV. Römischer Arbeitsbericht. – Latvijas Universitātes Raksti. Acta Universitatis Latviensis, Filoloģijas un Filosofijas Fakultātes serija, 1931–1933, 1, 4, 279–398.

Archives. Furthermore, his collection of nearly 1,000 photographs of Vatican documents – long believed to have been lost during the Second World War – was recently rediscovered within the same archives.[18]

Arbusow's work remains the most scientifically significant achievement in this field to date, despite the limited time at his disposal and the necessity of focusing on a restricted selection of sources. He provided an overview of the two major series of papal registers – the *Registra Vaticana* and *Registra Avenionensia* – for the period between 1198 and ca. 1378. Additionally, he consulted several archival units from the ancient Archive of Castel Sant'Angelo (*Archivum Arcis*), the original core of the Vatican Archives, along with volumes from the *Collectoriae* of the Apostolic Chamber, the Nunciature to Poland, and the supplication records from 1366 until the reign of Antipope Clement VII (1378–1394). His research focused primarily on the fourteenth century, the least known and researched portion of the Vatican's *Livonica* collections, and he dedicated significant attention to specific sources, such as the inventory of the belongings of the Archbishop of Riga, Friedrich von Pernstein († 1341).[19]

A JESUIT FROM RÄPINA: ALEKSANDER KURTNA IN THE VATICAN

The Estonian response to the Latvian project was spearheaded by Aleksander Kurtna (birth name Kurson), a flamboyant personality born in 1914 in Räpina, south-eastern Estonia, to an Estonian father and a Russian mother.[20] He hailed from roughly the same region as Georges Simenon's fictional *Peter the Lett* (born in Pskov), a crossroads of languages, cultures, and religions. Initially, Kurtna intended to become an Orthodox priest; however, in 1935, he came into contact with the Catholic Church, which was then pursuing an ambitious project to re-Catholicise Eastern Europe and bring Eastern Christians

18 See K. Neitmann. Das wissenschaftliche Lebenswerk Leonid Arbusows. Themen und Methoden seiner Forschungen zur Geschichte Livlands. – Leonid Arbusow (1882–1951) und die Erforschung des mittelalterlichen Livland. Hrsg. von I. Misāns, K. Neitmann. Böhlau, Köln, Weimar, Wien, 2014, 19–78.

19 L. Arbusow. Römischer Arbeitsbericht. I, 310. See also K. Toomaspoeg. Ühest Brugge kirstust, 206–211.

20 See his short biography and a description of his work in the Vatican based on Estonian sources in J. Kivimäe. Aleksander Kurtna ja Eesti Teaduste Akadeemia. – 50 aastat Eesti Teaduste Akadeemiat. Toim. A. Köörna. (Teaduslugu ja nüüdisaeg, 8.) Eesti Teaduste Akadeemia, Tallinn, 1989, 187–190.

into communion with Rome, primarily through the Jesuit Order and the Vatican Dicastery for the Eastern Churches. After converting to Catholicism and receiving foundational training in Poland, he travelled to Rome in 1936 as a seminarian at the Jesuit *Collegium Russicum*, founded by Pius XI in 1929 near the Basilica of Santa Maria Maggiore.

Our knowledge of Kurtna's life in Rome relies heavily on his own accounts: a fictionalised and often unreliable short autobiography published in Estonian periodicals in 1965 and 1966,[21] and a conference paper from 1966.[22] These may be supplemented by his substantial dossier compiled by the British Secret Intelligence Service (MI6) in 1944.[23] However, a comprehensive understanding of his personality and activities remains unattainable without access to the closed archives of the Italian and, more importantly, the Russian secret services.

According to Kurtna, his journey to Rome was motivated by a desire to study philosophy, and his entry into the Jesuit Order as a seminarian was merely a pretext. This claim, however, is dubious; Kurtna was, in fact, brought to Rome within a very specific ideological framework. His travel expenses were covered by the Catholic Bishop of Estonia, Eduard Profittlich, himself a Jesuit who had been trained for missionary work in the Soviet Union. Kurtna was one of two Estonian candidates sponsored by the bishop.[24] The *Russicum*, with its community of Russians and other émigrés from the former Russian Empire, served as a primary headquarters for the Vatican's project to re-evangelise the Soviet Union.

According to Kurtna, the primary purpose of his stay in Italy was to pursue studies at the Pontifical Gregorian University (*Università Gregoriana*) at Piazza della Pilotta. Lectures were conducted exclusively in Latin, a requirement that posed no difficulty for Kurtna, who was already fluent in at least five languages (Russian, Estonian, German, Polish, and Italian). Later, in November 1939, he gave an interview to the Estonian newspaper *Päewaleht*, providing a detailed description of the university's academic life.[25] He was not the only Estonian studying

21 A. Kurtna. Pronksvärava taga (Behind the Bronze Gate). – Kultuur ja Elu, 1965, 3–11, and Ühistöö. Rapla rajoonileht, 1966, 13, 15, 18, 20, 22, 25, 27, and 29 January; 1, 3, 5, 8, 9, 10, 12, 15, 17, 19, 22, and 24 February, 1, 3, 5, 8, 10, 12, and 15 March. In the present study, citations are taken from the edition published in Ühistöö, available online.

22 Tallinna Linnaarhiiv (Tallinn City Archives, TLA), TLA.R-27.1.937a.

23 The National Archives (Kew), KV 2/657.

24 M. Paas. Peapiiskop Eduard Profittlichi märterlusest. – Akadeemia, 2024, 10, 1779–1801, here 1782–1783, 1795. See also K. Laur. Tunnistaja. Hea Lugu, Tallinn, 2022 – a historical novel centred on Bishop Eduard Profittlich, though occasionally 'imaginative' with respect to Kurtna's figure.

25 Seal kus kõnekeeleks on ladina keel. – Päewaleht, 25 November 1939.

in this prestigious environment; however, the other young man sent to Rome by Bishop Profittlich, Johannes Raidal, abandoned his studies, and no further traces of him have been found.[26] The *Päewaleht* article reveals that by late 1939, Kurtna claimed to speak fifteen languages. It also noted that the Gregorian's student body comprised sixty different nationalities, all of whom benefited from a regime of extraterritoriality, as they were under the direct jurisdiction of the Vatican.

Kurtna was preparing a doctoral thesis in philosophy on the works of Lev Shestov, a Russian existentialist and critic of rationalism and positivism.[27] Shestov (who died in Paris in November 1938) was a Jewish thinker known for contrasting "Athens with Jerusalem" and for his meditations on Jesus' sleepless night in the Garden of Gethsemane.[28] While he was an ideal subject for study in this context, Shestov was also blacklisted by the Soviet authorities; consequently, Kurtna omitted any mention of this PhD topic from his later autobiography.[29] In any case, the thesis was never completed, and no records of it exist in the archives of the Gregorian University.

At this point, a further contradiction emerges between Kurtna's account and other sources. On 19 January 1939, the rector of the *Russicum*, Philippe de Régis, informed Bishop Profittlich that, in his opinion, Kurtna could no longer remain at the *Collegium*, as he was deemed unfit for the priesthood.[30] Consequently, some have supposed that he was expelled and returned to Estonia;[31] however, Kurtna himself claimed to have lived in Rome until November 1939, despite no longer residing at the *Russicum*.[32] In his memoirs, he somewhat glosses over this transition, which was not voluntary but rather the result of his "lack of vocation".

According to Kurtna, he received a small grant from the Estonian government in April 1939 to conduct research in the Vatican Archives.[33] On other occasions, however, he cited 16 May and 1 June 1939 as the start date of this funding.[34] He claimed the grant came as a complete surprise, resulting from the intervention of his father, Viktor Kurtna. Although

26 Raidal is presumed to have fallen in the war: J. Kivimäe. Minevik karussellil. – Keel ja Kirjandus, 2023, 66, 3, 341–344, here 342.
27 Päewaleht, 25 November 1939.
28 See, for example, A. Oppo. Lev Shestov: The Philosophy and Works of a Tragic Thinker. Academic Studies Press, Brookline (MA), 2020; M. Beaumont. Lev Shestov: Philosopher of the Sleepless Night. Bloomsbury Academic, London, 2021.
29 He did, however, refer to his thesis in a conversation with Jüri Kivimäe, see J. Kivimäe. Aleksander Kurtna, 187.
30 Archivum Romanum Societatis Iesu, Collegium Russicum, 1001: Ex Officio (1938–1939), quoted in M. Paas. Peapiiskop Eduard Profittlichi märterlusest, 1795.
31 M. Paas, *loc. cit*.
32 A. Kurtna. Pronksvärava taga, 3 February 1966.
33 A. Kurtna. Pronksvärava taga, 27 January 1966.
34 J. Kivimäe. Aleksander Kurtna, 187; Päewaleht, 22 November 1939.

a gardener by trade, Viktor was no ordinary labourer; he served as the superintendent of the park at Keila-Joa Manor, a summer retreat for many high-ranking government officials.[35] It is believed that Viktor spoke with his superior, Peeter Päts, the head of the State Parks Administration (*Riigiparkide Valitsus*) and brother of President Konstantin Päts.[36]

This account is intriguing, and although Kurtna reports it in his autobiography, it has been largely overlooked by historians. On the one hand, when he submitted a request in autumn for a more substantial and regular grant to study the Vatican Archives, he demonstrated a knowledge of the subject and provided detailed information on the materials to be examined; this suggests he must already have been working in the archives for some time. As he states, he obtained authorisation to enter the archives thanks to a recommendation from the papal nuncio in Estonia, Antonio Arata, who was an acquaintance of his.[37] On the other hand, if the initiative to research the Vatican Archives did not originate with Kurtna himself, how did it arise? It is also probable that after his expulsion from the *Russicum* – though he continued to associate with his fellow students and attend the Gregorian University as an unenrolled student (*fuori corso*) – Kurtna had already begun working as a library collaborator and language teacher, activities that later augmented his primary means of support.

As he himself noted, it was during this period that Kurtna became acquainted with Friedrich Bock, a staunch National Socialist who headed the German Historical Institute in Rome (*Deutsches Historisches Institut in Rom*). Bock was then managing the institute on behalf of its director, Edmund Ernst Stengel, who remained in Germany overseeing several research bodies for the Nazi government.[38] Another scholar Kurtna met in the Vatican was Gottfried Lang, who would later play a significant role in safeguarding Italian archives during the Second World War.[39]

35 P. Erelt. Ministrite suveidüll Keila-Joal. – Eesti Ekspress, 18 June 2011. Erelt appears to have consulted the records of the *Riigiparkide Valitsus* at the Estonian State Archives (RA, ERA.983). The State Parks Administration was created in April 1935 to administrate the parks of Kadriorg, Keila-Joa and Oru. Viktor Kurtna is likewise a figure of considerable interest, although less thoroughly documented in the sources. According to his military service record (RA, ERA.495.7.2302), he graduated from the Uspensk Agricultural School in 1907 and served as an officer in the First World War and in the Estonian War of Independence. According to the same source, he worked initially as a gardener, but in 1928 entered the state administration in Petseri. During his interrogation in 1944, Kurtna stated that his father was then an officer serving as an interpreter in the Soviet Army: The National Archives (Kew), KV 2/657, interrogation report, 3.

36 Kurtna himself stated as much: Pronksvärava taga, 27 January 1966.

37 A. Kurtna. Pronksvärava taga, 29 January 1966.

38 A. Hoffend. Zwischen Kultur-Achse und Kulturkampf. Die Beziehungen zwischen 'Drittem Reich' und faschistischem Italien in den Bereichen Medien, Kunst, Wissenschaft und Rassenfragen. Peter Lang, Frankfurt am Main, 1998, 270–274.

39 J. Klöckler. Verhinderter Archivalienraub in Italien. Theodor Mayer und die Abteilung

In his accounts of his archival work – though not elsewhere – Kurtna mentioned two other scholars who were undoubtedly relevant to his research: namely, the aforementioned Liisi Karttunen, now a member of the Finnish diplomatic corps, and the Jesuit Jean Chrétien Joseph Kleijntjens.[40]

Kurtna enrolled in the Vatican School of Palaeography, Diplomatics and Archive Administration, where he met its director, Giulio Battelli.[41] The school's curriculum was even more rigorous then than it is today; Kurtna supplemented these studies with unspecified courses at the German Historical Institute and later at the University of Rome (now *La Sapienza*), where he attended lectures by the Fascist philosopher Giovanni Gentile on the philosophy and methodology of history.[42]

As noted above, using his studies at the School and his growing familiarity with archival repertoires and indexes – where he encountered many of the same practical difficulties as Hermann Hildebrand before him – Kurtna submitted an official grant request. However, at this juncture, his autobiography begins to feature the series of contradictions, omissions, and exaggerations typical of his later persona as a spy. He claimed to have sent proposals not only to the Estonian government but also to the Academy of Sciences of the Soviet Union in Moscow.[43] Specifically, he allegedly offered to research documents concerning Peter the Great, building upon the earlier work of Yevgeny Shmurlo.[44] According to Kurtna's version of events, he even received a promise from Moscow that the Academy would begin financing his research in the second half of 1940. This version of events is highly implausible, as the Soviet coup d'état in Estonia in June 1940 was not a predictable event outside of Stalin's inner circle. In my view, the initial contact between Kurtna and the Soviet authorities must have occurred later; as will be discussed, he spent a significant period in Moscow between July and October 1940.

In any case, following a week of negotiations in Tallinn in November 1939, the Estonian state – already subject to increasing Soviet pressure, yet

"Archivschutz" bei der Militärverwaltung in Verona 1943–1945. – Quellen und Forschungen aus italienischen Archiven und Bibliotheken, 2006, 86, 491–537, here 513–518.

40 J. Kivimäe. Aleksander Kurtna, 188.

41 Battelli is considered a foundational figure in Italian palaeography and archival science; see, for example, G. Battelli. Lezioni di paleografia. Pontificia Scuola Vaticana di Paleografia e Diplomatica, Città del Vaticano, 1936.

42 On Gentile, the preeminent intellectual of Italian Fascism, see D. D. Roberts. Historicism and Fascism in Modern Italy. University of Toronto Press, Toronto, 2007.

43 A. Kurtna. Pronksvärava taga, 1 February 1966.

44 Е.Ф. Шмурло. *Петр Великий в оценке современников и потомства*. Сенат. тип., Санкт-Петербург, 1912.

still in existence – granted Kurtna a yearly stipend to research documents concerning Estonia in the Vatican Archives. Specifically issued by the Chancellery of the President of the Republic, this was a significant decision with scientific, cultural, and political weight, as reflected in contemporary press coverage.[45] The grant consisted of a substantial monthly sum of 1,500 – later increased to 2,000 – Italian lire (roughly equivalent to the same amount in euros today). Kurtna returned to Italy almost immediately, on 13 December 1939;[46] given the increasingly grim situation in Estonia, the move was probably strategic. From this point on, Kurtna's autobiography is filled with accounts of the *bella vita* in Rome, characterised by visits to restaurants, bars, and concerts. The months between late 1940 and June 1941 appear to have been the happiest period of his life. Despite his newfound financial security, the inexhaustible Kurtna continued to take on side work as a language teacher. It was through these lessons that he met Alma Hablitz, a young and talented soprano who had also received an Estonian government grant to study in Italy.[47] The couple began living together and were eventually married on 18 June 1942, at the Campidoglio in Rome.

Despite the distractions of his social life, Kurtna did indeed embark on an intensive period of research in the Vatican. By March 1940, he claimed to have recorded 600 documents; subsequently, in an unpublished 1966 conference paper, he asserted that during his tenure in Rome – which lasted until September 1944 – he had compiled a catalogue of approximately 7,000 sources.[48] The preliminary results of this research are preserved in the National Archives of Estonia in Tartu, which holds two reports from Kurtna and a folder containing the research project he proposed in March 1940 to the Chancellery of the President and the Estonian Academy of Sciences.[49] This documentation is highly relevant to our analysis and is complemented by the aforementioned 1966 conference paper, which provides a comprehensive overview of the Vatican Archives and their collections. In 1940, Kurtna highlighted several specific series which he had examined, chief among them the supplication

45 Päewaleht, 22 November 1939.
46 RA, EAA.661.1.489, 3.
47 E. Schütz. Alma Kurtna: unustamatu lauluõpetaja. Ilmamaa, Tartu, 2018; Eesti muusika biograafiline leksikon, I. Eesti Entsüklopeediakirjastus, Tallinn, 2007, 339. An intriguing report in Avvenire d'Italia, 14 February 1939, mistakenly states that the President of Latvia (!) had provided the grant for Alma Hablitz to perfect her art in Italy. Hablitz is likewise a figure of considerable interest; during the Second World War, she worked for the EIAR, Mussolini's broadcasting service.
48 T. Männik. Estica Vatikani arhiivis. Kokkuvõte A. Kurtna ettekandest. TLA.R-27.1.937a. This four-page summary is part of the Presentations in Archive Researchers' Conferences collection, dated 18 November 1966.
49 RA, EAA.661.1.489. See J. Kivimäe. Aleksander Kurtna.

registers, a collection then largely unknown in Baltic historiography (with the notable exceptions of the studies by Motzki and Arbusow Jr.). Following a detailed analysis and description of these registers, Kurtna drew attention to the *Ospizio dei Convertendi*. This remains a little-known archival collection containing records of individuals (primarily Protestants) who converted to Catholicism, including numerous cases of people from Livonia.[50]

In this manner, Kurtna chose to launch his research by focusing on two specific cases: a collection that had been only partially explored by Baltic historians – even adding corrections to Motzki's work – and an entirely unknown series. His long-term plan was eventually to expand his scope to the papal registers and the collections of the Apostolic Chamber (*Camera Apostolica*). An analysis of Kurtna's perspective on the archives, particularly through his 1966 conference paper, reveals that his approach was largely shaped by his own research interests, with significant attention devoted to medieval and early modern fonds. He gave particular prominence to the *Armaria* – which, according to Kurtna, consisted of eighty cabinets containing a vast number of documents accumulated haphazardly rather than through systematic selection – and to what he called the *Archivio Diplomatico*. In reality, the latter does not exist as an autonomous archival series; it probably refers to the collections now known as the *Archivum Arcis* and *Instrumenta Miscellanea,* which contain original documents on both parchment and paper.

Kurtna's research project outlined a systematic three-stage plan: the creation of a register or catalogue of sources, followed by their scholarly edition, and, ultimately, historical research on specific topics. The Estonian Academy of Sciences forwarded his proposal to the Director of the Estonian State Archives, Otto Liiv, for evaluation. Liiv, however, was a specialist in the Swedish period of Estonian history rather than the Middle Ages or the Vatican Archives. Nevertheless, his assessment of Kurtna's project – expressed in a letter dated 26 April 1940 – was entirely and even incomprehensibly negative.[51] In essence, Liiv deemed Kurtna unfit for the task, citing a perceived lack of knowledge regarding Estonian history. Despite having met Kurtna and exchanged correspondence with him since June 1939, Liiv remained unconvinced of his professional capacities. This dismissal may be interpreted as the

50 A. Raunio. L'Ospizio dei Convertendi, storie di conversioni miracolose e l'assistenza romana caratterizzata dallo spirito post-tridentino. – Proceedings of a Conference on Early Modern Rome, 1341–1667 *held on May 13–15, 2010 in Rome under the auspices of AACUPI*. Ed. by P. Prebys. Edisai, Ferrara, 2011, 510–517.

51 RA, EAA.661.1.489, 35–38.

reluctance of a bureaucrat to endorse an operative beyond his direct control, but political motives may have also played a role; Liiv later displayed certain sympathies for the Soviet authorities, whereas Kurtna – a former seminarian collaborating with clergy and Western scholars – was far from being a 'leftist' figure.[52]

KURTNA FINDS NEW SPONSORS

The Soviet coup d'état in Estonia on 21 June 1940 – which effectively ended the de facto existence of the Republic of Estonia – precipitated a major shift in Kurtna's life and research. This followed another pivotal moment: Italy's entry into the Second World War, on 10 June 1940, which led to restricted access to the Vatican Archives. However, Kurtna had secured a special arrangement with the young (42-year-old) Substitute for General Affairs at the Vatican Secretariat of State, Giovanni Battista Montini (the future Pope Paul VI), already a key figure within the Holy See.[53] Kurtna's relationship with Montini would play an essential role throughout the war years; in his later life, Kurtna often boasted of this friendship. The Russian linguist Julia Dobrovolskaya recounts a meeting with him in Moscow during the 1970s, where Kurtna wore a necktie that he claimed was a birthday gift from Paul VI.[54]

Montini allegedly granted Kurtna continued access to archival fonds even after the outbreak of war.[55] In my view – as can be deduced from his autobiography – Kurtna worked in close contact with the Second Section of the Secretariat of State, and it is likely that his career as an intelligence asset began here. A compelling argument in support of this theory is that when Kurtna was later arrested on suspicion of being a German or Soviet spy, he was promptly readmitted to the Vatican upon his release. According to his memoirs, every significant move he made was preceded by consultations with his contacts within the Holy See. In other words, while he was often suspected of spying *on* the Vatican, he was primarily a spy *for* the Vatican. In any case, Kurtna's activities as an intelligence officer began in earnest in June 1940. While some authors – without citing any evidence – claim that Kurtna was a "born" Soviet spy

52 On the political leanings of Liiv, which caused him difficulties during the German occupation, see P. Pillak. Otto Liiv 100. – Tuna, 2005, 5, 4, 152–154, here 154.
53 A. Kurtna. Pronksvärava taga, 12 February 1966.
54 J. Dobrovolskaja. Post Scriptum. Memorie. O quasi? Cafoscarina, Venezia, 2006, 45.
55 A. Kurtna. Pronksvärava taga, 12 February 1966.

sent to the *Russicum* specifically for that purpose,[56] there is absolutely no historical data to support such a theory.

At the end of July 1940, Kurtna travelled to Moscow, where his research in the Vatican Archives allegedly "aroused the interest of the comrades". Discrepancies exist between his autobiography and his later interrogation by British intelligence, during which he claimed the journey took place in September 1940. This trip was made possible by the unique geopolitical alignment of the Molotov–Ribbentrop Pact; at the time, the Soviet Union was an ally of Germany and maintained a similar, though unofficial, status with Italy.

Kurtna claimed that during this visit he dined with Andrey Vyshinsky as part of an exclusive group of eight people. However, this detail invites scepticism, as Vyshinsky was only appointed First Deputy People's Commissar for Foreign Affairs on 6 September 1940. At this juncture, a degree of caution is necessary: spies who survive long enough to write their memoirs are often prone to narcissism, tend to 'sugar-coat' their pasts, and frequently aggrandise their roles. A parallel can be found in Branko Bokun, another wartime operative in Rome and the Vatican. Although a figure of minor significance, Bokun published extensive – and often unreliable – accounts of his activities, including the highly improbable claim of having met the exiled Mussolini on the island of Ponza in late 1943.[57]

There is a striking detail regarding Bishop Eduard Profittlich: following his arrest by Soviet occupation authorities in June 1941, he was interrogated by NKVD agents. On 2 October, Profittlich was questioned specifically about Kurtna's departure from Soviet Estonia. The bishop was accused of facilitating (or, more precisely, failing to report) Kurtna's 'escape' abroad via Finland, a charge that contributed to his death sentence in November 1941.[58] This raises a logical contradiction: why would Kurtna need to 'escape' if the Moscow authorities had already authorised and financed his work? Kurtna himself claimed to have travelled from Moscow to Tallinn in late October 1940, where Profittlich allegedly assisted him through a "German Baltic commission" – probably the German diplomatic mission – to renew his visa. He then departed Estonia, entering Germany at Tilsit.[59] As suggested by Kurtna's 1944 interrogation report, this elaborate itinerary was a ploy designed to conceal his time in Moscow. While some authors suggest that Kurtna

56 V. Sibilio. Misteri vaticani. Independent edition, 2023.
57 B. Bokun. Spy in the Vatican. Praeger, New York, 1973, 135–143.
58 M. Paas. Peapiiskop Eduard Profittlichi märterlusest, 1794–1795.
59 The National Archives (Kew), KV 2/657, 30.

may have provided incriminating information about Profittlich, a more straightforward explanation exists: the NKVD agents who arrested the bishop were simply unaware that Kurtna was an asset of the GRU (then known as the Fourth Bureau).

While Kurtna returned to Rome in the autumn of 1940, his autobiography is notably silent regarding the period leading up to June 1941; this narrative hiatus leaves his actions during the months preceding the German invasion of the USSR entirely undocumented. It is certain, however, that by this time Kurtna had received specific instructions for his work with the GRU. The primary obstacle in tracing his activities as a Soviet agent remains the limited access to relevant archives. Nevertheless, Owen Matthews, author of the most recent biography of Richard Sorge,[60] has successfully consulted the archives of Soviet military intelligence in Podolsk (TsAMO RF, Central Archives of the Ministry of Defence of the Russian Federation). It is highly desirable that contemporary historians follow this precedent to further investigate Kurtna's clandestine role.

At this juncture, a crucial question arises: what, precisely, were Kurtna's directives? During his 1944 interrogation by British intelligence, he outlined several of his tasks. Primarily, his role involved gathering rumours and intelligence regarding shifts in Vatican policy, a mission that would become exceedingly demanding during the German occupation of Rome. Interestingly, the Soviets also assigned him minor, almost 'apprentice-level' tasks, such as transcribing historical documents the Vatican refused to release to the USSR. One such request concerned a "Greek monk called Skilitsa Kouropalatos", none other than John Skylitzes, the eleventh-century Byzantine historian.[61] Moscow's interest in Skylitzes was purportedly linked to a planned film on the Christianisation of Russia; according to Kurtna, the Vatican had withheld the sources because they contained "scandalous" details about the monk's private life. A far more significant objective, however, was lobbying the Holy See to establish formal diplomatic relations with the Soviet Union. Nevertheless, as the interrogation report clarifies, the Soviets intended to utilise Kurtna not only within the Vatican but also as an observer of German institutions across Rome.

Like any self-respecting spy, Kurtna maintained strategic female acquaintances. By chance, in 1940, he met two secretaries working for Otto von Bismarck, the head of the Political Office at the German

60 O. Matthews. An Impeccable Spy. Richard Sorge, Stalin's Master Agent. Bloomsbury, London, 2019.

61 J. Wortley. John Skylitzès: A Synopsis of Byzantine History, 811–1057. Cambridge University Press, Cambridge, 2010.

Embassy in Rome (at that time located at Villa Wolkonsky):[62] Fräulein Laubert and Fräulein Jäger. They provided him with information regarding their work, while he, in turn, fed them "false news" to mislead his German contacts.[63] His primary contact at the Soviet Embassy was an intelligence officer known to him as "Gorshkov", whom he had previously met in Moscow. This was indeed a real and significant figure: Nikolai Mikhailovich Gorshkov, who would later become a renowned Soviet operative active in Algeria during the Cold War. In 1941, Gorshkov would have been approximately twenty-nine years old.[64]

In this manner, Kurtna's research grant was extended in 1940 by the Soviet authorities; however, on 22 June 1941, he lost his new benefactor. When the personnel of the Soviet Embassy departed Rome following the invasion of the USSR, Kurtna was expected to operate independently, though he was left entirely without financial resources. According to his autobiography, he sought counsel from Montini and Cardinal Eugène Tisserant, the powerful Secretary of the Congregation for the Oriental Churches and a pivotal figure in the Holy See's Eastern European policy.

As Kurtna recounts, the Vatican authorities advised him to seek assistance from Father Robert Leiber, a professor at the Gregorian University and the Pope's private confessor. Leiber reportedly suggested that he consult Friedrich Bock at the German Historical Institute.[65] At this juncture, the narratives diverge: in his autobiography, Kurtna claims he only established contact with the Nazi authorities – specifically Herbert Kappler – in June 1942, upon Leiber's recommendation.[66] However, during his 1944 interrogation by British Secret Services, he admitted to having met Kappler as early as March 1942 through Friedrich Bock.[67]

Kurtna, who described Bock as a "hideous Nazi" in his Soviet-era memoirs, had on other occasions only praise for the historian. In reality, Bock – an avowed and proud Nazi – had actively supported Kurtna's Vatican research and eventually hired him at the library of the German Historical Institute in July 1941, with a good monthly salary of 2,000 lire.

62 On the German Embassy, its diplomats, and the intelligence services in Rome, see (with due caution), E. Dollmann. With Hitler and Mussolini: Memoirs of a Nazi Interpreter. Skyhorse, New York, 2017.
63 The National Archives (Kew), KV 2/657, 31.
64 B. Hautecloque. L'affaire Georges Pâques (1963–64). Un haut-fonctionnaire français au service des Soviétiques pendant toute la Guerre froide. – Nuova Antologia Militare, 2024, 5, 20, 515–542, here 520. Gorshkov's service in Italy is also corroborated by Russian biographical sources.
65 A. Kurtna. Pronksvärava taga, 12 February 1966; 15 February 1966.
66 A. Kurtna. Pronksvärava taga, 22 February 1966.
67 The National Archives (Kew), KV 2/657, 26.

Under this arrangement, Kurtna was officially tasked with researching the Great Schism in the Baltic region within the Vatican Archives. However, in February 1942, he lost this position due to a new directive: German historical institutes abroad were henceforth permitted to employ only German nationals. Kurtna still held his Estonian passport, to which he had promptly added Soviet citizenship in the summer of 1940, a detail he wisely chose to conceal at the time.

It was Bock – a "close friend" of Kappler – who introduced him to Kurtna, undoubtedly much earlier than March 1942.[68] Herbert Kappler would become a defining figure in Kurtna's career as a spy. At thirty-four, this SS *Obersturmbannführer* served as the head of the SD (the *Sicherheitsdienst*) and represented Heydrich's RSHA (Reich Security Main Office) in Rome. It was a position that remained largely in the shadows until the German occupation of the city in September 1943, an event that thrust Kappler into a direct operational role and established him, alongside his subordinate *Hauptsturmführer* Erich Priebke, as one of the major German war criminals in Italy. In his memoirs, Kurtna provides a vivid description of his encounter with Kappler: a "polite, blond, slightly overweight middle-aged man with a scar on his left cheek".

In summary, when the Soviet personnel evacuated their embassy in Rome, they failed to provide Kurtna with the necessary resources to sustain his activities. Consequently, he was drawn into the web of German intelligence, not through Admiral Canaris's Abwehr, but via the SD. Officially, his role was to provide the German Embassy with monthly reports on the diplomatic relations between the Third Reich and the Vatican, for which he received a salary of 2,000 lire.

The question that later preoccupied both British intelligence and the Soviet secret services was whether Kurtna continued to operate as a Soviet mole collecting data from both the Germans and the Vatican while disseminating disinformation. Kurtna lacked radio transmitters or other technical equipment, and with the Soviet Embassy in Rome closed, he had no direct channel to Moscow. Curiously, the Germans were fully aware of his previous associations with the Soviets, yet they appear to have had complete confidence in him. It was only later, in 1944, that Kurtna claimed to have re-established contact with Gorshkov via correspondence.

68 Kappler's correspondence with his superiors in Berlin indicates that he was already working with Kurtna as early as February; see D. Alvarez, R. A. Graham. Nothing Sacred, 137.

KURTNA'S JOURNEY TO ESTONIA AND HIS IMPRISONMENT

Between 23 June and 30 July 1942, Aleksander Kurtna travelled to Estonia. This journey remains perhaps the most enigmatic episode in his history, with strikingly different accounts appearing in his autobiography and the 1944 interrogation report. The stated motivations for the trip vary significantly between the two versions. In his memoirs, Kurtna claims he travelled at the invitation of the Slovak Ambassador to the Vatican, Karol Sidor, ostensibly to investigate the situation of Jews in Slovakia.[69] He also asserted that the Vatican Secretariat of State and Cardinal Tisserant had tasked him with meeting Jesuits in Prague and uncovering the fate of a Catholic priest missing in Estonia. To these official reasons, he added a personal motive: concern for two of his brothers, who had been arrested by the Germans.

In the 1944 version, his journey was motivated by a single, pragmatic reason: an attempt to renew his research grant in Tallinn by submitting a request to the *Eesti Omavalitsus* (Estonian Self-Government), the puppet administration under German control. His travel application would have been supported by both Bock's German Historical Institute and the Bibliotheca Hertziana in Rome. The 1966 version is demonstrably false, beginning with the departure date: Kurtna claimed to have left on 18 June, whereas the 1944 records confirm he departed on 23 June. In his later memoirs, he backdated his departure to accommodate a lengthy, fabricated account of a visit to Prague, a trip that, in reality, never took place. He introduced this fiction to reinvent himself as a witness concerned with the fate of Central European Jews, even claiming to have received early reports regarding the Auschwitz concentration camp.

It is highly probable that the British version of this journey is also partially inaccurate: Kurtna declared to MI6 officers that he had travelled directly to Estonia via Tilsit in East Prussia, Lithuania, and Latvia. However, in his memoirs, he inserts a visit to Berlin following the (fabricated) trip to Prague. According to this account, his journey was facilitated by Kappler in Rome – whom Kurtna introduces for the first time in his autobiography at this juncture – who supposedly referred him to the Gestapo in Berlin for further travel clearances.

69 Karol Sidor was a controversial figure. Though a staunch antisemite, he was not an active participant in the Holocaust; his concerns regarding the fate of Jews in Slovakia may have been known to Kurtna through other channels. See, for example (with some caution), K. Sidor. Denníky 1930–1939. Ústav pamäti národa, Bratislava, 2010.

In Berlin, at the RSHA headquarters on Prinz-Albrecht-Straße, Kurtna allegedly held a lengthy meeting with SS *Standartenführer* Friedrich Panzinger. The 39-year-old Panzinger was a pivotal and controversial figure in wartime espionage and counter-espionage, later serving West German intelligence and the CIA. At the time, he headed Division A of the Fourth Section (*Amtsgruppe IV A*) of the RSHA, the heart of the Gestapo.[70] Since Berlin was directly on his route, Kurtna visited both the Gestapo and Panzinger again during his return trip. It is difficult to believe that Kurtna would have been granted an audience with such a high-ranking officer merely to discuss a travel visa. Furthermore, the timeline of his journey is telling: reaching Tallinn from Rome in just three days (by 25 June), including a stopover in Berlin under 1942 conditions, strongly suggests he utilised military transport or possessed high-priority travel credentials.

Upon his arrival in Estonia, according to the interrogation report, Kurtna learned for the first time that one of his brothers was imprisoned and another had been executed by the Germans. In this version – though notably not in his autobiography – he identifies the officer who ordered the execution as SS *Obersturmbannführer* Georg Elling. This revelation opens a new and significant investigative path: Elling, a former friar who served in the SS in Lorraine and later the SD as a specialist in religious affairs, was ultimately sent as a secret agent to Rome and the Vatican, where he disappeared in 1944. My research confirms that he was an active member of Einsatzgruppe A in the Baltic states in 1941, a fact not widely documented.[71] This detail is particularly striking given that Kurtna would later work alongside Elling. Are these accusations accurate, and if so, why did Kurtna omit them from his memoirs? This silence raises the question of whether his later collaboration with Elling was a calculated act of insincerity or whether, in 1944, he was attempting to distance himself from a man whose crimes he knew all too well.

According to the British report, Kurtna was informed of these tragic circumstances by a schoolmate named Otto Tamm. While his autobiography mentions an encounter with Tamm, it provides a completely different version of the event. Crucially, Kurtna omitted a vital detail

70 I. Wirth. Friedrich Panzinger: Am Ende stand der Suizid. – Täter Helfer Trittbrettfahrer, XVI, NS-Belastete aus München. Hrsg. von W. Proske. Kugelberg Verlag, Gerstetten, 2023, 325–336.

71 Exhaustive data regarding Elling's career were located in a file prepared in 1965 for the prosecutor of the *Kammergericht* in Berlin, intended for a murder prosecution. Since it was discovered that Elling had died in 1952 in Kempten, the case was closed; the dossier contains general information and service records but lacks specific details on the crimes of which he was accused. See Landesarchiv Berlin, B Rep. 057-01, Nr. 959.

from his interrogation: after meeting Kappler in Rome and Panzinger in Berlin, he entered the grey building at Tõnismäe 16 in Tallinn, the headquarters of the SiPo (*Sicherheitspolizei*) and SD in Estonia, the local branch of the RSHA. There, he was received personally by the chief, the thirty-year-old SS *Sturmbannführer* Martin Sandberger, a notorious war criminal. Sandberger subsequently referred him to an officer whom Kurtna identified as "Untersturmbannführer Zefeldt". This individual was, in reality, SS *Untersturmführer* (Second Lieutenant) Gustav von Seefeld, who was in charge of two sections of the *Politische Polizei*: "Foreign Countries" and *Abwehr*. While Kurtna claimed the purpose of this visit was to seek news of his brothers – and that Seefeld informed him of Nikolai's execution – the fact remains that Seefeld was an intelligence officer, not a functionary responsible for such enquiries.

The British version of the return journey offers further contradictions: Kurtna claimed to have met Professor Juhan Vasar, a scholar of the modern period and deputy to the Director of Education for the Estonian Self-Government.[72] Although his request for a new grant was denied, he received 8,000 lire (a sum he notably converted into Italian currency) specifically to produce photographic reproductions of archival documents.

His return route is described differently in the two sources. In his statement to British intelligence, he admitted to stopping in Berlin but was suspiciously eager to specify that he remained there for only ten hours and met no one. Conversely, in the 1966 version, he portrays his journey as a harrowing escape, fraught with fear and potential danger, claiming that during this stopover, Panzinger and the Gestapo treated him with cold suspicion. Kurtna's journey to Estonia can be interpreted in several ways. While it probably served as a form of reward from the Germans, his meetings with intelligence officers – such as Panzinger, Sandberger, and Seefeld (and potentially others omitted from his autobiography) – were clearly professional in nature. These encounters undoubtedly facilitated a high-level exchange of information, solidifying his role as a trusted asset within the Nazi intelligence hierarchy.

Aleksander Kurtna returned to Rome on 30 July 1942, only to be immediately arrested at the Termini railway station. He was taken to Regina Coeli, Rome's primary pre-trial detention centre, a prison whose structure and history bear a striking resemblance to Patarei prison in Tallinn. Much like his Estonian journey, Kurtna's arrest remains a deeply

72 Eesti teaduse biograafiline leksikon. Toim. J. Aaviksoo, K. Siilivask. Kd. IV, Sar–Y. Lisa. Eesti Entsüklopeediakirjastus, Tallinn, 2013, 136–137.

controversial episode. According to his autobiography, he and his wife, Alma, were detained because his address was discovered on a captured member of the Resistance.[73] Kurtna identifies this man as a Slovenian, and from this point in his memoirs, he introduces the theme of the Yugoslav resistance, claiming to have fought alongside them. However, the British report paints a far more complex picture: the Italian Military Secret Service (SIM), while investigating a Swiss espionage ring, had arrested two Soviet agents on 13 June 1942: the Finn Holgar Tarvonen (alias Ernest Harn) and Fritz Schneider (alias Hermann Leo Marley, or "Paolo"). This led to the discovery of an extensive network of Italian accomplices.[74] Unlike Kurtna, Tarvonen and Schneider possessed a radio transmitter; the SIM intercepted and decrypted a message directed to them that began with the phrase "We are from the North", which contained specific instructions or information intended for Kurtna.

Colonel Manfredi Talamo, head of the SIM counter-espionage squad, and his team found themselves in a precarious and confounding situation regarding Kurtna.[75] From the outset, Kurtna did not hesitate to reveal his work for the German services, confessing to being a spy, but for Berlin, not Moscow. This admission did little to aid his case. The Italian secret services harboured a deep-seated loathing for their German counterparts, and the revelation only added to the prevailing confusion. According to Kurtna's own testimony, the Italians interrogated him approximately twenty times. In his memoirs, he even recounts a "mock execution" staged by Colonel Talamo in a desperate attempt to force a confession. It is essential to note the discrepancy in his narratives: in his autobiography, Kurtna omits any mention of his intelligence activities, instead portraying himself as arrested without cause. This stands in stark contrast to his 1944 interrogation, where he provided the British with detailed information regarding his operations for the Soviet Union.

Kurtna offers a detailed account of his time in prison, enriched by several compelling anecdotes. Most notably, he claims to have met the renowned writer Curzio Malaparte in Regina Coeli. Malaparte was indeed arrested in 1943, shortly after the fall of Mussolini's regime on 25 July, and spent a month in detention. According to Kurtna, the writer would read chapters of his work-in-progress, the novel *Kaputt*, to his

73 A. Kurtna. Pronksvärava taga, 3 March 1966.

74 A. Mantica, V. Fragalà. La dimensione sovranazionale del fenomeno eversivo in Italia. Studio sui collegamenti fra le organizzazioni terroristiche internazionali. – XIII Legislatura – Disegni di legge e relazioni – documenti. Senato della Repubblica, Camera dei Deputati, Roma, 2001, 1–246, here 59–60: a very curious document that does not mention Kurtna.

75 R. A. Graham. L'occhio del Sim sulla Città del Vaticano. – Gnosis. Rivista italiana di intelligence, 2022, 28, 2, 134–149, here 146–148.

fellow inmates. Decades later, in 1976, Kurtna achieved a remarkable editorial feat by managing to publish an extract from *Kaputt* in the prestigious *Loomingu Raamatukogu* series in Tallinn.[76] This was no small accomplishment: Malaparte was featured on the Soviet blacklist of Western writers, primarily due to his provocative 1932 novel, *Lenin, the Good Fellow* (*Le bonhomme Lénine*), which was viewed with extreme hostility by Moscow.[77]

From this point forward, Kurtna's autobiography veers into pure fantasy. He claims that on 10 September 1943 – as German forces occupied Rome following the Italian armistice and the flight of King Vittorio Emanuele III to Brindisi – a prisoner revolt at Regina Coeli led to his liberation. This account is entirely inaccurate. According to his British interrogation records, Kurtna was actually released on 30 September 1943. His liberation was orchestrated by Herbert Kappler, who had also moved to arrest Colonel Manfredi Talamo. This was more than mere retaliation for Kurtna's imprisonment; it was a strike against the Italian SIM, which had long hindered German intelligence operations in the capital. Talamo's fate was tragic: he was later executed by Kappler's SS during the Fosse Ardeatine massacre on 24 March 1944.[78]

Upon their release, Kurtna and his wife, Alma, received a substantial sum of money from Friedrich Bock and settled into a luxurious new apartment at Via Cola di Rienzo 28, in the elegant Prati neighbourhood. Curiously, in his autobiography, Kurtna claims to have lived in a much more modest location on Viale Regina Margherita. While his memoirs depict him as a "freedom fighter" actively participating in the defence of Rome against the German occupation, several intriguing figures from the Nazi hierarchy emerge in his narrative. Kurtna mentions encounters with two pivotal personalities in occupied Rome: the aforementioned Georg Elling and Eugen Dollmann, an honorary SS *Obersturmbannführer*. According to Kurtna, these men facilitated Alma's release, though the British interrogation report attributes the liberation to Bock and Kappler. Dollmann remains an enigmatic figure; while some authors have mistakenly labelled him the head of German intelligence in Rome,

76 C. Malaparte. Kriket Poolas. (Loomingu Raamatukogu, 38/39). Perioodika, Tallinn, 1976.
77 C. Malaparte. Le bonhomme Lénine. Bernard Grasset, Paris, 1932.
78 The Fosse Ardeatine massacre (24 March 1944) stands as the most heinous German war crime committed in Rome, alongside the deportation of the Roman Jewish community. The massacre was a reprisal for the partisan attack in Via Rasella the previous day, which killed 33 soldiers of the SS *Polizeiregiment Bozen*. Following Hitler's order to execute ten Italians for every German victim, the SS collected 335 prisoners from various Roman jails – including the Gestapo headquarters at Via Tasso – and murdered them in the Ardeatine caves. See A. Portelli. The Order Has Been Carried Out: History, Memory, and Meaning of a Nazi Massacre in Rome. Palgrave Macmillan, New York, 2003.

that role belonged strictly to Kappler. Dollmann's rank was largely ceremonial, yet he wielded immense influence as the primary liaison between the German command and Roman high society, bolstered by his personal connection to Heinrich Himmler.[79]

THE GERMAN OCCUPATION AND THE END OF THE WAR IN ROME

The German occupation of Rome lasted only nine months, yet for a historian, it remains an exceptionally dense and eventful period.[80] Vatican City had become a sanctuary for a vast number of diplomats, Allied officers, and opponents of the Third Reich, and the city was swarming with operatives from every intelligence service. For Kurtna, this was a season of intense activity, but also one of calculated repositioning. As it became increasingly clear that the Germans were losing the war – with the Allied arrival expected at any moment – he realised he needed to provide tangible services to Soviet intelligence to secure his future. At this juncture, Fräulein Laubert reappears, alongside a new contact: Fräulein Schwarzer, a secretary at the German Embassy. Kurtna first persuaded them to purge incriminating data from his SIM dossiers, which had fallen into German hands following the occupation. Subsequently, he began transmitting intelligence to the Soviets through the Siamese Legation (the current Thai Embassy at Villa Thai on Via Nomentana) and other clandestine channels. This strategic pivot probably saved his life when the war finally concluded.

Kurtna's most significant operational achievement was persuading Fräulein Schwarzer to defect to the Soviet cause. Her cooperation was partially rooted in her family's communist sympathies, but it was primarily secured through a calculated promise: with Soviet troops advancing towards her parents' home in Silesia, Kurtna guaranteed their protection. He later followed through on this by providing their names

79 E. Dollmann. With Hitler and Mussolini. On Dollmann's actual influence and his complex relationship with the SS hierarchy, see K. von Lingen. Allen Dulles, the OSS, and Nazi War Criminals. The Dynamics of Selective Prosecution. Cambridge University Press, Cambridge, 2013.

80 For the definitive accounts of this period, see R. Trevelyan. Rome '44. The Battle for the Eternal City. Secker & Warburg, London, 1981; R. Katz. The Battle for Rome. Simon & Schuster, New York, 2004; L. Klinkhammer. Zwischen Bündnis und Besatzung. Das nationalsozialistische Italien und die Republik von Salò, 1943–1945. (Bibliothek des Deutschen Historischen Instituts in Rom, 75). Niemeyer, Tübingen, 1993. For a topographical and visual history, see A. Majanlahti, A. Osti-Guerrazzi. Roma occupata 1943–1944. Itinerari, storia, immagini. Il Saggiatore, Milan, 2010.

to Soviet authorities to ensure their safety. Schwarzer's position granted Kurtna access to the German Embassy's most sensitive files, including copies of Tarvonen's confession, which contained a comprehensive list of active Soviet operatives. Simultaneously, Moscow utilised Kurtna to feed the Germans sophisticated disinformation, such as falsified reports on the outcomes of the Tehran Conference. However, Soviet directives were often inconsistent, forcing Kurtna to rely on his own ingenuity and imagination to execute his missions and maintain his cover.

Kurtna also played a role in assisting escaped Soviet prisoners of war and forced labourers, many of whom were hidden within the Siamese Legation. However, this effort was not a direct instrument of Soviet policy; rather, it was spearheaded by a local branch of the Mladorossy (Young Russians). While this movement was ideologically hostile to Communism, its members – led in Rome by Vassili Nikolaevich Dyukin – collaborated with the Soviet Union during the war out of nationalistic necessity. Paradoxically, Kurtna was actually reprimanded by his Soviet handlers for this involvement, as Moscow viewed such independent initiatives with deep suspicion.[81]

Even as he pivoted toward Moscow, Kurtna continued to serve German interests by assisting Georg Elling. Their meeting took place at the Albergo Maestoso, where Elling revealed that despite his long tenure in Berlin as the RSHA's Vatican specialist, he lacked practical field experience and could not even speak Italian. During this encounter, Kurtna received a payment of 6,000 lire. As the German occupation of Rome drew to a close, Kurtna became central to Elling's project of establishing a 'stay-behind' intelligence network intended to operate under Allied occupation. As the Germans evacuated, Kurtna agreed to remain in the city as an undercover agent, accepting a staggering sum of 100,000 lire to fund his future activities. Simultaneously, however, he obtained from Fräulein Schwarzer the complete list of agents that Elling was leaving behind in Rome, effectively compromising the entire German operation before it even began.

In his autobiography, Kurtna paints a radically different picture: he depicts himself as a dedicated Resistance member, sheltering Soviet officers in safe houses and enthusiastically welcoming the American liberators. However, 5 June 1944, the day after the Allies entered Rome, proved to be far from celebratory for him. Early that morning, Kurtna went to the Vatican to deliver a cache of documents obtained from

81 The National Archives (Kew), KV 2/657, 22.

Fräulein Schwarzer. He left them with one of Montini's secretaries, intended for ultimate delivery to the Soviets. Afterwards, he and his wife headed to Villa Thai to gauge the situation in the city. On their return journey, near Villa Nemorense, they were intercepted by the very same SIM agents who had arrested Kurtna in 1942: Tranfo and Di Bello.

Kurtna was returned to Regina Coeli, where he was interrogated by various U.S. and British intelligence officers. Finally, he was visited by Nikolai Gorshkov, who initially feigned ignorance of Kurtna's identity before engaging him in a private conversation. Gorshkov instructed him to withhold all information from the British and Americans. Seeking to protect himself, Kurtna requested and obtained a formal guarantee from the British Secret Service that his depositions would remain strictly confidential. On 21 September 1944, Kurtna and his wife were handed over to Soviet military authorities and transported to Moscow. There, he faced trial for treason and was sentenced to the Norilsk labour camp. Some years later, in 1948, his former fellow seminarian Walter Ciszek encountered him in Norilsk. To Ciszek's surprise, he found Kurtna well-fed and dressed in a white coat, working in the camp's medical service, a "warm and quiet place" far removed from the brutal labour of the other inmates.[82]

Existing studies on espionage in Rome and the Vatican rank Kurtna among the most eminent intelligence operatives of the era.[83] Conversely, Branko Bokun – a fellow spy in the Vatican who knew him well – offered a more biting assessment, describing him as "one of the most obscure double agents" and a "small, insignificant, and compliant mouse of a man who travelled paths laid out for him by others". For Bokun, Kurtna was a triple agent, working simultaneously for the Vatican, the Germans, and the Soviets.[84] Notably, Bokun is one of the few to highlight the connection with the Holy See, which, as this study suggests, is crucial to deciphering Kurtna's true role. The web of suspicion around him was broader still: in 1942, the Italian SIM even suspected him of collaborating with the French Deuxième Bureau.[85] Despite his historical importance, Kurtna's figure remains shrouded in mystification, some of it self-authored, and some the result of superficial

82 D. Alvarez. Spies in the Vatican: Espionage and Intrigue from Napoleon to the Holocaust. University Press of Kansas, Lawrence, 2002, 316–318.

83 See, for example, D. Alvarez, R. A. Graham. Nothing Sacred, 120, 121, 126, 127; O. Chadwick. Britain and the Vatican During the Second World War. Cambridge University Press, Cambridge, 1988, 176.

84 B. Bokun. Memoirs of A Nomadic Humorist. Irony, Pathos and Farce. Vita Books, London, 2007, 19–21.

85 R. A. Graham. L'occhio del Sim sulla Città del Vaticano, 148.

historical accounts. In Italian literature one encounters the recurring, inaccurate trope of the "Lithuanian priest" who spied for both Berlin and Moscow.[86]

A definitive judgement on Kurtna's activities as an intelligence officer remains elusive without access to the Soviet sources held in the Podolsk archives. Nevertheless, several key considerations can be made. Some authors have alleged that Kurtna was responsible for the arrest and subsequent death of numerous Jesuits and others involved in the Vatican's "re-evangelisation" project in Eastern Europe, as well as clergy opposed to the Nazi regime.[87] These claims, however, appear to lack a factual basis. Had Kurtna surrendered all the data in his possession to either the Soviets or the Nazis, it would probably have triggered hundreds of arrests. Yet, there is no documented evidence of anyone being detained specifically due to his denunciations. As the evidence suggests, both Moscow and Berlin were primarily interested in high-level political intelligence within the Vatican; they did not utilise Kurtna as a tool for dismantling missionary projects or for the systematic persecution of individual religious opponents.

Another complex issue lies in determining exactly for whom Kurtna worked. The evidence suggests he served all three powers – the Germans, the Soviets, and the Vatican – simultaneously and with equal dedication. While his German handlers initially held sway, by 1943–1944 it became clear that the Reich was losing the war, and his primary allegiance shifted decisively toward the Soviets. It is also noteworthy that Kurtna provided a vast amount of data to British intelligence – directly defying Gorshkov's orders – in an effort to demonstrate his willingness to collaborate with the Allies. It was for this reason that his 1944 interrogation report was eventually reviewed in London by Kim Philby, who sought to examine the extent of the Soviet intelligence network in Italy.

Ultimately, the Soviets appear to have accepted Kurtna's version of events. Had there been even a modicum of suspicion regarding his ultimate loyalty, he would probably have been executed immediately. Instead, he was sentenced to ten years in the Gulag – a remarkably lenient term given the political climate of the era. Another telling detail involves his wife, Alma. She was never arrested and, in February 1945,

86 G. Pillon. Spie per l'Italia. I Libri del No, Roma, 1968; C. De Risio. Generali, Servizi Segreti e Fascismo: la guerra nella guerra 1940–1943. Mondadori, Milano, 1978.

87 E. Caretto. Le spie in Vaticano. – Corriere della Sera, 13 August 2003. Caretto's assertions represent a trend in popular journalism that often prioritises sensationalism over archival verification.

gave birth to their daughter in Moscow. This implies she must have been pregnant since May 1944, during the final weeks of the German occupation of Rome. Curiously, Kurtna makes no mention of this in his autobiography, which concludes abruptly with his "happy" return to the Soviet homeland.

Throughout his years in Rome – with the exception of his imprisonment from July 1942 to September 1943 – Kurtna worked in the Vatican Archives. In his 1966 account, he claimed to have copied or catalogued approximately 7,000 sources. According to Kurtna, this extensive catalogue was confiscated by American forces upon his arrest on 5 June 1944.[88] In reality, however, he was apprehended by the Italian Military Intelligence Service (SIM).[89] As no trace of this catalogue has been found in the archives of the US Office of Strategic Services (OSS) or British intelligence, its ultimate fate remains a mystery. The records of the SIM are currently held by the Italian intelligence agencies (AISE) and remain strictly inaccessible to researchers; consequently, it is impossible to determine whether Kurtna's documents are still preserved within those files.

CONCLUSIONS: ARCHIVAL RESEARCH FROM HELK TO THUMSER AND A FINAL JUDGEMENT ON KURTNA

Kurtna's scholarly interests extended far beyond the Vatican Secret Archives – where he focused on the registers of supplications and papal letters – to the central archives of the Society of Jesus (ARSI). These Jesuit archives house a wealth of primary material concerning Livonia from the sixteenth century onward. Following the Second World War, these same sources were systematically examined by Vello Helk, the only Estonian historian to conduct extensive research in Roman ecclesiastical archives during that period. Helk's seminal studies on the Jesuit Order in Livonia drew upon the very foundations that Kurtna had once navigated, though under vastly different circumstances.[90]

88 TLA.R-27.1.937a, 2 (14).

89 The National Archives (Kew), KV 2/657, interrogation summary, 28 August 1944, 12.

90 See V. Helk. Die Jesuiten in Dorpat 1583–1625. Ein Vorposten der Gegenreformation in Nordosteuropa. Odense University Press, Odense, 1977; see also the unpublished annotations and copies in RA, EAA.5383. See also J. Undusk. How to Become a Perfect Danish-Estonian Historian: Homage to Vello Helk. – East and Central European History Writing in Exile 1939–1989. Ed. by M. Zadencka, A. Plakans, A. Lawaty. Brill, Leiden, Boston, 2015, 236–248.

Following Vello Helk's contributions, no systematic research on Livonica within the Vatican was undertaken until the recent editorial project led by Matthias Thumser and his colleagues. This major initiative focused on volumes 13 to 16 of the first series of the *Liv-, Est- und Kurländisches Urkundenbuch* (LEKUB). The project aims to incorporate Vatican documents covering the period from 1472 to 1494; volumes 13 and 14 have already been published; the very sources that Kurtna had once catalogued.[91]

In summarising the current state of research, the initial impression is one of significant complexity and fragmentation. The earlier volumes of the LEKUB were compiled before the formal opening of the Vatican Archives in 1881; it was only starting with Volume 7, edited by Hildebrand – covering the period from 1423 – that Vatican sources began to be utilised. However, these early volumes were edited according to outdated historiographical criteria, often failing to cite specific sources or provide diplomatic and palaeographic descriptions. Consequently, they offer limited assistance regarding Vatican materials. Even the editors of the most recent volumes (from Volume 13 onward) were able to consult only a limited selection of available papal records. As the editors themselves acknowledge, practical constraints meant that the *Vaticana* could only be partially incorporated. While materials from the German Historical Institute in Rome were used, they were incomplete at the time, meaning that many registers – beyond the intentional exclusion of minor judicial and gratial matters – appear only sporadically in the current editions.[92]

Any study of the medieval and early modern history of Livonia based solely on the LEKUB must be considered scientifically incomplete, given the substantial quantity of unpublished and uncatalogued sources yet to be integrated into these volumes. Before assessing the potential source material available in the Vatican, it is necessary to survey the documents already accessible within the Baltic region. Existing editions by Motzki, Arbusow, and others must be carefully considered, alongside a significant volume of unpublished transcriptions and registers. Of

91 LEKUB, Bd. 13, 1472–1479. Bearb. von M. Mahling, K. Neitmann, M. Thumser. Böhlau, Köln, Weimar, Wien, 2018; LEKUB, Bd. 14, 1480–1483. Bearb. von C. Gahlbeck, M. Mahling, K. Neitmann, M. Thumser. Böhlau, Köln, Weimar, Wien, 2020; volumes 15 and 16 are in preparation.

92 LEKUB, Bd. 13, 17: "Aus praktischen Gründen konnten die Vaticana nur teilweise berücksichtigt werden. Parallel zur Bearbeitung des 13. Bandes des LUB wurde das "Repertorium Germanicum" für Sixtus IV. erstellt. Zwar wurden von seiten des herausgebenden Deutschen Historischen Instituts in Rom Materialien zur Verfügung gestellt, doch waren diese seinerzeit noch nicht vollständig, sodass die vatikanischen Registerserien über den prinzipiellen Verzicht auf minder wichtige Justiz- und Gratialsachen hinaus nur punktuell in den vorliegenden Band eingeflossen sind."

primary importance are the papers of Liisi Karttunen, left in Estonia in 1919 and now preserved in Tartu, and those of Vello Helk, also held there. Any future research project in the Vatican should commence with the systematic cataloguing of these edited and unpublished sources. Ideally, a centralised database should be established to collate information from the LEKUB, the works of Motzki and Arbusow, and the foundational papers of Karttunen, Helk, and Kurtna.

Returning to Kurtna, he remains the scholar who devoted more time than any other to the study of Livonica within the Vatican Archives. The sheer brilliance of this man is well documented: he was a polyglot fluent in dozens of languages and, following his release from the Gulag, a prolific translator who introduced a vast corpus of world literature to the Estonian public. He was undoubtedly a controversial figure, often polarising those who knew him; in his later years, he probably served the Estonian KGB, though evidence suggests he did so without causing harm to others.[93] In my personal estimation, Kurtna's entry into the world of espionage in 1940–1941 was a pragmatic decision to finance his Vatican studies. His archival research was never a mere cover for his intelligence work; rather, it was his true vocation, for which espionage provided a dangerous but necessary source of funding.

„ME OLEME PÕHJAST": *LIVONICA* VATIKANI ARHIIVIS JA ALEKSANDER KURTNA ERIJUHTUM

Kristjan Toomaspoeg

Vatikani Apostellik Arhiiv on oma 85 riiulikilomeetri ja miljonite dokumentidega üks peamiseid keskaja Liivimaa (tänapäeva Eesti ja Läti) ajaloo allikate hoidlaid. Kuigi suur osa neist materjalidest on Baltimaade uurijatele tuntud, moodustavad need vaid murdosa sellest, mis on veel tundmata. Süstemaatiline teadustöö algas 1881. aastal, mil paavst Leo XIII avas arhiivi uksed uurijatele. Enne seda olid arhiivile erandkorras ligi pääsenud vaid üksikud õpetlased, nagu Aleksandr Turgenev 1840. aastatel ja arhiivi prefekt Augustin Theiner 1860. aastatel. Esimene balti ajaloolane, kes aastatel 1885–1886 Vatikani allikaid uuris, täiendades monumentaalset allikakollektsiooni „Liv-, Est- und Kurländisches Urkundenbuch"

93 Some Estonian intellectuals found reports from a certain "Agent Calvino" in their dossiers after the opening of the KGB archives and attributed them to Kurtna.

(LEKUB), oli Riia linnaarhivaar Hermann Hildebrand. Hiljem integreeriti Vatikani-uuringud rahvusvahelistesse projektidesse. Aastatel 1907–1915 kopeeris paavsti registreid soomlanna Liisi Karttunen, tema ümberkirjutused (aastatest 1219–1571) anti 1919. aastal üle Eesti diplomaatilisele delegatsioonile ja asuvad praegu Tartus Rahvusarhiivis. Vahetult enne teda kopeeris ja avaldas Preisimaa ajaloolane Arthur Motzki väärtusliku seeria Liivimaad puudutavaid palvekirju (*Registra Supplicationum*) aastatest 1342–1366. Läti valitsuse toel tegi silmapaistvaimad uuringud Leonid Arbusow noorem, kes aastatel 1926–1931 uuris põhjalikult paavsti 14. sajandi registreid ja kelle ulatuslik fotokogu leiti hiljuti taas üles Läti Riiklikust Ajalooarhiivist.

Eesti vastus Läti uurimisprojektidele kehastus Aleksander Kurtnas (snd 1914 Räpinas). Kurtna, kes valdas vabalt mitmeid keeli, pöördus 1935. aastal katoliiklusse – ajal, mil Vatikan püüdis jesuiitide ordu kaudu tugevdada katoliikluse mõju Ida-Euroopas. 1936. aastal saadeti Kurtna Rooma jesuiitide seminari *Collegium Russicum* õppima. Kuigi ta pidi 1939. aastal preestriametiks sobimatuse tõttu seminarist lahkuma, jäi ta Rooma, teenis elatist keeleõpetajana ja sai seejärel stipendiumi Vatikani arhiivide uurimiseks. Kurtna tegi Roomas intensiivset teadustööd, keskendudes palvekirjade registritele ja seni vähetuntud *Ospizio dei Convertendi* fondile, mis sisaldas katoliiklusse pöördunud liivimaalaste andmeid.

1940. aasta suvel toimunud Nõukogude riigipööre Eestis ja Itaalia astumine Teise maailmasõtta muutsid täielikult Kurtna elu. Ta sai Vatikani riigisekretäri asendajalt Giovanni Battista Montinilt (tulevaselt paavst Paulus VI-lt) eriloa arhiivitöö jätkamiseks. Samal perioodil algas tema spioonikarjäär. Autori hinnangul oli Kurtna eelkõige Vatikani spioon, kuid 1940. aasta sügisel sattus ta Moskvas tõenäoliselt Nõukogude sõjaväeluure (GRU) teenistusse. Tema ülesandeks oli koguda kuuldusi Vatikani poliitika kohta ja edastada luureinfo oma Rooma kontaktile Nikolai Gorškovile. Kui Nõukogude saatkond 1941. aasta suvel Roomast evakueeriti, jäi Kurtna rahata. Vatikani vaimulike soovitusel pöördus ta Saksa Ajalooinstituudi juhi Friedrich Bocki poole, kes andis talle tööd ja tutvustas teda Herbert Kapplerile – Roomas asuva natsliku julgeolekuteenistuse (SD) ja RSHA juhile. Nii tõmmati Kurtna Saksa luurevõrgustikku ja ta hakkas natsidele koostama igakuiseid raporteid Kolmanda Riigi ja Vatikani suhete kohta. Sakslased teadsid tema Nõukogude minevikust, kuid usaldasid teda täielikult. 1942. aasta suvel reisis Kurtna natside ja Vatikani toel korraks okupeeritud Eestisse, kohtudes teel Berliinis Gestapo ohvitseri Friedrich Panzingeriga ning

Tallinnas Martin Sandbergeriga – see kinnitab tema rolli usaldusväärse Saksa agendina. Eestis sai ta teada, et sakslased olid hukanud ühe tema vendadest. Naasnud juulis 1942 Rooma, arreteeris Itaalia sõjaväeluure (SIM) Kurtna, sest nad olid dešifreerinud Nõukogude spioonidele saadetud raadiosõnumi koodfraasiga „Me oleme Põhjast", mis sisaldas Kurtnale mõeldud juhiseid. Kurtna veetis aasta Regina Coeli vanglas, kus ta väidetavalt olevat kohtunud kirjanik Curzio Malapartega. 1943. aasta septembris, kui Roomas algas Saksa okupatsioon, vabastas Herbert Kappler Kurtna vanglast ja neutraliseeris Itaalia SIM-i. Sõja lõpukuudel Roomas viibides ja mõistes, et Saksamaa kaotab, asus Kurtna intensiivselt edastama infot Nõukogude luurele. Tema suurimaks saavutuseks oli värvata Saksa saatkonna sekretär Schwarzer, kelle kaudu sai ta juurdepääsu natside kõige salajasematele dokumentidele ja natside poolt Rooma jäetud salaagentide nimekirjale, mille ta kohe Nõukogude poolele reetis. 4. juunil 1944, päev pärast liitlasvägede saabumist Rooma, arreteerisid SIM-i agendid Kurtna uuesti. Pärast ülekuulamisi USA ja Briti luure poolt anti Kurtna ja tema naine Alma sama aasta septembris üle Nõukogude võimudele ja viidi Moskvasse. Kurtnale mõisteti riigireetmise eest kümme aastat Norilski vangilaagris – stalinistliku ajastu kohta üllatavalt leebe karistus, mis viitab sellele, et Moskva aktsepteeris tema versiooni sündmustest. Pärast vabanemist naasis Kurtna Eestisse, kus temast sai väljapaistev ja produktiivne tõlkija, kes vahendas eesti lugejatele suure hulga maailmakirjandust. Teaduslikust vaatenurgast on Aleksander Kurtna mees, kes on veetnud Vatikani arhiivides Liivimaa allikaid uurides rohkem aega kui ükski teine eesti õpetlane. Arhiivitöö oli tema tõeline kutsumus, mitte lihtsalt kate spionaažile, mis oli talle vaid viis oma Vatikani-uuringuid finantseerida.

Tulevaste Vatikani-projektide arendamiseks on hädavajalik esmalt süstemaatiliselt kataloogida ja digiteerida juba Baltikumis (Tartus ja Tallinnas) leiduvad avaldamata transkriptsioonid, eriti Liisi Karttuneni, Vello Helki ja Aleksander Kurtna isiklikud paberid. Ideaalis tuleks luua tsentraliseeritud andmebaas, mis koondaks LEKUB-i, Motzki, Arbusowi, Karttuneni, Helki ja Kurtna materjalid, et võimaldada tulevastele põlvedele täielik ülevaade Liivimaa ajaloost.





Design and layout copyright © Estonian Academy Publishers, 2026

EESTI TEADUSTE AKADEEMIA KIRJASTUS
Kohtu 6, 10130 Tallinn
Tel 6 454 504, faks 6 466 026
piret.lukkanen@eap.ee

ESTONIAN ACADEMY PUBLISHERS
Kohtu 6, 10130 Tallinn
Phone +372 6 454 504, fax +372 6 466 026
piret.lukkanen@eap.ee

Trükitud trükikojas Alfapress OÜ
Reti tee 8, 75312 Peetri, Estonia

Printed by Alfapress OÜ
Reti tee 8, 75312 Peetri, Estonia

ISSN 1406-2925

