

filij rectam. Et facta est habitacio eoru
 demissa pergentib; usq; sephar montem
 orientale. Isti filij leui secundu cognatio
 nes ⁊ linguas ⁊ regiones in gentib; suis.
 Hec familie noe iuxta populos ⁊ nat
 ones suas. Ab his diuise sunt gen
 tes in terra post diluuiū. Erat autem ter
 ra labij unius ⁊ sermonu eoru idē. Cum
 q; proficiscerentur de oriente inuenerunt
 campum in terra sennaar ⁊ habitau
 erunt in eo. Dixit q; alter ad proximu su
 um. Venite faciamus lateres ⁊ coqua
 mus eos igni. habuerūt q; lateres pro sa
 xis ⁊ bitumen pro cemento ⁊ dixerunt.
 Venite faciamus nob ciuitate ⁊ turrim
 cuius culmen ptingat ad celum ⁊ cele
 bremus nomen nrū. Antequam diuida

JXT

MEDIEVAL MANUSCRIPTS AND THE COMPUTATIONAL HUMANITIES

BIG DATA, SCRIBES, AND THE “PARIS BIBLE”

by **DAVID JOSEPH WRISLEY**
 and **ESTELLE GUÉVILLE**

ARC HUMANITIES PRESS



BOOK CULTURES, MEDIEVAL TO MODERN

Further Information and Publications

www.arc-humanities.org/series/book-series/

MEDIEVAL MANUSCRIPTS AND THE COMPUTATIONAL HUMANITIES

BIG DATA, SCRIBES, AND THE “PARIS BIBLE”

by

DAVID JOSEPH WRISLEY
and **ESTELLE GUÉVILLE**

ARCHUMANITIES PRESS

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library.

© 2026, Arc Humanities Press, Leeds



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Licence.

The author(s) assert(s) their moral right to be identified as the author(s) of their part of this work.

Permission to use brief excerpts from this work in scholarly and educational works is hereby granted provided that the source is acknowledged. Any use of material in this work that is an exception or limitation covered by Article 5 of the European Union's Copyright Directive (2001/29/EC) or would be determined to be "fair use" under Section 107 of the U.S. Copyright Act September 2010 Page 2 or that satisfies the conditions specified in Section 108 of the U.S. Copyright Act (17 USC §108, as revised by P.L. 94-553) does not require the Publisher's permission.

ISBN (Hardback): 9781802702439

ISBN (Paperback): 9781802704501

e-ISBN (PDF): 9781802704488

e-ISBN (ePUB): 9781802704495

DOI: 10.17302/mfsk1586

www.arc-humanities.org

Printed and bound in the UK (by CPIGroup [UK] Ltd), USA (by Bookmasters), and elsewhere using print-on-demand technology.

Publisher (manufacturer) details: Arc Humanities Press, 14 Clifton Moor Business Village, James Nicolson Link, York YO30 4XG, United Kingdom.

EU Authorized Representative details (for GPSR purposes): Amsterdam University Press, Nieuwe Prinsengracht 89, 1018 VR Amsterdam, The Netherlands. www.aup.nl

CONTENTS

List of Illustrations.....vii

Acknowledgements..... ix

Abbreviations..... xi

Introduction 1

Chapter 1. Do the Humanities Dream of Patterns?11

Chapter 2. Transcription, Scribal Modelling, and
Artificial Intelligence in Medieval Studies.....43

Chapter 3. Assessing Manuscript Co-Creation
Using Computational Methods83

Chapter 4. Towards a Future of Collaborative Medieval Studies..... 117

Conclusion..... 155

Bibliography 161

Index 179

LIST OF ILLUSTRATIONS

Figures

- Figure 1. A network visualization illustrating prologues shared (beyond the sixty-four common prologues) by seven manuscripts.34
- Figure 2. Density plots illustrating the frequency of use of ∂n^* opposed to ∂min^* in four manuscripts.....76
- Figure 3. Density plots illustrating the frequency of use of pze^* opposed to pre^* in four manuscripts.77
- Figure 4. Four images from a single manuscript, Mazarine, MS 6, comparing suspected shifts in scribe.....86
- Figure 5. Rolling stylometry analysis of the different hands identified in the manuscript CCC, MS 4995
- Figure 6. A 2D Principal Component Analysis plot of a TF-IDF weighted analysis of HTR-created transcriptions of two manuscripts. 100
- Figure 7. A 3D Principal Component Analysis plot of a TF-IDF weighted analysis of HTR-created transcriptions of three manuscripts. 101
- Figure 8. A 3D Principal Component Analysis visualization of a TF-IDF weighted analysis of character 4-grams containing special letterforms and abbreviations in three manuscripts 106
- Figure 9. A 3D Principal Component Analysis (PCA) of a TF-IDF weighted analysis of HTR-created transcriptions of six manuscripts. . 111
- Figure 10. (1) The impact of co-authorship on citation counts between the 1900s and 2024; (2) The impact of co-authorship on citation counts between 1990 and 2024; (3) The impact of co-authorship on citation counts per year between the 1930s and 2024; (4) The evolution of the number of authors between 1940 and 2024..... 127

Tables

Table 1. A list of the sixty-four common prologues, with the English names of the biblical books and the corresponding Stegmüller identifier, based on Lambeth MS 1364.25

Table 2. A selection of eighteen Paris bibles, their origins and dates, and a calculation of the percentage of prologues they contain of the most common sixty-four prologues.33

Table 3. The number of abbreviations and abbreviated words used in sample lines of text from the incunable Beinecke ZZi 56.69

Table 4. Three glyphs found in Beinecke, MS 387, their probable corresponding characters and sample Unicode codepoints we use to transcribe them.75

Table 5. List of the computational experiments in Chapter 3, organized by their appearance in the chapter.85

Table 6. Data from our visual analysis of the hand changes in CCC, MS 49, an English bible in the Paris bible style dated to ca. 1270–1280.93

Table 7. List of the pages that we automatically transcribed from three manuscripts using Transkribus for Experiment 3. 104

Table 8. An overview of previous scholarly attributions of scribal identity along with our assessment of the identification with Grusch as copyist for six manuscripts..... 113

Table 9. Data from thirty-seven journals in medieval studies. 125

ACKNOWLEDGEMENTS

THIS BOOK EMERGED from a collaborative effort made possible by the intellectual generosity and shared commitment of many individuals and institutions. First, since this co-authored book is the fruit of many years of sustained collaboration, we would like to thank each other for the engagement and enthusiasm with which we each approached this project.

We would also like to acknowledge the support of our current and past respective employers, Louvre Abu Dhabi, New York University Abu Dhabi, and Yale University, whose infrastructure and community provided resources and space for the reflection that gave rise to our work. This project was generously supported by research funds at New York University Abu Dhabi for infrastructure, manuscript digitization, a sabbatical that provided the time, travel, and data work required for our research, and funded research assistant positions. We are grateful to Niccolò Acram Cappelletto, Alice Fournier, Amanda Robin Hemmons, and Joshua Isaac, who served as our research assistants and for their spirit of dialogue that shaped every stage of this work. Open access publication of this book was also supported by institutional funding from New York University Abu Dhabi.

To the teams behind the handwritten text recognition software we used, Transkribus, we are grateful for the open, iterative ethos of digital scholarship that made this work possible. We relied on the sustained discussions with members of the Transkribus community: Nora Barakat, Tobias Hodel, Suphan Kirmizialtin, Annemieke Romein, and Melissa Terras. We are especially indebted to the librarians and curators at New York University Abu Dhabi and Yale University, as well as to Matthew Heintzelman at the Hill Museum and Manuscript Library and Charlotte Denoël at the Bibliothèque nationale de France, whose expertise in manuscript collections, digital collections, and preservation enriched both our analysis and our methods.

We would like to extend a special thanks to Frédéric Spagnoli of the Université Marie-et-Louis Pasteur in Besançon, France, and all the students from the Master's in "Rare Books and Digital Humanities" and "Editions Numériques et Patrimoine de l'Antiquité" who participated in the Paris Bible Correct-a-Thon we organized in January 2023: Robert Lloyd, Gauri Bhagwat, Alice Fournier, Amanda Robin Hemmons, Alexandre Keyes, Lucia Sol

Bezzecchi Petroff, Marie Noirot, Anna Chemisova, Benedicta Arthur, Sharon Hassive Guerra Álvarez, Nina Jacobson, Sumeyye Topkara, Sonaj Kailas, Kateri Soulard, Jesus David Macchi Franco, Elia Coulot, Serhat Acar, Úna Fallar, and Diego Rodríguez. We would also like to thank Pierre-Emmanuel Guilleray of the Bibliothèque d'étude et de conservation at the Bibliothèque municipale de Besançon, France, for opening the library's collections to our project and hosting our participants.

We would also like to acknowledge Laura Morreale for spearheading transcription challenges in the medievalist community. Our experience participating in them helped shape our own challenge and supported our thinking which resulted in the fourth chapter of this book.

We would like to thank our editors for their patience and skill in ushering the manuscript through to completion as well as the reviewers of the manuscript for their generous comments.

David would like to thank Jeff Richards and Ray Siemens, for their constant encouragement and intellectual exchange. Estelle would like to thank Jessica Brantley, Alex Gil, and Agnieszka Rec for their generous guidance, patience, and support, and for the freedom to pursue intellectual paths extending well beyond her dissertation.

Finally, we thank our partners and communities for their patience and encouragement during the long hours of writing, coding, and revising this manuscript.

ABBREVIATIONS

Aarau KB	Aarau, Aargau Kantonsbibliothek
BAV	Città del Vaticano, Biblioteca Apostolica Vaticana
Beinecke	New Haven, Yale University, Beinecke Library
BL	London, British Library
BnF	Paris, Bibliothèque nationale de France
BnP	Lisboa, Biblioteca nacional de Portugal
CCC	Cambridge University, Corpus Christi College
Cologne	Cologne, Fondation Martin Bodmer
DigiVatLib	Digitized Collections of the Biblioteca Apostolica Vaticana
Free Library	Philadelphia, Free Library of Philadelphia
HMML	The Hill Museum and Manuscript Library
HTR	handwritten text recognition
Lambeth	London, Lambeth Palace
LAD	Abu Dhabi, Louvre Abu Dhabi
Mazarine	Paris, Institut de France, Bibliothèque Mazarine
NT	New Testament
OT	Old Testament
Sarnen KB	Sarnen, Kollegiumsbibliothek, Stiftsarchiv Muri-Gries
Schaffhausen MB	Schaffhausen, Ministerialbibliothek
St. Gallen KB	St. Gallen, Kantonsbibliothek
Stegmüller	Stegmüller, <i>Repertorium biblicum medii aevi</i> , Vol. 1
UPenn	Philadelphia, University of Pennsylvania Libraries

