

# *Lectio*



STUDIES IN THE TRANSMISSION OF TEXTS & IDEAS

1

EDITOR IN CHIEF

Gerd VAN RIEL

EDITORIAL BOARD

Anthony DUPONT

Michèle GOYENS

Marleen REYNDERS

SUBMISSIONS  
SHOULD BE SENT TO

Marleen REYNDERS

[marleen.reynders@kuleuven.be](mailto:marleen.reynders@kuleuven.be)

© 2014 Brepols Publishers n.v., Turnhout, Belgium

All rights reserved.  
No part of this publication may be reproduced,  
stored in a retrieval system, or transmitted,  
in any form or by any means, electronic, mechanical,  
photocopying, recording, or otherwise,  
without prior permission of the publisher.

Cover illustration:  
*Bible of Anjou*, Maurits Sabbebibliotheek,  
KU Leuven (Belgium), fol. 309r.  
© KU Leuven, Maurits Sabbebibliotheek, photo by Bruno Vandermeulen



D/2014/0095/142

ISBN 978-2-503-55268-2

Printed on acid-free paper

# Analysis of Ancient and Medieval Texts and Manuscripts: Digital Approaches

Edited by  
Tara ANDREWS and Caroline MACÉ



BREPOLS



# TABLE OF CONTENTS

Tara ANDREWS – Caroline MACÉ <i>Introduction</i>	11
<i>Acknowledgements</i>	15

## SECTION 1 STEMMATOLOGY

Tuomas HEIKKILÄ <i>The Possibilities and Challenges of Computer-assisted Stemmatology. The Example of Vita et miracula s. Symeonis Treverensis</i>	19
Philipp ROELLI <i>Petrus Alfonsi, or: On the Mutual Benefit of Traditional and Computerised Stemmatology</i>	43
Jean-Baptiste CAMPS – Florian CAFIERO <i>Genealogical Variant Locations and Simplified Stemma: A Test Case</i>	69
Alberto CANTERA <i>The Problems of the Transmission of the Avestan Texts and the Tools for Avestan Text Criticism (TATEC)</i>	95

# TABLE OF CONTENTS

## SECTION 2 STATISTICS AND STYLISTICS

Armin HOENEN	
<i>Simulation of Scribal Letter Substitution</i>	119
Karina VAN DALEN-OSKAM	
<i>Authors, Scribes, and Scholars. Detecting Scribal Variation and Editorial Intervention via Authorship Attribution Methods</i>	141
Francesco STELLA	
<i>Generic Constants and Chronological Variations in Statistical Linguistics on Latin Epistolography</i>	159

## SECTION 3 INTERTEXTUALITY

Linda SPINAZZÈ	
<i>Intertextual Research with Digital Variants in Musisque Deoque: A Case Study</i>	183
Samuel RUBENSON	
<i>A Database of the Apophthegmata Patrum</i>	203
Charlotte TUPMAN – Anna JORDANOUS	
<i>Sharing Ancient Wisdoms across the Semantic Web Using TEI and Ontologies</i>	213
Maxim ROMANOV	
<i>Toward the Digital History of the Pre-Modern Muslim World: Developing Text-Mining Techniques for the Study of Arabic Biographical Collections</i>	229

## SECTION 4 SCRIPT ANALYSIS

Ainoa CASTRO CORREA	
<i>Palaeography, Computer-aided Palaeography and Digital Palaeography: Digital Tools Applied to the Study of Visigothic Script</i>	247

TABLE OF CONTENTS

Eugenio R. LUJÁN – Eduardo ORDUÑA <i>Implementing a Database for the Analysis of Ancient Inscriptions: New Developments in the Hesperia Electronic Corpus of Palaeohispanic Inscriptions</i>	273
---	-----

SECTION 5  
CODICOLOGY

Ira RABIN <i>Ink Identification to Accompany Digitization of Manuscripts</i>	293
---	-----

Patrick ANDRIST <i>Going Online Is Not Enough! Electronic Descriptions of Ancient Manuscripts, and the Needs of Manuscript Studies</i>	309
---	-----

Joris J. VAN ZUNDERT <i>By Way of Conclusion: Truly Scholarly, Digital, and Innovative Editions?</i>	335
---	-----





## INTRODUCTION

Within the field of digital textual studies there is a stereotype, unjust but tenacious, that for all the work of image scanning, transcription, and creation of ever larger online archives of text, there is little that can be done with them that is novel – that, for all the digitization work of the past few decades, the result has been little more than the ability to make better concordances. How (if at all), ask the skeptics, has the digital turn truly changed the nature of our research, particularly in the field of medieval scholarship where our collections are almost never large enough to justify the term ‘big data’? In short, when the transcription is made and the edition published online, to what research use can our texts be put?

This collection of articles aims to give an up-to-date overview of the use of computer-assisted methods in several fields of scholarship dealing with ancient and medieval texts and manuscripts (from codicology and palaeography to textual criticism and literary or historical studies), across the boundaries of language and period. In moving away from theoretical debates about what the field of ‘digital humanities’ is or should be, we hoped to get a clearer picture of what textual scholars can achieve when they use computers for their research needs and purposes, and what their expectations may be in terms of the technology and developments in computational methodology.

The overview we are offering here is far from complete – that would be impossible, and it was not our goal – but in many respects we find it very stimulating. All kind of new avenues of

research are emerging, thanks to the creativity of scholars and to their interest in what digital means can offer. The research landscape that we are depicting here would have been unthinkable even 10 years ago, and it is still a dynamic landscape whose transformations are difficult to predict. It is only possible to draw some general trends from the very diverse contributions to this volume, all of which testify to the dynamism of textual scholarship.

A first trend we can discern is the implementation of methods of analysis taken from other fields of science (chemistry, physics, biology), in order for example to study the physical characteristics of manuscripts (Rabin) or to identify relationships between manuscripts that contain versions of the same text. Experiments in the field of stemmatology are the subject of several papers within this volume (Heikkilä, Roelli, Camps and Cafiero, Cantera), and each of these articles offers a quite different approach. This shows that, far from being an old-fashioned and moribund philological activity, this sub-field of textual scholarship is flourishing in the digital environment.

Statistical methods of various sorts have been applied to textual scholarship for a long time; this is by no means a new phenomenon. Even so, we see an ever-increasing sophistication of the use of statistics, not only to formulate questions that have easy mathematical answers, but more and more to address questions that have always been of interest to textual scholars but often considered impossible to answer. Several articles in this volume illustrate many ways in which statistics can fruitfully be used to interrogate textual data, from paleography and epigraphy (Castro, Hoenen, Luján and Orduña) to authorship attribution (Van Dalen-Oskam, Stella), robustly linked data and intertextuality (Andrist, Spinazzè, Rubenson, Tupman and Jordanous), and historical research (Romanov). It has become very clear that, in several fields of textual scholarship as well, we need even larger corpora of data than exist as present and appropriate tools with which to explore them. This may be the key to how we can revive and surpass the scholarship of the nineteenth century. In many ways, twentieth-century scholarship has been dominated by theory, rightly casting into doubt many of the (false) certainties built by nineteenth-century positivism. Yet today we see that

very few of those theories have survived the last millennium, whereas we may be in a position now to revise and renew our visions of the past thanks to a new way of looking at data.

At the heart of science perhaps lies a constant effort to free ourselves as much as possible from all kinds of fascinations, not only ideological illusions but also from a specific and peculiar fascination with the object under study or with the tools used in its study. In that respect we can offer here a quite mature picture of science in our field. The time when scholars were doing statistical analysis of their texts for the sake of statistics may be considered definitely over. For the same reason this volume showcases comparatively fewer tools, which quickly become outdated, than other comparable collections of papers. Our authors focus rather on practices and methodologies – the sort of analysis that can be done once the tools and data are in place.

One of the new ways of looking at data is enhanced by the possibility of linking it. This is an intellectual activity that has always been key to scholarship: making meaningful links between isolated facts, isolated texts, isolated languages. In the past those links existed in scholars' minds, to be expressed in carefully-formulated arguments within the pages of academic journals. The direct connection and presentation of linked data, and the scholastic openness and accessibility it implies, is one of the most interesting novelties in the world of digital scholarship.

Finally, another tendency is clear: through digital forms of publication we are moving away from traditional forms of dissemination of our research through print-based media, wherein academic publishers are the gatekeepers. But here we need to be careful: although some have argued convincingly for the shortcomings of the current system of research publication, it is too easy to discount the fact that academic publishers fulfill several functions for which we have not yet found a workable or convincing replacement. First and most directly, publishers handle the layout, typography, and physical dissemination of works of scholarship; at present the digital alternative usually requires the scholar to become expert at (or hire experts for) the design, implementation, and maintenance of the necessary infrastructure. Second is the sociological distinction between what is of an adequate academic or scholarly standard and what is not. This is a

debate in its own right – there are as many ways to argue that the categories are unjustified as there are rationales for their maintenance – but the arguments for and against an office of ‘gate-keeper of academic quality’ must be heard in their own right, independent of the question of new technological forms of publication. These are the two points that are the most crucial to understanding how we can make a successful shift from traditional modes of academic publication to freely-accessible resources on the web, and what role academic publishers will play in that shift.

We are very much aware that our book does not cover all possible aspects of digital textual scholarship. We hope nevertheless that this overview of some of the research projects that currently exist and are still evolving may stimulate further developments and encourage junior (and not so junior) researchers to carry on, and to add new results and new questions to those which have been presented here.



## ACKNOWLEDGEMENTS

The articles contained in the present book were presented during a workshop organized on April 2-3, 2012 in Leuven and Brussels. This workshop was financially supported by the *Tree of Texts* project, a CREA ('creative research') project (3H100334), funded by the KU Leuven from 1/10/2010 to 30/9/2012), by the COST Action *Interedition* (<http://www.interedition.eu>), by the ESF Research Networking Programme COMSt (Comparative Oriental Manuscript Studies), by the Flemish Royal Academy of Belgium (KVAB – contactforum) and by the Faculty of Arts of the KU Leuven.

In addition to the funding bodies, we would also like to thank the members of our scientific committee, who actively participated in the process of selection of the proposals and in the evaluation of the papers: Aurélien Berra (Université Paris-Ouest), Thomas Crombez (Universiteit Antwerpen), Juan Garcès (Göttingen Centre for Digital Humanities), Tuomas Heikkilä (University of Helsinki), Torsten Schaßan (Herzog August Bibliothek Wolfenbüttel), Frederik Truyen (KU Leuven), Dirk Van Hulle (Universiteit Antwerpen). We are also grateful to Brian Garcia (KU Leuven), who assisted us with the copy-editing of the proceedings, and Thomas Leibundgut (Universität Bern), who assisted with the preparation of the final volume.



PATRICK ANDRIST

## GOING ONLINE IS NOT ENOUGH! ELECTRONIC DESCRIPTIONS OF ANCIENT MANUSCRIPTS, AND THE NEEDS OF MANUSCRIPT STUDIES\*

*How important is the date of a manuscript for scholarship?*

While preparing this paper, I asked a few colleagues – philologists, historians, codicologists, but also managers of manuscript collections – what information they most want to find when looking at an online catalogue of ancient manuscripts. The question was: ‘When you look at such a catalogue, which are the five pieces of information you would most like to find?’ Depending on their need or their specialty, some mentioned the bindings or the materials, others the decorations, the miniatures, or the availability of reproductions. The content was always important, even if one respondent did not mention it. But the one piece of information that was consistently mentioned, one way or another, was the date of the manuscript or, more precisely, the date of the manuscript part one was working on.

In a spirit of humour, I sometimes asked: ‘Do you just need a date? Or do you need a correct date?’ In other words, how important is it that the date of a manuscript is more or less correct? For manuscript scholars, the answer is crystal clear. As everyone knows, in most cases it is crucial to know or at least to have a good idea of when a manuscript was copied or painted because most of the manuscript-related disciplines – if not all of them – are historical ones.

\* I am very thankful to the organisers of the conference for their initiative, particularly to Tara Andrews, who very nicely reviewed the language in depth; to the ESF and the Interedition for their financial support; to Marilena Maniaci and André Binggeli for usefully discussing several points with me. Of course, I alone assume responsibility for the many shortcomings of this article.

For example, for a book historian studying the development of book techniques, it is obviously useful to know when the books are dated. For an art historian studying paintings at the court of Otto I the Great, it certainly makes a difference if a miniature was painted in the tenth or the eleventh century. If one is editing ancient texts and there are several witnesses to this text, the dates of the manuscripts play a major role in sorting out the variants one way or another. It is even more critical if one for example studies the reception of a work in the fourteenth century. In a word, as soon as there is some kind of historical approach, the date of the object is significant for the overall scholarly appreciation of the codex.

(1) Thus, our first, obvious claim is:

*Manuscript studies need information that is as correct as possible, particularly about the dates of the manuscripts.*

... but this is not enough!

A situation scholars are frequently faced with is the fact that many manuscripts do not have one, but several dates. This is not a question of undated manuscripts because one can often figure out more or less when they were copied or painted by analysing the materials or the script. Nor is it a question of bindings, which are almost always much younger than the written or painted part of the volume. Rather, it is a question of volumes made up of several parts, with not all of them produced at the same time or in the same place.

The difficulty of cataloguing manuscripts in a way that complex manuscripts are also satisfactorily described is not new; scholars have often pointed this out and made suggestions to solve the problem.<sup>1</sup> This paper extends the debate to online descriptions and discusses several ways to prepare them, in order to present clearly and unambiguously relevant chronological information. This is not a comparison of database formats (EAD, Marc, etc.), specific software or databases, even though some software solutions make it easier to unambiguously describe a codex. This is about ways to organise the information visually – and, when

<sup>1</sup> Thorndike 1946, p. 94–95; Robinson 1980, p. 61–69; Derolez 1995, p. 381–383; Gumbert 1997; Munk Olsen 1998, p. 128–129; Andrist 2006; Gumbert 2009 etc.



possible, structurally – so that the users of the online description can (1) immediately and easily understand which data (concerning content, writing support, layout, ink, etc.) belong to the same production unit, and do not confuse them with data belonging to other production units,<sup>2</sup> and can (2) automatically retrieve correct information when including a date in their search queries.

### *The case of the multi-date codex Barocci 33*

To illustrate the problem of multi-date manuscripts, let us consider a very clear example, albeit one already used elsewhere.<sup>3</sup>

*Barocci 33* is a Greek manuscript in the Bodleian Library in Oxford. It is made up of three main parts, which can be called ‘Production Units’ (or PUs).

The oldest PU is B, the second part of the current volume. It was copied by Michael Lygizos in the second half of the fifteenth century (see figure 1 below).

A, the PU that makes up the first part of the volume, was copied in the first half of the sixteenth century by two scribes, one of whom was called Kōnstantios, and was prepended to the older PU. It may even have been copied with this intention, but this is difficult to demonstrate.

Finally, in 1595, Nicolaos Labros copied C, the third PU, and added a colophon at the end, including his name and the date. Eventually, or perhaps even immediately, it was appended to the other PUs of the current volume.

A. s. XVI <sup>1/2</sup>	A. (f. 1–116) hands a–b.: incl. Kōnstantios Text 1: M. Blastaris
B. s. XV <sup>2/2</sup>	B. (f. 117–244) hand c.: Michael Lygizos Texts 2–6: incl. Georgius Gemistius Plethon, <i>opera misc.</i>
C. a. 1595	C. (f. 245–(421)) hand d.: Nikolaos Labros Texts 12–13 f. 418v: colophon including the date and the scribe name

FIGURE 1: Overview of Barocci 33.

<sup>2</sup> On the concept of Production Units, see Andrist, Canart & Maniaci 2010, and 2013, especially p. 59–61. See also Gumbert 2004. Summarised explanations in English can also found in Andrist 2011.

<sup>3</sup> Andrist 2010, p. 19–21.

The result is a volume that ends with the name of a scribe and a date. By browsing the bibliography of the codex, it becomes apparent that more than one scholar believed the date applied to the entire manuscript, and mistakenly attributed information from PU A and PU B to 1595. What difference does it really make? Does it make a difference if these texts are linked to the date of PU B or PU C for a philologist preparing the edition of Plethon? Again, the answer is obvious. In the first case, this manuscript would have been copied a few years after the author's death in 1452; in the second case, it would have been copied some 150 years later.

In reality, these scholars are victims of the description of Henry Coxe in 1853 (see figure 2 below)<sup>4</sup>, who plainly states at the beginning of his description that the codex was written in 1595. One cannot, however, say that his date is wrong because it can be found in the manuscript. In reality, this is a problem of correctly linking the information about the date with other pieces of information from the codex, such as the content, the scribes, and so forth.

(2) Our second claim is:

*Manuscript studies need the various types of information in a description to be correctly linked to each other.*

... it is already good, but it could even be better!

### 33.

Chartaceus, in 4to minori, ff. 418, sec. xvi. exeuntis :  
manu sacerdotis ejusdam anno 1595 scriptus.

1. ΜΑΤΘΙΑΙ Hieromonachi adversus Judæos libri quinque, prævia cuique capitulorum tabula. fol. 1.

Tit. τοῦ σοφωτάτου ἐν ἱερονομάχοις Ματθαίου  
κατὰ Ἰουδαίων λόγος πρῶτος.

Sequuntur versus isti iambici,  
Ματθαῖος εἶργει τῶν Ἰουδαίων θράσος,  
ὥσπερ χαλινούς πέντε φιμώσας λόγοις,  
ὅστις δὲ τούτων τὴν ἐπίρρητον πλάνην  
πλάνην ἀτεχνῶς ἐλέγξει τῷ λόγῳ,  
ἄρδην ἀπάσας συγκαθεῖλεν αἰρέσεις,  
μήτηρ γὰρ αὐτῶν ἡ θεοκτόνων ἔρις.

Tit. cap. i. ὅτι τρισυπόστατον καὶ ἡ παλαιὰ τὸν  
Θεὸν κηρύττει γραφὴ ἐν Πατρὶ καὶ Υἱῷ καὶ ἁγίῳ

FIGURE 2: Beginning of Henry Coxe's description, in 1853.

<sup>4</sup> Coxe 1853.

How frequently do these kinds of situations occur? Where Greek manuscripts are concerned, experience shows that manuscripts quite frequently contain multiple PUs comprising their main written or painted parts.<sup>5</sup> When I was curator of the Bongarsiana in Bern, I took some time to look into Arabic, Hebrew, German and old Latin manuscripts, and found that many of them were affected by the same phenomenon. As far as medieval manuscripts are concerned, as soon as one starts looking for these kinds of situations, one seems to find them almost everywhere.

*Online description type 1a:  
A bad interpretation of the mono-layer model*

Before considering a third claim, let us see how the information about Barocci 33 might be correctly presented in an online description. We will consider the two most widely used models for preparing description patterns – although several other models are possible.

The first one is the standard mono-layer model,<sup>6</sup> which has been in use in printed catalogues since the middle of the twentieth century and is the most frequently used model on the Internet today. In this model, the various descriptive categories follow each other, each one in one and only one field per description (cf. below).

Let us consider four different ways to work with such a model, taking Barocci 33 as an example. In the following (incomplete) simulations, the same exact pattern, based on the mono-layer model, is completed in four different ways.

The first way represents the worst solution and should never be used:

In this first example, the fields were completed following the same principle as the catalogue of Coxe (but including some extra information). Failure to mention the applicable limits of the given date and to give a correct dating for PU A and PU B makes this description very flawed. The structure of the manuscript can

<sup>5</sup> This means excluding the notes and the bindings; otherwise about 100% of the manuscripts would be concerned.

<sup>6</sup> I no longer use the expression 'analytic model' because of the expression 'catalogo analitico' in Italian, which is used to refer to a "full-scale" catalogue; see Petrucci 2001, p. 93-104.

1a. - A bad example
Signature
<b>Barocci 33</b>
Extent
(3) 421 (1) folios
Date
March 14, 1595
Content
1. (f. 1r-116r sup.) Matthaeus Blastaris, <i>Adversus iudaeos</i> 2. (f. 117r-209v sup.) Gennadius Scholarius, <i>Adversus iudaeos</i> 3-5. (f. 210r-230v) Georgius Gemistus Plethon, <i>opera misc.</i> , cf. cat. 6. (f. 231r-241v) Isocrates, <i>Ad Demonium oratio paraenetica</i> 7-11. (f. 244v) notes, by Andreas Dónos, cf. cat. 12. (f. 247r-414v) Gennadius Scholarius, <i>Apologia pro quinque cyp. concili Florentini</i> 13. (f. 415r-418v sup.) <i>Definitio synodi Florentini</i>
Material
Several kinds of paper: including watermarks "Cardinal Hat", "Pair of scissors" ...
Scripture and hands
a. (f. 1-24) unknown hand b. (f. 25-116) hand of <Kónstantios> c. (f. 117-244) hand of <Michael Lygizos> d. (f. 247-421) hand of <Nicolao Labros> (see colophon, f. 418v)
Etc.

FIGURE 3: An unsatisfactory method to describe Barocci 33 using the mono-layer model.

only be deduced if the reader notices that both the hands and the texts change at the same point in the manuscript. Only readers who know the activity dates of the scribes can form an idea about the chronological distance between these parts.

Unfortunately, a fair amount of examples of this description type can still be found online, as well as in printed catalogues.



*Online description type 1b:*

*An acceptable interpretation of the mono-layer model.*

A different method to complete the same simulated scheme in an acceptable way, which is not software-dependent and can be set

up in any XML or SQL type of database, is offered below. It is simply a matter of awareness, standards and systematics.

1b. - An acceptable example	
Signature	
<b>Barocci 33</b>	
Extent	
(3) 421 (1) folios	
Date	
(f. 1-116) S. XVI 1/2 ; (f. 117-244) S. XV 2/2 ; (f. 245-421) a. 1595, 14th of March	
Content	
1. (f. 1r-116r sup.) Mattheus Blastaris, Adversus iudeos 2. (f. 117r-209v sup.) Gennadius Scholarius, Adversus iudeos 3-5. (f. 210r-230v) Georgius Gemistus Plethon, opera mis, cf. cat. 6. (f. 231r-241v) Isocrates, Ad Demonium oratio paraenetica 7-11. (f. 244v) notes, by Andreas Dônos, cf. cat. 12. (f. 247r-414v) Gennadius Scholarius, Apologia pro quinque cyp. concili Florentini 13. (f. 415r-418v sup.) Definitio synodi Florentini	
Material	
Several kinds of paper: including watermarks "Cardinal Hat", "Pair of scissors" ...	
Scripture and hands	
a. (f. 1-24) unknown hand b. (f. 25-116) hand of <Kônstantios> c. (f. 117-244) hand of <Michael Lygizos> d. (f. 247-421) hand of <Nicolao Labros> (see colophon, f. 418v)	
Etc.	

FIGURE 4: An acceptable way of describing Barocci 33 following the standard mono-layer model.

In this second example, the three PUs and their extent are mentioned in the 'date' field. The reader can immediately understand the basic structure of the codex. He must then remember this structure when reading the other fields of the description; he should not, for example, take for granted that the text of Mattheus Blastaris was originally copied into this same manuscript

with the texts of Plethon, even though they are bound together today.<sup>7</sup>

Thankfully, a fair number of such acceptable descriptions can already be found on the Internet.

*Online description type 1c:  
An improved example*

Provided that the programmers did not limit the available space of the database field, the same pattern can be used slightly differently to better underline the natural structure of the manuscript. Because it implies that one reads, understands and reflects on the physical features of the codex, this constitutes a so-called ‘syntactical descriptions’.<sup>8</sup> Again, this improved method is not software-dependent and can be set up in any XML or SQL type of database.

In this third example, not only are the extent and the dates of the PUs mentioned at the beginning, but the separate parts are also numbered from A through C. To prevent mistakes and confusion, this numbering is repeated in every field and used as a way to structure the information. As a result, readers can always very clearly know which PU in the manuscript should be associated with the given information.

In spite of this important improvement, this type of description still has some disadvantages.

First, repeating the structure in every field takes up time, energy and space. And there is always the risk of mistakes because of imperfectly repeated information. One can easily imagine what happens if there are 20 PUs in a manuscript.<sup>9</sup>

Second, it does not allow the reader to take in all the characteristics of a PU at once. One has to browse all the description

<sup>7</sup> Another ‘danger’ this type of description poses – though not illustrated by the example above – is the tendency of the cataloguers to mention two subsequent PUs from the same date in one shot, even if they were produced in a very different context.

<sup>8</sup> ‘Syntactical description type C’ in Andrist 2015; on the concept of ‘syntax of the codex’, see Andrist, Canart & Maniaci 2013, p. 9.

<sup>9</sup> In the below scheme, the displays of the ‘content’ and ‘material’ fields were limited by the programmer, but this kind of problem does not normally occur in standard Internet software. In the above example, it was compensated through use of a scrollable field.

## 1c. - An improved example

Signature
<b>Barocci 33</b>
Extent
421 f. = (3) 116 + 128 + 177 (1) folios
Date
A. (f. 1-116) S. XVI 1/2 B. (f. 117-244) S. XV 2/2 C. (f. 245-421) a. 1595
Content
A. (f. 1-116) 1. (f. 1r-116r sup.) Matthaeus Blastaris, Adversus iudeos B. (f. 117-244) 2. (f. 117r-209v sup.) Gennadius Scholarius, Adversus iudeos 3-5. (f. 210r-230v) Georgius Gemistius Plethon, opera mis 6. (f. 231r-241v) Isocrates, Ad Demonium oratio paraenetica
Material
Paper A. (f. 1-116) Watermarks: 2 series of "Cardinal's hat" with countermark "B" ... B. (f. 117-244) <del>Watermarks: 2 series of "Cardinal's hat" with countermark "B" ...</del>
Scripture and hands
A. (f. 1-116) a. (f. 1-24) unknown hand b. (f. 25-116) hand of <Kōnstantios> B. (f. 117-244) c. (f. 117-244) hand of <Michael Lygizos> C. (f. 245-421)
Etc.

FIGURE 5: An improved way to describe Barocci 33 following the standard mono-layer model.

fields if he wants to link the date, the hand, the layout, the decoration, etc., of the PU he is interested in.

I have not been able to find examples of systematic use of this model online, although it has been successfully used in printed catalogues.<sup>10</sup> There are, however, hybrid solutions between type 1b and 1c, both online and in print, where the PUs are systematically and clearly distinguished only in the content field.<sup>11</sup>

<sup>10</sup> See, e.g., Kerstin 2003; also Kouroupou & Géhin 2008.

<sup>11</sup> See, e.g., the description of Zürich Rh. 97 in Mohlberg 1952, p. 206, or Charfet, Rahmani 151 (Catalogue SONY 775) in the database e-ktobe of Syriac manuscripts (<http://www.mss-syriaques.org/>).

*Online description type 1d:  
Another improved way of interpreting the mono-layer model*

Another way to clearly distinguish the PUs of a manuscript, which does not require illustration here, is to prepare one description per PU, instead of one description for the entire manuscript. This solution entirely avoids the risk of ‘mislinking’ pieces of information that do not belong together, without overcharging the fields. In the case of Barocci 33, this would mean writing three descriptions. This is another type of ‘syntactical description’.<sup>12</sup>

Where should one then describe the common elements like bindings, running notes, common history, bibliography, etc.? Several options are possible:

- An ill-advised solution would be to repeat the general information in every description.
- A better solution would be to for example add this to the first or the last description in the series. One must then design or complete the scheme in such a way that users clearly know they are not accessing the entire description, but just one PU, even though all the fields are completed.
- A third solution would be to create an extra description specifically for the common elements. The reader would then only see partly completed descriptions – either with information about the volume or about the individual PUs.

In any case, a one-to-one relationship between the number of descriptions and the number of volumes being described would be lost. In many situations, however, this need not pose a problem.

More problematically, it is then much more difficult to get an overview of the manuscript, particularly when the online description was obtained through a search query. If there is no internal link to the other description records within the series that make up the manuscript, there is no way of accessing them without completing a new search. Again, this is very much tied to how the database was designed in the first place.

<sup>12</sup> ‘Syntactical description type B’ in Andrist 2015.



This type of solution has sometimes been used for online catalogues, in particular for descriptions of volumes containing fragments, but also occasionally for other types of multi-PU volumes.<sup>13</sup> For example, interesting cases can be found in the Marburger Repertorium, a specialised catalogue dedicated to German manuscripts from the thirteenth and the fourteenth century, where fragments currently found in the same binding are described separately from each other but, when relevant, together with the other fragments of the same original manuscript now dispersed among different libraries.<sup>14</sup>

### *Introduction to the bi-layer model*

As we just saw, satisfactory descriptions can be achieved through proper use of the mono-layer model, no matter what kind of database is being used. However, each way of representing the structure of the manuscript results in some visibility loss of information, and/or a loss of time. To move beyond these shortcomings, one needs another model based on a different approach to the description, such as the bi-layer model.

Conceptually, it is an archaeological or stratigraphic approach: the manuscript is considered a physical object made up of several other embedded physical objects, which must be separately described. In the upper level, there is only one object: the current volume, which is also a Circulation Unit. In the second level, there are either Production Units or older Circulation Units which, again, can comprise simple Production Units or even older Circulation Units, etc. Of course, the reality can be more complex, but this simplified explanation satisfactorily underlines the two main objects of a complex manuscript: the current vol-

<sup>13</sup> For a very interesting, systematic and compact use of this approach in a printed catalogue, see Gumbert 2009a and 2009b. For a hybrid solution between type 1d and 2a in the database e-ktobe, see below footnote 18.

<sup>14</sup> <http://www.mr1314.de/>; see for example the descriptions of Berlin, Staatsbibl., mgf 734 in nine entries. When users click the first one (mgf 734 Nr. 1), they get the independent description of the first fragment. However, when they click the second entry (mgf 734 Nr. 2), they very interestingly receive a description of three fragments of the same codex, today in Berlin, Nürnberg and Stuttgart. For another description of the same current volume in Berlin, see below, footnote 18; concerning the Marburger Repertorium, see Andrist 2010, p. 22–26.

ume and the several embedded Production Units. Each of them is then separately described.

In practice, the result is to set up two description patterns: one pattern for the volume and its general information, the other pattern for the PUs. Consequently, a codex description comprises a single volume description using the volume pattern and one PU description for each PU in the manuscript, making repeated and subsequent use of the PU pattern. Information about the production date and contents are given in the PU layer only. In a sense, the PUs – not the manuscript – are the main entities to be described.

Of course, one can think of an improved multi-layer approach, in which each level is separately described, no matter how many there are. So far, I have seen no Internet or print solution do this, even though recursion is very common in computer programming. Multi-layered descriptions are somewhat easier to prepare for a printed description, but not necessarily always easier to understand.

I have also done some tests with a five-pattern solution, including one specific pattern for the binding, directly linked to the volume layer; another for the texts, underneath the PU layers; and a fifth one for supplementary elements such as notes or added slips, also linked to the volume layer.<sup>15</sup> However, this has often resulted in online descriptions that are difficult to set up and understand.

It is worth again emphasising that this bi-layer model is not tied to any specific software or database. Although, of course, not every software package or database allows for it. It is entirely independent of implementation technology and can be achieved through storing descriptions in a relational database, or through extensive XML tagging. It is very suitable for any type of electronic approach and, indeed, it has already been used online with many different software packages in many different ways – some examples are given below.

It is indeed noteworthy that the rules for the software Manuscriptum XML, dedicated to the description of manuscripts, en-



<sup>15</sup> Even though the binding and the notes can be considered separate levels, this is not the case for texts. It was convenient, however, to describe them with their own pattern. For ways to treat notes in a bi-layer approach, see below p. 326–327.

tirely integrate the concept of multi-part manuscripts, and that the software provides means to describe complex manuscripts at various levels through the notion of 'Blocks'.<sup>16</sup> Let us hope that cataloguers will use it with the same understanding of a multi-part manuscript and, as a result, produce coherent multi-layer descriptions!

*Online description type 2a:  
An all-inclusive interpretation of the 'bi-layer' model*

In this first type, all the information relating to one PU is simply grouped in one occurrence of the PU pattern, including all its texts, its paintings, its scripts, its material aspect, etc. This is the most simple and basic type of syntactical description.<sup>17</sup>

There are then two basic ways of linking the information about a PU and its containing volume:

- They can all be displayed in the same window, one after the other. If the software allows for this, the PU information can even be embedded into the general information, so that the applicable parts of the general description (e.g. the signature, the material information and the binding) come at the beginning; followed by the PUs and with the remaining general information such as history and bibliography at the end.
- Alternatively, the catalogue can apply a 'click-and-jump' technique: one obtains the description of a particular PU by clicking the appropriate link in the volume description or in another PU description. This technique is often easier to realise, but it is then more difficult to keep an overall view of the current codex. In any case, it is clearly useful to repeat the codex signature at the top of each PU description.

For example, conceptually, an all-inclusive bi-layer description of Barocci 33 would look like this:

<sup>16</sup> Riecke 2009 (available online at <http://www.manuscripta-mediaevalia.de/hs/handbuch.pdf>). On this project, funded by the Deutsche Forschungsgemeinschaft (DFG), see Giel 2010.

<sup>17</sup> 'Syntactical description type A' in Andrist 2015.

Layer 1

## General information

*Signature:* Barocci 33  
*Binding:* ...  
*History:* ... bought by the library in 1629 ...  
*Bibliography:* cat. Coxe ...  
 etc.

Layer 2

PU 2  
(f. 117-244)

s. XVI ½  
*Content:*  
 1.(f. 1r-116r) M.  
 Blastaris ...  
*Scribes:*  
 a. (f. 1r-24v) aunknown  
 hand  
 b. (f. 25r-116r)  
 <Kōnstantios>  
*Material:*  
 Paper: watermarks...  
 etc.

PU 1  
(f. 1-116)

s. XV 2/2  
*Content:*  
 2.(f. 117r-209v) Gennadius...  
 3-5. (f. 210r-230v)  
 G.Gemistius Pletho...  
 6.(f. 231r-241v) Isocrates, ...  
 7-11. ...  
*Scribes:*  
 c. Michael Lygizos...  
 etc.

PU 3  
(f. 245-421)

1595  
 etc.

FIGURE 6: The basic structure of a bi-layer stratigraphic description of Barocci 33.

An online description could then look like this:

First, the overall volume description:

## 2a – Overall volume description

Signature

**Barocci 33**

Extent

421 f. = (3) 116 + 128 + 177 (1) folios

A. (f. 1-116) Matthaeus Blastaris

B. (f. 117-244) Gennadius Scholarius; Isocrates

C. (f. 245-421) Gennadius Scholarius; Definitio synodi Florentinae

Etc.

(including the description of the Binding, Notes, History, Bibliography...)

Then the three descriptions of the three Production Units follow:

## 2a – Description of Production Unit A

Heading
<b>A. (f. 1-116) Matthaëus Blastaris</b>
Date
(f. 1-116) S. XVI 1/2
Content
1. (f. 1r-116r sup.) Matthaëus Blastaris, Adversus iudeos
Material
Paper with watermarks: 2 series of "Cardinal's hat" with countermark "B" ...
Scripture and hands
a. (f. 1-24) unknown hand b. (f. 25-116) hand of <Kōnstantios>
Etc.

## 2a – Description of Production Unit B

Heading
<b>B. (f. 117-244) Gennadius Scholarius; Isocrates</b>
Date
S. XV 2/2
Content
2. (f. 117r-209v sup.) Gennadius Scholarius, Adversus iudeos 3-5. (f. 210r-230v) Georgius Gemistius Plethon, opera mis 6. (f. 231r-241v) Isocrates, Ad Demonium oratio paraenetica 7-11. (f. 244v) originally empty; today with notes, cf. infra Supplementary elements
Material
Paper with watermarks: 2 series of "pair of scissors" ...
Scripture and hands
c. Hand of <Michael Lygizos>
Etc.

2a – Description of Production Unit C	
Heading	<b>C. (f. 245-421) Gennadius Scholarius; Definitio synodi Florentinae</b>
Date	a. 1595 (see colophon, f. 418v)
Content	12. (f. 247r-414v) Gennadius Scholarius, Apologia pro quinque cpp. concilii Florentini 13. (f. 415r-418v sup.) Definitio synodi Florentinae
Material	Paper with watermarks: not recognised
Scripture and hands	d. Hand of <Nicolaos Labros> (see colophon, f. 418v)
Etc.	

FIGURE 7: A stratigraphical way to describe Barocci 33 following the bi-layer model.

The result is very clear, and it is difficult to mix information that is not contextually relevant – as is also illustrated by the following description. It was created in the database ‘Archives et manuscrits’ of the Bibliothèque nationale de France for codex Paris gr. 1823.<sup>18</sup>

<sup>18</sup> (<http://archivesetmanuscrits.bnf.fr/>); unfortunately, there is no permanent link to these descriptions. See also, on the website for the project ‘Bibliotheca Laureshamensis-digital’ (<http://www.bibliotheca-laureshamensis-digital.de/de/index.html>), the description of codex Vaticanus Pal. lat. 554, in two ‘Faszikel’. Remarkably enough, this project entirely takes into account the multi-part manuscripts, as explained in the online presentation: ‘Handelt es sich um einen Codex, der aus mehreren ursprünglich selbständigen Teilen zusammengesetzt ist, wozu auch Fragmente gezählt werden, weichen die Beschreibungen im Aufbau etwas ab [...]’ (<http://www.bibliotheca-laureshamensis-digital.de/de/projekt/erschliessung.html>). Let us again hope that these principles will be applied in a very systematic way! A convincing hybrid solution between type 2a and 1d is sometimes found in the database e-ktobe (see above footnote 11), where the two layers are in fact realised with the same pattern; see for example the description of manuscript Paris BnF syr. 434. An interesting use of the bi-layer model can also be found in the presentation of Berlin, Staatsbibl., mgf 734 in Manuscripta Mediaevalia (<[www.manuscripta-mediaevalia.de](http://www.manuscripta-mediaevalia.de)>; permalink: <http://www.manuscripta-mediaevalia.de/dokumente/html/obj31250686,T>). The layers for each fragment are made by repeatedly using two embedded standard fields and increasing their numbering. Concerning Manuscripta Mediaevalia, see Andrist 2010, p. 27–31.

## Olympiodore/Damascius + Psellos + Timée de Locres + Théophraste

Paris BnF Grec 1823

[II] + 176 + (6) + [III]

345 × 235 × 50 (feuillet : 328 × 230)

Manuscrit formé par la réunion de quatre parties dues à des copistes différents.

Toutes les parties du recueil sont constituées de papiers filigranés à fines vergeures et pliées in-folio.

Reliure François Ier à couverture de maroquin vert sombre virant au brun sur le dos. Technique alla greca : dos plat, coiffes saillantes sur les tranche-filés, pas de chasse, chants rainurés, vestiges de quatre fermoirs montant du plat inférieur vers le plat supérieur (deux en gouttière, un en tête et en queue). Décoration des plats : combinaison d'un décor imprimé à froid de filets dessinant trois rectangles emboîtés et d'un décor doré s'insérant dans ces rectangles : entrelacs, fleurs de lys, monogramme F surmonté d'une couronne et, au centre des plats, grand fer aux armes de France entourée du collier de l'ordre de Saint-Michel et avec la salamandre dans un brasier, emblème de François Ier ; la dorure se présente en deux couleurs, l'une plus jaune l'autre plus grise, les éléments

Informations sur le traitement

Notice rédigée par Jocelyn Groisard (août-septembre 2007).

## Olympiodore/Damascius

I = ff. 1-130

1536

13 cahiers, tous des quinions. Signatures de α' à ιγ' de première main dans l'angle inférieur externe du premier recto du cahier. Réclame verticale dans l'angle inférieur interne du dernier verso du cahier.

Filigranes : I.1) Deux flèches croisées surmontées d'une étoile : cahiers 1-5 ; paire distincte aux ff. 10 et 25 ; = Harlfinger Flèche 21 (et non Flèche 19 comme indiqué par erreur dans ce répertoire), papier utilisé par le même copiste en 1538 et 1539 ; I.2) Deux flèches croisées surmontées d'une étoile, différent du précédent : cahier 6 ; paire distincte aux ff. 56 et 60 ; I.3) Ancre inscrite dans un cercle surmonté d'une étoile : cahier 5 ; paire distincte aux ff. 41 et 42 ; probablement identique à Harlfinger Ancre 12, papier attesté en 1534 à Venise ; I.4) Ancre : cahiers 7, 9 et 11 ; paire distincte aux ff. 61 et 63 ; I.5) Échelle inscrite dans un cercle surmonté d'une

Toutes les notes marginales sont des scholies copiées de première main à l'encre rouge, ou de plus rares corrections également de première main mais à l'encre noire (par ex. ff. 29r, 44v, 50r, 59v, 76r).

Présentation du contenu

(ff. 1r-29r) Olympiodore, Commentaire sur le Phédon de Platon, éd. L. G. Westerink, *The Greek Commentaries on Plato's Phaedo. Volume I : Olympiodorus*, Amsterdam-Oxford-New York, 1976 : [titre dans marge externe] Ὀλυμπιόδωρου φιλοσόφου σχόλια εἰς τὸν πλάτωνος φαίδωνα, suivi de l'indication d'une lacune initiale de 6 feuillets : λείπει δὲ τοῦτοις τὰ ἐξ ἀρχῆς φύλλα ἕξ ; [explicit f. 29r, l. 17] οὐ ζωοῦσιν αὐτά.

(ff. 29r-82r) <Damascius>, Commentaire sur le Phédon de Platon, éd. L. G. Westerink, *The Greek Commentaries on Plato's Phaedo. Volume II : Damascius*, Amsterdam-Oxford-New York, 1977, pp. 25-285 :

## Psellos

II = ff. 131-142

Début des années 1540

2 cahiers : 1 × 8 (138). 1 × 4 (142). Signatures de α' à β' au milieu de la marge inférieure du premier recto du cahier.

Filigrane : II) Deux flèches croisées orientées vers le bas et surmontées d'une grande étoile, avec la contre-marque 3M surmontée d'un trèfle ; paire distincte aux ff. 142 (contre-marque f. 139) et 131 (contre-marque f. 138) ; = Harlfinger Flèche 24a-b, papier attesté à Venise en 1542.

Un seul copiste dont l'écriture présente quelques similitudes avec celle de Georges Basilikos (à qui cette partie

etc.

FIGURE 8: An online description of Paris gr. 1823 following the all-inclusive interpretation of the bi-layer model.

In the above example, the general information comes first. There is no date at the beginning, but a note explains that the codex comprises four parts copied by three scribes. Each of the four parts is then described separately, with the date included toward the beginning of the description. In printed catalogues, this type of solution has been advocated as early as 2003<sup>19</sup> and is becoming more and more widespread.<sup>20</sup>

In the last few years, with the help of some colleagues, I have had the opportunity to test different systems. I was able to write syntactical descriptions using very different databases such as HAN, based on a Marc format system; scopeArchiv, which is close to the EAD philosophy; Nuova Biblioteca Manoscritta (NBM), a MySQL database with a Java front end; or e-codices, which strictly follow the TEI XML standards.<sup>21</sup>

It is important to stress that, contrary to a widespread idea, a bi-layer syntactical description is not necessarily a long one.<sup>22</sup> It is simply a matter of organising what little information there might be in a meaningfully structured way.

There is always some question about where to describe any reader and owner notes. Since they are always physically written on folios that belong to a PU (or to elements of the binding), should they be put in the PU layer (or with the binding)? This would usually make sense for notes relating to a particular text or to a particular owner of the PU before it was bound with the other PUs. Or should they be described in the volume layer, since they usually were not produced at the same time as the texts and because they often relate to the current volume and not to a

<sup>19</sup> Andrist 2003.

<sup>20</sup> See, for example, Andrist 2007, Stutzmann & Tylus 2007, Bobichon 2008, Dukan 2008, Di Donato 2011, Del Barco 2011.

<sup>21</sup> Unfortunately, most of these tests are no longer available online. Concerning these systems: HAN: [http://aleph.unibas.ch/F?con\\_lng=GER&func=file&local\\_base=DSV05&file\\_name=verbund-han](http://aleph.unibas.ch/F?con_lng=GER&func=file&local_base=DSV05&file_name=verbund-han); scopeArchiv: for example in the Burgerbibliothek Bern, <http://katalog.burgerbib.ch/suchinfo.aspx>; on this experiment, see Andrist 2010, p. 40–43; NBM: [www.nuovabibliotecamanoscritta.it](http://www.nuovabibliotecamanoscritta.it). For more about NBM, see Andrist 2010, p. 35–40. A few syntactical descriptions are available in e-codices, for example in Basel, Universitätsbibliothek, Cod.A.VII 3: <http://www.e-codices.unifr.ch/en/description/ubb/A-VII-0003>.

<sup>22</sup> See for example, the description of Bodm. 115 on e-codices, <http://www.e-codices.unifr.ch/fr/description/cb/0115> (soon to be replaced by a longer description).



specific PU? One could even think of describing them as independent Production Units (but this mostly gives them a undue importance), or to present some notes at one level, and some other notes at the other level, depending on their nature and content (it is then frequently difficult for readers to find them). For practical reasons, I usually describe them in one specific field in the volume layer only.

Before presenting another very important advantage of type 2a below, as far as searching is concerned (see p. 329), another type following the bi-layer model must be presented. It was conceived by people interested in texts or paintings, who are sometimes unsatisfied not to find all their description in the same place and are thus unable to at once see the volume contents in their last configuration.

*Online description type 2b:  
A content-focused interpretation of the 'bi-layer' model*

How can one give a unified overall presentation of the contents in an online description of the codex that outlines the individual PUs? The easy solution is to move the contents field from the PU layer to the volume layer.<sup>23</sup> In the case of Barocci 33, the overall volume description would then look like in figure 9 below.

The PU descriptions would look like in figure 7, without the Content field. This description result more than sufficiently meets the basic conditions for a clear representation of the codex, provided the PUs are also distinguished in the content field. Otherwise, the repeated call of specialists to clearly show which texts are on which part of the manuscript<sup>24</sup> will not be heeded, and there will be a big risk that the readers do not correctly link the contents and the other types of information, like in the 1b type described above (see above, p. 315–316). Besides, setting apart

<sup>23</sup> This could be done for any feature the cataloguers want to stress. For example, in a catalogue of illustrated manuscripts, one could be tempted to group all the paintings in one field only.

<sup>24</sup> See above p. 310 footnote 1. For an example of model 2b in printed catalogue, see the description of cod. Quart. Eccl. Slav. 17 in Cleminson, Moussakova & Voutova 2006, p. 97.

## 2b – Second part: Material aspects of Unit A

Heading
<b>A. (f. 1-116) Material Description</b>
Date
(f. 1-116) S. XVI 1/2
Material
Paper with watermarks: 2 series of "Cardinal's hat" with countermark "B" ...
Scripture and hands
a. (f. 1-24) unknown hand
b. (f. 25-116) hand of <Kōnstantios>
Etc.

## 2b – First part: Description of all the content

Signature
<b>Barocci 33</b>
Extent
421 f. = (3) 116 + 128 + 177 (1) folios
Content
<p>A. (f. 1-116)</p> <p>1. (f. 1r-116r sup.) Mattheus Blastaris, Adversus iudeaos</p> <p>B. (f. 117-244)</p> <p>2. (f. 117r-209v sup.) Gennadius Scholarius, Advrsus iudaeos</p> <p>3-5. (f. 210r-230v) Georgius Gemistius Plethon, opera misc.</p> <p>6. (f. 231r-241v) Isocrates, Ad Demonicum oratio paraeneti</p> <p>7-11. (f. 244v) originally empty; today with notes, cf. infra Supplementary elements</p> <p>C. (f. 245-421)</p> <p>12. (f. 247r-414v) Gennadius Scholarius, Apologia pro quinque cpp. concilii Florentini</p> <p>13. (f. 415r-418v sup.) Definitio synodi Florentinae</p>

FIGURE 9: Beginning of the description of Barocci 33 following the content-focused interpretation of the bi-layer model.

the texts makes the relation between text and script, or text and layout, much less spontaneously visible.

However, the main problem of this approach to online descriptions becomes apparent with the need to search for and retrieve information. We will now turn to this problem.

*A goal for future databases: the possibility to always automatically retrieve correctly linked information*

Several types of description were presented above, including various forms that allow for easy (and correct) linking of information pieces that belong together. However, most of the time, while

the correct links are visible to the reader, there is no such internal linking at the computer database level. Consequently, if a database user searches a specific text (or a painting, scribe, etc.) from a specific time, the results usually include incorrectly linked information in the case of multi-date manuscripts<sup>25</sup>.

(3) Our third claim is:

*Manuscript studies require that online searches offer complete and correct results, particularly when chronological limits are entered, so that only information within the given time range is retrieved.*

In a standard mono-layer model, it is not difficult to understand why the computer cannot easily discern, for example, which texts belong to which dates: all the dates are in one field and all the texts in a different one. The same problem applies to a bi-layer model of type 2b, because the content information is split from the date information and grouped in a single field. A powerful parsing method could conceivably compensate for this situation, especially if the PUs are clearly indicated in each of the relevant fields (type 1c above), but this has not been the case so far, and it would not be a reliable solution because the ability to parse individual PU information would then depend on the internal layout of the field freely set up by the cataloguers.

Theoretically, of the description types presented above, only type 2a can easily solve the problem. This is why use of type 2a is strongly recommended.

As we saw, catalogues based on a type 2a model should be able to display correct information when a user enters multiple search criteria. But is this the case with the existing online databases following this model? A series of tests yielded somewhat disappointing results. The tests were based on searches for texts recurring in several manuscripts, including multi-part ones. With the text and a date range, the search function should find those copies of the text that were copied within the given date range – and only those! Unfortunately, all but one of the tested databases failed to give correct results, with NBM the exception (see above). In the other cases, the search results usually offered all the manuscripts with the given text and a date specification falling into the given

<sup>25</sup> See also Gippert 2015.

date range, even if this date had nothing to do with the PU containing the searched text. For example, in the case of the Barocci 33, the text of Plethon would be retrieved even when the search query was limited to the sixteenth century. This, of course, is not satisfactory and shows that correct retrieval of information, which is structurally easy to implement, should in any case be included in the instructions for the database programmers. This is theoretically the case, for example, with the German software Manuscriptum XML.<sup>26</sup> It raises quite a few hopes among scholars, but this potential for scholarly research will only be realised if the software is always used in a coherent way, and when the resulting structured descriptions are not mixed in a single database with unstructured ones (or can be isolated from them).

Readers should understand that these remarks do not aim to criticise or discourage the efforts of those who put this information on the web. On the contrary, as a user, I am very thankful to those who continue this work. I wish only to underline one common limitation of the current databases and thereby point to one major improvement that should be included in the next generation of databases.

### *Conclusion*

As explained above, manuscript studies need:

(1) *Online information that is correct: this is a question of the basic ethics of scholarship.*

This does not rule out that less reliable pieces of information are also given as pieces belonging to the ‘historiography’ of the manuscript, such as data from old or outdated publications/descriptions. But they should be clearly separated from the systematic and more reliable information, and not interfere with this. For the same reason, there is no space in scholarly databases for



<sup>26</sup> See above p. 320–321. See Riecke 2009, p. 113: ‘Dabei sollte in MXML die Struktur der Handschrift durch die Hierarchisierung der Blöcke möglichst genau abgebildet werden, um bei Recherchen genau zu dem relevanten Faszikel oder Text bzw. zu dem bestimmten Faszikel- oder Textblock leiten zu können.’

obviously wrong or 'quick and dirty' types of information or descriptions.

*(2) Online information that is correctly linked: a user of a manuscript description must always be able to correctly relate information pieces that belong to the same PUs, particularly the dates of the PUs.*

As explained above, there are multiple, very simple ways in which these goals can be satisfactorily achieved, no matter which software is being used. They do not necessarily require much time and are also perfectly suitable for 'small-scale' descriptions. No matter whether the software allows for stratigraphic descriptions and no matter what software is being used, the most important consideration is the capacity and willingness of the cataloguer to see the significant discontinuities in a manuscript and to report them to users. Again, this is not primarily a question of software, but of awareness, organisation and systematics.

*(3) Online information that can be completely and correctly retrieved: this criterion primarily depends on the software and the database design. Unfortunately, this condition is currently only very rarely met.*

Historiographical pieces of information should not be included in standard 'factual' searches; or there should at least be an easy way to exclude them.

Imagine what power manuscript sciences could have if every on-line database were to allow a correct retrieval of information that could be accessed through a portal, so that it became possible to retrieve correct information from all the databases at once? This is a dream, but as we saw, it is a dream that can be realised.

There is an old adage in computer science: 'Garbage in... garbage out'! Manuscript databases are not exempted from this rule – the more correct and correctly linked information is entered, the better the database will meet the needs of manuscript studies.

Merely going online is definitely not enough.

## Bibliography

- P. Andrist (2003), ‘*Catalogus codicum graecorum Helveticorum. Règles de catalogage, élaborées sous le patronage du Kuratorium „Katalogisierung der mittelalterlichen und frühneuzeitlichen Handschriften der Schweiz“*, version 2.0’, Bern: Burgerbibliothek Bern, reviewed by P. Augustin, *Scriptorium*, 58 (2004), p. 122-127.
- P. Andrist (2006), ‘La descrizione scientifica dei manoscritti complessi: fra teoria e pratica’, in *Segno e testo*, 4, p. 299-356, + 8 plates.
- P. Andrist (2007), *Les manuscrits grecs conservés à la Bibliothèque de la Bourgeoisie de Berne – Burgerbibliothek Bern. Catalogue et histoire de la collection*, Dietikon-Zürich: Urs Graf.
- P. Andrist (2010), ‘La description des manuscrits médiévaux sur Internet: un regard critique’, in E. Crisci, M. Maniaci & P. Orsini (eds.), *La descrizione dei manoscritti: esperienze a confronto*, Cassino: Università degli Studi di Cassino (Studi e ricerche del Dipartimento di Filologia e Storia, 1), p. 19-45.
- P. Andrist (2011), ‘The Physiognomy of Greek contra Iudaeos Manuscript Books in the Byzantine Era. A Preliminary Survey’, in R. Bonfil, O. Irshai, G. G. Stroumsa et al. (eds.), *Jews in Byzantium. Dialectics of Minority and Majority Cultures*, Leiden, Boston: Brill (Jerusalem Studies in Religion and Culture, 14), p. 549-585.
- P. Andrist, (2013), *La syntaxe du codex. Essai de codicologie structurale*, Turnhout: Brepols (Bibliologia 32).
- P. Andrist (2015), ‘Syntactical description of manuscripts: a powerful tool for understanding, communicating and searching ancient books’ in A. Bausi et al. (eds.), *Comparative Oriental Manuscript Studies. An Introduction*, Hamburg.
- P. Andrist, P. Canart & M. Maniaci (2010), ‘L’analyse structurale du codex, clef de sa genèse et de son histoire’, in A. Bravo García & I. Pérez Martín (eds.), *The legacy of Bernard de Montfaucon: Three Hundred Years of Studies on Greek Handwriting. Proceedings of the Seventh International Colloquium of Greek Palaeography (Madrid - Salamanca, 15-20 September 2008)*, Turnhout: Brepols (Bibliologia 31A), p. 289-299.
- P. Bobichon (2008), *Hébreu 669 à 703. Manuscrits de théologie*, Turnhout: Brepols (Manuscrits en caractères hébreux conservés dans les bibliothèques de France. Catalogues, 1).
- R. Cleminson, E. Moussakova & N. Voutova (2006), *Catalogue of the Slavonic Cyrillic Manuscripts of the National Széchényi Library*, Budapest, New York: Central European University Press (CEU Mediaevalia, 9).
- H. Coxe (1853), *Catalogus Codicum Manuscriptorum Bibliothecae Bodlei-*

- anae Pars Prima Recensionem Codicum Graecorum Continens*, Oxford; reprinted with handwritten corrections from N. Wilson (1969), Oxford, col. 50–52.
- J. Del Barco (2011), *Hébreu 1 à 32*, Turnhout: Brepols (Manuscripts en caractères hébreux conservés dans les bibliothèques de France. Catalogues, 4).
- A. Derolez (1995), ‘La codicologie et les études médiévales’, in J. Hamesse (ed.), *Bilan et perspectives des études médiévales en Europe. Actes du premier Congrès européen d’études médiévales, Spoleto, 27-29 mai 1993*, Louvain-la-Neuve, p. 371–386.
- S. Di Donato (2011), *Hébreu 214 à 259: Commentaires Bibliques*, Turnhout: Brepols (Manuscripts en caractères hébreux conservés dans les bibliothèques de France. Catalogues, 3).
- M. Dukan (2008), *Fragments bibliques en hébreu provenant de Guenizot*, Turnhout: Brepols (Manuscripts en caractères hébreux conservés dans les bibliothèques de France. Catalogues, 2).
- R. Giel (2010), ‘Cataloging for the Web. Manuscriptum XML and Manuscripta Mediaevalia’, in E. Crisci, M. Maniaci, P. Orsini (coord.), *La descrizione dei manoscritti: esperienze a confronto*, Cassino: Università degli Studi di Cassino (Studi e ricerche del Dipartimento di Filologia e Storia, 1), p. 47–60.
- J. Gippert (2015), ‘Catalogues and cataloguing of Oriental manuscripts in the digital age’, in A. Bausi et al. (eds.), *Comparative Oriental Manuscript Studies. An Introduction*, Hamburg.
- J. P. Gumbert (1997), ‘Codicological Descriptions – Why Short, and Why Long?’, in J. Vodopivec & N. Golob (eds.), *Book and Paper Conservation. Proceedings*, Ljubljana, p. 51–62.
- J. P. Gumbert (2004), ‘Codicological Units: Towards a Terminology for the Stratigraphy of the Non-Homogeneous Codex’, in E. Crisci & O. Pecere (eds.), *Il codice miscellaneo, tipologia e funzioni. Atti del convegno internazionale (Cassino, 14-17 maggio 2003)*, Cassino: Università degli studi di Cassino (Segno e testo, 2), p. 17–42.
- J. P. Gumbert (2009a), *Illustrated Inventory of Medieval Manuscripts. Leiden, Universiteitsbibliotheek, BPL: Hilversum*.
- J. P. Gumbert (2009b), ‘IIMM – A completely new Type of Manuscript Inventory’, in *Gazette du livre médiéval*, 55, p. 43–46.
- K. Hajdú (2003), *Katalog der griechischen Handschriften der Bayerischen Staatsbibliothek München*, Bd. 3. Codices graeci Monacenses 110–180, Wiesbaden: Harrassowitz.
- M. Kouroupou & P. Géhin (2008), *Catalogue des manuscrits conservés dans la Bibliothèque du Patriarcat Œcuménique: les manuscrits du monastère de la Panaghia de Chalki*, 2 vols., Turnhout: Brepols.

- L. C. Mohlberg (1952), *Katalog der Handschriften der Zentralbibliothek Zürich 1, Mittelalterliche Handschriften*, Zürich: Berichthaus.
- B. Munk Olsen (1998), 'L'élément codicologique', in P. Hoffmann (ed.), *Recherches de codicologie comparée. La composition du codex au Moyen Âge, en Orient et en Occident*, Paris (Collection Bibliologie), p. 105-129.
- A. Petrucci (2001), *La descrizione del manoscritto: storia, problemi, modelli*, Roma: Carocci, p. 93-104.
- A.-B. Riecke (2009) *Die Erstellung von Handschriftenbeschreibungen nach den Richtlinien der DFG mit Hilfe von Manuscriptum XML*, Version 1.1, <<http://www.manuscripta-mediaevalia.de/hs/handbuch.pdf>>.
- P. R. Robinson (1980), 'The "Booklet": A Self-Contained Unit in Composite Manuscripts of the Anglo-Saxon period', in A. Gruys & J. P. Gumbert (eds.), *Codicologica 3, Essais typologiques*, Leiden (Litterae textuales), p. 46-69.
- D. Stutzmann, P. Tylus (2007), *Les manuscrits médiévaux français et occitans de la Preussische Staatsbibliothek et de la Staatsbibliothek zu Berlin – Preussischer Kulturbesitz*, Wiesbaden: Harrassowitz (Staatsbibliothek Preussischer Kulturbesitz. Kataloge der Handschriftenabteilung; Reihe 1. Handschriften; Bd. 5).
- L. Thorndike (1946), 'The Problem of the Composite Manuscript', in *Miscellanea Giovanni Mercati*, VI, Città del Vaticano (Studi e testi, 126), p. 93-10.

