

ISSN 2410-0951

COMSt  
Comparative Oriental Manuscript Studies  
COMSt  
Bulletin

Volume 4 • Number 1 • Special Issue • Spring 2018

Edited by  
Alessandro Bausi, Paola Buzi, Marilena Maniaci,  
Zisis Melissakis, Laura E. Parodi, Eugenia Sokolinski

COMSt

*Comparative Oriental Manuscript Studies Bulletin* is the biannual on-line and print-on-demand journal of the European research network Comparative Oriental Manuscript Studies. Born in 2009 as a European Science Foundation Research Networking Programme, the network has been affiliated to the Centre for the Study of Manuscript Cultures (<http://www.manuscript-cultures.uni-hamburg.de/>) since 2016.

*Comparative Oriental Manuscript Studies Bulletin* welcomes articles, project descriptions, conference reports, book reviews and notes on all topics connected with the written cultures of the Mediterranean Near and Middle East and related traditions. Contributions should be sent to Comparative Oriental Manuscript Studies, Hiob Ludolf Centre for Ethiopian Studies, Asia Africa Institute, Hamburg University, Alsterterrasse 1, 20354 Hamburg, Germany; [eugenia.sokolinski@uni-hamburg.de](mailto:eugenia.sokolinski@uni-hamburg.de). For submission guidelines and former issues visit <<https://www.aai.uni-hamburg.de/en/comst/publications/bulletin.html>>.

#### **Editorial Board**

Alessandro Bausi, Hamburg, Germany

Paola Buzi, Rome, Italy

Marilena Maniaci, Cassino, Italy

Zisis Melissakis, Athens, Greece

Laura E. Parodi, Genoa, Italy

#### **Editorial Secretary**

Eugenia Sokolinski, Hamburg, Germany

#### **Advisory Board**

Tara Andrews, Bern, Switzerland

Patrick Andrist, Munich, Germany

Ewa Balicka-Witakowska, Uppsala, Sweden

Malachi Beit-Arié, Jerusalem, Israel

Pier Giorgio Borbone, Pisa, Italy

Françoise Briquel Chatonnet, Paris, France

Valentina Calzolari, Genève, Switzerland

Ralph Cleminson, Winchester, UK

Matthew Driscoll, Copenhagen, Denmark

Alessandro Gori, Copenhagen, Denmark

Dickran Kouymjian, Fresno, USA – Paris, France

Marco Palma, Cassino, Italy

Ira Rabin, Berlin – Hamburg, Germany

Lara Sels, Leuven, Belgium

Gaga Shurgaia, Venice, Italy

Jan-Just Witkam, Leiden, The Netherlands

Copyright © COMSt (Comparative Oriental Manuscript Studies) 2018

Publication date July 2018

This journal is published under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International license (CC BY-NC-ND 4.0) <[www.creativecommons.org](http://www.creativecommons.org)>.



ISSN 2410-0951

# Contents

<i>Linking Manuscripts from the Coptic, Ethiopian and Syriac Domain: Present and Future Synergy Strategies. Preface to the Special Issue</i> (Alessandro Bausi, Paola Buzi, Pietro Liuzzo, and Eugenia Sokolinski)	5
<b>Part 1 Project Notes</b>	11
<i>Beta maṣāḥəft: Manuscripts of Ethiopia and Eritrea</i> (Dorothea Reule)	13
<i>IslHornAfr and its Database of Islamic Literary Production from the Horn of Africa</i> (Alessandro Gori)	29
<i>The TEI-XML Architecture of Ethiopian Manuscript Archives: Respecting the Integrity of Primary Sources and Asserting Editorial Choices</i> (Anaïs Wion)	33
<i>The ‘PATHs’ Project: an Effort to Represent the Physical Dimension of Coptic Literary Production (Third–Eleventh centuries)</i> (Paola Buzi, Julian Bogdani, and Francesco Berno)	39
<i>The ‘TraCES’ Project: Towards a New Approach to Studying the Gə‘əz Language</i> (Eugenia Sokolinski)	59
<b>Part 2 Case Studies</b>	67
<i>Dealing with the Stratigraphy of Coptic Codices: the Cases of MSS Pierpont Morgan Library M578 and Coptic Museum, inv. 13446</i> (Nathan Carlig)	69
<i>Die Quellen von August Dillmanns Lexicon linguae Aethiopiae: Anmerkungen zu den Prolegomena und den verwendeten Sigla</i> (Wolfgang Dickhut und Andreas Ellwardt)	79
<i>Describing the Complex: the Multiple Dimensions of a Relational Database</i> (Sara Fani)	89
<i>Multi-level Digital Annotation of Ethiopic Texts</i> (Susanne Hummel, Vitagrazia Pisani, and Cristina Vertan)	97
<i>The CMCL Clavis Coptica. On Producing a Standardized List of (Coptic) Works and Manuscripts</i> (Tito Orlandi)	107
<i>Some Remarks about Coptic Colophons and Their Relationship with Manuscripts: Typology, Function, and Structure</i> (Agostino Soldati)	115
<i>Encoding and Annotation of Ancient Places in Ethiopia</i> (Solomon Gebreyes Beyene and Pietro Maria Liuzzo)	121
<i>Encoding Strategies and the Ethiopic Literary Heritage: The Physiologus as a Case Study</i> (Massimo Villa)	143
<b>Index</b>	151

# **Special Issue**

## **Proceedings of the Workshop *Linking Manuscripts from the Coptic, Ethiopian, and Syriac Domain: Present and Future Synergy Strategies***

**Hamburg, 23 and 24 February 2018**



# **Linking Manuscripts from the Coptic, Ethiopian, and Syriac Domain: Present and Future Synergy Strategies:**

## **Preface to the Special Issue**

*Alessandro Bausi, Paola Buzi, Pietro Liuzzo,  
and Eugenia Sokolinski*

On 23 and 24 February 2018, a two-day workshop took place at Universität Hamburg, dedicated to ‘Linking Manuscripts from the Coptic, Ethiopian, and Syriac Domain: Present and Future Synergy Strategies’.

Several projects based at the Hiob Ludolf Centre for Ethiopian Studies at Hamburg were behind the workshop. These included the Union of the German Academies-sponsored *Beta maṣāḥaft: Manuscripts of Ethiopia and Eritrea* (see the project note, pp. 13–27) and the ERC-funded *TraCES: From Translation to Creation: Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages* (see the project note, pp. 59–66), both chaired by Alessandro Bausi. The Landesforschungsförderung Hamburg provided significant funds for academic exchange between Ethiopianists in Hamburg and Copenhagen within the framework of the project *Transmission of Knowledge in the Red Sea Area*. The initiative was co-sponsored by the ERC-funded project *PATHs: Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature. Literary Texts in their Geographical Context. Production, Copying, Usage, Dissemination and Storage* based at Sapienza Università di Roma and headed by Paola Buzi (see the project note by Paola Buzi, Julian Bogdani, and Francesco Berno, pp. 39–58). Cooperating projects included the projects *CMCL: Corpus dei Manoscritti Copti Letterari* (Rome / Hamburg), *Syriaca.org* (Vanderbilt University, Nashville), *IslHornAfr: Islam in the Horn of Africa: A Comparative Literary Approach* (Copenhagen), and *EMA: Ethiopian Manuscripts Archives* (Paris). Local organization was in the hands of Pietro Liuzzo and Eugenia Sokolinski.

The aim of the workshop was to have an informal exchange of practices accepted by the different research initiatives working with manuscripts from similar backgrounds and having a strong digital dimension. The participants had the chance to compare the challenges faced and the outcomes achieved or expected. The discussion focused on such points as (1) interactions between projects of digitization of catalogues of manuscripts from the Christian Orient, (2) alignment of authority lists for clavis identifiers, ancient places,

and ancient people, (3) standards for the reuse of primary canonical texts, (4) exploitation of common metadata standards for further outputs, and finally (5) future development perspectives for digital resources in the field.

The workshop was organized in three sessions, dedicated, respectively, to dealing with places, literary works, and manuscripts.

The first session (chaired by Paola Buzi) focused on the various strategies employed by the different projects in identifying and encoding information on (primarily historical) places in the respective research environments. Solomon Gebreyes Beyene illustrated how the *Beta maṣāḥaft* project deals with encoding and annotating places on the example of the digital edition of an Ethiopic historical text, the *Chronicle of Galāwdewos* (see the contribution by Solomon Gebreyes Beyene and Pietro Liuzzo in this issue, pp. 121–141). Among other things, he illustrated how TEI-XML standards, and adaptation of existing ontologies, can help interoperability and visibility of data. In their paper on ‘Texts and contexts: an effort to link Coptic literary texts to their archaeological context’, Paola Buzi and Julian Bogdani showed, among other things, the approach adopted by the project *PAThs* to classifying and describing places relevant for the Coptic literary tradition (see the *PAThs* project note, pp. 39–58, §§1–2). The practices adopted by the *Syriaca.org* project for their Gazetteer (which, among other things, was also a source of inspiration for the *Beta maṣāḥaft*) were presented (via Skype) by David Allen Michelson, who underlined the importance of assigning stable Universal Resource Identifiers (URI) to provide a digital structure for linking, data, aggregation and search functionalities. A special attention has been paid to the geographical fuzziness of historical places, a point the participants returned to during the discussion. The places (and other entities) recorded by the *Islam in the Horn of Africa* project were at the centre of the talk of Alessandro Gori (see the *IslHornAfr* project note, pp. 29–32).

The second session (chaired by Alessandro Bausi) focused on the Literary Works. The very first presentation, ‘Identification of *one* work’ by Tito Orlandi, formulated the main challenges scholars have to deal with when building up repertories, or claves, of works in a certain (in this case, Coptic) tradition (see the contribution by Tito Orlandi in this issue, pp. 107–114). Massimo Villa in his talk ‘Encoding the Ethiopic literary heritage: issues and case studies’ showed how this and other problems are being dealt with by the *Beta maṣāḥaft* project (see the contribution by Massimo Villa in this issue, pp. 143–149). David Allen Michelson and Nathan P. Gibson (via Skype) spoke of the ‘New Handbook of Syriac Literature ([Syriaca.org/nhsl](http://Syriaca.org/nhsl))’ and of the ‘Guide to Syriac Authors’, other two core components of the *Syriaca.org* portal with data available in TEI and RDF which use stable URIs. They stressed the im-

portance of alignment to other authorities like Virtual International Authority File (VIAF). Paola Buzi and Francesco Berio returned to the Coptic tradition in their paper ‘Coptic literature: authors, works, collections’ (see the *PATHs* project note, pp. 39–58, §3).

The complexity of a relational database developed for the description of Islamic written production from the Horn of Africa in the *IslHornAfr* project was addressed in the presentation by Sara Fani ‘Describing the Complex: the Multiple Dimensions of a Relational Database’ (see this issue, pp. 89–96). In his talk ‘Relation Labeling: The Case of Islamic Manuscripts’, Michele Petrone focused on the solutions offered by the *IslHornAfr* project to managing the data elicited from endowment notes (*waqf*) on the example of an additional note in a manuscript from the collection of *ṣayḥ* Kamāl from Agaro (Gurage zone, Ethiopia). The *waqf* lists a number of entities assigned specific, fixed roles: founder, object of endowment, beneficiary, and controller. The approach chosen implies a full transcription of the *waqf*, related in the database to the specific codicological unit; the owners, founders, and beneficiaries (all being distinct entities in the database) are related to the paratext.

Finally, possibilities of linguistic study of digitally encoded texts were explored in the presentations on ‘Multilayered digital annotation of Ethiopic texts’ by Susanne Hummel, Vitagrazia Pisani, and Cristina Vertan (see their contribution in this issue, pp. 97–106) and on ‘The digital Dillmann and a corpus-based lexicon’ by Wolfgang Dickhut and Andreas Ellwardt (see this issue pp. 79–88).

The third session (chaired by Angela Bernardo) was dedicated to (digital) scholarship of manuscripts. Tito Orlandi opened the session by addressing the questions of how we define what is a manuscript, considering its changing and living nature, in his paper ‘Identification of *one* manuscript’ (see the contribution by Tito Orlandi in this issue, pp. 107–114). Dorothea Reule and Denis Nosnitsin offered some practical solutions to the various challenges posed by electronic cataloguing in the talk ‘Encoding of Manuscripts in *Beta Maṣāḥāfi*’ (see the *Beta maṣāḥāfi* project note, pp. 13–27). Anaïs Wion spoke of ‘Corpus, manuscript, document: the basic XML-TEI architecture of Ethiopian Manuscript Archives (EMA) project and why manuscript matters’ (see the *EMA* project note in this issue, pp. 33–38). Nathan Carlig in his paper ‘Dealing with Coptic codices stratigraphy: two case studies’ illustrated how the recently promoted ‘archaeological’ approaches to codicology has been applied to the description of manuscripts in the *PATHs* project database (see this issue, pp. 69–77). ‘Coptic colophons and their relationship with manuscripts: typology, function, and structure’ was the title of the paper of Agostino Soldati (published on pp. 115–119). Finally, Orhan Toy offered an insight into

the design of ‘The *Islam in the Horn of Africa* Database’. The PostgreSQL database is a relational database management system (RDBMS) using Ruby on Rails as the general web application framework. The search is enabled by the ElasticSearch open-source engine.

Cooperation and interaction of the involved projects resulted in a number of immediate consequences. The direct contact allowed hands-on exploration of the possibilities of alignment of data architecture to a degree sufficient to allow for a maximum of data linking, sharing, and exchange. A significant degree of interoperability was achieved between the *Beta maṣāḥaḥft* web application and the *EMA* and *IslHornAfr* databases. Thus, the *Beta maṣāḥaḥft* schema was expanded to accommodate the typologies proposed by the *EMA* project, so that the *EMA* XML files could be directly imported and are now searchable and viewable on the *Beta maṣāḥaḥft* portal. The *IslHornAfr* data, encoded in a different format, could be converted and fully integrated with the manuscripts / works / places / persons entities of the *Beta maṣāḥaḥft* model. Members of the *Beta maṣāḥaḥft* and *PATHs* projects also extensively discussed the respective clones alignment as well as common strategies to dealing with ancient places and manuscript stratigraphy, resulting in set ups of interoperations among the active applications. *Syriaca.org* and *Beta maṣāḥaḥft* also joined efforts to produce accessible Linked Open Data aligning the ontologies used and setting up the possibility for cross-federated searches from the respective websites. The discussion also touched upon some future possibilities of joint academic initiatives, especially focused on production and exposition of Linked Open Data to allow queries across related datasets.

On a more general level, the discussion revolved around fundamental issues which have been central to the *COMSt* (*Comparative Oriental Manuscript Studies*) approach and community since the very inception of the work in 2009, and this makes it particularly reasonable that this collection of contributions appears as a monographic issue of this *COMSt Bulletin*: sharing and interoperability of data and resources, sustainability in time of generated data and later reutilization within a different framework, issues of common languages and formats (in cataloguing, describing, editing, annotating etc.), and in general, the search for common answers to common problems.

One should maybe also underline two further essential points:

(1) The projects involved in the workshop already gave important evidence and examples of concrete forms of mutual cooperation, as detailed above, but one should never forget that for most of them (*Beta maṣāḥaḥft*, *TraCES*, *PATHs*), their existence is deeply rooted in the work that was carried on for years, and even for decades, by former projects, the data of which it was possible to recover and re-utilize fully years later: this is the case, as far as

projects in Ethiopian studies are concerned, for the recovery of massive essential topographical, prosopographical, literary, codicological, and textual data from the *Encyclopaedia Aethiopica* and *Ethio-SPaRe: Cultural Heritage of Christian Ethiopia. Salvation, Preservation, Research* projects, and other and minor ones; and this is also the case of the *PATHs* project, with an even deeper chronological background, for the reutilization of codicological, literary and textual data of the *CMCL (Corpus dei manoscritti copti letterari)* project.

(2) If comparison on a larger scale is a fascinating and challenging task, the comparison on a smaller scale, as is possible for areas that are homogeneous from the point of view of their geographical distribution and cultural coherence and consistency, like the domain of Coptic, Ethiopic (including Ethiopian Islamic), and Syriac manuscripts, still offers a privileged vantage point of observation and experimentation, that cannot be replaced by comparison on a global and larger scale: in this small scale there is still a huge unexploited potential of cooperation that can contribute to optimize and make progress small fields at a fast pace by joining forces together, in a perfect ‘*COMSt spirit*’, building up more solid premises for broader cooperations and large-scale comparisons.

\*

The contributions in this Special Issue all originate in the papers presented during the ‘Linking Manuscripts’ workshop. The articles are grouped in two chapters, Project Notes, introducing the research initiatives involved in varying degree of detail, and Case Studies, illustrating individual solutions to particular problems. Within each chapter, the papers are arranged alphabetically, by project title in the Project Notes and by the author’s name in the Case Studies.

An alphabetical index including proper names (persons, places, projects, institutions) and literary works completes the volume.



**Part 1**  
**Project Notes**





# *Beta maṣāḥaft:* Manuscripts of Ethiopia and Eritrea

*Dorothea Reule, Universität Hamburg*

The project *Beta maṣāḥaft: Manuscripts of Ethiopia and Eritrea* creates a portal that manages the entry, presentation, and analysis of data related to the manuscript tradition of the Ethiopian and Eritrean Highlands. This means encoding and semantically relating descriptions of manuscripts, works, places, and persons. With a simple network of referenced entities in TEI as background architecture, the project produces data at several different levels and serves it in multiple ways, displaying the contents and the semantic relations encoded in it.

## **1 Introduction**

With Ethiopian manuscripts as its starting point, the project *Beta maṣāḥaft: Manuscripts of Ethiopia and Eritrea*<sup>1</sup> creates a research environment by encoding and semantically relating descriptions of manuscripts, works, places, and persons.<sup>2</sup>

As yet, there is no comprehensive prosopography, gazetteer, or clavis of the Ethiopian tradition, and descriptions of manuscripts are scattered in various catalogues, some of which are difficult to access. The publication of the *Encyclopaedia Aethiopica*<sup>3</sup> has meant a considerable advance in the knowledge organization in Ethiopian studies. *Beta maṣāḥaft* has inherited its digitized indexes, curated by the project coordinator, Eugenia Sokolinski. They constitute the backbone of our work, place, and person authority files. Starting from these records, as well as from existing claves covering parts of the Ethio-

- 1 The project (officially ‘Schriftkultur des christlichen Äthiopiens und Eritreas: Eine multimediale Forschungsumgebung’) is funded within the framework of the Academies’ Programme (coordinated by the Union of the German Academies of Sciences and Humanities) under survey of the Akademie der Wissenschaften in Hamburg. It is headed by Prof. Alessandro Bausi and coordinated by Eugenia Sokolinski. The technical lead is Pietro Liuzzo. A video presentation of the project is available at <<https://youtu.be/bI950izCu2E>>. *Beta maṣāḥaft* is a TEI (Text Encoding Initiative, <[www.tei-c.org](http://www.tei-c.org)>) project. We use GitHub (<<https://github.com/>>) to store our data and we deploy our data with eXist-db (<<http://eXist-db.org>>) software for databases. The project is extremely indebted for their continued support to the TEI and eXist-db communities. We would also like to thank the team at the Hiob Ludolf Centre for Ethiopian Studies and the University of Hamburg for their support to the project.
- 2 TEI Guidelines, <<http://www.tei-c.org/Vault/P5/3.0.0/doc/tei-p5-doc/en/html/>>. For an introduction to TEI, see <<http://www.tei-c.org/Vault/Tutorials/mueller-index.htm>>.
- 3 *E Ae* I–V.

pic literary heritage,<sup>4</sup> we continue to structure and produce information about primary and secondary sources for the study of the Ethiopic literary tradition and manuscript culture.

Among the expected project outcomes are descriptions of a significant share of Ethiopic manuscripts available for research, a *clavis* of the Ethiopic literary tradition (the *Clavis Aethiopica*), a gazetteer, and a prosopography. In the first two years of the project, progress has been made in all focus areas. The encoding of eight historical catalogues is in progress,<sup>5</sup> and more will follow. Records for all works, persons, and places mentioned in these catalogues have been created or developed, if already inherited from the *Encyclopaedia Aethiopica* index. In addition, when describing a work contained in a manuscript that is being encoded, its author and other works he composed can be described simultaneously, or stub records can be created with basic information.

This note will briefly present the project's background and data architecture, the encoding of manuscripts, works, and places as its main areas of focus, going into more detail in the description of our encoding of manuscripts.<sup>6</sup>

## 2 History

Writing existed in Ethiopia since the first millennium BCE, and following the official conversion of the country to Christianity in the fourth century, the first biblical and other Christian texts were translated from Greek to Gə'əz, or Classical Ethiopic.<sup>7</sup>

Despite the more numerous amount of manuscripts that must have been produced in the Aksumite period, only the famous 'Abbā Garimā Gospels,<sup>8</sup> radiocarbon dated from the fourth to the seventh century,<sup>9</sup> are preserved from this period. Manuscripts become more abundant from the fourteenth and fifteenth century onwards, and manuscript production continues to this day,

4 The *Clavis Apocryphorum Veteris Testamenti* (CAVT), the *Clavis Apocryphorum Novi Testamenti* (CANT), the *Bibliotheca Hagiographica Orientalis* (BHO), the *Bibliotheca Hagiographica Graeca* (BHG), Kinefe-Rigb Zelleke's repertory of hagiographical texts (KRZ).

5 Dillmann 1847, 1848; Goldschmidt 1897; Turaev 1906; Grébaud and Tisserant 1935; Ullendorff 1951; Wright 1877; Zotenberg 1877.

6 During the workshop *Linking Manuscripts from the Coptic, Ethiopian and Syriac domain*, issues concerning the encoding of manuscripts for the *Beta maṣāḥəft* were co-presented by myself and Denis Nosnitsin.

7 Bausi 2015, 47.

8 McKenzie and Watson 2016, 40–41.

9 On the importance of cataloguing see Bausi 2007; on the Ethiopic manuscript tradition see Bausi 2014, 2015, 2016; on codicological characteristics see also Nosnitsin 2015 and 2016.

especially in rural regions, where books remain more expensive than manuscripts.

The exact amount of Ethiopian manuscripts has been estimated to be around 200,000, of which *c.*20,000 are accessible, mainly in Europe and the USA, and only around 10,000 to 12,000 were described in some form of catalogue.<sup>10</sup> A growing number of Ethiopian manuscripts is accessible online.<sup>11</sup> Some digital editions of Ethiopian texts are also available; the highly valuable work of Ran HaCohen on the Ethiopic Bible has to be particularly mentioned.<sup>12</sup>

### 3 Data structure

The data architecture uses XML as a data entry format.<sup>13</sup> We have records for manuscripts, works, persons and places which we connect with one another and validate to our schema, which is a customization of TEI (Text Encoding Initiative). Because so much data is entered, we use various shortcuts for the data entry.<sup>14</sup> For example, we enter only IDs instead of URIs and only a `<ptr>` element with a `@target` instead of a full bibliographic reference.

To ensure the homogeneity of the entered data, all project members contribute to the project guidelines, which are publicly accessible.<sup>15</sup> This practice ensures that all project members share the same level of information, can immediately verify any encoding practices and initiate discussions on encoding solutions for newly encountered phenomena if more specificity is needed in comparison to the TEI Guidelines.

## 4 Manuscripts

### 4.1 Introduction

*Beta maṣāḥaft* aims to be a comprehensive research environment containing data about as many known Ethiopian manuscripts available for description as possible. The priority of the first project phase (2016–2021) is the encoding

<sup>10</sup> See Bausi 2007, p. 92–93.

<sup>11</sup> We maintain a list of digitized Ethiopian manuscripts at <https://github.com/BetaMasaheft/Manuscripts/wiki/list-of-available-images-of-manuscripts>, building upon the list published by Ted Erho, [https://www.academia.edu/24013825/Digitized\\_Ethiopian\\_manuscripts\\_online](https://www.academia.edu/24013825/Digitized_Ethiopian_manuscripts_online). At <http://www.menestrel.fr/spip.php?rubrique565&lang=fr>, a list of catalogues containing Ethiopian manuscripts and descriptions of the collections can be found.

<sup>12</sup> See <https://www.tau.ac.il/~hacohen/Biblia.html>, last accessed 1 August 2018.

<sup>13</sup> Different formats are then served via our data API, but the unified data model helps the learning process for users considerably.

<sup>14</sup> The customization of the schema is minimal, but many rules are added to enforce project decisions and support consistency. Lists of values are available with their definitions.

<sup>15</sup> See <http://betamasaheft.eu/Guidelines/>, last accessed 1 August 2018.

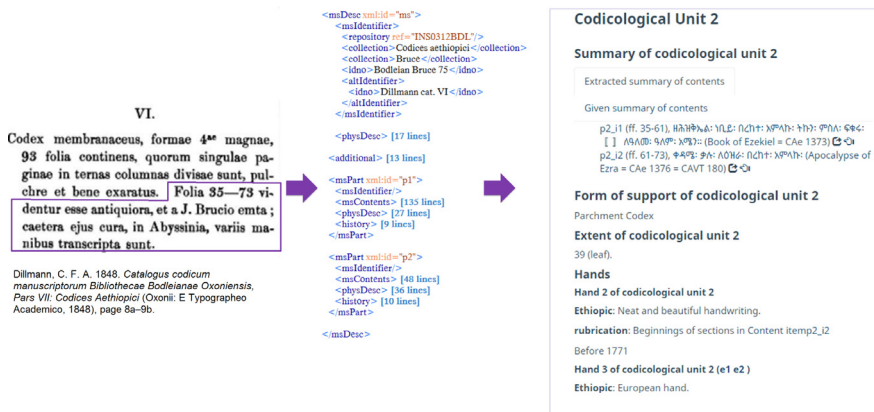
of the manuscript descriptions already present in existing catalogues, starting from the nineteenth-century catalogues of European libraries. On the whole, the sources of our records are varied—besides catalogues spanning two centuries we also describe for the first time previously unknown manuscripts.

In the first project phase our focus lies on the description of codicological features, while text transcriptions and deeper literary and philological analyses will follow in the later project phases. However, so as not to lose any information given in the catalogues on which we base our description, any literary and philological content provided there is encoded as well.<sup>16</sup> Thus all types of *Beta maṣāḥaft* records related to one entry are edited in parallel.

## 4.2 Structure

Our records follow the TEI Guidelines for the encoding of manuscripts.<sup>17</sup> All records contain the elements `<msIdentifier>`, `<history>`, `<physicalDescription>` and `<additional>`. In the `<msIdentifier>` element, all shelfmarks and other identifiers under which a manuscript is known are given. The bibliographical sources of the record are encoded in `<additional>`.

Manuscript records contain in the `<msContent>` element `<msItem>`s, in which the intellectual content is encoded. Records for composite manuscripts contain as many `<msPart>`s as the manuscript has codicological units. Each `<msPart>` can contain in its turn the elements `<msIdentifier>`, `<history>`, `<physicalDescription>` and `<additional>`, containing information pertaining to this particular unit (see Ex. 1).



Ex. 1. Identification of codicological units from the catalogue description, encoding in XML as two `<msPart>`s and visualization on the web application.

16 On the importance of the textual information provided in catalogues, see Bausi 2007, especially 107–108.

17 <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/MS.html>.

Our schema is flexible and allows the encoding of all available information without forcing non-optimal decisions. The flexibility of the XML and TEI allows us to do justice to the heterogeneity of sources without losing data and to surface additional computable information, otherwise inaccessible in printed catalogues, like the univocal and seamless identification of each descriptive feature.

### 4.3 Attribution of statements

The sources of our manuscript records are not limited to the historical catalogues which we use as a starting point. Taking into consideration all literature on the manuscripts we are working on, and examining the manuscripts ourselves whenever possible, the information which we can provide is often much more detailed and up to date. This entails working with many sources and accommodating conflicting statements. Therefore, it is crucial that any statement given in our records can be immediately attributed to its source.

This is made possible by the use of the `@resp` attribute, which can contain a reference to a project editor, a cited publication, or another person for which we have an internal or an external authority file. It is thus always possible to determine from our records which statement is based on the existing catalogue description, on secondary sources or a project editor's further research (see Ex. 2).

```
<title type="complete" ref="LIT2246salotz" xml:lang="gez">
  ጸሎታ፡ ለእግዚአብሔር፡ ማርያም፡ ዘጸሎት፡ አመ፡
  <choice>
    <sic resp="PRS3562Dillmann">ጸወጃለወርሐ፡</sic>
    <corr resp="MV">ጸወጃለወርሐ፡</corr>
  </choice>
  ስኔ፡ በደብረ፡ ጎልጎታ፡ ዝውኣቲ፡ ሙቃብረ፡ እግዚአብሔር፡
</title>
```

Ex. 2. IDs of the original cataloguer and a project editor as values of `@resp` in a `<choice>` element.

### 4.4 Intellectual content

The intellectual content of each manuscript is described within `<msContent>`, where a hierarchic series of `<msItem>`s can be organized.<sup>18</sup> All works contained in the manuscript are referred to by their ID to (1) identify each content piece and (2) maintain for each intellectual content a separate file. IDs can also be created if the work is not known yet (and becomes the official *Clavis Aethiopica* numeric ID).<sup>19</sup> We encode whether the work is contained in its entirety in the manuscript or not using a `@type` attribute in the `<title>` element. The intellectual structure of works is reflected in the structure of the `<msItem>`s, which, as all XML elements, can also be nested. Any textparts of

18 <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/MS.html#msco>.

19 On the encoding of works, see the contribution by Massimo Villa in this volume.

works are referred to, if possible, by their exact ID to be able to compare the actual sequences as they appear in the manuscripts to the standardized one stored in the work record (see Ex. 3 and Ex. 4).

```

<msItem xml:id="ms_i1.5">
  <locus from="151v"/>
  <title type="incomplete" ref="LIT2509Weddas"/>
  <msItem xml:id="ms_i1.5.1">
    <title type="complete" ref="LIT2509Weddas#Monday"/>
  </msItem>
  <msItem xml:id="ms_i1.5.2">
    <title type="complete" ref="LIT2509Weddas#Tuesday"/>
  </msItem>
  <msItem xml:id="ms_i1.5.3">
    <title type="complete" ref="LIT2509Weddas#Wednesday"/>
  </msItem>
  <msItem xml:id="ms_i1.5.4">
    <title type="incomplete" ref="LIT2509Weddas#Thursday"/>
    <note>The end of the hymn is missing.</note>
  </msItem>
  <msItem xml:id="ms_i1.5.5">
    <title type="incomplete" ref="LIT2509Weddas#Saturday"/>
    <note>A part of the hymn is missing.</note>
  </msItem>
  <msItem xml:id="ms_i1.5.6">
    <title type="incomplete" ref="LIT2509Weddas#Sunday"/>
    <note>The end of the hymn is missing.</note>
  </msItem>
  <note>The hymn for Friday,
  <ref type="work" corresp="LIT2509Weddas#Friday"/>,
  is missing.</note>
</msItem>
</msItem>

```

Ex. 3. Organization of a MS's intellectual content in nested msItems, BNFet12.

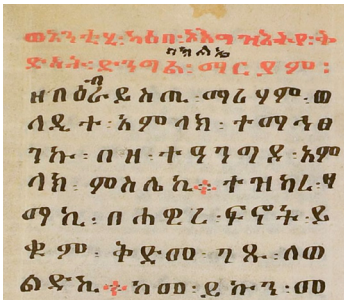
```

<div type="edition">
  <div type="textpart" subtype="chapter" xml:id="Sunday"/>
  <div type="textpart" subtype="chapter" xml:id="Monday"/>
  <div type="textpart" subtype="chapter" xml:id="Tuesday"/>
  <div type="textpart" subtype="chapter" xml:id="Wednesday"/>
  <div type="textpart" subtype="chapter" xml:id="Thursday"/>
  <div type="textpart" subtype="chapter" xml:id="Friday"/>
  <div type="textpart" subtype="chapter" xml:id="Saturday"/>
  <div type="textpart" subtype="chapter" xml:id="Zekra"/>
  <label xml:lang="gez">ዝገዳ ውዳሌ፡ ስለ ግዝእት፡ ግርጌም፡ ወላዲተ፡
  አምላካ </label>
</div>
</div>

```

Ex. 4. Organization of the same work in its work record, LIT2509Weddas.

In each single <msItem>, the text of titles, *incipit*, *explicit*, colophons, and other adscriptions is given whenever possible (see Ex. 5).



```

<incipit xml:lang="gez">
  <locus target="#22ra"/>
  <hi rend="rubric">ወላዲተ፡ ካዕብ፡ አሉግዝእት፡ ቅድስት፡ ድንግል፡</hi>
  <add place="above">ብከልኢ</add>
  <hi rend="rubric">ግርጌም፡</hi>
  ዘሰዕ
  <add place="above">ብ</add>
  ራይሰጢ፡ ግርጌም፡ ወላዲተ፡ አምላካ፡ ተግንግኑ፡ በገ፡ ተግንግዶ፡ አምላካ፡ ምስሊኪ
  <hi rend="rubric">፡፡፡</hi>
  ተዝካራ፡ ምክኒ፡ በካዲራ፡ ፍፍት፡ ይቆም፡ ቅድመ፡ ገደ፡ ለወልድኪ
  <hi rend="rubric">፡፡፡</hi>
</incipit>

```

Ex. 5. *Incipit* of a textual unit in MS BnF Éthiopien 102, f. 22ra (© gallica.bnf.fr / Bibliothèque nationale de France) and its encoding in XML: The manuscript's main physical features—rubrication and textual additions—are all rendered.

The thus constantly increasing amount of searchable Gə‘əz text contributes considerably to the identifiability of yet undescribed works. The thorough description of the intellectual content of manuscripts is heavily interconnected with the parallel development of the work records.

#### 4.5 Physical description

For the correct description of the manuscripts or codicological units (<msPart>), values corresponding to all attested book and written artefact forms in Ethiopia have been added to the schema as values of the attribute @form of the element <objectDesc>: codex, scroll, inscription,<sup>20</sup> leaf, leporello, and notebook. Values such as book and photograph have also been added as such objects form part of the Ethiopian material tradition preserved in libraries. Likewise, all needed values for the material, such as parchment and paper, have been added to the schema. Dimensions can be given, depending on the catalogue, in millimetres or inches. Loci specifying the position of the described item in the manuscript can be added as needed in the <extent> element, including the case of several codicological units (see Ex. 6).

```

<supportDesc>
  <support>
    <material key="parchment"/>
  </support>
  <extent>
    <measure unit="leaf">111</measure>
    <locus from="1" to="108"/>
    <locus from="117" to="119"/>
    <dimensions type="outer" unit="in">
      <height>8</height>
      <width>7.25</width>
    </dimensions>
  </extent>
</supportDesc>
...
<supportDesc>
  <support>
    <material key="parchment"/>
  </support>
  <extent>
    <measure unit="leaf">8</measure>
    <locus from="109" to="116"/>
    <height unit="in">2.5</height>
    <width unit="in">1.75</width>
  </extent>
  <collation>
    <note>
      <locus from="109" to="116"/>
      did not originally belong to the manuscript
    </note>
  </collation>
</supportDesc>

```

Ex. 6. Object description in the <supportDesc> elements of the two units, BDLaethe15.

```

<layoutDesc>
  <layout>
    <ab type="ruling">A mistāra was used for ruling.</ab>
  </layout>
  <layout>...</layout>
  <layout columns="1">
    <note corresp="#textarea #margin">
      Data on text area and margin dimensions taken from
      <locus target="#11"/>
    </note>
    <dimensions unit="mm" xml:id="textarea">
      <height>87</height>
      <width>55</width>
    </dimensions>
    <dimensions type="margin" unit="mm" xml:id="margin">
      <dim type="top">12</dim>
      <dim type="bottom">15</dim>
      <dim type="right">12</dim>
      <dim type="left">10</dim>
    </dimensions>
  </layout>
  <layout>...</layout>
  <layout>...</layout>
</layoutDesc>

```

Ex. 7. Layout description in <layoutDesc>, DSEthiop12.

20 Inscriptions are catalogued here together with manuscripts because of their importance in linguistic studies.



In the `<layoutDesc>` element, all available information for the layout is added, usually containing at least the number of columns and lines, but including much more exact data if available. Information on ruling, pricking and punctuation is also recorded here within an `<ab>` element using values for the attributes `@type` and `@subtype` which have been defined in the schema (see Ex. 7).

The quire structure is indicated in the `<collation>` element, listing each quire as an item containing information on its dimension (`<dim>`) and position in the manuscript (`<locus>`). Information about quire marks and decoration is added to the `<signatures>` element (see Ex. 8 and 9).

The visualization on the web application is not limited by the restraints of printed publications. This allows us to visualize quire diagrams for all manuscripts and to include different quire formulas. The visualization of the quire structure is obtained with VisColl<sup>21</sup> (see Ex. 10).

In `<handDesc>` element any available information on the manuscript's handwriting is described (see Ex. 11).

Any later additions<sup>22</sup> to the manuscript are listed within the `<additions>` element. Adding various records pertaining to the institution holding

```
<item xml:id="q3">
  <dim unit="leaf">8</dim>
  <locus from="17r" to="24v"/> III(8; s.l.: 4, stub after 5; 5, stub
after 4/fols. 17r-24v)
</item>
```

Ex. 8. Quire structure of quire III in the `<collation>` element, ESagm002.

```
<signatures>
  <note>Quire marks have been recently written with Arabic numerals in
  pen in the lower margin of the first leaf of each quire.</note>
</signatures>
```

Ex. 9. Information on quire marks in the `<signatures>` element, BAVet106.

III(8; s.l.: 4, stub after 5; 5, stub after 4/fols. 17r-24v)



Ex. 10. Visualization of the quire structure of quire III, ESagm002 (cp. Ex. 8).

<sup>21</sup> See Porter et al. 2017.

<sup>22</sup> As additions, we identify elements described by Andrist 2015, 511–513 as constituting quaternary strata.



the manuscript is a characteristic feature of Ethiopian manuscripts, thus these additions constitute important historical documents. In order to classify and search these additions, a list of values for the @type attribute is maintained in the schema. When available, the addition's text or part of it is added and marked up in a <q> element. Otherwise, only the language of the addition is indicated (Ex. 12).

```
<handDesc>
  <handNote xml:id="h1" script="Ethiopic">
    <seg type="script">15th-century script </seg>
    <seg type="ink">Black, red</seg>
    <date notBefore="1508" notAfter="1540"/>
    <desc>The handwriting displays some old paleographic features, e.g., ሰ
    has the loop marking the 7th order attached directly to the body of
    the letter, without the linking line. The crux ansata is used for
    ornamental purpose in the margins and in the intercolumn (cp. <bibl>
      <ptr target="bm:Uhlig1988Palaographie"/>
      <citedRange unit="page">329</citedRange>
      </bibl>). </desc>
    <seg type="rubrication"> Nomina sacra; titles and initial lines of the
    incipit pages; cruces ansatae. </seg>
  </handNote>
</handDesc>
```

Ex. 11. Hand description in the <handDesc> element, BAVet6.

```
<additions>
  <list>
    <item xml:id="a1">
      <locus target="#207v"/>
      <desc type="OwnershipNote">Written in <ref target="#h2"/>. </desc>
      <q xml:lang="gez">ዝዕዳ ሐፍ ዘ <persName ref="PRS11419WaldaMaryam"
        role="owner">ወልደ ማርያም</persName>
      </q>
    </item>
  </list>
</additions>
```

Ex. 12. Description of an ownership note in the <additions> element, BDLaethd4.

## 5 Works

We elaborate the records of all works contained in the manuscripts we are encoding, adding titles, statements on the authorship, an abstract, a keyword of the literary period in which the work was created and relations to other entities within the project. Transcriptions, editions and translations of texts are to become the focus of a later project phase, but are already integrated in the research environment whenever they become available. Among these, one of our current best examples is the text of the *Chronicle of King Galāwdewos* (CAe 3122, record ID: LIT3122Galaw), edited and curated in its XML version by Solomon Gebreyes Beyene.<sup>23</sup>

23 Solomon Gebreyes Beyene 2016.



Place records contain the attested names of the place in local languages and translation, including possible variants, as well as any information available on the foundation of the place, its existence and development. Coordinates can be added or will be retrieved if a reference to the place's Wikidata ID is given (see Ex. 15).

```
<place type="library" subtype="institution" sameAs="Q213678">
  <placeName xml:lang="it" xml:id="n1">Biblioteca Apostolica Vaticana </placeName>
  <placeName xml:lang="en" type="alt">Vatican Library </placeName>
  <country ref="Q237">Vatican </country>
  <location>
    <geo>41.90076 12.47447 </geo>
  </location>
  <listBibl type="catalogue">
    <bibl>
      <ptr target="bm:GrebTiss1935Codices"/>
    </bibl>
  </listBibl>
</place>
```

Ex. 15. Repository record of the Vatican Library with reference to wikidata and coordinates.

## 7 Persons

As for works and places, the backbone of our person records is formed by the digitized index of the *Encyclopaedia Aethiopica*.<sup>29</sup> Our focus is the work with persons related to the Ethiopian manuscript culture, which includes but is not limited to rulers, religious authorities, scribes, donors, and commissioners. New records, particularly for persons mentioned in manuscripts, are constantly created. All records contain the person's original and transliterated names, specified by `@type` attributes and basic information on their life and occupations as well as a reference to their Wikidata ID, if existing (see Ex. 16).

```
<person sex="1">
  <persName xml:lang="gez" xml:id="n1" type="birth">ዖግቤ ጸዮን </persName>
  <persName xml:lang="gez" corresp="#n1" type="normalized">Yāgbā Ṣayon </persName>
  <persName>
    <roleName type="title">ጳጳሴ </roleName>
  </persName>
  <persName xml:lang="gez" xml:id="n2" type="regnal">ሰሎሞን </persName>
  <persName xml:lang="gez" corresp="#n2" type="normalized">Salōmon </persName>
  <occupation type="ruler">Emperor </occupation>
  <floruit when="1285" evidence="internal">He succeeded his father <persName
    ref="PRS10303Yekunno">
    <roleName type="title">ጳጳሴ </roleName> Yəkunno ጳጳሴ </persName> in
    <date>1285 </date> . </floruit>
  <death when="1294" evidence="internal"/>
</person>
```

Ex. 16. The `<person>` element, PRS10215Yagbas.

29 The source data was the general index which included all proper names and work titles mentioned anywhere in the five volumes of the *EAE*, some of which may not be relevant for the project at the moment.

As in all other types of entities, relations between the subject of the record and other entities are crucial to our connected research environment, as illustrated by Ex. 17.

```
<listRelation>
  <relation name="saws:isAttributedAuthorOf" active="PRS1980Aphrahat"
    passive="LIT1275Dersan"/>
  <relation name="saws:isAttributedAuthorOf" active="PRS1980Aphrahat"
    passive="LIT2041Nagarb"/>
  <relation name="skos:exactMatch" active="PRS1980Aphrahat"
    passive="http://syriaca.org/person/10"/>
</listRelation>
```

Ex. 17. Relations in PRS1980Aphrahat.

### 8 Relations between entities

We express relations between different entities using the `<relation>` element throughout the file.

The list of relations in use is available in the project guidelines. We use values for the `@name` attribute of the `<relation>` element which are available in existing ontologies such as SAWS or SNAP<sup>30</sup> whenever possible (see Ex. 17 above) and create new values for types of relations specific to the Ethiopian tradition (see Ex. 18).

```
<listRelation>
  <relation name="bm:isGoldenGospelOf" active="BLorient508"
    passive="INS0448HamaraN"/>
</listRelation>
```

Ex. 18. Relation in BLorient508, declaring that the MS is a ‘Golden Gospel’ of Ḥamara Noh.

### 9 Conclusion

The data encoded in all types of the project’s entities increases daily. All data being searchable and interrelated, its increase leads to a cumulative increase in possible research questions and has equally allowed from the very beginning of the project new discoveries such as the identification of works described in differing ways in several catalogues or previously unknown connections between persons, manuscripts and places (Ex. 19).

Over the time span of the project, this will contribute considerably to our knowledge of the literary heritage, culture and history of Ethiopia as well as its relations with other cultures.

30 SNAP:DRGN *Standards for Networking Ancient Prosopographies: Data and Relations in Greco-Roman Names*, <<http://snapdrgn.net/ontology>>.



- Multiple Text Manuscripts*, Studies in Manuscript Cultures, 9 (Berlin and New York: De Gruyter, 2016), 111–153.
- , P. G. Borbone, F. Briquel-Chatonnet, P. Buzi, J. Gippert, C. Macé, Z. Melissakes, L. E. Parodi, W. Witakowski, and E. Sokolinski, eds, *Comparative Oriental Manuscript Studies: an Introduction* (Hamburg: Tredition, 2015).
- BHG. Socii Bollandiani 1895, 1909, 1957, 1969, 1984. *Bibliotheca hagiographica graeca seu elenchus vitarum sanctorum graece typis impressarum*, Subsidia Hagiographica, 8 (Bruxelles: Société des Bollandistes, 1<sup>st</sup> edn: 1895, 2<sup>nd</sup> edn: 1909; 3<sup>rd</sup> edn: ed. F. Halkin, I–III, 1957); *ibid. Auctarium*, ed. F. Halkin, Subsidia Hagiographica, 47 (Bruxelles: Société des Bollandistes, 1969); *ibid. Novum Auctarium*, ed. F. Halkin, Subsidia Hagiographica, 65 (Bruxelles: Société des Bollandistes, 1984).
- BHO. Socii Bollandiani 1910. *Bibliotheca hagiographica orientalis*, Subsidia Hagiographica, 10 (Bruxelles and Beyrouth: Société des Bollandistes and Imprimerie Catholique, 1910).
- CANT. Geerard, M. 1992. *Clavis apocryphorum Novi Testamenti*, Corpus Christianorum (Turnhout: Brepols, 1992).
- CAVT. Haelewyck, J.-C. 1998. *Clavis apocryphorum Veteris Testamenti*, Corpus Christianorum (Turnhout: Brepols, 1998).
- Dillmann, C. F. A. 1847. *Catalogus codicum manuscriptorum orientalium qui in Museo Britannico asservantur*, Pars III: *Codices Aethiopicos amplectens* (London: British Museum, 1847).
- 1848. *Catalogus codicum manuscriptorum Bibliothecae Bodleianae Oxoniensis*, Pars VII: *Codices Aethiopici* (Oxonii: E Typographeo Academico, 1848).
- EAE. S. Uhlig, ed., *Encyclopaedia Aethiopica*, I: A–C; II: D–Ha; III: Ha–N; in cooperation with A. Bausi, ed., IV: O–X; A. Bausi in cooperation with S. Uhlig, ed., V: Y–Z. *Supplementa. Addenda et Corrigenda. Maps. Index* (Wiesbaden: Harrassowitz Verlag, 2003, 2005, 2007, 2010, 2014).
- Goldschmidt, L. 1897. *Die Abessinischen Handschriften der Stadtbibliothek zu Frankfurt am Main (Rüppell'sche Sammlung) nebst Anhängen und Auszügen* (Berlin: Calvary, 1897).
- Grébaut, S. and E. Tisserant 1935. *Bybliothecae apostolicae Vaticanae codices manu scripti recensiti iussu Pii XI Pontificis maximi. Codices Aethiopici Vaticani et Borgiani, Barberinianus orientalis 2, Rossianus 865, I: Enarratio codicum* (In Bybliotheca Vaticana, 1935).
- KRZ. Kinefe-Rigb Zelleke 1975. 'Bibliography of the Ethiopic Hagiographical Traditions', *Journal of Ethiopian Studies*, 13/2 (1975), 57–102.
- Liuzzo, P. M. 2017. 'Encoding and Representation Challenges of the Project Beta maṣāḥəft: Manuscripts of Ethiopia and Eritrea', in *Proceedings of Balisage*:



- The Markup Conference 2017, Washington, DC, August 1-4, 2017*, Balisage Series on Markup Technologies, 19 (Washington, D.C.: Balisage, 2017), <doi:10.4242/BalisageVol19.Liuzzo01>.
- McKenzie, J. S. and F. Watson 2016. *The Garima Gospels: Early Illuminated Gospel Books from Ethiopia*, Manar al-Athar Monograph, 3 (Oxford: Manar al-Athar, University of Oxford, 2016).
- Nosnitsin, D. 2015. ‘Pricking and Ruling in Ethiopic Manuscripts: an Aid for Dating?’, *Comparative Oriental Manuscript Studies Bulletin*, 1/2 (2015), 94–108.
- 2016. ‘Lesser Known Features of the Ethiopian Codex’, in E. Ficquet, Ahmed Hassen, and T. Osmond, eds, *Movements in Ethiopia, Ethiopia in Movement. Proceedings of the 18<sup>th</sup> International Conference of Ethiopian Studies*, I (Addis Ababa: CFEE–IES and Los Angeles, CA: Tsehai Publishers, 2016), 75–90.
- Porter, D., A. Campagnolo, and E. Connelly 2017. ‘VisColl: A New Collation Tool for Manuscript Studies’, in H. Busch, F. Fischer and P. Sahle, eds, *Kodikologie und Paläographie im Digitalen Zeitalter 4 – Codicology and Palaeography in the Digital Age 4*, Schriften des Instituts für Dokumentologie und Editorik, 11 (Norderstedt: Books on Demand, 2017), 81–100.
- Solomon Gebreyes Beyene 2016. *Chronicle of King Gälawdewos (1540-1559): Critical Edition and Annotated Translation*, PhD thesis (Universität Hamburg, 2016).
- and P. M. Liuzzo 2017a. ‘Report – Ancient Places in the Chronicle of King Galāwdewos’, <<http://commons.pelagios.org/2017/08/report-ancient-places-in-the-chronicle-of-king-galawdewos/>>.
- 2017b. ‘Final Report: Some Interesting Historical Places in the Chronicle of Galāwdewos’, <<http://commons.pelagios.org/2017/10/final-report-some-interesting-historical-places-in-the-chronicle-of-galawdewos/>>.
- Turaev, B. 1906. *Ėfiopskija rukopisi v S.-Peterburgě*, Pamjatniki ėfiopskoj pis'menosti, 3 (St Petersburg: Tipografija imperatorskoj Akademii Nauk, 1906).
- Turska, M., J. Cummings, and S. Rahtz 2016. ‘Challenging the Myth of Presentation in Digital Editions’, *Journal of the Text Encoding Initiative*, 9 (2016) = M. Müller, ed., *Selected Papers from the 2014 TEI Conference*, <doi:10.4000/jtei.1453>.
- Ullendorff, E. 1951. *Catalogue of Ethiopian Manuscripts in the Bodleian Library* (Oxford: Clarendon Press, 1951).
- Wright, W. 1877. *Catalogue of the Ethiopic Manuscripts in the British Museum Acquired since the Year 1847* (London: British Museum, 1877).
- Zotenberg, H. 1877. *Catalogue des manuscrits éthiopiens (gheez et amharique) de la Bibliothèque nationale*, Manuscrits Orientaux (Paris: Imprimerie nationale, 1877).





# **IslHornAfr**

## **and its Database of Islamic Literary Production from the Horn of Africa\***

*Alessandro Gori, University of Copenhagen*

The project IslHornAfr deals with the Islamic manuscript and literary heritage from the Horn of Africa, primarily Ethiopia, but also Eritrea, Somalia/Somaliland, and Djibouti. It has been developing a relational database to incorporate all existing and new data on the works circulating in the area, whether local or imported/translated, the authors of these works, whether local or foreign, and the carrier media, whether manuscripts or printed books.

Literary heritage of sub-Saharan African Islamic societies as well as the local manuscript production has only relatively recently started attracting scholarly attention, the lion's share of which went to the western African areas. East Africa, in particularly the Horn, with Ethiopia at its core, has been more often than not perceived as the ancient Christian domain (as several projects discussed in this issue show), the Islamic studies having been considerably marginalized. Little systematic knowledge is available on the literature read and produced by the Muslims in the area. Against this backdrop, the project *IslHornAfr: Islam in the Horn of Africa, A Comparative Literary Approach* was conceived as a contribution to both African and Islamic studies, with the aim of producing a critical and comprehensive picture of the Islamic literary history of the Horn of Africa.<sup>1</sup>

The project explores the Islamic cultural history in the Horn of Africa as it is reflected in the literary traditions of the region (Ethiopia, Eritrea, Djibouti, Somalia/Somaliland). It considers primarily the manuscript tradition but also printed texts. Combining known sources and the new discoveries made during a series of field missions, the project has already considerably expanded our knowledge of the texts composed, translated, or copied by the Muslims in the Horn of Africa.

In the nearly five years of the project run, the team has been able to survey well over 2,000 manuscripts and books contained in 23 collections. Descriptions of varying degree of detail have been produced for all of them, which also meant filling in the relevant related tables in the project database, programmed by Orhan Toy.<sup>2</sup>

\* The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme, grant agreement no. 322849 (ERC Advanced Grant IslHornAfr, 2013–2018).

1 <<http://www.islhornafr.eu>>. See also Gori 2015.

2 See the current version at <<http://islhornafr.tors.sc.ku.dk/backend>>.

The database foresees distinct tables for (1) manuscripts and books themselves, each identified by a unique identifier provided by the project, which contains a reference to its current collection; (2) respective collections; for (3) persons (authors, copyists, sponsors, custodians, owners, etc.), (4) relevant places (locations of collections, places of manuscript production or ownership transferral, learning centres, places of activity of registered persons), and (5) texts transmitted in the manuscripts and books. Besides, controlled authority files establish taxonomy of genres, languages, and scripts. A separate database section is dedicated to the bibliography, subdivided in books, articles, book sections, and dissertations.

All texts (main texts and additional texts in the manuscripts, documents, books, whether in Arabic or in local languages) can be therefore classified with the help of the project database according to their genres, contents, titles, authors, places of creation, number of witnesses, distribution of witnesses, and linguistic and graphic peculiarities. So far, records for well over 4,000 texts have been created, many of those of local production and therefore previously unknown.<sup>3</sup>

In combination with the data collected on persons and places, the research into the literary production has already produced a new understanding of the various routes and mechanisms behind the spread of the Islamic culture in Ethiopia and surrounding countries. A special attention has been given to the application of Arabic script to local languages when writing down theological treatises, poetry, or documentary texts.<sup>4</sup>

The challenges encountered by the project are manifold. Some are due to the particular geographical setting: the instability of the political situation in the Horn has strongly influenced the choices when defining the field mission research areas. Some important sites could not be reached so far during episodes of turmoil, or due to the lack of cooperation from the local authorities and/or library custodians. Some are due to the complexity of the research material: when dealing with manuscripts from Ethiopia, we have to deal with texts written, alongside Arabic, in a variety of local languages, both Semitic (such as Amharic) and Cushitic (Afar, Oromo, Somali). These language competences are rarely combined in one person. Besides, the texts in local languages are usually transcribed using the Arabic script (*‘aḡamī*), applied in an unstandardized improvised manner by each author and/or copyist, making it often impossible to reconstruct the actual transcription and consequently understand the meaning. Identifying texts implies therefore a considerable effort.

3 Cp. e.g. Petrone 2015, 2018.

4 Cp. e.g. Fani 2017; Hernández 2017.

Finally, there are the formalization challenges similar to those encountered by the projects working in other domains, such as the definition of a work. In the project's logic, the minimal level segmentation was accepted, and every textual unit distinct from the point of view of content was treated as a separate entity, irrespective of the form of presentation, i. e. whether it is a main text in a manuscript, or an addition, even a marginal note, as each of them can have been authored separately and can potentially be transmitted in a different constellation. Disambiguation of works, often sharing similar titles, is an ongoing work. Since we are lucky to be able to consult most of the manuscripts (or their images), in most cases we manage to identify each work with some precision.

Devotional litanies and hymns are particularly challenging in this respect, as they can circulate as components of a more or less established and structured textual constellation but also as independent textual units. This is the case for example of the *Mawlid* collection in Harar, or of *ṣayḥ* Hāšim's *Faṭḥ al-Raḥmānī*.<sup>5</sup> The new material discovered by the project will allow a better understanding of the formation and the circulation of these two puzzling collections of texts.

Another challenge is the definition, and description of, a place. The easiest to encode have been the places the team could visit during the field missions. In these cases, precise coordinates could be recorded with the help of a GPS device, all names used to refer to the place could be registered in situ, and the geographical setting described with precision. When in addition to that we have historical sources referring to the place, an extensive database record can be produced. (Needless to say, manuscripts and persons related to the place are visualized automatically with each record.) Quite different are places that may be mentioned in manuscripts but we can only very approximately pinpoint on the map. In this case, we try to provide some very general indications (e.g. the district, or region, or a nearby town) to help the users orient themselves. For the moment, the database contains 325 place records (many of which lie naturally beyond the research area, as some of the manuscripts now in the Horn of Africa may have been produced e.g. in Yemen, or relevant persons may have been active in other areas of Arabia or North Africa).

Naturally, homonymy and the resulting disambiguation problem exists also for places and persons. Just as with texts, in case of doubt, separate records are created that can be merged later.

The *IslHornAfr* project aims at a longest possible sustainability. Thus, thanks to the cooperation with the project *Beta maṣāḥaṭ* (see the project note

5 Gori 2016a, 2016b. Cp. also the paper by Sara Fani in this issue.

in this issue) and its openness to active interoperability, the project data shall be incorporated with the *Beta maṣāḥaḥft* portal. A first attempt has already been successful, and much of the *IslHornAfr* material can be also accessed from <<http://betamasaheft.eu/>>. <sup>6</sup>

### References

- Fani, S. 2017. ‘Scribal Practices in Arabic Manuscripts from Ethiopia. The ‘Aḡamization of Some Scribal Practices in Fuṣḥā and ‘Aḡamī Manuscripts from Harar’, *Islamic Africa*, 8 (2017), 144–170.
- Gori, A. 2015. ‘IslHornAfr: Islam in the Horn of Africa. A Comparative Literary Approach’, in A. Bausi, A. Gori, D. Nosnitsin, and E. Sokolinski, eds, *Essays in Ethiopian Manuscript Studies. Proceedings of the International Conference Manuscripts and Texts, Languages and Contexts: the Transmission of Knowledge in the Horn of Africa, Hamburg, 17–19 July 2014*, Supplement to *Aethiopia*, 4 (Wiesbaden: Harrassowitz Verlag, 2015), 17–20.
- 2016a. ‘Some Observations on the Text of ṣayḥ Hāšim’s *Faṭḥ al-rahmānī*’, *Aethiopia*, 19 (2016), 135–148.
- 2016b. ‘Text Collections in the Arabic Manuscript Tradition of Harar: the Cases of the *Mawlid* Collection and of ṣayḥ Hāšim’s *al-Faṭḥ al-Raḥmānī*’, paper presented at the conference *The Emergence of Multiple-Text Manuscripts, Centre for the Study of Manuscript Cultures*, Hamburg, 12 November 2016.
- Hernández, A. 2017. ‘The Ajamization of Islam in Ethiopia through Esoteric Textual Manifestations in Two Collections of Ethiopian Arabic Manuscripts’, *Islamic Africa*, 8 (2017), 171–192.
- Petrone, M. 2015. ‘Devotional texts in Ethiopian Islam: a *munāḡāh* invoking the intercession of prophets, male and female saints and ‘ulamā’’, in A. Bausi, A. Gori, D. Nosnitsin, and E. Sokolinski, eds, *Essays in Ethiopian Manuscript Studies. Proceedings of the International Conference Manuscripts and Texts, Languages and Contexts: the Transmission of Knowledge in the Horn of Africa, Hamburg, 17–19 July 2014*, Supplement to *Aethiopia*, 4 (Wiesbaden: Harrassowitz Verlag, 2015), 259–272.
- 2018. ‘Sufism and Textual Practices in 20<sup>th</sup> Century Ethiopia’, in M. Lafkioui and V. Brugnattelli, eds, *Written Sources about Africa and their Study – Le fonti scritte sull’Africa e i loro studi*, *Africana Ambrosiana*, 3 (Milano: Biblioteca Ambrosiana – Centro Ambrosiano, 2018), 313–337.

6 This was possible thanks to the flexibility of the *Beta maṣāḥaḥft* schema and did require considerable adjustment to the project data. Besides, the bibliography had to be imported into Zotero database maintained by the Hiob Ludolf Centre for Ethiopian Studies; some further corrections may still be needed.

# **The TEI-XML Architecture of *Ethiopian Manuscript Archives*: Respecting the Integrity of Primary Sources and Asserting Editorial Choices**

*Anaïs Wion, Centre National de la recherche scientifique, Institut des Mondes Africains*

The paper illustrates the choices made by the *Ethiopian Manuscript Archives* collaborative project when defining the TEI-XML schema and architecture for the encoding of charters and similar documentary texts transmitted within Ethiopian manuscripts.

The *Ethiopian Manuscript Archives* (EMA) is a collaborative project carried out by historians and philologists working on manuscript documents produced by the Ethiopian Christian kingdom between the tenth and the twentieth centuries. It was developed at the Institut de Recherche et d'Histoire des Textes (IRHT, Paris)<sup>1</sup> thanks to a grant from the Agence nationale de la recherche (ANR-Cornafrique) between 2010 and 2012.

***Medieval and early modern Ethiopian archives and electronic editing tools***  
Ethiopian manuscript archives is a general term encompassing administrative, juridical, and historical texts, which were produced by the Ethiopian political and religious authorities to proclaim their laws, rules and traditions. The term ‘archives’ is to be thought of in a very wide sense—practical writings, legal and pseudo-legal writings, local or would-be ‘universal’ historiography—and also as standing in juxtaposition to religious and literary texts. The producers of these documents were the royal, and, to a lesser degree, religious administrations. Private acts were issued comparatively late, from the mid-eighteenth century or a little earlier. Several thousands, perhaps even hundreds of thousands, of such documents of diverse character, constitute a coherent corpus of primary sources, so far largely under-exploited. One of the reason for this is the fact that these documents are spread in blank spaces within the liturgical and biblical manuscripts of monasteries and churches of the Ethiopian highlands, as well as in the Ethiopian manuscripts collections of the Western libraries. Ethiopian ancient archives are literally dissolved inside the libraries.<sup>2</sup>

Establishing ways of publishing and analysing these documents is thus part of an approach, innovative in so far as it draws on digital technologies,

1 <<http://www.cn-telma.fr/>>, accessed 29 May 2018.

2 See the special issue of *Northeast African Studies*, 11/2 (2011) on the topics of Ethiopian archives, edited by Anaïs Wion and gathering articles by Donald Crummey, Manfred Kropp, Claire Bosc-Tiessé and Marie-Laure Derat, Deresse Ayenachew, Anaïs Wion, and Paul Bertrand.

```

<locus target="#58va"/>
<title xml:lang="fra">
  Clause d'immunité émise par le roi Galāwdēwos (1540-59) en faveur du monastère de Wāldēbbā
</title>
<note type="résumé">...</note>
<note type="scientifique">...</note>
<date type="production" notAfter="1559" notBefore="1540">1540-59</date>
<q xml:lang="gez">
  <seg ana="#invocation" type="interpretation">በእኩቴተ፡ አብ፡ መወልድ፡ መማገፊ፡ ቅዱስ፡</seg>
  <seg ana="#provision" type="interpretation">እ፡ ህራዕቲ፡ ወላውገዝቲ፡ ለሰባ፡ ገዳም፡ ዘዋልድብ፡</seg>
  :
  <seg ana="#suscription" type="interpretation">እ፡ ገገሥ፡ ገላውዲዎስ፡ ዘተሰመደኡ፡ አጽና፡ ሰባ፡</seg>
  <seg ana="#clauses" type="interpretation">
    ከመ፡ ኢይቅረብ፡ ይ፡ ዘኢአበረ፡ ውሰተ፡ ገዳም፡ በሌ፡ አብ፡ መወልድ፡ መማገፊ፡ ቅዱስ፡ መሰፊ፡ ገወጂ፡ ያገፊ፡ ወገፊ፡ ለዋርቅ፡ ወበሌ፡ ኡግዝተ፡
    ማርያም፡ ወላዲተ፡ አምላክ፡
  </seg>
  <seg ana="#clauses" type="interpretation">ከመ፡ ኢይባላ፡ ገልጎ፡ ወጎዝ፡ ይኩን፡ በሥልጣኛመ፡ ለልጥርብ፡ ወጸውሰሰ፡</seg>
  <seg ana="#clauses" type="interpretation">ወከመ፡ ኢይፍላቅ፡ ዘገን፡ ዘተደሰፊ፡ ጎብ፡ ዝገን፡ መጽሐፈ፡ መገሣገሥ፡።።።</seg>
</q>
<q xml:lang="fr">
  À la gloire du Père, du Fils et du Saint Esprit. J'ai légiféré sous peine d'excommunication pour les gens du monastère de
  <placeName type="religious-institution" ref="place.xml#I.002">Wāldēbbā</placeName>
  , moi,
  <persName ref="pers.xml#p002">
    <name>Galāwdēwos</name>
    qui ait pris le nom d'
    <addName>Asnāf Sagad</addName>
  </persName>
  , afin que celui qui ne résiderait pas dans ce monastère n'en approche pas par la bouche du Père, du Fils et du Saint Esprit,
  et par la bouche des quinze prophètes et par la bouche des douze apôtres et par la bouche de Notre Dame Marie mère de
  Dieu. Afin qu'il ne pénètre pas dans ce
  <term>g'elt</term>
  qu'il soit excommunié par le pouvoir de Pierre et de Paul. Et afin que l'on n'efface pas ceci qui a été écrit dans ce
  <term>maṣḥafa mangēst</term>
  <ptr target="#note3-1"/>
</q>

```

Fig. 1. XML mark-up of a document (BNFabb152#a2).

and classical by its situation within the tradition of diplomatics. The electronic publication of these documents has a number of objectives in mind: gathering, editing, translating, annotating, analysing these primary sources. EMA privileges the use of the TEI (Text Encoding Initiative) mark-up, a versatile language that can report on the different states of a text: its content, its materiality, and also the intention, notes, comments of all the authors, scribes, readers, other users, including the scholars participating to the digital edition. Indeed, one of the main scientific assets of TEI mark-up is that data can be extracted directly from the document. Encoding preserves the integrity of the text while generating layers of enriched data.<sup>3</sup> It offers also the possibility of interoperability, and guarantees that this publication of sources will evolve in harmony with tools and practices in international use. A starting collaboration with the project *Beta maṣāḥaft* at Hamburg has proven that interoperability is not an empty word.

### ***Metadata architecture: sticking to the facts and promoting scientific choices***

Structuring the metadata is the first scientific choice, in the limit of what the standard allows, of course. Resulting from a collaboration between myself, Anaïs Wion (IMAF-CNRS), Cyril Masset (IRHT-CNRS) and Lou Burnard (formerly consultant for TGE Adonis at CNRS, now TGIR Huma-Num), the

3 Always worth reading: Burnard 2014.



Fig. 2. View of a charter with transcription, translation, and commentary, the formulaic elements are highlighted (as appearing on <http://betamasasheft.eu> portal).

architecture chosen for the XML files in the EMA project respects the materiality of the documentation and allows the authors to create their own scientific collections. TEI deals with each level of the project, which are: text, manuscript, and corpus.

The textual document is the main item. It is the smallest semantic unit. It is encoded with a <msItem> and receives an @xml:id. Such a ‘textual document’ can be for instance a land charter, a list of kings, a private transfer of land property and so on. Each document is transcribed and translated. Semantic encoding (dates, place-names, names of persons, and so on) is strictly done on the text itself and not on the paratextual elements (see fig. 1, the semantic mark-up is conducted on the translation). The aim is to extract data from the primary source only, or at least to allow the readers to search through indexes and understand if the information is extracted from the primary sources or, in some cases, from elements added by the external sources. Another type of encoding wants to highlight the structure of the diplomatic discourse, using some <seg> tags. It concerns mainly the charters, which are very formalized documents (see fig. 1 for the XML mark-up on transcription and fig. 2 for the possible visualization). Then, the textual document is described by two different types of notes. One is a simple summary, with no other information than the one provided by the document itself (following the diplomatic classical tradition of the ‘regest’). The other type, often much more elaborate, is a scientific note. This is where the analysis of the author can be displayed (fig. 3).

We decided not to impose any taxonomy for categorization (using labels such as ‘charter’, ‘list’, ‘contract’ and so on), but instead to mark up systematically the ‘technical vocabulary’ used by the documents themselves. Terms used to refer to land, to transaction, to taxation, to legal and archival practices, etc. are therefore encoded and serve as index keys to navigate the corpus (fig. 4). Here again, the philosophy is to reduce as much as possible





Fig. 3. View of the reproduction, the summary, and the scientific note accompanying a land charter (as appearing on the EMA portal).

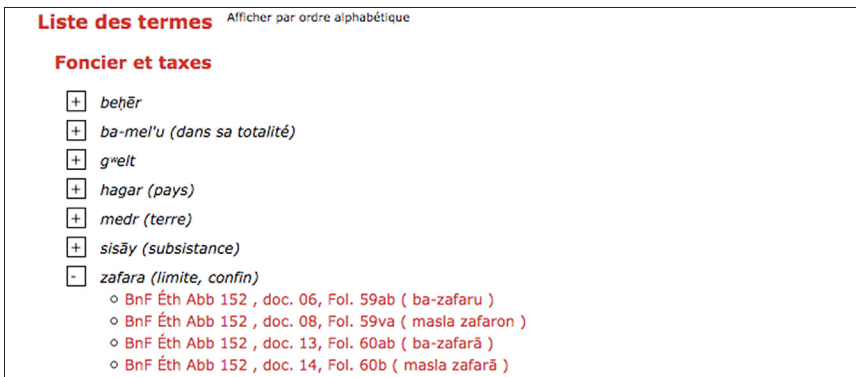


Fig. 4. The EMA filters for the ‘technical vocabulary’.

external description and let the primary source display the material for their understanding.

Each of those textual items remains linked to the manuscript in which it is copied. Each manuscript is described in a separate XML file, the digital unit corresponding to the material unit. This metadata architecture was selected to stress that archival documents are spread across the manuscripts. Archival documents in Christian medieval and early modern Ethiopia have been preserved because churches and monasteries played the role of archival centres, most of the time for their own benefit but sometimes also at a regional level,



and the main media for copying and preserving pragmatic documents were pre-existing manuscripts. This feature has to be better understood, and for that it should be painstakingly documented. So far, the choices regarding the architecture of the metadata have scrupulously followed the materiality of the documentation. To put it simply: no document can be separated from the codex in which it is copied, and the codex is considered the main documentary entity.

Being an editorial project, EMA had to provide a solution going beyond the scrupulous respect of the factual reality and offer the liberty of choices and self-determination to the scholars working in it. They are considered as authors and editors of their own corpus: they gather different manuscripts with their documents, arrange them in a collection, and are in charge of their own scientific project (fig. 5).

<b>Archives of Atkana Giyorgis parish (Begamder)</b>	Anais Wion	The church of Atkana Giyorgis, in Bagemder, was founded in 1596 by Räs Walda Krestos, general of King Sarša Dengel (1563-97) and betwaddad. During the 18th century, the local elite has produced numerous documents to administrate their lands on the parish and beyond. This corpus is made of some sixty Amharic documents, preserved in two manuscripts. It is the first attempt to publish exhaustively such a corpus of lay administrativ acts in Amharic and therefore to document this great change that occur in all Christian Ethiopia at this period.	<ul style="list-style-type: none"> <li>• création du fichier électronique et responsable de l'édition électronique Cyril Masset Anais Wion</li> <li>• Révision du fichier électronique (avril 2011) Lou Burnard</li> </ul>
<b>Archives of Hayq</b>	Manfred Kropp	This is a series of documents inserted in the so-called Golden Gospel or Gospel of Iyasus-Mo'a, belonging to and housed in the church of the monastery Qaddus 3estfanos (St. Stephen) in the lake of Hayq. The documents are dating from the late 13th to the beginning of the 16th century. As for the contents they represent historical narratives, covenants between the abbot of Hayq and the Ethiopian emperors, acts of privileges, donations and tax exemptions for the monastery, and, in number most important, regulations about the tázkar (memorial feast) for different persons, always linked to rights and usage of plots of land given to the monastery or belonging to it.	<ul style="list-style-type: none"> <li>• création du fichier électronique et responsable de l'édition électronique Cyril Masset Anais Wion</li> <li>• Révision du fichier électronique (avril 2011) Lou</li> </ul>

Fig. 5. Description of archives crediting the authorship (detail, as appearing on <<http://betamasaheft.eu>> portal).

### ***TEI-XML is much more than an editing tool***

Scientific expectations from EMA, which we hope will be revived by the collaboration with the *Beta maṣāḥaft* project, are numerous. The first one is to edit documents that have been used extensively for a long time, often without thinking sufficiently about what they really are, and often using a small fraction of the huge existing body of documentation. Yet, simple editing and index generation would not be a sufficient reason in order to engage in such a long and sometimes complicated process of encoding in TEI-XML. The potential this encoding offers is far too often under-used, and the wide possibilities are worth thinking about. For instance, using the tags related to the identity of persons and linking those tags with a time-line, with a space dimension, and

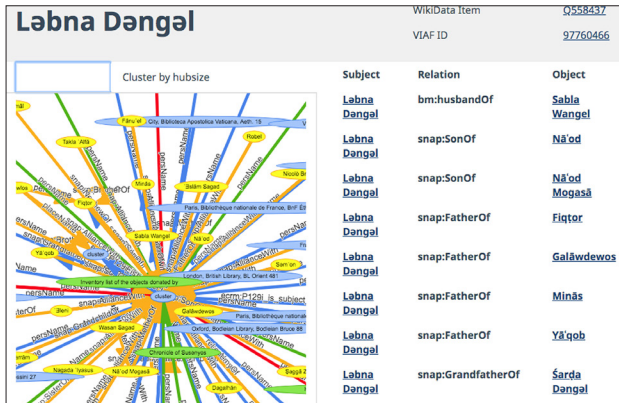


Fig. 6. Explicit and implicit relations of 'Ləbna Dəngəl', including those provided by the EMA project (as appearing on <http://betamasheft.eu> portal).

with social networks, one could significantly improve the prosopography of Ethiopian medieval and early modern society (fig. 6). Mapping the territories that the charters and other land documents mention would also be of great help to understand Ethiopian land tenure and the political and economic issues at stake when controlling the land. But here, again, first of all one has to encode strictly what the documents say, and thus gradually gather enough data on the type of lands and tenure throughout time and space.

Generic tools for manipulating encoded texts are so far scarce, and not every scholar wants to learn XPath and XQuery in order to interrogate his or her own material. Yet there is little doubt that encoding is a rewarding intellectual activity and shall be even more so in the future. The multiplication of experiences and the dialogue between projects is for sure a sign of dynamism and it creates the conditions for building a common culture in digital humanities.

## References

- Burnard, L. 2014. *What is the Text Encoding Initiative? How to add intelligent markup to digital resources* (Marseille: OpenEdition Press, 2014) <http://books.openedition.org/oepr/426>.
- Wion, A., ed., 2011. *Northeast African Studies*, 11/2 = Special Issue: *Production, Preservation, and Use of Ethiopian Archives (Fourteenth–Eighteenth Centuries)* (2011) <http://msupress.org/journals/issue/?id=50-21D-55C>.

# The ‘PATHs’ Project: an Effort to Represent the Physical Dimension of Coptic Literary Production (Third–Eleventh centuries)

*Paola Buzi, Julian Bogdani, and Francesco Berno,  
‘Sapienza’ Università di Roma*

*PATHs – Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature. Literary Texts in their Geographical Context. Production, Copying, Usage, Dissemination and Storage* is an ambitious digital project based in Rome, working towards a new historical and archaeological geography of the Coptic literary tradition. This aim implies a number of auxiliary tasks and challenges, including classification of authors, works, titles, colophons, and codicological units, as well as the study and wherever possible exact mapping of the relevant geographical sites related to the production, circulation, and storage of manuscripts.

## ***1. An introduction (Paola Buzi)***

The ‘PATHs’ project—where ‘PATHs’ is an acronym, or better an abbreviation, for ‘Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature. Literary Texts in their Geographical Context. Production, Copying, Usage, Dissemination and Storage’—aims to provide an in-depth diachronic understanding and effective representation of the geography of Coptic literary production, which is the corpus of writings, with almost exclusively religious content, produced in Egypt between the third and the eleventh centuries in Coptic (i.e. the last phase of Ancient Egyptian language).<sup>1</sup>

The double nature of the numerous scientific disciplines involved in ‘PATHs’—philology, codicology, and literature on the one side and archaeology, and geography on the other—is well represented by its logo (fig. 1), which is inspired by the devotional footprints inscribed by pilgrims, monks, and devout people, not necessarily Christian, all over Egypt in Late Antiquity, in order to mark their presence in places that were considered important for the religious life (PATHs < πατ = ‘footprint’, ‘trace’).



Fig. 1. PATHs project logo

1 This article is one of the scientific outcomes of the ERC Advanced project ‘PATHs – Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature. Literary Texts in their Geographical Context: Production, Copying, Usage, Dissemination and Storage’, funded by the European Research Council, Horizon 2020 programme, project no. 687567 (PI: Paola Buzi, Sapienza Università di Roma), <<http://paths.uniroma1.it>>.

Exactly like these devotional footprints, Coptic literature in its physical dimension—that is, Coptic books as material artefacts—left a real and concrete trace in the Egyptian landscape. ‘PATHs’ intends to investigate the relationship between settlements, as revealed by the archaeological investigations, and intellectual production, as revealed by manuscripts, and provide a new comprehensive perspective on the spread and development of Coptic literature and manuscript culture.

While the main product of ‘PATHs’ shall be a digital archaeological atlas of Coptic literature,<sup>2</sup> the project also aims at creating a series of new scientific tools that have the ambition to become pivotal for Coptic studies.<sup>3</sup>

From the first steps of the project, it has been decided that the work of ‘PATHs’ should be based on sharing ideas, achievements and results, in order to create a true collaboration network involving projects with similar or complementary purposes, and to encourage the contribution of other scholars and researchers. For this reason, all the data contained in the ‘PATHs’ database and atlas will be freely and easily accessible, reusable, and exportable.

Moreover, since there is no reason to retrace the research trails already successfully explored by others, whenever possible, ‘PATHs’ takes advantage of the results achieved by other projects and initiatives (*Corpus dei Manoscritti Copti Letterari* (CMCL), Trismegistos, Pleiades, Virtual International Authority File (VIAF), etc.), by integrating their results without redundant overlapping. This allows, through a mutual exchange of data, to plan and build new routes into unexplored territories, such as, for example, the relationship between Coptic literary manuscripts and the physical space (i.e. archaeological context) where they were produced, circulated, stored, and finally discovered. Connection points with other projects—such as e.g. specific IDs attributed to places, works, or manuscripts—are of course always clearly acknowledged.

With no claim to have the last word—which would be pretentious and not plausible—‘PATHs’ hopes to provide useful methods, models, and tools on which further historical, literary, codicological, and archaeological research related to Christian Egypt may be based. It can well be that not all the scientific choices will be ‘approved’ by other scholars, but they are always based on long and meticulous reflection and regard any aspect of the project, and the ‘PATHs’ team does its best to thoroughly explain and motivate them, so that users may decide if they are convincing enough to be adopted.

In the first eighteen months of the project the following results have been achieved:

2 Bogdani 2017, 59-69.

3 For more details on the project, see Buzi 2107, 507-516; Buzi et al. 2017.

- Conception and implementation of the relational database, on which the archaeological atlas of Coptic literature is based, and setting up of the GIS to be used for the geographical representation of the atlas.<sup>4</sup>
- Complete classification of Coptic literary works (c.1,200) by a systematic attribution of a *Clavis Coptica* (CC) identification number, integrating the work already done in this field by CMCL.
- Complete classification of 114 'Coptic authors', through the attribution of stable identifiers. Moreover, a detailed form of description for each author has been elaborated. This includes a biographical profile and a classification according to the following categories: original author (no matter if Greek or Coptic); stated author (by titles, colophons, tradition, etc.); author of the master work.<sup>5</sup>
- Complete classification of Coptic titles (c.650) through the attribution of a CC identification number.
- Complete digital edition (with English translation) of the entire corpus of Coptic titles (third to eleventh centuries).
- Complete classification of Coptic colophons or scribal subscriptions by means of a stable identifier (c.180).<sup>6</sup>
- Ongoing complete digital edition (with English translation) of the entire corpus of Coptic colophons.
- Complete classification of 6,135 Coptic manuscripts (or better codicological units), by means of the attribution of stable identifiers, in order to have univocal coordinates of reference to the entire Coptic book production. Such a classification is progressively expandable as soon as new manuscripts are discovered or identified.
- Elaboration of a protocol of detailed (digital) codicological description to be gradually applied to all collected manuscripts<sup>7</sup> (at the moment between 10% and 15% of the corpus has been described in detail).
- Mark-up (in TEI XML) of a selected corpus of literary works that are consistent in terms of their area of production and intellectual milieu, to be used for a tentative identification of places and geographical areas where specific works and literary genres were conceived. This corpus is stored online at <<https://github.com/paths-erc/coptic-texts>>.
- Complete census of the relevant sites (c.320 until now), known as places where single manuscripts (such as codices buried with a body, as a funerary kit) or entire 'collections' (for example a monastery library) have

4 See § 2 below.

5 See § 3 below.

6 See the article by Agostino Soldati below.

7 See the article by Nathan Carlig below.

been found or produced, or important for reconstructing the cultural and religious landscape of late antique and medieval Egypt.

- Elaboration of an accurate form of description of the classified places, including a summary archaeological description, precise coordinates, information on more ancient and more recent phases of occupation and usage, on the eventual function of the site as episcopal see, etc.

Any scientific result achieved until now is based on a long and ongoing reflection within the team and on discussions with other scholars. The musings on the classification of authors and works have been particularly complex. ‘Coptic authors’ are very often not ‘Coptic’ at all and frequently are fictitious or semi-fictitious. While the phenomenon of pseudoepigraphy is not exclusive to Coptic literature, it is much more pronounced here than in other ancient oriental literary traditions and takes often unexpected forms. The situation with works is equally complex, since the re-assembling of several original works—mainly Greek, but also Coptic—to create a new textual product is extremely frequent and strongly characterizes the fluid nature of Coptic literature, where authorship is often a weak point of reference.<sup>8</sup>

Dealing with Coptic works, therefore, it becomes crucial to try to answer questions like: What is an author? How to define a work? How to identify a work (being aware that the titles are not trustworthy ‘coordinates’)? How to describe and classify a textual collection? What is the relation between a work and the physical place(s) related to its production and dissemination? This is exactly what ‘PATHs’ is trying to do, always linking the intellectual production to its physical context.

## **2. Places. A georeferenced database dedicated to ‘Coptic’ Egypt (Julian Bogdani)**

It should be stressed from the beginning that ‘PATHs’ is not the first online database aimed at providing a gazetteer of relevant ancient Egyptian places. Trismegistos is by far the most famous and important online database providing information about texts originating from Egypt.<sup>9</sup> It was developed to collect papyrological resources of Egyptian origin and soon expanded to host epigraphical sources as well; its geographical scope was also expanded to include virtually the entire ancient world.<sup>10</sup> This database is designed to collect Texts, Collections, Archives, People, Networks, Authors, Editors, and, most relevantly here, Places; as of May 2018, 11,596 Egyptian places have been catalogued, related to texts by *provenance* or *attestation*. Trismegistos

8 See also the contribution by Tito Orlandi in this issue.

9 See Verreth 2013.

10 <[https://www.trismegistos.org/about\\_coverage.php](https://www.trismegistos.org/about_coverage.php)>.

assigns to each place a unique identifier, called TM Geo ID, a stable URI,<sup>11</sup> and provides a series of place names in several languages (Ancient Greek, Latin, Egyptian, Coptic, Arabic) and for each language in several variants attested in manuscripts. If available, the adjectival forms (called Ethnics) are also provided, completed with the main reference bibliography. The section dedicated to Places lists all the toponyms related to Egypt mentioned both in documentary and literary sources and for this reason must be considered not as a complete list of actual places, but as a catalogue of toponym attestations, disambiguated, clustered, and, when possible, georeferenced. In fact, well-known places are provided with geographical coordinates, although the geographical position is not a prerequisite for their inclusion in the dataset. Many places whose names are known from documentary or literary manuscript sources still lack a precise or vague localisation, and are therefore not available in the geographical visualization tool implemented in their website.

Place name disambiguation, at a larger scale, is among the goals of another project, Pleiades,<sup>12</sup> a community-based digital gazetteer of ancient place names built upon the *Barrington Atlas of the Greek and Roman World*<sup>13</sup> and soon enhanced to include multiple sources. It has now grown to become the reference gazetteer for the Mediterranean world and beyond. Pleiades provides for each site unique identifiers, stable URIs, geographical coordinates, cross references to other online databases and/or traditional publications, etc.<sup>14</sup> The Pleiades gazetteer provides the 'shared referencing system to enable connectivity through common references',<sup>15</sup> which is the foundation upon which the Pelagios Commons community rests. Pelagios Commons provides a collaborative online tool aimed at connecting and putting together multiple and different web resources dedicated to the ancient world (Mediterranean and

11 E.g. TM Geo ID 1341 stands for Memnoneia–Djeme (Thebes west), <[www.trismegistos.org/place/1341](http://www.trismegistos.org/place/1341)>.

12 <<https://pleiades.stoa.org/>>.

13 Talbert and Bagnall 2000.

14 Pleiades exposes its data in multiple formats, both human and machine readable (see, for instance, Simon, Barker, and Isaksen 2012; Isaksen et al. 2014; Simon et al. 2016). Alternative machine readable representations (such as Atom, JSON, KML, RDF+XML, Turtle) for each site and the general data export available at the downloads page (<<https://pleiades.stoa.org/downloads>>) provide an extraordinary tool for the implementation of (spatially enabled) databases and facilitate the reuse of the data, distributed with a Creative Commons Attribution (CC BY) license. Pleiades also automatically harvests and displays contents from other distributed general purpose or specialized databases and web platforms, such as Flickr (<<https://www.flickr.com>>) for photographs (the use of specifically created tags is highly recommended, to create meaningful clusters of images) or Pelagios.

15 Simon et al. 2016, 5.



beyond) by using the RDF Linked Open Data technology.<sup>16</sup> In simple words, the core mission of the project is to create a distributed network of resources, composed of stand-alone archives and databases maintained by projects who implement their own data structure. A central place in this architecture is reserved to Places which act as the common reference that all resources must implement. Places are assigned URIs, and are being referred by the unique identifiers given by the reference database they are described in, and disambiguated by references to major gazetteers such as Pleiades, Geonames, or Wikidata. Pelagios inherits from the Pleiades gazetteer a rather complex network of connection between multiple datasets focused on the ancient Mediterranean, such as Vici.org,<sup>17</sup> the Digital Atlas of Roman Empire (DARE),<sup>18</sup> etc., representing an important connection hub for the academic community.

This was the starting point for *Places*—the entity in ‘PATHs’ information system responsible for listing, cataloging, classifying, and describing sites and centres that are known to be active in the reference chronological period, i.e. the third to eleventh centuries CE. The *Places* file branched the initial dataset from the above mentioned databases, inheriting the existing connection network, limited to the Egyptian territory,<sup>19</sup> with a special focus on the (late) Roman and Medieval period.

From the very first moment of the initial branch, some important differences from the previous gazetteers arose; most importantly, it is the very concept of *place*, beyond the name or label we use for it, that bears important differences and needs to be further clarified. In the ‘PATHs’ database, *Places* are, first of all, archaeological sites, regardless of the kind of documentation we may have on them. The main aim of the project is not only the attestation of

16 Simon, Barker, and Isaksen 2012; Isaksen et al. 2014; Simon et al. 2016.

17 ‘Vici.org is the archaeological atlas of Classical Antiquity. It is a community driven archaeological map, inspired by and modelled after Wikipedia’, <<https://vici.org/about-vici.php>>.

18 DARE (<<http://dare.ht.lu.se/>>) is a webGIS project at the Department of Archaeology and Classical History, Lund University, Sweden, in collaboration with Pleiades. It was initially meant to provide a tiled base map of the Roman Empire and later became a full featured webGIS, inspired by the Barrington Atlas (Talbert and Bagnall 2000) but with a higher level of accuracy thanks to the integration of digital resources such as satellite imagery, national topographic maps, source texts, other source material, and scholarly literature.

19 Asswan and the First Cataract are traditionally considered to mark the southern border of Egypt, since predynastic times (Baines and Málek 2000, 20). It is also the conventional southern boundary of our research, even though it is not a very strict demarcation. Sporadic ‘intrusions’ south of the First Cataract might not be uncommon. The other boundaries are marked by natural features, such as the Sahara Desert, the Mediterranean and the Red Sea.



*place names* and their possible variants, connected or not to an actual position by geographical coordinates, but the concrete existence of a certain context in a given and well determined topographical space<sup>20</sup> during a predefined chronological range. This does not mean that toponyms and their variants are not being recorded, but that there are many cases of 'anonymous' places, of which no ancient name can be reasonably provided, which we are however able to describe on an archaeological basis. The challenge is to compile a complete and detailed repertory of all sites known to have been active from the third to the eleventh century CE, in order to obtain an overall and detailed picture of the Egyptian geography of the time, with a particular focus on centres of production, storage, and circulation of Coptic literary manuscripts. Close collaboration with active archaeological missions and steady monitoring of new publications ensure high quality and up-to-date information on research conducted in Egypt.

Particular attention is given to sites that are known to have provided manuscripts or manuscripts fragments or that are somehow related to the manuscript creation, storage, or circulation.<sup>21</sup> These *Places* usually receive more detailed analysis in order to obtain the best reconstruction that it is possible to recover of the archaeological context of the discovery.

Finally, since the majority (if not the entirety) of literary works written in Coptic have religious content, the religious geography is of particular interest. Consequently, the third type of *Places* that deserve special attention to our purposes are the bishoprics, since they mark a very concrete network of relevant locations from a religious, literary, and political point of view.

20 'Well determined' does not mean that we are always able to provide reliable geographic coordinates or clear boundaries; sites with an unknown or unclear location are also included in our file. The topographical space of an ancient place is not invalidated by our inability to reconstruct it, due to lack of research or a more generic lack of information that has come down to us. Our ability to locate, more or less precisely, the position of a place is measured by other means and clearly stated in the descriptive protocol. Traditionally, the geography of the 'Coptic' Egypt has been tightly conditioned from what can be recovered from written documents (Amélineau 1893; Timm 1984–1992) with a scarce attention to the actual archeological situation. These most important contributions are still the foundations on which any attempt to suggest new reconstructions lies. And yet the archaeological documentation of phases later than Hellenistic and Roman Egypt has grown enormously, and this documentation cannot be ignored. New studies have acknowledged this fact and provide important syntheses to the academic community (O'Connell 2014).

21 Precious information on these processes are typically yielded by paratexts, such as colophons, that sometimes accompany the texts. For a more detailed and documented description of these aspects see the article by Agostino Soldati below.

The gazetteer is based almost exclusively on the available bibliography and on the accessible archive documentation. The research is backed up and supported by field missions—when and where allowed by the general security conditions—aimed at verifying the present state of conservation of sites and monuments, excavated and documented several decades ago, collecting new photographic documentation, and permitting a better understanding of the topography of the archaeological context through an accurate analysis. A complete re-assessment of the Egyptian archaeological geography during the late antique and medieval period on the basis of previous study and autoptic analysis would be impossible due to the limited and often denied access to archaeological areas, the very large geographical size of Egypt, and the richness of its archaeological heritage. What ‘PATHs’ is trying to achieve is to provide a bibliography-based catalogue of the state of the art, with some in-depth analysis based on fieldwork for select areas of particular interest.

With these premises in mind, it is possible to make a brief introduction to the detailed protocol applied in the description and classification of the sites, implemented by ‘PATHs’.<sup>22</sup> All available toponyms and toponym variants are carefully collected and registered for each site, trying to cover the broader timespan. One of them is chosen to be the main name of the site—its principal label—and is given in its English form. There is not a fixed rule for this choice: usually the best-known form is chosen, the one most commonly used in the available bibliography. It can bear traces of the Greek name of the place (e.g. Dionysias), from the westernized Arabic form (e.g. Luxor from the Arabic الأقصر, *al-Uqşur*) or derived from a more strict transcription (e.g. Manqabad from منقباد). Sometimes translations or transcriptions of toponyms directly derived from the Coptic tradition are being used (e.g. Monastery of Saint Phoibammon) or, finally, where no ancient or modern toponym is available, naming conventions deriving from archaeological surveys and catalogues, like the names of the graves of the Valley of the Kings (KV + progressive number), Valley of the Queens (QV + progressive number), or other Theban Tombs (TT + progressive number)<sup>23</sup> are adopted. The other variants are given with no claim to completeness, but with the sole desire of providing a richer and more dynamic search experience. One particular exception regards toponyms and place names collected from colophons: these are the subject

22 For an introductory description of the general structure of ‘PATHs’ information system, see Bogdani 2017. Technical details and full database schema documentation are regularly published and maintained up to date on ‘PATHs’ documentation repository on GitHub (<<https://github.com/paths-erc/paths-docs>>). All information is made available with open-source MIT license.

23 KV, QV, and TT naming systems were introduced by J.G. Wilkinson (1835) and continued by other archaeologists and scholars working in the area.

of detailed individual studies aimed at providing a qualified and meaningful connection between paratexts (namely colophons), manuscripts that are associated with these documents, and, finally, actual places.

The abundance of names in different languages called for a more abstract identification and naming system, capable of ensuring disambiguation. This is obtained by assigning a unique incremental numeric identifier to each site.<sup>24</sup> The uniform resource name (URN) is composed by the project acronym (`paths`), the entity name (in this case `places`), and a numeric identifier. An incomplete example of the disambiguation and alignment of various toponyms is as follows:

```
URN: paths.places.23
Site name: Abydos
Egyptian transliteration(s): ꜥbdw, ĩbd, ĩbt, ꜥbt, pꜥ tš
n ĩbt?
Greek name: Ἀβυδοῦς
Coptic name variant #1: ⲉⲃⲟⲩ
Coptic name variant #2: ⲁⲃⲟⲩ
Arabic name variant #1: العرابة المدفونة
Arabic name transliteration variant #1: al-‘Arabat al-Mad-
fūnah
Arabic name variant #2: البليتا
Arabic name transliteration variant #2: al-Balyanā
```

In order to credit the sources and to supply fast and easy mapping<sup>25</sup> with provenance database (when available), other identifiers are provided, primarily TM Geo ID (a domain specific database) and Pleiades ID:<sup>26</sup>

```
TM Geo Id: 34
Pleiades Id: 756512
```

- 24 This identifier is automatically assigned by the database engine. The ID field is also the table's primary key.
- 25 This is meant not only for consultation but also for the programmatic treatment and publication of the information. By exposing unambiguous reference links to other data providers, automatic interconnection is made easier.
- 26 As already explained, Pleiades is an impressive linking hub, and connections to other important databases such as DARE, DARMC (<<http://ags.cga.harvard.edu/darmc/>>), Wikipedia (and Wikidata), etc., are easily inferred from it, both manually and programmatically. Linking all these datasets already referenced in Pleiades would be therefore redundant. External IDs are only referenced if there is a relationship of identity (or near identity) between a 'PATHs' place and a Trismegistos or Pleiades place. If a 'PATHs' place is only a part of a Pleiades place, the reference is not provided. E.g. the tombs of Western Thebes are recorded individually by 'PATHs' but clustered under one ID by Pleiades. In this case, each Theban tomb in 'PATHs' is referred to the higher level place, Western Thebes, which is in turn directly linked to Pleiades ID 786067.

Places are then grouped into nine larger conventional areas that do not represent an actual administrative division, neither ancient nor modern. The grouping is used solely for a very rough organization of the descriptive protocol. The clusters created so far are:

- Lower Egypt (Delta)
- Memphis
- Fayyum
- Middle Egypt
- Northern upper Egypt
- Western Thebes
- Eastern Thebes
- Southern upper Egypt
- (Nubia)

As far as the ancient geography is concerned, the name of the nome where each place is located is filed;<sup>27</sup> if the place is known to have been the capital of a nome, this is clearly stated. Moreover, information about bishoprics is listed: whether or not a place has been an episcopal see, and, if yes, the year from which the bishopric is attested. This information is not easily recoverable but it has great importance for the reconstruction of the religious geography of Late Antiquity and Middle Ages.

Archaeological or textual studies sometimes do offer clues suggesting a more precise relationship between two or more sites, defining an actual hierarchy. Such information is not systematic, because it depends to the highest degree on chance in archaeological research, yet when available it is extremely important to be recorded. In this case, a special field retains upward-only topographical relationships.<sup>28</sup>

A few words must be added on the problem of the classifications of sites by function (typology) and their chronology.

Typology definition follows standards defined by other projects, namely Pleiades.<sup>29</sup> It was highly difficult to decide not to implement a tailor-made vocabulary and support previously built ones, but we believe this will enhance future dialogue and interoperability between data providers, and make this dataset easily exploitable by other partners and users. The place types taxonomy adopted by Pleiades has been filtered and limited only to site typologies

27 The nomes list adopted by Trismegistos (see Verreth 2013, 9; 11–13) is being used.

28 One example may be enlightening: that of three, probably temporary, hermitage units—paths.places.315 (C 6), paths.places.318 (C 7), and paths.places.320 (C 7)—referred by recent studies to the monastery of Deir el-Shelwit (paths.places.281), which in this case plays a central role in the area (Delattre and Lecuyot 2016, 715).

29 <<https://pleiades.stoa.org/vocabularies/place-types>>.

identifiable in Late Antique Egypt. The following typologies are currently available:<sup>30</sup>

- acropolis
- agora, forum, plaza
- theater, amphitheater
- aqueduct
- architectural complex
- basilica
- bath
- bridge
- catacomb, cemetery, necropolis
- cave
- castrum
- church
- monastery
- circus
- cistern
- estate, villa
- fort, tower
- hermitage unit(s)
- mine, quarry
- monument
- mosque
- port
- production center (manufacturing, fishing)
- salinae
- sanctuary (religious center)
- settlement
- settlement-modern
- shrine
- station (road or coastal)
- temple
- temporary military installation or camp
- tomb
- tumulus
- tunnel
- undefined
- well

As for the chronological description: when dealing with a multimillennial civilization, the late antique phases cannot be fully comprehended in isolation from earlier history. The preceding periods—Dynastic, Ptolemaic, and Roman—left behind a monumental legacy that significantly shaped the late antique landscape. It is therefore fundamental to provide a brief (without the ambition of being thorough) diachronic description of each place in order to fully contextualize its later, Christian, phases.

30 The vocabulary remains open and can be enriched at anytime, if necessary.

The chronology of the ‘Coptic’ period is given in full detail, with the starting and ending years. This is by no means an exact chronological representation: in most cases datings of archaeological sites are loose, based on few stylistic patterns, and only in rare cases in rigorous stratigraphical sequences. Years are therefore intended in a very symbolic way.<sup>31</sup>

In describing the earlier historical phases, the main aim is to offer a succinct overview of the history of a site, stressing transformations and changes in function and typology, and reuses of monuments over long timespans. These earlier phases are described on their own, and different typologies can be filed for each site in different epochs. The typology shares the vocabulary described above, while the chronology follows a more schematic and conventional classification into macro-periods. Again, the chronological scheme has been borrowed from Pleiades,<sup>32</sup> in turn shaped after shared standards defined by well-known publications, collected by PeriodO, a gazetteer of period definitions for linking and visualizing data.<sup>33</sup> Once again, the adoption of shared vocabularies is justified by the focus of our project—which is not a rigorous organization of Egyptian settlements over many millennia—and we hope that it will ease future collaboration and cross-references with other Egyptological databases and projects.

Further information, such as graphical documentation (photographs, plans, sketches, etc.), topographical surveys, satellite imagery, topographical maps, digital elevation models, all of different scales, and geographical coverage are being collected in a conjunct GIS platform, able to perform rather complex spatial queries and accurately interconnect data of different provenance. The GIS is also the test bench of the geographical representation of the data that will be made available to the academic community through the Internet by the Archaeological Atlas.

### ***3. Coptic literature: authors, works, and textual corpora. Some methodological notes and case studies (Francesco Berno)***

It is obvious that the categories ‘authors’ and ‘works’ are closely related within the ‘PATHs’ database. In turn, they are connected to the ‘manuscripts’ category, since our text-oriented analysis has inevitably the concrete manuscript evidence at its core.

31 E.g., ‘starting year 501’ and ‘ending year 550’ stands for ‘first half of the sixth century’. This is a conventional manner to represent both undefined periods and very exact dates and make them easily searchable and comparable. A query for ‘starting date equal or bigger than 501 and ending year equal or smaller than 550’ will match ‘first half of the sixth century’, ‘first (or second) quarter of the sixth century’, but also exact ‘year 532’.

32 <<https://pleiades.stoa.org/vocabularies/time-periods>>.

33 <<http://perio.do/>>.

Indeed, 'PATHs' classification of an author is primarily based on his relationship with the works which are ascribed in different ways to his name, contributing to form a figure that may be real or fictitious, or, as in most cases, a mixture of reliable pieces of biographical information and later additions, often with evident apologetic or defamatory purposes.<sup>34</sup> Conversely, the identification of a work<sup>35</sup> cannot be separated from its connection with the tradition(s)—and I shall use this notion with great care, since it threatens to overexpose elements of ostensible continuity—that preserved the text, and with the environment(s) in which it was produced, copied, and transmitted.

'PATHs' has developed the following system to define any virtual kind of relationship between a work and a so-called 'author'. An 'author' can be linked to a single work as:

- *stated author*: the name to which a work is attributed in Coptic;
- *work's creator*: actual and original author of the work in its original language, either Greek or Coptic;
- *author of the master work*: author of a work that has lost its original authorship, but continues to circulate in the Coptic literary tradition under a pseudoepigraphical attribution and in such a redrafted redaction that it has to be considered as a work independent of the original one. The most striking example is the Coptic rewriting of Plato's *Republic* (IX 588b–589b) preserved in NHC VI, 5.

These identifications can be built on textual (the content itself), paratextual (titles, colophons, other scribal subscriptions) and/or 'external' evidence (other works, historiographical traditions, etc.), and this information is clearly recorded and made available in our database in a specific field.

34 Especially in the case of notable ecclesiastical figures at the origins of the Coptic church (Athanasius, Cyril, etc.), whose (pseudo-)biographical dossiers increased over the course of the development of Coptic literature. This led to the formation of the so-called 'cycles', that is, groups of works composed between the seventh and the eighth centuries and devoted to the lives of fourth- and fifth-century Fathers (or falsely attributed to their names). On this characterizing phenomenon of Coptic writing activity, see Orlandi 1986. It is also interesting to note a quite opposite trend, namely the continued use of the figures of notable heresiarchs, in order to denote their heresy over many centuries. See, for instance, the use of the name of Nestorius in the homily *On the Nativity* (CC 0129), attributed to Demetrius of Antioch.

35 By 'work' I mean here what Orlandi 2013 defines 'textual unit', that is, what is 'identified in modern scholarship by means of author and title [...], but also specifying the literary genre' (91); thus, a work can be preserved by different 'codicological units'. I use 'text'—which is an intrinsically wider and more vague term—to refer to the concrete dictate of a work, that is the combination of words, grammatical structures, and sentence patterns composing a work.

‘PATHs’ record for each author includes an internal numerical identifier (ID), the VIAF identifier, as well as different names and designations attributable to the author<sup>36</sup> and any religious or ecclesiastical title. This is followed by a short biography of the author composed with a special interest in the role he and his production have played in the development of Coptic literature and, in most cases, with a discussion of his more relevant extant works. This section ends with a thorough register of his extant literary production, that is, a list of all the literary works that are ascribed to the author, both in Coptic and in near linguistic environments.

For the Greek ‘authors’,<sup>37</sup> we exclusively record the works that were actually transmitted in Coptic, and this last section is customarily organized in two parts, the first covering the authentic works, the second the pseudoepigraphical ones. For the Coptic ‘authors’,<sup>38</sup> we record both the works originally composed in Coptic (*genuina* and *dubia/spuria*<sup>39</sup>) and their translations in other languages. Thus, in this case, the repertory is threefold, since for obvious reasons the authenticity is indicated solely for the Coptic works.

It is not uncommon for a work to disclose and combine more than one category of ‘author’, or several authors in the same category. The homily *Quod deus non est auctor malorum* (CC 0081<sup>40</sup>) is an effective example of both these phenomena. In fact, this is an authentic homily by Basil of Caesarea, the Greek manuscript tradition being quite unanimous in such identification (CPG 2853).<sup>41</sup> Basil’s authorship must have been known also in Coptic, since a Sahidic translation—whose *inscriptio* is now missing—is included in MONB.GS<sup>42</sup> (CLM

36 English name, Italian name (assigned by the CMCL), Greek name and, obviously, Coptic name(s) (as it/they actually appear(s) in the manuscript tradition).

37 By ‘Greek’ I mean authors whose works were likely composed originally in Greek.

38 By ‘Coptic’ I mean authors whose works were likely composed originally in Coptic. The problematic cases of authors who seem to have written both in Greek and Coptic will be shown under this latter category.

39 However, we are fully aware of the provisional and forced nature of this conceptual framework, all the more so in a complex literary tradition such as the Coptic. On this topic, see Mayer 2017, especially 979–985.

40 The *Clavis Coptica* (CC) is the standard system developed by the *Corpus dei Manoscritti Copti Letterari* (CMCL) to identify a Coptic work. Their number is constantly growing because of new textual discoveries and identifications.

41 For a recent survey on the Coptic textual transmission of Basil’s *corpus*, see Suciù 2017, especially 65–67, with selected bibliography.

42 This is the standard system developed by the CMCL to identify complete or reconstructed Coptic manuscripts. The *sigla* consists in a first section of four letters—which depends on the original provenience of the codex—followed by a two-letter progressive code.



414<sup>43</sup>), a codex entirely consisting in (authentic) homilies by the Cappadocian Father. However, at least two more Coptic testimonies of this work are transmitted. The first one, an almost complete<sup>44</sup> Bohairic version preserved in Vatican City, Biblioteca Apostolica Vaticana, Vat. copt. 57 (MACA.AC; CLM 72), is ascribed, according to the *inscriptio*, to John Chrysostom, who is said to have authored all the homilies collected in the manuscript. Lastly, the chapter of the *History of the Patriarchs of Alexandria* devoted to Athanasius attests that the bishop wrote a treatise 'in which he proves that evil comes from the devil [...] and that there is no evil at all with God';<sup>45</sup> a treatise which, beyond reasonable doubt, can be identified with our homily.<sup>46</sup>

Thus, in the Coptic literary tradition, three 'names'<sup>47</sup> are said to have authored this work. As a result, our entry will display two 'stated authors' (Chrysostom and Athanasius) and a 'work's creator' (Basil), with the aim of providing a clear and comprehensive overview of the work's Coptic reception and legacy.<sup>48</sup>

Moreover, there are cases in which an author falls under two or three categories at the same time, that is to say, is indicated in Coptic as the author both of works that can be certainly or reasonably ascribed to him and of works whose authorship is uncertain or wrongly attributed.

43 The *Coptic Literary Manuscript* (CLM) identifier is a univocal numerical identifier attributed by 'PATHs' to all complete, reconstructed, and fragmentary preserved Coptic manuscripts.

44 Despite the title, which specifies, in accordance with all the other *inscriptiones* of the codex, 'εβολ ηεν ηεηηοικον' ('[taken] by its *ethikon*' [that is, by the moral section that closes Chrysostom's homilies]). This makes us wonder about the actual Coptic perception of the Greek text. For a textual, paratextual, and codicological analysis of Vat. copt. 57, see Berno et al.2018.

45 V, 67, 397. See Orlandi 1968, 67–68.

46 It may be of interest to note that this elliptical reference appears immediately after a quick hint at the relationship between Athanasius and the bishop of Caesarea: '[a]nd he used to write to Basil, and Basil used to answer his letters, and used to address him as My Father'.

47 I keep using such a peculiar designation in order to stress in the strongest possible terms the problematic biographical consistency of these attributions and, more broadly, of the notion of 'authorship' itself in a Coptic environment. It is highly likely that, in most cases, the 'author' attributed to a work was just a mask, a name, without any (or with very limited) personal and biographical substance.

48 Although the complexity of the redactional history of CC 0081 is particularly remarkable, this is not an isolated case. See, for instance, the Melito's homily *De anima et corpore* (CC 0223). This text, lost in Greek, is attributed in Coptic to Athanasius, and goes in Syriac under the name of Alexander of Alexandria. On this problematic attribution, cf. Orlandi 2003.

In this context, I would like once more to take John Chrysostom as a privileged case study. The case of the bishop of Constantinople presents an even more challenging situation. Besides the subdivision just mentioned, we find that two works ascribed in Greek to Chrysostom are attributed in Coptic to other authors.<sup>49</sup> Nevertheless, we felt the necessity to store in our database this fundamental information, which otherwise would have been lost, by adding a new subset, namely that of ‘Literary works attributed in Coptic to other authors, while in Greek to Chrysostom’. Thus, ‘PAThs’ record will provide a concise but exhaustive list of the extant literary production connected to his name, without limitation to the works which are explicitly ascribed to Chrysostom and regardless of the authenticity of this attribution.

Sure enough, such a classification relies on and is made possible by a coherent identification-system of any and every single textual unit—with the obvious restriction to the literary works—by providing it with a stable identifier (the CC entry) and by relating it with the literary traditions and languages in which the same work is preserved. Therefore, an essential step is the presentation to each CC entry of other relevant *claves*<sup>50</sup> that, when available, might offer a map of the work’s dissemination and legacy in other linguistic, geographical and cultural backgrounds.

The modern conventional ‘titles’ attributed in our database on the basis of the designations provided by CMCL or other resources, as well as the ancient *inscriptiones/subscriptiones* ascribed in the manuscripts to their related works, are not sufficient to identify a text in a consistent and methodologically satisfying way<sup>51</sup>, and this indicates the need for an in-depth analysis of the contents of each manuscript evidence of the Coptic textual tradition. Our ‘works’, marked with an univocal CC entry, are always linked to and identified by the manuscript(s) in which they are preserved.

As for the relationship with other literary traditions and as for the numerous cases in which we have evidences of different Coptic redactions of the same work, this results in another cluster of questions. In fact, their exact overlapping is just a million-to-one shot. This difficulty applies to a wide range of textual and literary circumstances by addressing the relationship be-

49 Homilies in Mt 15, 21 (CC 0147; attributed in Coptic to Eusebius of Caesarea) and *De poenitentia* (CC 0166; attributed in Coptic to Severian of Gabala). On the Coptic reception of Chrysostom’s *corpus*, see Voicu 2011 and 2013.

50 Namely, *Clavis Patrum Graecorum* (CPG), *Bibliotheca Hagiographica Graeca* (BHG), *Bibliotheca Hagiographica Orientalis* (BHO), *Bibliotheca Hagiographica Latina* (BHL), *Clavis Apocryphorum Veteris Testamenti* (CAVT), *Clavis Apocryphorum Novi Testamenti* (CANT), *Clavis Aethiopica* (CAe).

51 On the variety and heterogeneity shown by the use, the position, the structure, and the role of Coptic titles, I refer to Buzi 2004, 2005, and 2011.

tween a Coptic original work (or, more frequently, a Coptic translation of a Greek model), on the one hand, and, on the other, previous or later reworked versions of the same text transmitted in other languages, or again two or more Coptic redactions or translations of the same work<sup>52</sup>.

We therefore need to establish a common 'degree of flexibility/fluidity' – which has to maintain a dialectical relationship with a more traditional 'idea of textual stability' associated with the notion of literature<sup>53</sup>—within which to set the decision concerning whether two texts can be taken as the same work or not. In other words: what is the threshold of redrafting and amending beyond which two or more versions of the same model can no longer be considered the same, resulting in two independent works? The main concern of 'PATHs' is to offer as much reliable and complete information as possible and, at the same time, to give our database user the possibility of making her/his own choice (which could be different from ours). Consequently, our option has been to provide the references to all the *claves* that either point to previous versions contributing to the formation of the Coptic work or indicate following translations which are in various ways related to and dependent upon the Coptic text. This means that we have opted for an hourglass-model. Predictably, the most noticeable implications of this option involve those genres that are more exposed to rewriting and emending processes, such as hagiographical and martyr tales<sup>54</sup>.

A last issue has to be briefly discussed, namely the presence of a CC for both single work and textual *corpora*. Currently, for instance, one can find a *clavis* for the *De cella* by Agathonicus, and a different one for the *Agathonicaeum corpus*, that is, for a collection of works by Agathonicus, including the *De cella* itself. Another, maybe more obvious instance is provided by the *Letters of Paul*. Just as an example, CC 0699 identifies the *Letter to the Galatians*, while under CC 0724 we find the *Pauline Epistles* as a whole.

This feature – which is largely dependent upon the CMCL classification, and undoubtedly deserves greater attention and further investigation – is not due to an inconsistency of our taxonomy, since it complies with the necessity

52 I am hinting, as mere instances, at the relationships between CC 0423 and CC 0633 (*Vita Moysis Archimandritae*), and between CC 0187, CC 0233 and CC 0549 (*Vita Pisentii*). Cp. also the contribution of Tito Orlandi in this issue.

53 As for the Coptic *milieu*, I would like at least to mention Lundhaug 2017.

54 I would point out the following *clavis* entries as specific case studies: CC 0227, CC 0229, CC 0231, CC 0232, CC 0233, CC 0234, CC 0236. In these cases, we provide the reference to all the Greek textual traditions that led to the Coptic text as we know it, as well as the reference to all the translations that derived from the Coptic text, regardless of the often significant textual elaborations.

of safeguarding the unity of a Coptic work/translation as it was actually perceived in Coptic religious life.

Some evidence<sup>55</sup> leads us to maintain that a textual *corpus* might be conceived as ‘something more’ or ‘something different’ than the single works included in it and, as such, had an independent circulation. In short, significant *corpora* stand as works in their own right.

This is information that must be stored and adequately appreciated in an analysis of Coptic literature.

### References

- Amélineau, É. 1893. *La géographie de l'Égypte à l'époque copte* (Paris: Imprimerie nationale, 1893).
- Baines, J., and J. Málek. 2000. *Cultural Atlas of Ancient Egypt*, revised ed. (New York: Facts On File, 2000).
- Berno, F., P. Buzi, A. Soldati, and F. Valerio 2018. ‘Vat. copt. 57. A Codicological, Literary and Paratextual Analysis’, *COMSt Bulletin*, 4/2 (2018), forthcoming.
- Bogdani, J. 2017. ‘The Archaeological Atlas of Coptic Literature. A Question of Method’, *Vicino Oriente*, 21 (2017), 59–69.
- Buzi, P. 2004. ‘Titles in the Coptic Manuscript Tradition: Complex Structure Title and Extended Complex Structure Titles’, in M. Immerzeel and J. van er Vliet, eds, *Coptic Studies on the Threshold of a New Millennium, Proceedings of the Seventh International Congress of Coptic Studies. Leiden, August 27–September 2, 2000*, *Orientalia Lovaniensia Analecta*, 133 (Leuven–Paris–Dudley, MA: Peeters, 2004), 309–316.
- 2005. *Titoli e autori nella tradizione copta. Studio storico e tipologico*, Biblioteca degli Studi di Egittologia e di Papirologia, 2 (Pisa–Roma: Giardini Editori e Stampatori in Pisa, 2005).
- 2011. ‘Miscellanee e florilegi. Osservazioni preliminari per uno studio dei codici copti pluritestuali: il caso delle raccolte di *excerpta*’, in P. Buzi and A. Camplani, eds, *Christianity in Egypt; Literary Production and Intellectual Trends in Late Antiquity. Studies in honor of Tito Orlandi*, *Studia Ephemeridis ‘Augustinianum’*, 125 (Roma: Institutum Patristicum Augustinianum, 2011), 177–203.
- 2017. ‘Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature. Literary Texts in their Geographical Context. Production, Copying, Usage, Dissemination and Storage (‘PATHs’). A New Project on Coptic Literary Texts’, *Early Christianity*, 8 (2017), 507–516.
- Buzi, P., J. Bogdani, N. Carlig, M.C. Giorda, and A. Soldati 2017. ‘Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature. Literary Texts in their Geographical Context. Production, Copying, Usage, Dissem-

55 See Orlandi 2018.

- ination and Storage': A New International Project on Coptic Literature and the Role of the Coptic Papyrus Codices of the Museo Egizio for Its Development', *Rivista del Museo Egizio* 1 (2017) <doi:10.29353/rime.2017.656>.
- Delattre, A. and G. Lecuyot 2016. 'À qui et à quoi servaient les 'ermitages' des vallées sud-ouest de la montagne thébaine?' in P. Buzi, A. Camplani, and F. Contardi, eds, *Coptic Society, Literature and Religion from Late Antiquity to Modern Times. Proceedings of the Tenth International Congress of Coptic Studies, Rome, September 17<sup>th</sup>–22<sup>nd</sup>, 2012 and Plenary Reports of the Ninth International Congress of Coptic Studies, Cairo*, Orientalia Lovaniensia Analecta, 247 (Leuven–Paris–Bristol: Peeters, 2016), 709–715.
- Isaksen, L., R. Simon, E.T.E. Barker, and P. de Soto Canameres 2014. 'Pelagios and the Emerging Graph of Ancient World Data', in *ACM Web Science Conference, WebSci '14, Bloomington, IN, USA, June 23–26, 2014* (Bloomington, Indiana: ACM, 2014), 197–201 <doi:10.1145/2615569.2615693>.
- Lundhaug, H. 2017. 'An Illusion of Textual Stability: Textual Fluidity, New Philology, and the Nag Hammadi Codices', in L. Ingeborg Lied and H. Lundhaug, eds, *Snapshots of Evolving Traditions. Jewish and Christian Manuscript Culture, Textual Fluidity, and New Philology* (Berlin: De Gruyter, 2017), 20–53.
- Mayer, W. 2017. 'A Life of Their Own. Preaching, Radicalisation, and the Early ps.-Chrysostomica in Greek and Latin', in F.P. Barone, C. Macé and P. Ubierna, eds., *Philologie, hermenéutique et histoire des textes entre Orient et Occident. Mélanges en hommage à Sever J. Voicu* (Turnhout: Brepols, 2017), 977–1004.
- O'Connell, E.R. 2014. *Egypt in the First Millennium AD: Perspectives from New Fieldwork*. (Leuven: Peeters, 2014)
- Orlandi, T., 1968. *Studi copti. 1. Encomio di Marco. 2. Le fonti copte di Severo di Ashmunein. 3. La leggenda di S. Mercurio* (Milano–Varese: Istituto Editoriale Cisalpino, 1968).
- 1986. 'The History of Coptic Literature', in B. A. Pearson and J. Goehring, eds, *The Roots of Egyptian Christianity* (Philadelphia: Fortress Press, 1986), 51–81.
- 2003. *Omelia De anima et corpore: clavis coptica 0223, CPG 2004*, Testo, traduzione, introduzione e note (Roma: CIM, 2003).
- 2013. 'Terminology for the Identification of Coptic Literary Documents', *Journal of Coptic Studies*, 15 (2013), 87–94.
- 2018. 'Once Again on the Coptic Dossier of Agathonicus of Tarsus', in H. Behlmer, U. Pietruschka, and F. Feder, eds, *Ägypten und der Christliche Orient. Peter Nagel zum 80. Geburtstag* (Wiesbaden: Harrassowitz Verlag, 2018), 177–186.
- Simon, R., E. Barker, and L. Isaksen 2012. 'Exploring Pelagios: A Visual Browser for Geo-Tagged Datasets', paper presented at the *International Workshop*

- on Supporting Users' Exploration of Digital Libraries, 23–27 September 2012, Paphos, Cyprus*, <<http://oro.open.ac.uk/34496/>>.
- Simon, R., L. Isaksen, E. Barker, and P. de Soto Cañamares 2016. 'The Pleiades Gazetteer and the Pelagios Project', in M. Lex Berman, R. Mostern, and H. Southall, eds, *Placing Names: Enriching and Integrating Gazetteers* (Bloomington: Indiana University Press, 2016), 97–109.
- Suciu, A. 2017. 'The Coptic Versions of *De ascetica disciplina* Attributed to Basil of Caesarea (CPG 2890). With an Appendix Containing the Armenian Version', *Revue des Études Byzantines*, 75 (2017), 65–100.
- Talbert, R.J.A., and R.S. Bagnall 2000. *Barrington Atlas of the Greek and Roman World* (Princeton, N.J.: Princeton University Press, 2000).
- Timm, S. 1984–1992. *Das Christlich-Koptische Ägypten in Arabischer Zeit. Eine Sammlung Christlicher Stätten in Ägypten in Arabischer Zeit, unter Ausschluss von Alexandria, Kairo, des Apa-Mena-Klosters (Dēr Abū Mina), der Skētis (Wādi N-Naṭrūn) und der Sinai-Region*, I–VI (Wiesbaden: Reichert, 1984–1992).
- Verreth, H. 2013. *A Survey of Toponyms in Egypt in the Graeco-Roman Period*, Trismegistos Online Publications, 2 (Köln–Leuven: Trismegistos Online Publications, 2013), <<https://www.trismegistos.org/top.php>>.
- Voicu, S. 2011. 'Per una lista delle opere trasmesse in copto sotto il nome di Giovanni Crisostomo', in P. Buzi and A. Camplani, eds, *Christianity in Egypt; Literary Production and Intellectual Trends in Late Antiquity. Studies in honor of Tito Orlandi* (Roma: Institutum Patristicum Augustinianum, 2011), 575–610.
- 2013. 'John Chrysostom in the Oriental Languages', *COMSt Newsletter*, 5 (2013), 41–46.
- Wilkinson, J.G. 1835. *The Topography of Thebes and General View of Egypt* (London: J. Murray, 1835).

# The ‘TraCES’ Project: Towards a New Approach to Studying the Gə‘əz Language

*Eugenia Sokolinski, Universität Hamburg*

The project TraCES introduces a novel research infrastructure for the computer assisted study of the Gə‘əz language as it is transmitted in written sources from various periods. This newly created corpus of annotated Ethiopic texts opens new horizons to linguists and philologist and can be subjected to an array of scientific queries.

Gə‘əz, frequently referred to as (Classical) Ethiopic, a Semitic language of the South Semitic branch, offers a particular case study for a Semitic language that developed on African soil, with remarkably little influence from non-Semitic Afro-Asiatic languages (Cushitic). The language can boast extensive documentation in a vast array of written sources going back two millenia. The earliest extensive witnesses are monumental inscriptions, but the overwhelming bulk is formed by literary texts transmitted in manuscripts, the oldest being datable to Late Antiquity and the most recent to the twentieth or even twenty-first century. The texts include early translations from Greek, later translations from Arabic, and local compositions. This situation calls for a diachronic approach to language study, rooted in a thorough morphological and lexical analysis of the literary sources.

To set up the tool box for such analysis has been, since 2014, the task of the project *TraCES: From Translation to Creation: Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages* at the Hiob Ludolf Centre for Ethiopian Studies of Universität Hamburg.<sup>1</sup>

## ***The texts***

At the core of the TraCES project is a corpus of literary texts that have been judged representative of a historical period or type of transmission, corresponding thus to a particular layer of Gə‘əz written heritage.<sup>2</sup>

- 1 This project was funded by the European Research Council (ERC) under the European Union’s Seventh Framework Programme (grant agreement no. 338756, PI: Alessandro Bausi, Universität Hamburg), <<https://www.traces.uni-hamburg.de/>>. See also Bausi 2015a.
- 2 On digitization approach to the texts, see Sokolinski 2016.



For the current project phase, the selected layers have been: (1) Gə'əz of the Aksumite period;<sup>3</sup> (2) Gə'əz of the Aksumite translations from Greek;<sup>4</sup> (3) Gə'əz of mediaeval translation from Arabic;<sup>5</sup> (4) original texts composed in Gə'əz (even if these may as well be partially inspired by translated models).<sup>6</sup>

Among the original texts, the *Kəbra nagašt* represents a special case of a text outstanding for its peculiar literary development in romance form of a subject well-known in several traditions. Most probably compiled and fixed in this form in the first half of the fourteenth century, with possibly at least partial 'Vorlage' in Copto-Arabic milieu (depending on the degree of trust one places in the 'colophon'), it incorporates elements from various traditions of the Near East (Christian, but also Jewish, with Islamic parallels) for which expression had to be sought in Gə'əz. While historians of Late Antiquity tend to highlight the significance of *Kəbra nagašt* for early relationship and acquaintance with Christian Near Eastern history and literatures,<sup>7</sup> others have remarked the ambiguity of the textual evidence.<sup>8</sup> The lexical and grammatical features might be pointing to a complex and varied translation history.

### **Research prerequisites**

As a first step, a research infrastructure had to be created.

An XML database, set up in cooperation with another Hamburg-based project, *Beta maṣāḥəft: Manuscripts of Ethiopia and Eritrea*,<sup>9</sup> hosts the authority files associated with the texts and the named entities occurring in the texts.

The extensive metadata on texts includes structured information on the titles and labels under which the works may be circulating in Ethiopian tradition; on the authorship and attributions; on the existing editions and translations as well as secondary publications discussing the works in question; on the Vorlage, recensions, and versions existing in other languages. Each text is

3 This layer is illustrated by vocalized inscriptions RIÉ 187, RIÉ 188, RIÉ 189, RIÉ 193I, RIÉ 193II, RIÉ 195I, RIÉ 195II, and RIÉ 232. The edition is based on the transcription in the *RIÉ* volumes.

4 This layer is represented by the texts of the Gospel of Matthew, see Zuurmond 2001, the *Testamentum Domini*, see Beylot 1984, and the *Epistle of Eusebius to Carpianus*, see Bausi 2015b.

5 E.g. the *Life of Secundus the Silent*, ed. Heide 2014.

6 Here, three very different text types have been selected, for the moment all from the earlier phase of Ethiopic literature: the national epos, the *Kəbra nagašt* (ed. Bezold 1905), the theological anti-heretic *Book of Mysteries* (ed. Ḥəruya 'Ērməyās 2007), and the historiographic *Chronicle of 'Amda Šəyon* (ed. Marrassini 1993);

7 See e.g. Shahīd 1976; Debié 2010.

8 See e.g. Bausi 2010, 2016; Piovaneli 2013.

9 See the project note by Dorothea Reule in this issue.



assigned a unique ID consisting of the alphanumeric sequence corresponding to the number associated with the work in the *Clavis Aethiopica* (CAe, created primarily by the *Beta maṣāḥaft* project), prefixed by 'LIT' and followed by a short title.<sup>10</sup> Each work description is linked to the detailed description of manuscript testimonies to the work as well as to records dedicated to the author(s), translator(s), commissioner(s), and other persons involved in the creation and transmission of the text.

These latter records are part of the extensive authority files repertory containing data on persons and places relevant for the Ethiopian manuscript culture, including all those mentioned in the texts that are being annotated in the *TraCES* project. In this way, each personal or place name found in the text can be directly linked to the record containing extensive metadata on the person or place in question, including name variants, historical background, geographical data, etc.<sup>11</sup>

At the core of the research infrastructure is the digital dictionary of Gə'əz.<sup>12</sup> The dictionary had as its base the text of the *Lexicon linguae Aethiopicae* by Dillmann.<sup>13</sup> The entire text has been retyped and normalized (by Alessandro Bausi and Andreas Ellwardt), and each lemma was (automatically) supplied with a unique ID (by Cristina Vertan). Subsequently, the TXT file was upconverted from string to XML using regex and XSLT (by Pietro Liuzzo). Each lemma is encoded as an extensive TEI-XML file (using TEI Dictionary module);<sup>14</sup> further encoding is possible with the help of a web-based editor. The dictionary is visualized online with the help of an application based on the eXist-db, developed by Pietro Liuzzo. The word list has been significantly enlarged having included lemmata not listed by Dillmann but registered in the texts processed by the *TraCES* project, or in other dictionaries or indices of the Ethiopic language. The dictionary is linked to the text corpus placed online by the *Beta maṣāḥaft* project, meaning that in addition to the examples provided by the lexicographers, all occurrences of the

10 E.g., the CAe number for the *Kəbra nagašt* is 1709; its ID in the database is LIT1709Kebran, the record can be viewed at <<http://betamasaheft.eu/LIT1709Kebran>> with the source code being <<http://betamasaheft.eu/LIT1709Kebran.xml>>. The CAe number for the *Chronicle of 'Amda Šəyon* is 4275; its ID is LIT4275ChronAmdS (<<http://betamasaheft.eu/LIT4275ChronAmdS>>, etc.

11 See e.g. PRS1854Amdase, the record on King 'Amda Šəyon, or LOC6136Waj, the record on the region of Wağ.

12 <<http://betamasaheft.eu/Dillmann/>>, last accessed 20 June 2018.

13 Dillmann 1865. See also the paper by Wolfgang Dickhut and Andreas Ellwardt in this issue.

14 See e.g. the lemma ቡባኤ ተ, <<http://betamasaheft.eu/Dillmann/lemma/L8f5fa9c212e24d4e99b1a64b7d1e4205.xml>>.

lemma in actual texts can be easily visualized.<sup>15</sup> While currently only the exact forms can be searched for, an engine is being developed by Stefan Druskat to include a morphological parser and a set of grammatical rules that will enable the web applications to simultaneously search for all declined and/or conjugated forms of a certain lemma.

The extensive morphological tag set for annotating Ethiopic, developed within the *TraCES* project framework, is another important prerequisite (and a by-product in its own right). The annotators can select among 33 distinct part-of-speech tags, each supplied, wherever necessary, by a set of additional grammatical features (e.g., for the Verb, the features include Person, Gender, Number, Tense/Aspect/Mood; the Infinitive takes the nominal features of Case and State).<sup>16</sup>

### **Work tools**

The project would not be possible without the help of a vast array of electronic tools assisting the researchers.

The main tool is the GeTa annotator, a Java application developed by Cristina Vertan specifically for the *TraCES* project.<sup>17</sup> The tool allows automated and synchronized Latin transcription of Ethiopic texts to be manually corrected and then used as the basis for mark-up at several levels. Transcribed graphic units can be split into minimal meaningful morphological units (tokens); each token is then associated with a part-of-speech tag (with features) and linked to the relevant lemma in the electronic dictionary (fig. 1).

At the same time, the tool allows for a basic philological annotation, including text divisions (chapters, sections, verses, lines, sentences, etc.), line and page breaks, as well as marking up the Named Entities, including persons, places, titles, offices, and dates. Personal names and toponyms are linked to the aforementioned authority files, enabling full indexing and disambiguation of terms. Comments can be introduced at all levels.

While the tool stores the data in JSON format, various converters allow exporting data into an array of other formats, ensuring the sustainability of the project data and interoperability with other projects.

Thus, thanks to the TEI-XML converter, the *TraCES* texts with philological mark up can be visualized on the *Beta maṣāḥəft* web application.

15 For ቡባሪ ፣ see the visualization at <<http://betamasasheft.eu/Dillmann/lemma/L8f5fa9c212e24d4e99b1a64b7d1e4205>>.

16 See Hummel and Dickhut 2016 for an introduction into the tag set, with a focus on the annotation of nouns.

17 See Vertan 2016 for the preliminary description of the tool functionalities as well as the paper by Susanne Hummel, Vitagrazia Pisani, and Cristina Vertan in this collection.



Thanks to the GeTa-to-ANNIS ‘Pepper’ converter, implemented by Stefan Druskat, the linguistic annotation can be now stored, visualized, searched, and analyzed online, with the help of the web-based tool for linguistic mark up, ANNIS (ANNOtation of Information Structure; fig. 2).<sup>18</sup> Texts annotated in the *TraCES* project, alongside the relevant metadata, shall be hosted in the sustainable portal of the Hamburger Zentrum für Sprachkorpora (HZSK).<sup>19</sup> Users can then search and analyze texts using the local ANNIS instance maintained by HZSK.

### **Outlook**

With the help of the annotated corpus and the search and visualization tools it shall be possible for the first time to attempt a new diachronic picture of the language which, even if not spoken for most of the past millennium, has remained the only medium of written expression until well into the nineteenth century. For the first time, frequency analysis can be carried out for lexical and grammatical phenomena; changes in linguistic patterns, in word use, and stylistic choices can be first analysed and accounted for on an objective basis.

### **References**

- Bausi, A. 2010. ‘The Massacre of Najrān: The Ethiopic Sources’, in J. Beaucamp, F. Briquel-Chatonnet, and C. J. Robin, eds, *Juifs et chrétiens en Arabie aux V<sup>e</sup> et VI<sup>e</sup> siècles. Regards croisés sur les sources*. Le massacre de Najrān, II, Centre de recherche, d’histoire et civilisation de Byzance, Monographies, 32 (Paris: Association des Amis du Centre d’Histoire et Civilisation de Byzance, 2010), 241–254.
- 2015a. ‘TraCES: From Translation to Creation: Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages’, with an annex by C. Vertan, in A. Bausi, A. Gori, D. Nosnitsin, and E. Sokolinski, eds, *Essays in Ethiopian Manuscript Studies. Proceedings of the International Conference Manuscripts and Texts, Languages and Contexts: the Transmission of Knowledge in the Horn of Africa, Hamburg, 17–19 July 2014*, Supplement to *Aethiopica*, 4 (Wiesbaden: Harrassowitz Verlag, 2015), 11–15.
- 2015b. ‘La versione etiopica della Epistola di Eusebio a Carpiano’, in R. Zarzeczny, ed., *Aethiopia Fortitudo ejus: Studi in onore di Monsignor Osvaldo Raineri in occasione del suo 80° compleanno*, *Orientalia Christiana Analecta*, 298 (Roma: Pontificio Istituto Orientale, 2015), 107–135.
- 2016. ‘La leggenda della Regina di Saba nella tradizione etiopica’, in F. Battiatto, D. Hartman, and G. Stabile, eds, *La Regina di Saba: Un mito fra Oriente e*
- 18 <<http://corpus-tools.org/annis/>>. See also Druskat and Vertan 2018.
- 19 <<https://corpora.uni-hamburg.de/>>. The launch of the TraCES corpus online is scheduled for autumn 2018.

- Occidente, Atti del Seminario diretto da Riccardo Contini, Napoli, Università 'L'Orientale', 19 novembre 2009–14 gennaio 2010*, Archivio di Studi Ebraici, 8 (Napoli: Centro di Studi Ebraici, Dipartimento Asia, Africa e Mediterraneo, Università degli Studi di Napoli 'L'Orientale', 2016), 91–162.
- Bausi, A. and E. Sokolinski, eds 2016. *150 Years after Dillmann's Lexicon: Perspectives and Challenges of Gə'əz Studies*, Supplement to Aethiopica, 5 (Wiesbaden: Harrassowitz Verlag, 2016).
- Beylot, R., ed., tr. 1984. *Testamentum Domini éthiopiens. Édition et traduction* (Louvain: Peeters, 1984).
- Bezold, C., ed., tr. 1905. *Kebrā Nagast. Die Herrlichkeit der Könige, nach den Handschriften in Berlin, London, Oxford und Paris zum ersten Mal im Äthiopischen Urtext hrsg. und mit dt. Übers. versehen*, Abhandlungen der Philosophisch-Philologischen Klasse der Königlichen Bayerischen Akademie der Wissenschaften, 23/1 (München: Verlag der K. Akademie der Wissenschaften in Kommission des G. Franz'schen Verlags, 1905).
- Debié, M. 2010. 'Le Kebrā Nagast ou La Gloire des rois: une réponse apocryphe aux événements de Najran?', in J. Beaucamp, F. Briquel-Chatonnet, and C. J. Robin, eds, *Juifs et chrétiens en Arabie aux V<sup>e</sup> et VI<sup>e</sup> siècles. Regards croisés sur les sources*. Le massacre de Najrân, II, Centre de recherche, d'histoire et civilisation de Byzance, Monographies, 32 (Paris: Association des Amis du Centre d'Histoire et Civilisation de Byzance, 2010), 255–278.
- Dillmann, C.F.A. 1865. *Lexicon linguae Aethiopiae, cum indice Latino. Adiectum est vocabularium Tigre dialecti septentrionalis compilatum a W. Munziger* (Lipsiae: T.O. Weigel, 1865; repr. New York, 1955, and Osnabrück: Biblio Verlag, 1970).
- Druskat, S. and C. Vertan 2018, 'Nachnutzbarmachung von Forschungsdaten und Tools am Beispiel altäthiopischer Korpora', in G. Vogeler, ed., *Kritik der digitalen Vernunft. Abstracts zur Jahrestagung des Verbandes Digital Humanities im deutschsprachigen Raum, 26.02.–02.03.2018 an der Universität zu Köln*, (Köln: Universität zu Köln, 2018), 270–273.
- Heide, M., ed., tr. 2014. *Secundus Taciturnus. Die arabischen, äthiopischen und syrischen Textzeugen einer didaktischen Novelle aus der römischen Kaiserzeit*, Mit einem Beitrag von S. Weninger zur *Übersetzungssprache der äthiopischen Version*, Äthiopistische Forschungen, 81 (Wiesbaden: Harrassowitz Verlag, 2014).
- Ḥərūya 'Ermyās, ed. 2015. መጽሐፈ ፡ ምሥጢር ፡ ዘአባ ፡ ጊዮርጊስ ፡ ዘጋሥጫ ፡ (*Maṣḥafa məšṭir za-<sup>3</sup>abbā Giyorgis za-Gaśəččā*, 'The Book of Mystery of <sup>3</sup>Abba Giyorgis of Gaśəččā') (Addis Ababa: Alem Printing House, 2007).
- Hummel, S. and W. Dickhut 2016. 'A Part of Speech Tag Set for Ancient Ethiopic', in Bausi and Sokolinski 2016, 17–30.

- Marrassini, P., ed., tr. 1993. *Lo scettro e la croce. La campagna di 'Amda Šeyon contro l'Ifāt (1332)*, Studi Africanistici, Serie Etiopica, 4 (Napoli: Università degli Studi di Napoli l' 'Orientale', 1993).
- Piovanelli, P. 2013. 'The Apocryphal Legitimation of a 'Solomonic' Dynasty in the Kəbrä nāgāšt— A Reappraisal', *Aethiopica*, 16 (2013), 7–44.
- RIÉ. É. Bernard, A.J. Drewes, and R. Schneider, *Recueil des Inscriptions de l'Éthiopie des périodes pré-axoumite et axoumite*, introduction de Fr. Anfrays, I: *Les documents* (Paris: Diffusion de Boccard, 1991).
- Shahîd, I. 1976. 'The 'Kebra Nagast' in the Light of Recent Research', *Le Muséon*, 89 (1976), 133–178.
- Sokolinski, E. 2016. 'The TraCES Project and Gə'əz Studies', in Bausi and Sokolinski 2016, 13–16.
- Vertan, C. 2016. 'Bringing Gə'əz into the Digital Era: Computational Tools for Processing Classical Ethiopic', in Bausi and Sokolinski 2016, 31–42.
- Zuurmond, R. 2001. *Novum Testamentum Aethiopice*, III: *The Gospel of Matthew*, Aethiopistische Forschungen, 55 (Wiesbaden: Harrassowitz Verlag, 2001).

**Part 2**  
**Case Studies**





# Dealing with the Stratigraphy of Coptic Codices: the Cases of MSS Pierpont Morgan Library M578 and Coptic Museum, inv. 13446\*

*Nathan Carlig, Sapienza Università di Roma*

This paper proposes a codicological analysis of two Coptic manuscripts, illustrating how the accurate protocol of codicological description developed within the ERC Advanced Grant project ‘PATHs’, combined with the concepts and the models of the codex stratigraphy described in the recent book *La syntaxe du codex*, can lead to innovative results and new perspectives in understanding Coptic book production.

## **Introduction**

A new detailed protocol of description of Coptic manuscripts has been developed within the ‘PATHs’ project.<sup>1</sup> This protocol takes into account all the physical features of a manuscript, such as the content, the language and/or dialect, the book form and writing support, the number of original and preserved leaves or fragments, the main characteristics of a page (columns, lines per column, characters per line), the dimensions, the quire collation, the presence of quire signatures, pagination, foliation (both ancient and modern), the writing, the manufacture of papyrus codex from papyrus rolls, the binding, the presence of decorations, the inks, ancient or modern restorations, etc. For almost every information recorded, there is the possibility to explain in detail the sources on which this is based and to provide bibliography. Moreover, every manuscript record can be linked to one or many works, identified by means of a *Clavis Coptica* (CC) ID,<sup>2</sup> places, titles, and colophons, all recorded in specific database tables developed by members of the ‘PATHs’ team.

A special section is also dedicated to the codex ‘stratigraphy’, in order to apply to Coptic manuscripts the last tendencies in Greek and Latin codicology summarized in *La syntaxe du codex. Essai de codicologie structurale* published in 2013 by Patrick Andrist, Paul Canart, and Marilena Maniaci. In the book, they describe a new method of genetic and stratigraphic analysis

\* The research leading to these results has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 687567, PI: Paola Buzi, Sapienza Università di Roma), <<http://paths.uniroma1.it>>.

1 See the project note by Paola Buzi, Julian Bogdani, and Francesco Berno in this issue, § 1.

2 The *Clavis Coptica* (CC) or *Clavis Patrum Copticorum*, created by Prof. T. Orlandi on the model of the *Clavis Patrum Graecorum* (CPG), aims at providing a complete set of univocal identifiers for Coptic works. It is now fully available online at <[www.cmcl.it](http://www.cmcl.it)>.

of complex manuscripts, that is to say manuscripts in codex form that underwent transformations related to their content, writing support, and/or binding during their history.<sup>3</sup> The sources for this study are the very numerous Greek manuscripts, of which the three authors are renowned specialists.

In this book, new concepts have been proposed, such as that of ‘Unité de production’ (UniProd), which is defined as:

l’ensemble des codex ou des parties de codex qui sont le résultat d’un même acte de production. L’acte de production est l’ensemble des opérations, délimitées dans le temps et dans l’espace, qui créent un ou plusieurs objets ou parties d’objets, dans notre cas, un ou plusieurs codex ou parties de codex<sup>4</sup>

and the ‘Unité de circulation’ (UniCirc), defined as:

l’ensemble des éléments qui constituent un codex à un moment déterminé. Elle peut équivaloir à une UniProd ou / et être le résultat d’une transformation.<sup>5</sup>

Another important concept introduced is that of the ‘models of transformations’ that a manuscript may undergo. These are classified into two main categories, simple and multiple. A model of simple transformation can be identified as one operation aimed at increasing or reducing a codex by adding or removing content and/or writing material, joining together previous UniCircs to make a new larger UniCirc, destroying part(s) of a codex or dividing it into more new UniCircs. A model of multiple transformations can be identified as a succession of simple transformations.<sup>6</sup>

In this paper, I will apply both the ‘PATHs’ protocol and the concepts and the models of transformation of the codex defined in *La syntaxe du codex* to two parchment manuscripts: New York, Pierpont Morgan Library, M578 (CLM<sup>7</sup> 231) and Cairo, Coptic Museum, inv. 13446 (CLM 3469 + 6293). The double aim is to show 1) to what extent the stratigraphic method of analysis of *La syntaxe du codex*, which is based, as already said, on Greek manuscripts, can be applied to Coptic manuscripts, and 2) how this combined analysis can lead to innovative results regarding Coptic manuscripts and codicology.

3 Andrist, Canart, and Maniaci 2013, 7–9. In this sense, most manuscripts are complex in one or the other way.

4 Ibid. 59. See also pp. 59–60.

5 Ibid. 59. See also p. 61.

6 For a complete description see Andrist, Canart, and Maniaci 2013, 61–70 (models of simple transformations) and 70–79 (models of multiple transformations).

7 The Coptic Literary Manuscript (CLM) number is the identifier used in the ‘PATHs’ database to record Coptic manuscripts. On the IDs assigned to MS M578, see also note 16 below.

### § 1. New York, Pierpont Morgan Library, M578

MS New York, Pierpont Morgan Library, M578 comes from the Monastery of the Archangel Michael of the Fayyum and is dated probably to the tenth century.<sup>8</sup> It contains three works on 130 leaves: 1) Isaac of Kalamon, *Vita Samuelis Archimandritae* (CC 0216) on ff. 1r–68r, 2) Ephrem the Syrian, *In Ioseph patriarcham* (CC 0138) on ff. 69r–97r, and 3) the apocryphal *Paralipomena Ieremiae* (CC 0576) on ff. 97v–130v. F. 68v is blank. The manuscript is recorded in L. Depuydt's *Catalogue of the Coptic Manuscripts in the Pierpont Morgan Library*,<sup>9</sup> and as MICH.BF in T. Orlandi's *Corpus dei Manoscritti Copti Letterari*.<sup>10</sup> Thanks to a critical reading of Depuydt's description by Paola Buzi, Agostino Soldati, and myself, as well as the analysis of the facsimile<sup>11</sup> and the examination of the original manuscript in New York in July 2017 by Maria Chiara Giorda, I will show that this codex consists actually of two distinct ancient UniCircs each corresponding to a single UniProd, the first going from f. 1 to f. 68, the second, from f. 69 to f. 130.

In fact, a series of codicological features are indicative (see table 1). First, we must have a look at the quire collation. The first UniProd (quires I–IX, ff. 1–68) is composed of eight quaternions (quires I–VIII, ff. 1–64) followed by one binion (quire IX, ff. 65–68). The second UniProd (quires X–XVII, ff. 69–130) is composed of six quaternions (quires X, XII, and XIV–XVII) and two quires of seven leaves (quires XI and XIII). The pagination and the quire signatures are reinitialized in the second unit. The pagination runs first from  $\alpha$  (f. 1r) to  $\rho\lambda\epsilon$  (f. 68r, while f. 68v is left completely blank). Then, the pagination starts again from  $\alpha$  (f. 69r) and goes to  $\rho\kappa\Delta$  (f. 130v). In a similar way, the quire signatures are usually (but not always) written on the first and on the last page of each quire. They run from  $\alpha$  (f. 8v) to  $\theta$  (f. 68r, since f. 68v is blank), then start again from  $\alpha$  (f. 76v) and go to  $\theta$  (f. 130v). The end of the first series of page and quire numbers corresponds to the end of the *Vita Samuelis Archimandritae* (f. 68r), followed by a blank page (f. 68v). On the other hand, the *Paralipomena Ieremiae* start on f. 97v, that is, immediately after the end of Ephrem the Syrian's homily on f. 97r. The decoration of the page numbers on ff. 1–68 always follows a specific pattern, while another pattern is used on ff. 69–97. The writing is a bimodular upright majuscule (sloping for the titles) showing a contrast between thick and thin strokes. However the contrast looks

8 The date 'between April 14, ad 891 and August 29, 893' for M578 stated in Depuydt 1993, 357 is actually based on the date of the colophon written on the leaf New York, Pierpont Morgan Library, MS 3815, reused as lower pastedown of M578. It therefore cannot be taken as a sure dating of M578.

9 Depuydt 1993, 357–359 (no. 173).

10 Available at <<http://www.cmcl.it>>.

11 Hyvernat 1922.

Quires	Folia	Content	Pagination	Decoration of page nos	Quire numbering (first–last page)	Lots at the time of acquisition
I (ff. 1–8)		Isaac of Kalamon, <i>Vita Samuelis Archimandritae</i> (CC 0216)	α–παε	pattern 1	α (only on the last page)	Lot A
II (ff. 9–16)	β–β					
III (ff. 17–24)	γ–γ					
IV (ff. 25–32)	Δ–Δ					
V (ff. 33–40)	ε–ε					
VI (ff. 41–48)	ς–ς					
VII (ff. 49–56)	ζ–ζ					
VIII (ff. 57–64)	η–η					
IX (ff. 65–68)	θ (only on the first page)					
	68v	Blank	Blank		Blank	
X (ff. 69–76)		Ephrem the Syrian, <i>In Ioseph Patriarcham</i> (CC 0138)	α–ρκλ (with some errors)	pattern 2	α (only on the last page)	Lot B (first part)
XI (ff. 77–83)					Δ–β	
XII (ff. 84–91)	90r				γ–γ	Lot C
XIII (ff. 92–98)		<i>Paralipomena Ieremiae</i> (CC 0576)			Δ–Δ	Lot B (second part)
	97r					
	97v					
XIV (ff. 99–106)					ε–ς	
XV (ff. 107–114)					ζ–ζ	
XVI (ff. 115–122)					η–η	
XVII (ff. 123–130)					θ–θ	

Table 1. Discontinuities identified in MS New York, Pierpont Morgan Library M578.

sharper after f. 69r. The ruling type in the whole manuscript is a very simple one (Leroy 00A2, Muzerelle 1-1-11/0/0/A).

Finally, the modern history summarized in the library catalogue<sup>12</sup> confirms the fact that there were originally two distinct manuscripts. The leaves were acquired in three separate lots. Lots A (ff. 1–68) and B (ff. 69–89 and 97–130) were bought as two distinct codicological units in Paris in 1911, while lot C (ff. 90–96) was acquired in Cairo in 1912. During their restoration in Rome in the 1910s, lot A was named M578, while lots B and C were joined

12 Depuydt 1993, 359.

together as M601. In Hyvernat's *Facsimiles*,<sup>13</sup> the three parts are joined under the shelfmark M578 as one manuscript, which is still the case, while shelfmark M601 was reassigned to another manuscript.<sup>14</sup>

Following the observations made above, two records were initially created in the 'PATHs' codicological database. Ff. 1–68, corresponding to the first UniProd/UniCirc, with Isaac of Kalamon's *Vita Samuelis Archimandritae*, were recorded as CLM 231. Another record, CLM 712, was created for ff. 69–130, corresponding to the second UniProd/UniCirc, with Ephrem the Syrian's homily *In Ioseph patriarcham* and the *Paralipomena Ieremiae*.

This example (fig. 1) shows how the model of transformation A4—characterized by a union of two codices—works, as described in *La syntaxe du codex*.<sup>15</sup> It focuses on the ancient history of the manuscript, before the division into three lots. The 'PATHs' database always tries to go up to the most ancient UniCirc.<sup>16</sup>

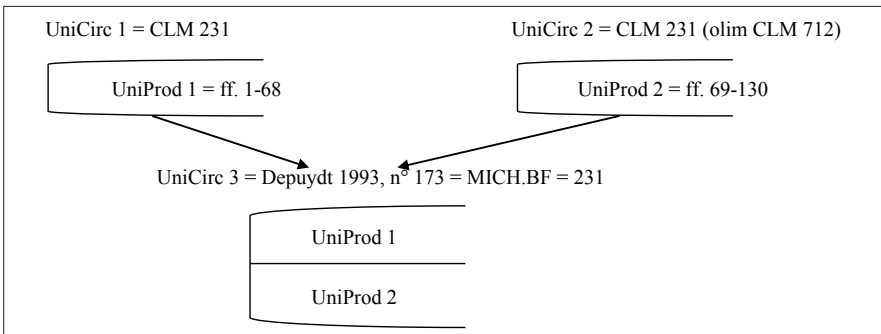


Fig. 1. A4 model of transformation of MS New York, Pierpont Morgan Library, M578.

13 Hyvernat 1922.

14 CLM 256 = MICH.CE. See Depuydt 1993, 62–64 (no. 45).

15 Andrist, Canart, and Maniaci 2013, 66.

16 I thank Francesco Valerio, who told me, while this article was already in press, about other manuscripts from the Monastery of the Archangel Michael near Phan-toou, that can be similar to M578 in their manufacture. According to his analysis for the 'PATHs' project, these manuscripts are surely composed of one UniProd, while some of them show a double pagination and quire numbering, such as CLM 203, 218, 221, 222, 228, and 241, and some others, only a double pagination, such as CLM 225 and 237. An accurate serial analysis of all these manuscripts showing double pagination and/or quire numbering, could certainly lead to more nuanced results, since double pagination and quire numbering cannot be alone a sure element to identify an unification of two previous UniCirc. Nevertheless, thanks to this observation, there is a possibility that M578 was already one UniCirc in ancient times. Consequently, for now, in the framework of the 'PATHs' project, we decided to gather both UniProd/UniCirc in one record (CLM 231) and to cancel CLM 712.

## § 2. *Cairo, Coptic Museum, inv. 13446*

The manuscript now kept in Cairo, Coptic Museum, inv. 13446 was discovered in the Theban hermitage MMA 1152 in March 2005,<sup>17</sup> together with two papyrus manuscripts, one containing Pseudo-Basil's *Canons* (CC 0090),<sup>18</sup> the other containing the *Encōmion* of Pisenthius, bishop of Coptos (CC 0238).<sup>19</sup> The datation is debated. Some scholars point to a datation to the ninth or tenth century, while, for others, the manuscript can be dated to the seventh or eighth century.<sup>20</sup> The manuscript consists of a set of 52 parchment leaves, originally kept between two wooden boards. The quire collation cannot be described so far. The parchment is considerably damaged and seems to have been exposed to high temperatures. Moreover, all the leaves are mutilated, mostly in the inner part, i.e. the part near the binding (if there was one). On ff. 1–50, the manuscript contains the last part of the Old Testament book of *Isaias* (CC 0739), corresponding to chapters 47–66.<sup>21</sup> The text is written in two columns per page in Biblical majuscule and ends with a final title. The last two leaves, ff. 51–52, combine decorations and drawings (frame, cross, and birds) with text (*Acta Petri* = CC 0026)<sup>22</sup> and request particular attention.

Leaf 51 has the same dimensions as the previous leaves and also shows the same pattern of ruling and pricking, but no pagination. On the recto, an interlaced frame containing an interlaced cross was drawn in red and green. The four rectangles left blank inside the frame were later filled by a second hand with a small informal sloping majuscule more recent than the elegant Biblical majuscule of the *Isaias* text. The text is identified as the *Acta Petri*. On the verso of the leaf, we see a bird and another drawing, around which the continuation of the *Acta Petri* was written. The second hand continues the copy of the *Acta Petri* on f. 52r and v. Leaf 52 however presents some phys-

17 On the discovery, see Górecki 2007.

18 Cairo, Coptic Museum, inv. 13448 = CLM 713.

19 Cairo, Coptic Museum, inv. 13447 = CLM 714.

20 On this debate, see Boud'hors 2017, 195.

21 The comparison of the very badly preserved first pages of the codex with other manuscripts bearing the text of *Isaias* lead us to identify the beginning of the text in this codex with chapter 47. It is therefore no coincidence if another testimony of *Isaias*, P.Bodmer XXIII = CLM 40 = DISH.AH (end of the fourth or first half of the fifth cent.; see Kasser 1965), a fully preserved parchment codex coming from the so-called 'Bodmer Papyri' or 'Dishna Papers', hosts the text of *Isaias* from ch. 47 to the end. As the codex bears the title 'The third part of the book of Isaias the Prophet' (π[μεζ]ωμον τ η μ[ερο]ς η π[ω]με η [ησα]ας πεπροφητης) on f. 2v, it indicates that the division of *Isaias* into three parts, the third one corresponding to ch. 47–66, was common. The *Isaias* text of MS Cairo, Coptic Museum, inv. 13446 is being edited by A. Suciú (Göttingen, Germany).

22 The *Acta Petri* are being edited by P. Piwowarczyck (Katowice, Poland).

ical features that make it different from the previous ones. It is a bit smaller than leaves 1–51 (14 × 9 cm against 14 × 11 cm)<sup>23</sup> and has no ruling, pricking, or pagination. It is therefore possible that f. 52 was added later to complete the writing space necessary to finish the copy of the *Acta Petri* (see table 2).

f.	r/v	F/H	Pagination	Text
	r	H	No pag.	Cross and frame + <i>Acta Petri</i>
51	v	F	No pag.	Bird + second drawing + <i>Acta Petri</i>
	r	F	No pag.	<i>Acta Petri</i>
52	v	H	No pag.	<i>Acta Petri</i>

Table 2. Content of the two last leaves of MS Cairo, Coptic Museum, inv. 13446.

Thus, we can conclude that the manuscript is composed of two UniProds. UniProd 1 consists of ff. 1–51, with the *Isaias* text written in Biblical majuscule (ff. 1–50), the decorated frame with the cross (f. 51r), and the drawings (f. 51v), while UniProd 2 corresponds to the copy of the *Acta Petri* by the second hand on the space left blank on f. 51r and 51v and continued on f. 52, which was added for this specific purpose. Thanks to this analysis, UniProd 2 can be further identified as a UniProd-C-MC, since new content on previous material (C) and new content on new material (MC) are added.<sup>24</sup>

The manuscript further corresponds to two UniCircs. UniCirc 1 is composed of ff. 1–51 before the writing of *Acta Petri*, and is later transformed into UniCirc 2, after the copy of the *Acta Petri* and the addition of f. 52. It therefore fits perfectly in model of transformation A3 described in *La syntaxe du codex*, which is characterized by adding a new content both on blank parts of the manuscripts and on new material added at the end on the manuscript.<sup>25</sup> Moreover, in this case we clearly observe, on f. 51, that the end of UniProd 1 (frame, cross and drawings) and the beginning of UniProd 2 (*Acta Petri*) are interlacing or overlapping, since both production acts were executed on the same leaf.

In the ‘PATHs’ codicological database, two records have been created for Cairo, Coptic Museum inv. 13446, each corresponding to one UniProd. UniProd 1 is described in CLM 3469 and UniProd 2 is recorded as CLM 6293. In this specific case, the field ‘Codex stratigraphy’ is very helpful to describe the relationship of both records and UniProds, so as to highlight the his-

23 I thank Paola Buzi for taking the measurements of the codex during a research mission to the Coptic Museum in June 2018.

24 On this more precise typology, see Andrist, Canart, and Maniaci 2013, 60.

25 Andrist, Canart, and Maniaci 2013, 65.

tory of the manuscript and its transformations. Fig. 2 summarizes the above observations.

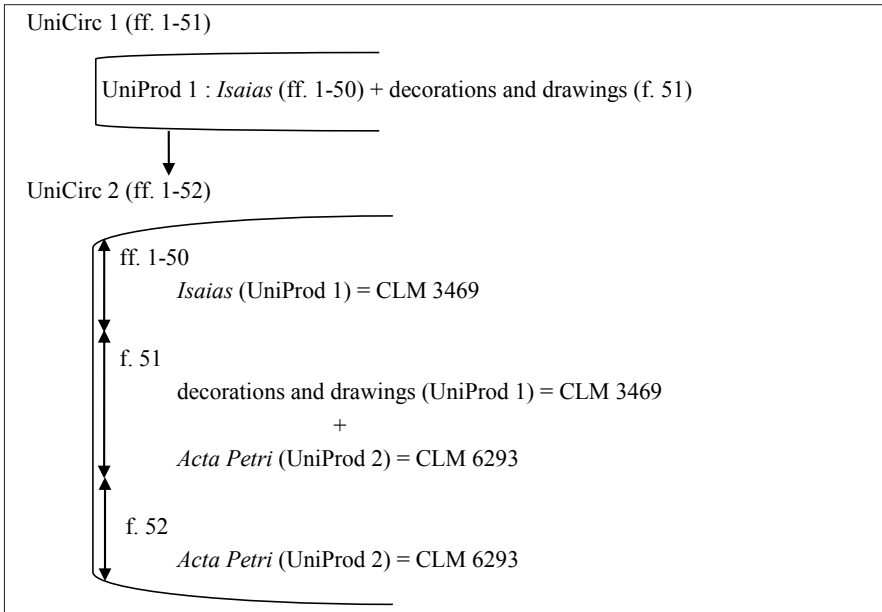


Fig. 2. A3 model of transformation of MS Cairo, Coptic Museum, inv. 13446

The application of an accurate description protocol to Coptic manuscripts like the one developed in the ‘PATHs’ project leads to new conclusions about already known manuscripts, such as identifying two manuscripts in what were thought to be one original manuscript, or distinguishing two phases of production of a manuscript. Moreover, although the conclusions reached in *La syntaxe du codex*, which rely only on Greek manuscripts, are sometimes very sophisticated, and seem hardly applicable to the Coptic manuscripts (mainly because of their poor state of conservation and their dispersal throughout the world, which make their reconstruction very difficult), our case studies show that it is worth making an attempt. As we have seen above, we can for now conclude that models of simple transformation fit well with what we can obtain from the study of well-preserved Coptic manuscripts. Trying to apply these concepts and models more systematically to Coptic manuscripts will therefore allow us to shed a new light on Coptic book production.

## References

Andrist, P., P. Canart, and M. Maniaci 2013. *La syntaxe du codex. Essai de codicologie structurale*, Bibliologia, 34 (Turnhout: Brepols, 2013).



- Boud'hors, A. 2017. 'À la recherche des manuscrits coptes de la région thébaine', in D. Brakke, S.J. Davis, and S. Emmel, eds, *From Gnostics to Monastics. Studies in Coptic and Early Christianity in Honor of Bentley Layton*, *Orientalia Lovaniensia Analecta*, 263 (Leuven: Peeters, 2017), 175–212.
- Depuydt, L. 1993. *Catalogue of Coptic Manuscripts in the Pierpont Morgan Library*, *Corpus of Illuminated Manuscripts*, 4 (Leuven: Peeters, 1993).
- Górecki, T. 2007. 'Sheikh Abd el-Gurna (Hermitage in Tomb 1152). Preliminary Report, 2005', *Polish Archaeology in the Mediterranean*, 17 (2007), 263–274.
- Hyvernat, H., ed. 1922. *Bibliothecae Pierpont Morgan codices coptici photographice expressi*, XXXI: *Vita Samuelis abbatis monasterii Virginis in monte Kalamon auctore Isaaco monacho, Ephremi anachoretæ sermo in patriarcham Ioseph, Paralipomena Ieremiae Prophetæ Sahidice* (Romae, 1922; repr. New York: Pierpont Morgan Library, 1958).
- Kasser, R. 1965. *Papyrus Bodmer XXIII. Esaïe XLVII, 1 – LXVI, 24* (Cologne-Genève: Bibliotheca Bodmeriana, 1965).



# Die Quellen von August Dillmanns *Lexicon linguae Aethiopicae*:

Anmerkungen zu den Prolegomena und den verwendeten Sigla\*

Wolfgang Dickhut und Andreas Ellwardt,  
Universität Hamburg

Dillmann's *Lexicon linguae Aethiopicae* has timeless value in various ways, nevertheless it needs a kind of user's manual. Even apparently obvious references require an explanation on several levels. Here we give a sample of such explanations. A completed list will eventually be provided via an online application.

Der bis heute unverminderte Wert von Dillmanns *Lexicon* besteht im Wesentlichen darin, dass er ein echtes Belegwörterbuch verfasst hat, er mithin also jedes Lemma durch mindestens eine Quelle belegt. Dillmann greift hierfür sowohl auf Handschriften als auch auf gedruckte und zum Teil von ihm selbst edierte Werke zurück. In seinen Prolegomena gibt er einige Informationen zu den benutzten Werken und schließt dies mit einer Tabelle zu den „Sigla in nostro opere adhibita“ ab.<sup>1</sup>

Die Benutzung dieser Informationen stellt sich im Detail jedoch gelegentlich etwas schwierig dar. Dies ist nicht nur dem Umstand geschuldet, dass sie in lateinischer Sprache gegeben werden, wenn dies auch die Benutzung für viele nicht gerade erleichtert. Im Folgenden sollen an einigen Stichproben gewisse problematische Gegebenheiten dargestellt werden; dies als Vorarbeiten zu einem vollständigen Verzeichnis der Quellen, der verwendeten Abkürzungen nebst Auflösung und dem Verweis auf die derzeit verfügbaren Editionen.

Dillmann beginnt mit der Bibel, indem er zunächst die Problematik der unterschiedlichen äthiopischen Bibelausgaben beschreibt; bei einigen Büchern kann er dadurch Abhilfe schaffen, dass er bereits selber eigene Ausgaben veröffentlicht hat, und wo dies der Fall ist, nutzt er selbige auch. Für einige biblische Bücher greift er auf verschiedene Handschriften zurück, benennt diese allerdings in der Einleitung lediglich summarisch und weist die konkrete Handschrift gegebenenfalls bei den einzelnen Lemmata nach. Verwirrend wird für manch einen das Ordnungsprinzip seiner Darstellung sein, er führt nämlich die einzelnen Bücher nach der Anzahl der von ihm kollationierten

\* Dieser Beitrag entstand in dem vom Europäischen Forschungsrat im Rahmen des Siebten Rahmenprogramms der Europäischen Union geförderten Projekt „TraCES: From Translation to Creation. Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages“ (ERC Advanced Grant „TraCES“; Grant Agreement Nr. 338756).

1 Dillmann 1865 (im Folgenden DL), V–XII, XXVII–XXX.

Handschriften auf (und während man noch einigermaßen sechs Weisheitsbücher zusammenbekäme, fragt man sich spätestens bei vermeintlich zwei Büchern Judith, welche eigenartige Bibel in Äthiopien wohl in Gebrauch wäre...): „In caeteris libris ad Codices Mss. recurrendum erat, quorum inter se conferendos habebamus [...] Jobi octo; III et IV Regum et Danielis septem; Jesaiae, Dodecapheton, Proverbiorum, Ecclesiastae, Sapientiae sex; Paralipomenon, I et II Ezrae, Esther quatuor; Jeremiae cum Threnis, Hezekielis, Siracidae tres; Baruch, Epistolae Jer., Tobith, Judith duos.“

Für den christlichen Orient ist allgemein die Bucheinteilung nach der Septuaginta gebräuchlich, was bekanntlich vor allem bei den vier Königsbüchern ins Auge fällt. So zitiert auch Dillmann die entsprechenden Passagen nach „1 (2 / 3 / 4) Reg.“. Allerdings finden sich auch vier Einträge mit „1 Sam.“ und zwei mit „2 Sam.“. Wenn Dillmann col. 606 s.v. **ላላት** :<sup>2</sup> 1 Sam. 13,6 anführt, ist das noch durch den Bezug auf das hebräische **לילך** nachvollziehbar. Ähnlich verhält es sich col. 465 s.v. **ቁድሮስ** :,<sup>3</sup> wo er eine Fehlesung in der äthiopischen Bibel, nämlich „Zeder“ für „Kidron“, dokumentiert („*Cedron torrens inter Hierosolyma et montem olivarium* (לילך): **ፈለገ** : **ቁድሮስ** : Joh. 18,1 (**ቁድሮስ** : **ፈለገ** : **አርዛ** : Platt, coll. 2 Sam. 15,23 al.)“). Bei allen anderen Fundstellen bleibt die Sinnhaftigkeit verborgen.

Unter „Kid.“ bietet Dillmann die wohlbekanntete Folge von Testamentum Domini, Testament in Galiläa und Epistula Apostolorum aus der Handschrift Stuttgart, Württembergische Landesbibliothek, or. 49, ohne zwischen den drei Texten zu unterscheiden. Ergo sind die Zahlen von 3ra bis 31rb TD, von 31rb bis 33va TG und von 33va bis 43vc EpAp. Hierbei fällt besonders auf, dass Dillmann die Handschrift weder nach recto-verso noch nach den drei Spalten weiter spezifiziert. Zusätzlich zum vollständigen Testamentum-Text zieht er auch zwei Texte heran, die in Äthiopien viel weitere Verbreitung gefunden haben. Dies ist zum einen das Morgenlob **ኪዳን** : **ዘንግሠ** : , aus dem er unter anderem „gratificator, largitor, benefactor, de Deo, qui nominatur **ባቋ** : **ለኩሉ** : Kid. f. 22 et **ባቋ** : **ተደላ** : **ዘኢይመውት** : Prec. Mat.“ anführt;<sup>4</sup> die identische Passage findet sich auf f. 13rb des „Kid.“. Ähnliches gilt für col. 888, s.v. **ወላዲ** : , hier führt er nur Prec. Mat. an,<sup>5</sup> aber auch hier wäre derselbe Text in seiner Stuttgarter Handschrift zu finden. Eine weitere Doppelung zum Testamentum Domini ist die Anaphora unseres Herrn Jesus Christus.

Mit „Lit.“ ohne weitere Zusätze bezeichnet Dillmann die drei in der römischen Ausgabe des äthiopischen NT<sup>6</sup> auf den Folien 158–176 abgedruck-

2 <<http://betamasaheft.eu/Dillmann/lemma/L5f5eca381f4941fbb8537c6d245dd24f>>

3 <<http://betamasaheft.eu/Dillmann/lemma/L22dfeb25df1f4b15b657886c241ed5c0>>.

4 <<http://betamasaheft.eu/Dillmann/lemma/Lba6afd79537a4cc7890d11f098db97cb>>.

5 <<http://betamasaheft.eu/Dillmann/lemma/L901846c68ff64a58a7b2da0367fdc709>>.

6 Petrus Ethiops 1548.

ten liturgischen Stücke, nämlich den Canon generalis, die Anaphora unseres Herrn Jesus Christus und die Anaphora unserer Herrin Maria. Er verweist darauf, dass die Anaphora unseres Herrn Jesus Christus auch in Ludolfs *Commentarius* gedruckt vorliegt<sup>7</sup>, was wohl darauf schließen lässt, dass dieses Werk für die seinerzeit von Dillmann im Blick gehabte Leserschaft des *Lexicons* leichter zugänglich war. Unter letztgenanntem Aspekt ist es hilfreich, die drei genannten Stücke genauer zu lokalisieren, da sowohl das äthiopische Missale als auch das Testamentum Domini Nostri Jesu Christi in modernen Editionen greifbar sind. Dillmann verweist auf sie mit der Sigel „Lit.“ gefolgt von der Foliennummer, dieser folgt durch Komma abgetrennt eine Zahl zwischen 1 und 4 in der Bedeutung 1 = ra, 2 = rb, 3 = va, 4 = vb. Eine Zahl zwischen 158,1 und 167,4 meint also den Canon generalis, zwischen 168,1 und 169,4 die Anaphora unseres Herrn Jesus Christus und zwischen 170,1 und 177,4 die Anaphora unserer Herrin Maria. So gilt also für beispielsweise das Lemma **ዐጽፍ ፡** (col. 1024) Ähnliches wie oben: Der Beleg **አዕጽፍተ ፡ ብርሃኑ ፡**, von Dillmann als Lit. 168,3 bezeichnet, könnte ebensogut auf Kid. 10r verweisen.

Es wurde bereits an anderer Stelle darauf hingewiesen, dass Dillmanns Erwähnung des *Maṣḥafa Məstir* („M.M.“) den Benutzer vor gewisse Probleme stellt.<sup>8</sup> Dies wiegt umso schwerer, als Dillmann in nicht weniger als 674 Einträgen auf dieses Werk Bezug nimmt. Seine Feststellung, er hätte einen Codex aus Basel von 372 Folien benutzt, kann dahingehend präzisiert werden, dass dies eine der Handschriften meinen muss, die Johann Ludwig Krapf nach Tübingen gebracht hatte, die von dort jedoch „an das Missionshaus in Basel geschenkt“ wurde, wie Ewald schreibt.<sup>9</sup> Nachforschungen vor Ort, sowohl in Basel als auch in Tübingen, blieben jedoch ergebnislos, die Handschrift scheint verschollen. Eine zusätzliche Frage ergibt sich aus dem Umstand, dass Dillmanns Verweise auch auf höhere Foliennzahlen als 372 gehen, so hat er beispielsweise col. 152 s.v. **ተአምላክ ፡** „Dei honore affici: **ለይሱሳ ሳሎ ፡ ስሙ ፡ ለእግዚአብሔር ፡ ስቡሕ ፤ ይትአምላክ ፡ (ስሙ ፡) ለእግዚአብሔር ፡ አምላክ ፡** M.M. f. 391“.

Über die von Dillmann gelesene Handschrift des *Gadla Lālibalā* hat Nafisa Valieva unlängst berichtet.<sup>10</sup> Hier ist hauptsächlich Dillmanns Formulierung über einen Kodex „Gadela Lalibela qui quondam in Bibl. Societ. Ind. Orient. asservabatur“ zunächst schwer verständlich, hat doch Enrico Cerulli eine Handschrift dieses Textes aus der Bibliothek des India Office beschrieben. Aus Valievas Ausführungen geht hervor, dass besagte Handschrift von

7 Ludolf 1691, 341–345.

8 Ellwardt 2016, 195.

9 Ewald 1844, 194 Anm. 1.

10 Valieva 2017.

Johann Ludwig Krapf nach London geschickt wurde, wo sie 1842 im India Office ankam. Dort hat Dillmann sie eingesehen, hat aber allem Anschein nach zum Zeitpunkt der Abfassung seines *Lexicon* erfahren, dass sie nicht mehr auffindbar wäre. Dies legt neben seiner Wortwahl „quondam“ auch der Umstand nahe, dass William Wright sie in seinem 1886 entstandenen handschriftlichen Katalog nicht erwähnt.<sup>11</sup>

Dillmann zitiert in 243 Einträgen aus dem *Maṣḥafa Dəggʷa*, allerdings ohne jegliche Stellenangabe. Die Prolegomena klären uns darüber auf, dass er hierzu seine eigenen Notizen benutzt, die er anhand einer Handschrift der Bibelgesellschaft in London angefertigt hat. Von dieser Handschrift berichtet Löfgren allerdings, dass sie verschollen sei.<sup>12</sup>

Die praktische Benutzbarkeit von Dillmanns *Lexicon* kann bereits dadurch wesentlich gesteigert werden, dass die von ihm im Abkürzungsverzeichnis gegebenen Auflösungen mit den in den Prolegomena gegebenen Informationen verknüpft werden, darüber hinaus aber auch mit Hinweisen auf den heutigen Stand der Disziplin, wie er sich etwa in den Bänden der *Encyclopaedia Aethiopica* niederschlägt.

In einem ersten Schritt müssen deshalb sämtliche tatsächlich von Dillmann verwendeten Abkürzungen aufgelöst und auf eine Weise erläutert werden, dass auch fachfremden Benutzerinnen und Benutzern etwa aus der Theologie, der Semiotik oder Klassischen Philologie unmittelbar klar wird, was Dillmann jeweils gemeint hat. Aus dieser Zielsetzung heraus sind den einzelnen Werken oder Sammlungen (etwa dem *Qerällos* oder dem *Haymānota ʾAbaw*) ggf. auch kurze und prägnante Charakterisierungen mitzugeben, die für Kennerinnen und Kenner der äthiopischen Literatur überflüssig sind. Im Folgenden sind als Muster die Erläuterungen zu vier „Kreisen“ mit den von Dillmann verwendeten Abkürzungen gegeben: die drei großen Sammlungen dogmatischer und kanonistischer Texte *Qerällos*, *Haymānota ʾAbaw* und *Senodos* sowie die von Dillmann in den Prolegomena (Abschnitt 1, Punkt m) zusammengestellten poetischen Texte. Das Gesamtverzeichnis wird später in englischer Sprache online<sup>13</sup> zur Verfügung gestellt werden.

11 Die Geschichte setzt sich in eigenartiger Weise bis in die Gegenwart fort: Cerulli konnte die Handschrift 1946 beschreiben, aber als Valieva sie 2016 einsehen wollte, wurde sie zunächst als unauffindbar gemeldet und nur Valievas Beharrlichkeit ist es zu verdanken, dass erfolgreiche Anstrengungen zu ihrer Wiederauffindung unternommen wurden.

12 Löfgren 1974, 75.

13 Die aktuelle Arbeitsversion befindet sich unter <<http://betamasasheft.eu/Dillmann/abbreviations>>. Über die *Beta maṣāḥaft* Plattform (BM) s. Reule in diesem Band.

**I. *Qerillos***

Der *Qerillos* ist eine Sammlung von größtenteils direkt aus dem griechischen Original übersetzten Vätertexten zu den christologischen Streitfragen im Umkreis des Konzils von Ephesus im Jahre 431.<sup>14</sup> Sie beginnt mit drei klassischen Schriften Kyrills von Alexandrien, nach welchem sie insgesamt benannt ist. Dillmann benutzte einen Tübinger Kodex (Universitätsbibliothek Tübingen, Ms. Orient Ma IX 28), auf den auch seine Folio-Angaben verweisen.<sup>15</sup> Zum Vergleich zog er Abschriften heran, die er sich bei einem Aufenthalt in London aus dortigen Handschriften gemacht hatte.<sup>16</sup>

Cyr. — Belege aus dem *Qerillos*, mit Verweis auf die Tübinger Handschrift. Fünf Schriften Kyrills, die Teil der Sammlung sind, zitiert Dillmann mit jeweils eigenem Siglum, nämlich:

Cyr. ad Theod. (= ad Theodosium) — Prosphonetikos „Über den rechten Glauben“ an Kaiser Theodosius II.; in der Tübinger Handschrift ff. 6r–25r.<sup>17</sup>

Cyr. ad Reg. (= ad Reginas) — Prosphonetikos „Über den rechten Glauben“ an Arkadia und Marina (Schwestern von Kaiser Theodosius), und zwar die „dogmatische Einleitung“; in der Tübinger Handschrift ff. 25v–31r.<sup>18</sup>

Cyr. de r. fid. (= de recta fide) — das ist der Hauptteil der eben genannten Schrift, der „hauptsächlich aus einem immensen Schriftstellenmosaik zur Darstellung der rechten Lehre über Christus [besteht]“<sup>19</sup>; in der Tübinger Handschrift bis f. 65v.<sup>20</sup>

Cyr. c. Pall. (= colloquium cum Palladio) — der berühmte Dialog „Dass Christus einer ist“; in der Tübinger Handschrift ff. 65v–98v.<sup>21</sup>

Cyr. [ho.] de Melch. (= [homilia] de Melchisedech) — die sieben Belege, die Dillmann mit diesem Siglum anführt, stammen alle aus der im *Qerillos* enthaltenen Ersten Melchisedekhomilie Kyrills; in der Tübinger Handschrift f. 127r–129r.<sup>22</sup>

14 Bausi 2010a.

15 Vgl. die Beschreibung bei Weischer 1973, 37f. („T“).

16 Nach Weischer 1973, 39 wohl die Handschrift London, British Museum, Add. 16200 („X“). Dillmann spricht in den Prolegomena von Abschriften „e codicibus, qui in Museo Britannico et in Bibliotheca Societatis Indiae Orientalis asservantur“; das mag außerdem auf die Weischer 1973, 34 beschriebene Handschrift London, India Office, Cod. Nr. 3 („I“) hindeuten.

17 Edition: Weischer 1973. Die Folio-Angaben hier und im Folgenden (außer zu Cyr. ad reg.) nach Six 2000.

18 Edition: Weischer 1993.

19 Weischer 1993, 14.

20 Eine „Übersicht über die Zitate der Schriftstellen“ bei Weischer 1993, 101–114.

21 Edition: Weischer 1977.

22 Edition: Weischer 1980, 82–97; vorher bereits Dillmann 1866, 88–93.

Voc. ad Cyr. (= Vocabularium ad Cyrillum bzw. ad Cyrilli opera) — in manchen Handschriften des *Qerāllōs* finden sich als *Sawāsaw* bezeichnete Glossare angehängt, in denen schwer oder gar nicht mehr verständliche Begriffe erläutert werden.<sup>23</sup> Dasjenige der Tübinger Handschrift ist von Dillmann ausgewertet worden.

## II. *Haymānota 'Abaw*

Unter dem Namen *Haymānota 'Abaw* („Bekenntnis der Väter“) läuft die im 16. Jahrhundert angefertigte äthiopische Übersetzung einer gleichnamigen arabischen Anthologie aus dem 11. Jahrhundert, *I'tirāf al-'ābā'*, in der sich patristische Texte vor allem zu trinitarischen und christologischen Fragestellungen zusammengestellt finden.<sup>24</sup> Für beide Fassungen fehlen bisher kritische Editionen. Die umfassendste Studie zu den verschiedenen Texten, aus denen sich das arabische Original zusammensetzt, legte 1937 Georg Graf vor.<sup>25</sup>

Haim. Ab. — Das *Haymānota 'Abaw* hat Dillmann nicht vollständig berücksichtigen können, sondern nur diejenigen Teile, von denen er sich in London eine Abschrift angefertigt hatte. Aus zwei Abschnitten seines Londoner *Haymānota 'Abaw* zitiert er dabei gesondert, nämlich:

M. Tom. – Unter dem Namen *Maṣḥafa Ṭomār*, d. h. „Buch des Briefes (nämlich eines vom Himmel herabgekommenen)“ läuft eine in der gesamten Christenheit populäre Einschärfung der Sonntagsheiligung. Der Text findet sich zumeist in Handschriften, die das *Haymānota 'Abaw* enthalten.<sup>26</sup>

Athan. de Trin. (Athanasius de Trinitate) – Hierbei handelt es sich um einen dem Athanasius zugeschriebenen Text zur Dreifaltigkeit, den Graf nicht zuweisen konnte.<sup>27</sup>

Fid. Jac. Bar. (Fides Jacobi Baradaei) – Über das Arabische ist ein „Glaube(nsbekenntnis) des Jakob Baradaeus“ (*Hāymānotu la-Yā'qob 'Elberād'i*) ins Äthiopische gelangt, das dem Namensgeber der „Jakobiten“ wohl bloß zugeschrieben wird und das sich in vielen *Haymānota 'Abaw*-Handschriften findet.<sup>28</sup>

## III. *Senodos*

Der *Senodos* ist die wichtigste Textsammlung für Kirchenrecht und Liturgie der äthiopischen Kirche. Es besteht weitgehend aus Übersetzungen aus dem

23 Vgl. Weischer 1973, 21–24, Weischer 1980, 136–141 und allgemein Meley Mulgeta 2010.

24 Wion/Fritsch 2005.

25 Graf 1937.

26 Witakowski 2010.

27 Witakowski 2003; Graf 1937, 368 (Nr. 22): „Unbekannt.“

28 Tubach 2007; Graf 1951, 10 (Nr. 125,3); für den arabischen Text Kleyn 1882, 110ff.



Arabischen und hat in seiner bis heute typischen Form bereits in Handschriften des 15. Jahrhunderts Spuren hinterlassen.<sup>29</sup>

Syn. (Synodus) – Unter diesem Siglum werden die Stellen angeführt, die nicht unter eines der nachfolgenden fallen. Außerdem gibt Dillmann in der Form „Lud. e Syn.“ häufig von ihm aus Ludolf geschöpfte Belege aus dem *Senodos*. Abt. [I Abt.] / II Abt. – Der Begriff „Abtəlis“ findet sich mehrfach in äthiopischen Kirchenrechtssammlungen zur Bezeichnung bestimmter Kanones. Im *Senodos* sind damit gemeint zwei in weiten Teilen identische Zusammenstellungen von 81 bzw. 80 (in der Tübinger Handschrift jeweils 81) angeblich apostolischen Kanones.<sup>30</sup>

Can. (Anc./Ant./Ap./Gangr./Laod./Neoc./Nic./II Nic./Sard.); Const. Ap. / II Const. Ap. / III Const. Ap.; Stat. Ap. – Unter diesen Abkürzungskombinationen zitiert Dillmann verschiedene im *Senodos* zusammengefasste Kanonesammlungen.<sup>31</sup>

Exp. Dec. (Expositio decalogi) – Eine dem Johannes Chrysostomus zugeschriebene Erläuterung der Zehn Gebote.<sup>32</sup>

Exp. Fid. Nic. (Expositio fidei Nicenae) – Die „Lehre der 318 Väter von Nizäa“, enthaltend einen dogmatischen und einen praktisch-monastischen Teil.<sup>33</sup>

#### **IV. Poetische Texte**

Deg. – *Dəgg<sup>wa</sup>* ist der gebräuchlichste, in seiner Etymologie unklare Name für das äthiopische Hymnenbuch.<sup>34</sup> Das erste vollständige Exemplar ist spätestens aus dem 15. Jahrhundert erhalten; es handelt sich bei den enthaltenen Hymnen nicht um Übersetzungen oder Abwandlungen griechischer oder arabischer Originale, sondern um eigene äthiopische Schöpfungen.<sup>35</sup> Dillmann benutzt die Exzerpte, die er sich aus einer Londoner Handschrift erstellt hatte (siehe oben).

29 Bausi 2010b. Zur von Dillmann benutzten Tübinger Handschrift 7 vgl. Six 2010, 8f.

30 Bausi 2010b, 623b; Kleiner 2003, dort auch zur Etymologie des Wortes aus gr. *apotélesma* (so Ludolf) oder gr. *títlos* über ar. *at-titlus* (so Dillmann). Ediert in Bausi 1995.

31 Für die einzelnen Titel vgl. Six 2000 und Bausi 2010b, 623b–624a; in der Online-Fassung auf der Plattform *Beta maṣāḥaft* sollen gerade diese Texte detailliert aufgeschlüsselt und genau zugeordnet werden.

32 Bausi 2010b, 623b: „an *Interpretation of the Decalogue* attributed to John Chrysostom“; Witakowski 2007, 295a (Nr. 48); Lusini 1989, 490f.; in der Tübinger Handschrift f. 293v–309r.

33 Bausi 2004; in der Tübinger Handschrift f. 281r–286v.

34 Nach DL s. v. **ድግጋ** : auch für den „cantus ecclesiasticus (Kirchengesang)“ selbst.

35 Habtemichael 2005.

Hymn. Mus. Brit. LIV, 7. – Es handelt sich um acht Hymnen auf ff. 52–63.<sup>36</sup> Machb. Meem. – Das nach den Anfangsworten *Māḥbara Məʿamanān* genannte Kommuniongebet der äthiopischen Messe<sup>37</sup> zitiert Dillmann nach einer Oxforder und einer Londoner Handschrift.<sup>38</sup>

Mavās. – Als *Mawāsəʿt* wird das äthiopische Antiphonar bezeichnet, das die vom Chor zu singenden Teile der Liturgie bietet.<sup>39</sup> Dillmann zieht eine Tübinger Handschrift heran.<sup>40</sup>

Tab. Tab. – *Ṭabiba Ṭabibān* heißt nach der ersten Anrufung Gottes darin ein von den Mönchen außerhalb der Liturgie gesungener Hymnus von hundert Strophen zu je fünf Versen.<sup>41</sup> Dillmann benutzt drei Handschriften aus Oxford und London, nach denen er den Text in seiner *Chrestomathia* auch herausgegeben hat.<sup>42</sup>

Die vollständige Aufarbeitung aller Dillmannschen Quellen ist zeit- und arbeitsaufwendig und gelegentlich auch zunächst erfolglos; sie ist dennoch unabdingbar, um die äthiopische Lexikographie weiter voranzutreiben. Die elektronische Plattform des *Lexicon* (<<http://betamasaheft.eu/Dillmann/>>) soll in der Zukunft nach und nach mit den jeweils aktuell erarbeiteten Informationen vervollständigt werden.

### Literatur

Bausi, A. 1995. *Il Sēnodos etiopico. Canoni pseudoapostolici: Canoni dopo l'Ascensione, Canoni di Simone il Cananeo, Canoni Apostolici, Lettera di Pietro*, Corpus Scriptorum Christianorum Orientalium, 552, 553, Scriptores Aethiopi- ci, 101, 102 (Lovanii: Peeters, 1995).

— 2004. ‘La versione etiopica della Didascalia dei 318 niceni sulla retta fede e la vita monastica’, in E. Lucchesi and U. Zanetti, eds, *Ægyptus Christiana. Mélanges d'hagiographie égyptienne et orientale dédiés à la mémoire du P. Paul Devos bollandiste*, Cahiers d'Orientalisme, 25 (Genève: Patrick Cramer, 2004), 225–248.

— 2010a. ‘Qerəllos’, in *EAE*, IV (2010), 287a–290a.

— 2010b. ‘Senodos’, in *EAE*, IV (2010), 623a–625a.

36 Dillmann 1847, 52b–53a die incipits.

37 Tedros Abraha 2007.

38 Vgl. Dillmann 1847, 54a und Dillmann 1848, 34ab; der äthiopische Text ediert in Dillmann 1866, 131–136.

39 Habtemichael Kidane und Red. 2007.

40 Nr. 27, vgl. Six 2000, 29–31.

41 Pietruschka 2010.

42 Dillmann 1866, 108–131.

- Dillmann, C. F. A. 1847. *Catalogus codicum manuscriptorum orientalium qui in Museo Britannico asservantur*, Pars III: *Codices Aethiopicos amplectens* (London: British Museum, 1847).
- 1848. *Catalogus codicum manuscriptorum Bibliothecae Bodleianae Oxoniensis*, Pars VII: *Codices Aethiopici* (Oxonii: E Typographeo Academico, 1848).
- 1865. *Lexicon linguae Aethiopiae, cum indice Latino. Adiectum est vocabularium Tigre dialecti septentrionalis compilatum a W. Munziger* (Lipsiae: T.O. Weigel, 1865; repr. New York, 1955, and Osnabrück: Biblio Verlag, 1970).
- 1866. *Chrestomathia Aethiopica edita et glossario explanata*, ed. A. Dillmann (Lipsiae: T. O. Weigel, 1866).
- E Ae*. S. Uhlig, ed., *Encyclopaedia Aethiopica*, I: A–C; II: D–Ha; III: Ha–N; in cooperation with A. Bausi, ed., IV: O–X; A. Bausi in cooperation with S. Uhlig, ed., V: Y–Z. *Supplementa. Addenda et Corrigenda. Maps. Index* (Wiesbaden: Harrassowitz Verlag, 2003, 2005, 2007, 2010, 2014).
- Ellwardt, A. 2016. ‘Beyond Dillmann’s *Lexicon* – Towards Digital Lexicography: Lessons from Syriac’, in A. Bausi and E. Sokolinski, eds, *150 Years after Dillmann’s Lexicon: Perspectives and Challenges of Gə‘əz Studies*, Supplement to *Aethiopica*, 5 (Wiesbaden: Harrassowitz Verlag, 2016), 185–199.
- Ewald, H. von 1844. ‘Ueber die Aethiopischen Handschriften in Tübingen’, *Zeitschrift für die Kunde des Morgenlandes*, 5 (1844), 164–201.
- Graf, G. 1937. ‘Zwei dogmatische Florilegien der Kopten. B. Das Bekenntnis der Väter’, *Orientalia Christiana Periodica*, 3 (1937), 345–402.
- Graf, G. 1951. *Geschichte der christlichen arabischen Literatur*, IV, Studi e testi, 147 (Città del Vaticano: Biblioteca Apostolica Vaticana, 1951).
- Habtemichael Kidane 2005. ‘Dəgg<sup>w</sup>a’, in *E Ae*, II (2005), 123a–124b.
- und Red. 2007. ‘Mäwašə’ət’, in *E Ae*, III (2007), 877a–878a.
- Kleiner, M. 2003. ‘Abtəlis’, in *E Ae*, I (2003), 50b–51a.
- Kleyn, H. G. 1882. *Jacobus Baradaeüs. De stichter der Syrische Monophysietische kerk*, PhD thesis, Rijksuniversiteit te Leiden (Leiden: Brill, 1882).
- Löfgren, O. 1974. *Katalog über die äthiopischen Handschriften in der Universitätsbibliothek Uppsala, Sowie Anhänge über äthiopische Handschriften in anderen Bibliotheken und in Privatbesitz in Schweden*, Acta Bibliothecae R. Universitatis Upsaliensis, 18 (Uppsala: Almqvist & Wiksell, 1974).
- Ludolf, H. 1691. *Ad suam Historiam Aethiopicam antehac editam Commentarius: in quo multa breviter dicta fusius narrantur, contraria refelluntur, Atque hac occasione praeter res Aethiopicas multa Autorum, quaedam etiam S. Scripturae loca declarantur...* (Francofurti ad Moenum: Sumptibus Johannis David Zunneri, 1691).

- Lusini, G. 1989. 'Appunti sulla patristica greca di tradizione etiopica', *Studi classici e orientali*, 38 (1989), 469–494.
- Meley Mulugetta 2010. 'Säwasəw', in *E Ae*, IV (2010), 562b–564a.
- Nosnitsin, D. 2010. 'Sälam', in *E Ae*, IV (2010), 484a–484b.
- Petrus Ethyops 1548. *Testamentum Novum: Cum Epistola Pauli Ad Hebreos tantum, cum concordantijs Evangelistarum Eusebii & numeratione omnium verborum eorundem. Missale cum benedictione incensi cerae et c. Alphabetum in lingua ግእዝ ፡ gheez, idest libera quia a nulla alia originem duxit, & vulgo dicitur Chaldea, Quae omnia Fr. Petrus Ethyops auxilio piorum sedente Paulo. III. Pont. Max. & Claudio illius regni Imperatore imprimi curavit* (Romae: Impensis Petri Comos Ethiopis, Valerius Doricus, 1548).
- Pietruschka, U. 2010. 'Ṭābibā ṭābiban', in *E Ae*, IV (2010), 799b–800a.
- Six, V. 2000. *Äthiopische Handschriften der UB Tübingen* (Hamburg, 2000; pro manuscripto; als PDF unter <<http://www.ub.uni-tuebingen.de/ueber-uns/bibliothekbestand/altbestand/handschriften/handschriftenkataloge.html>>).
- Tedros Abraha 2007. 'Maḥbärä mə'əmānan', in *E Ae*, III (2007), 652a–653a.
- Tubach, J. 2007. 'Jacob Baradaeus', in *E Ae*, III (2007), 261a–262a.
- Valieva, N. 2017. 'Ms Ethiopic 4 of the Collection of the India Office. A Strayed Manuscript of *Gadla Lālibalā*', *Aethiopica*, 20 (2017), 190–201.
- Weischer, B. M. 1973. *Qērellos I: Der Prosphonetikos über den rechten Glauben des Kyrillos von Alexandrien an Theodosios II.*, *Äthiopianistische Forschungen*, 7 (Glückstadt und Hamburg: Verlag J.J. Augustin, 1973).
- 1977. *Qērellos III: Der Dialog 'dass Christus einer ist' des Kyrillos von Alexandrien*, *Äthiopianistische Forschungen*, 2 (Wiesbaden: Franz Steiner Verlag GMBH, 1977).
- 1980. *Qērellos IV 3: Traktate des Severianos von Gabala, Gregorios Thaumaturgos und Kyrillos von Alexandrien*, *Äthiopianistische Forschungen*, 7 (Wiesbaden: Franz Steiner Verlag GMBH, 1980).
- 1993. *Qērellos II: Der Prosphonetikos über den rechten Glauben des Kyrillos von Alexandrien an Arkadia und Marina, I: Einleitung mit christologischem Florilegium*, *Äthiopianistische Forschungen*, 31 (Stuttgart: Franz Steiner Verlag GMBH, 1993).
- Wion, A. and E. Fritsch 2005. 'Haymanotä abäw', in *E Ae*, II (2005), 1073a–1075b.
- Witakowski, W. 2003. 'Athanasius', in *E Ae*, I (2003), 392a–394b.
- 2007. 'John Chrysostom', in *E Ae*, III (2007), 293a–295b.
- 2010. 'Ṭomar: Mäṣḥäfä ṭomar', in *E Ae*, IV (2010), 970a–971a.

# Describing the Complex: the Multiple Dimensions of a Relational Database\*

*Sara Fani, Copenhagen University*

The paper shows, on the example of manuscripts from Ethiopia containing the *Kitāb al-farā'id* which were surveyed by the *IslHornAfr* project, how a relational database can manage data on complex (composite and multiple-text) manuscripts.

One of the main aspects observed in manuscripts retrieved within the framework of the project *Islam in the Horn of Africa*<sup>1</sup>—and possibly the most challenging to deal with during the conception of the descriptive database of the project—is the codicological and textual complexity of this manuscript heritage. Most of the codices were non-homogeneous, and their complex textual and codicological stratigraphy had to be duly reflected in the relational database.

Needless to say, such complexity is a feature not limited to the manuscripts in Arabic script produced or circulated in the Horn of Africa, nor to the manuscripts in Arabic script in general. In fact, fruitful reflections on multiple-text manuscripts and composite manuscripts have developed in the last years, and important contributions on the subject have been published extending concepts and terminology previously used for western manuscripts also to different cultural contexts.<sup>2</sup>

As in other spheres, manuscripts in Arabic script from the Horn of Africa can be complex from the point of view of their material history (composite manuscripts) and from that of their content (multiple-text manuscripts). In his 2016 contribution,<sup>3</sup> Alessandro Gori showed that in the Islamic context of the Horn of Africa, composite manuscripts (with one codex made up of several codicological units which were formerly independent) are mostly represented by manuscripts assembled by Europeans travellers and scholars. Even if compiled locally, they were often bound or otherwise put together once they had to be stored in European collections. As for multiple-text manuscripts (single codicological/production units with two or more texts), we can distinguish between (a) closed, or canonized, collection of texts, such as, for example, 'liturgical' collections of devotional poems and litanies, always copied together, and (b) instable, or open, collections that group various texts usually

\* The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme, grant agreement no. 322849 (ERC Advanced Grant IslHornAfr).

1 See also the project note by Alessandro Gori in this issue.

2 See e.g. the contribution by Nathan Carlig in this issue.

3 Gori 2016.

on the same or related topic (sometimes also on different topics), most typically collections of texts related to the various fields of the traditional Islamic knowledge for teaching and learning.

In this paper, I would like to illustrate how these complexities of the material and textual structure of this manuscript heritage have been dealt with in the *IslHornAfr* database. A preliminary remark has to be done regarding the aim of the project, which has mainly a literary perspective. This is reflected in the description of the book heritage and in the structure of the database: in fact, the priority has been given to text entities which have been selected as the central reference element to be described.

### ***The identification of texts in the Kitāb al-farā`id***

As a case study, I have chosen one of the most representative works of the Islamic literary tradition of the Horn of Africa, the *Kitāb al-farā`id*, an extensive work in Old Harari, written in Arabic characters. Its title can be translated as ‘The Book of the Obligations’, or rather ‘The Book of Obligatory Portions’, if we want to maintain a closer correspondence to the meaning of these words in the Arabic tradition, which specifically refers to the law of inheritance. The numerical portions are indeed a constant motive in the explanations of moral and religious duties introduced in the work. For example:

God said ‘With three things you should persuade me, for three things you should fear me, for three things you should worship me.’

God said ‘With three things you should persuade me’: God said: ‘You should persuade me with your prayer, your fasting and your alms’.<sup>4</sup>

The case of the *Kitāb al-farā`id* was particularly challenging when trying to describe it in a database that has to account for textual and material evidence, as the work as we know it is actually a combination of different (three, in fact) textual entities not easily identifiable in manuscripts, as we shall see below.

The first to describe the *Kitāb al-farā`id*, using just one witness available at the time (now Vatican City, Biblioteca Apostolica Vaticana, Cer. et. 325<sup>5</sup>) was Enrico Cerulli.<sup>6</sup> He offered an Arabic vocalized transcription, a transliteration, and an Italian translation.<sup>7</sup> According to this edition, the text is divided into two sections which are reflected in a shift of content. The first section falls into the genre of wisdom literature and contains sayings and maxims with moral and religious subject or pertaining to ritual obligations. The second section falls into the genre of religious teaching, containing a cate-

4 Banti 2009–2010.

5 Raineri 2004, 232–233.

6 Cerulli 1936, 282–343.

7 Cerulli also used the text as a basis for his grammar and glossary of Old Harari, *ibid.* 344–437.

chism-like text, structured in a question and answer alternation, in which the articles of faith are first given in Arabic and then explained in Harari. In his introduction to the edition, Cerulli noticed that the colophon ‘oddly’ appears mid-text on f. 7 (the text covering ff. 1v–17v) and mentions a certain Ṭayyib al-Wanāğī al-Šadrī, to whom he ascribed the entire work.<sup>8</sup>

After more manuscripts containing the *Kitāb al-farā'id* came to light, it became obvious that the label is conventionally applied to several distinct works, written by different authors, which are usually copied together.

Thus, Ewald Wagner<sup>9</sup> could show that the two sections of Cerulli’s edition correspond to two different works, which explains the difference in content. The first work is a *Kitāb al-farā'id* ascribed to āw ‘Abd al-Raḥmān al-‘Arāšī (hereafter KF<sub>1</sub>), and the second is a *Kitāb al-farā'id* ascribed to *faqīh* Ṭayyib al-Wanāğī al-Šadrī (hereafter KF<sub>2</sub>). Wagner also noted that ‘the IES owns a third *Kitāb al-farā'id* written by a certain ‘Abd Allāh ‘Umar b. Ġibrīl al-Aswām al-Sāḥitī’ (hereafter KF<sub>3</sub>).<sup>10</sup> A closer analysis, drawing also on the manuscripts which clearly show all the text boundaries and ascriptions, has revealed, however, that Cerulli’s edition actually contained all the three texts (with the ‘oddly’ placed colophon appearing at the end of KF<sub>2</sub>).

Wagner explained Cerulli’s failure to identify the two sections as KF<sub>1</sub> and KF<sub>2</sub> with ‘a lacuna [...] at the passage of the first to the second work’. But in fact there is no textual lacuna, and the passage from KF<sub>1</sub> to KF<sub>2</sub> is evident in the edition<sup>11</sup> thanks to the presence of a typical *explicit* formula (*Tammāt Kitāb al-farā'id bi-kalām al-Ḥabaš aḥrağahu faqīh Ṭayyib al-Wanāğī al-Šadrī, naḥa ‘anā Allāh bihum amīn*,<sup>12</sup> with a verb, *aḥrağā*, normally used in Arabic for the compilations of collections of *ḥadīths* on special themes). The *explicit* is then followed by a *basmala*. Yet, while the passage from KF<sub>1</sub> to KF<sub>2</sub> is quite evident in the edition, the shift from KF<sub>2</sub> to KF<sub>3</sub> is in fact oblique: there is no *explicit* or colophon, nor any other kind of textual boundary.

### ***Textual boundaries in the manuscript witnesses***

As Giorgio Banti noted, ‘in most mss. one or even the two junctures between the three different texts have been skipped over, and only one or two of the three ascriptions remains’.<sup>13</sup>

8 Ibid. 282–283.

9 Wagner 1989.

10 Wagner 2005, 492. In addition to the description of MS Staatsbibliothek zu Berlin, Hs. Or. 10455 in Wagner 1997, 6–8, 118–119, see also MSS Addis Ababa, Institute of Ethiopian Studies (IES) 257 and 268 described in Gori et al. 2014, 1, 5. Cf. ALA<sup>3a</sup>, 23, 33–34.

11 Cerulli 1936, 290–291.

12 ‘It is concluded the Book of Obligations in the language of the Abyssinians which Ṭayyib al-Wanāğī al-Šadrī extracted...’, *ibid.* 290.

13 Banti 2009–2010, 168.



The quite old (early nineteenth century?) manuscript Addis Ababa, IES 256,<sup>14</sup> is an excellent witness as it clearly shows all the boundaries and is one of the more concluding for the textual identification. The first section (KF<sub>1</sub>, ff. 1v–4r) has an *explicit* with attribution to al-‘Arāšī; the second section (KF<sub>2</sub>, ff. 4r–7r) begins with an attribution to al-Wanāgī al-Šadrī, and has only *tammāt* as the *explicit*; in the third section (KF<sub>3</sub>, ff. 7r–10v), an *incipit* with attribution to al-Sāhītī was added and partially written on the margin; the *explicit* confirms the attribution to al-Sāhītī with the common formula (see fig. 1a, 1b).

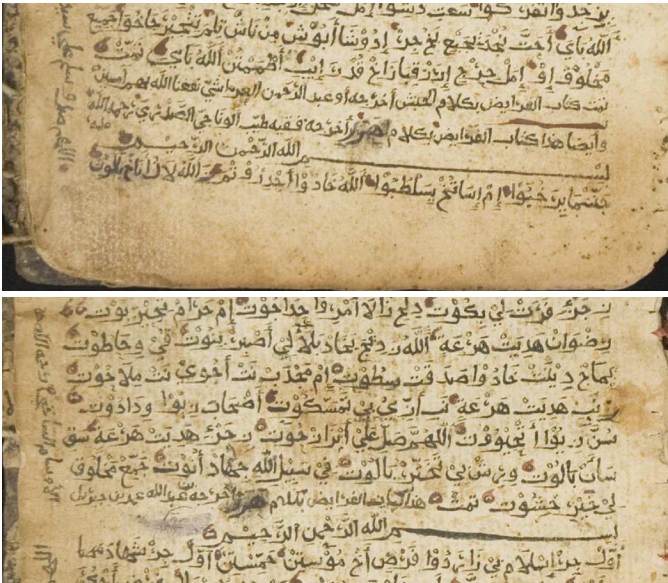


Fig. 1. Ms IES 256, text boundaries KF<sub>1</sub>–KF<sub>2</sub> (a: f. 4r, detail) and KF<sub>2</sub>–KF<sub>3</sub> (b: f. 7r, detail).

In other cases the textual distinctions and attribution are more complex. In fact, often we do not have anything testifying the textual boundaries. Where boundaries and attributions are evident, the *explicit*/colophon between KF<sub>1</sub>–KF<sub>2</sub> and KF<sub>2</sub>–KF<sub>3</sub> has sometimes been wrongly referred to the following text as opening formula or *vice versa* (as in the case of Cerulli’s edition). But quite regularly the *explicit* of the last text correctly ascribes the work to al-Sāhītī.

In a manuscript from ‘Abd Allāh Šarīf collection in Harar (MS Harar, ‘Abd Allāh Šarīf, 191), the text boundary between KF<sub>1</sub> (which is acephalous) and KF<sub>2</sub> is marked by the word *tamma* and a circle with a dot inside (fig. 2), which was traditionally used as a collation mark and *iğāza* (that is approval

14 Gori et al. 2014, 1.



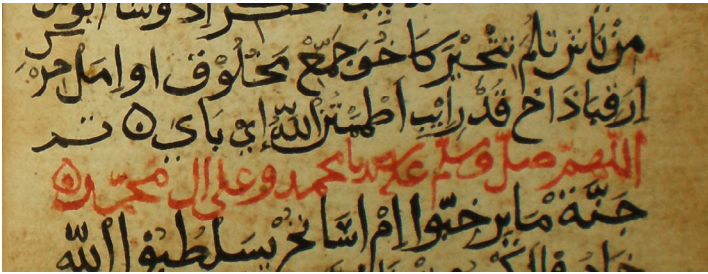


Fig. 2. Ms ASH 191, text boundary KF<sub>1</sub>-KF<sub>2</sub> (f. 7v, detail).

for transmission) but also as a separation of individual textual units (especially between *ḥadīṭ*). There is instead no text boundary between KF<sub>2</sub> and KF<sub>3</sub>, while the latter is correctly attributed in the colophon to al-Sāḥiṭī.

In manuscript Addis Ababa, IES 306,<sup>15</sup> KF<sub>1</sub> is copied on ff. 1r-6r and is acephalous; there is no text boundary within the scribal frame, but on the margin there is a note stating that ‘in a copy, after the eulogy on the Prophet, there is a *basmala*’ which is in fact a very common mark of textual boundary and a testimony to the manuscript having been collated with another witness of the text (fig. 3). In IES 306, the boundary between KF<sub>2</sub> and KF<sub>3</sub> is not marked at all, but, meaningfully, after the *explicit* of the last textual section the work is attributed to all three mentioned authors.<sup>16</sup>

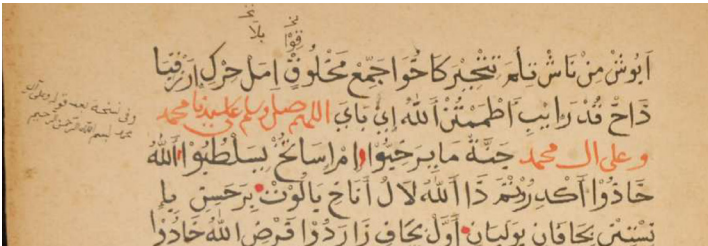


Fig. 3. Ms IES 306, text boundary KF<sub>1</sub>-KF<sub>2</sub> (f. 6r, detail; marginal note: *wa-ḥī nushāṭin ba‘da qawlihi wa-‘alā ahl Muḥammad, bi-‘sm Allāh al-Raḥmān al-Raḥīm*).

While the codicological evidence of textual boundaries is confusing (with the lack of textual boundaries being more common between KF<sub>2</sub> and KF<sub>3</sub>), the threefold textual identification can be confirmed by the clear thematic division between KF<sub>2</sub> and KF<sub>3</sub>. At a first glance, both fit into the general class of catechism-like texts, but the KF<sub>2</sub>, ascribed to al-Wanāḡī al-Ṣadīrī, deals primarily with ‘*aqīda*’ (creed) and general Islamic dogmas and beliefs (for example the

15 Ibid. 16.

16 One could infer that the attribution to the three authors is based on the witness used for the collation, rather than coming from the antigraph.

eschatological ones), and the KF<sub>3</sub>, ascribed to al-Sāḥitī, focuses on *fiqh* (law) issues, following the order and subjects of the Šāfi‘ite manuals.

### Data organization

In the *IslHornAfr* database, the *Kitāb al-farā‘id* is encoded as a set of entities. There is a general record for the entire work as it is traditionally transmitted and known, with its three sections (fig. 4).

The screenshot displays a detailed record for the text 'Kitāb al-farā‘id'. The title is shown in Arabic as 'كتاب الفرائض'. The record is organized into several sections: 'Sections' lists three sub-works: 'Kitāb al-farā‘id li-l-‘Arāšī', 'Kitāb al-farā‘id li-l-Wanāḡī al-Sadrī', and 'Kitāb al-farā‘id li-l-Awsām al-Sāḥitī'. 'Languages' is set to 'Old Harari' and 'Arabic'. 'Genres' include 'Belief', 'Ethics', and 'Jurisprudence'. 'Formal labels' are marked as 'Prose'. 'Bibliographical references' list four scholarly works discussing the text. 'References' show three entries indicating that the main text includes these three sub-works. 'Manuscript parts' are listed as 'ASH00118, 83v - 87r' and 'DWL00012, 1r - 11v'.

Fig. 4. Database record for the *Kitāb al-farā‘id*, detail (August 2018).

At the same time, each section is encoded as a textual entity in its own right, as they are attributed (sometimes also explicitly in the manuscripts) to

different authors.<sup>17</sup> Such encoding is also necessary as the single sections may have circulated separately. This is illustrated by MS Addis Ababa, IES 265, where KF<sub>1</sub> is copied on its own in an independent monomeric (homogeneous, containing a single codicological unit and text unit) manuscript (fig. 5).

Manuscripts & Prints / AAIE00265

Identifier

- Collection [IES - Institute of Ethiopian Studies](#) →
- ID AAIE00265
- Signature IES00265

Number of folios

10

Cataloguer

SF

Comments

ff. 7v-10v blank

**Bibliographical references**

- [A handlist of the manuscripts in the Institute ...](#) → on page xi, xviii, xlix, l, lix, xcii, 4
- [Arabic Literature of Africa. Volume 3 Fascicle ...](#) → on page 23, 33

**Parts**

- 1v - 7r

كتاب الفرع العرشي  
Kitāb al-farā'id li-al-'Arāšī

**Paratexts**

- 1r, [Previous shelfmark](#)

Harar 25

Fig. 5. Database record for MS Addis Ababa, IES 265, detail (August 2018).

This method of structuring data gives project researchers the opportunity to account both for the codicological and textual complexity of this manuscript heritage. The codicological homogeneity can be a connecting element among the different textual entities, as well as the textual homogeneity and continuity can link different codicological units.

There are many other cases of textual constellations similar to the *Kitāb al-farā'id*. One of them, also related to the city of Harar, is the work known as the *Mawlid šaraḥ al-‘ālamīn*, a constellation of texts including the *‘Unwān al-šarīf*, the *Tahmīs al-Fayyūmī ‘alā al-Burda*, and a connective group of *du‘ā* and doxologies.<sup>18</sup> Other cases include various collections of *du‘ā*, for

17 Cp. also the contributions by Massimo Villa and Tito Orlandi in this issue.

18 Gori 2010.

example the rather unstable one compiled by Kabīr Ḥamza b. Kabīr Maḥmūd b. Kabīr ‘Alī al-Ḥarallī al-Awsī al-Ḥanafī (1211–1279H/1796–1862AD) and commonly known under the title *Fawātiq al-falāḥ wa-bawāriq al-ṣalāḥ fī ḍikr mawlid al-nāṭiq bi-l-naḡāḥ*—which will hopefully be object of further enquiries within the framework of the project *Islam in the Horn of Africa*.

### References

- ALA<sup>3a</sup>. O’Fahey, R.S. 2003. *The Writing of the Muslim People of Northeastern Africa*, Arabic Literature of Africa, III A = Handbuch der Orientalistik, 1, 13 (Leiden, Boston: Brill 2003)
- Banti, G. 2009–2010. ‘The Literature of Harar until the End of the 19<sup>th</sup> Century’, in A. Gori and B. Scarcia Amoretti, eds, *L’Islam in Etiopia. Bilanci e prospettive*, *Civiltà del Mediterraneo*, 16–17 (Roma: Edizioni di Storia e Letteratura, 2009–2010), 171–172.
- Cerulli, E. 1936. *Studi Etiopici*, I: *La lingua e la storia di Harar*, Pubblicazioni dell’Istituto per l’Oriente (Roma: Istituto per l’Oriente, 1936), 282–343.
- Gori, A. 2010. ‘Texts in the Mawlid Collection in Harar: Some First Critical Observations’, in B.M. Tarsitani and M. Shigeta, eds, *Preserving Local Knowledge in the Horn of Africa: Challenges and Prospects for Collaborative Research in Oral Literature, Music and Ritual Practices*, African Study Monograph, Suppl. 41 (Kyoto: Kyoto University, 2010), 51–62.
- 2016. ‘Some Observations on Composite and Multiple-Text Manuscripts in the Islamic Tradition of the Horn of Africa’, in M. Friedrich and C. Schwarke, eds, *One-Volume Libraries: Composite and Multiple-Text Manuscripts*, *Studies in Manuscript Cultures*, 9 (Berlin and Boston: De Gruyter, 2016), 155–169.
- et al. 2014. *A Handlist of the Manuscripts in the Institute of Ethiopian Studies*, II: *The Arabic Materials of the Ethiopian Islamic Tradition*, Ethiopian Manuscripts, Texts and Studies Series, 20 (Eugene, OR: Pickwick Publications, 2014).
- Raineri, O. 2004. *Enrico Cerulli. Inventario dei manoscritti Cerulli etiopici*, *Studi e Testi*, 420, *Cataloghi sommari e inventari dei fondi manoscritti*, 8 (Città del Vaticano: Biblioteca Apostolica Vaticana, 2004).
- Wagner, E. 1989. ‘Bemerkungen zu zwei Handschriften des K. al-Farā’id’, in M. Macuch, C. Müller-Kessler, B.G. Fragner, eds, *Ḥoḳmot bāntā bētāh. Studia Semitica necnon Iranica. Rudolpho Macuch Septuagenario ab Amicis et Discipulis Dedicata* (Wiesbaden: Otto Harrassowitz, 1989), 389–395.
- 1997. *Islamische Handschriften aus Äthiopien*, *Afrikanische Handschriften*, 2 = *Verzeichnis der Orientalischen Handschriften in Deutschland*, 24, 2 (Stuttgart: Franz Steiner, 1997).
- 2005. ‘Farā’id, Kitāb al-’, in S. Uhlig, ed., *Encyclopaedia Aethiopica*, II: *D–Ha* (Wiesbaden: Harrassowitz Verlag, 2005), 491–492.

# Multi-level Digital Annotation of Ethiopic Texts\*

*Susanne Hummel, Vitagrazia Pisani, and Cristina Vertan,  
Universität Hamburg*

The GeTa tool has been developed at Hamburg to address the challenge of tokenization and multi-level annotation of Ethiopic texts, with the aim of further computer-assisted analysis of the morphology and lexicography of the Gə'əz language. The paper illustrates the workflow of linguistic annotation with the help of the tool.

## **1. Introduction**

Although of major importance for the study of Christian Orient, the Gə'əz language (also known as Classical Ethiopic) has been so far neglected by the new research trends in Digital Humanities. While some Gə'əz texts exist in digital form, there are no tools to assist their linguistic analysis. The project *TraCES: From Translation to Creation: Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages* aims at addressing this desideratum by the development of a complex annotation tool which allows the production of coherent, reliable, and extensive linguistic data. The tool (called GeTa for Gə'əz Text Annotation) is used to annotate a pre-selected corpus of texts: several texts belonging to different periods and genres of Ethiopic literature, original and translated have been singled out.<sup>1</sup> Each text is (in full or in part) annotated at different levels. The main level is formed by the detailed linguistic (part-of-speech) annotation ('deep annotation' in the project's terminology), where each word is linked to the corresponding dictionary entry. We also annotate named entities such as persons, places, dates, titles of work, or offices. Furthermore, we mark up the text structure (e.g. parts, chapters, sentences, verses). Special features related to the edition, like editorial intervention such as conjectures, are marked upon occurrence.

The GeTa tool and the data will be made freely available to enable a systematic, diachronic analysis of the Gə'əz language, including its lexicography, morphology, and style.

In this paper we focus on the workflow of linguistic annotation, and discuss the requirements and challenges posed by the annotation process for the tool development. We also briefly present the tool's components and the underlying data structure.

\* The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme, grant agreement no. 338756 (ERC Advanced Grant TraCES).

1 See the project note by Eugenia Sokolinski in this issue.

## 2. Challenges of the Gə'əz language for digital tools

As of today, from the computational linguistics point of view, the Gə'əz language belongs to the group of 'very low-resource languages',<sup>2</sup> i.e. languages with a significant lack of resources (corpora, lexicons, terminological data bases, semantic networks) and tools. Often, low-resource languages can be helped by adapting tools and materials existing for other languages within the same family. In the case of Gə'əz, this was not possible. Better-resource Semitic languages (such as Hebrew or Arabic) use a different writing system (right-to-left consonantal writing against the left-to-right syllabic writing for Ethiopic). Amharic uses the same writing system, but its morphological structure differs in many aspects from that of Gə'əz.<sup>3</sup> Besides, all these languages are still low resource, and the available tools and data are very limited.

A number of tools claim to be language independent. They incorporate data from very large language corpora, so that linguistic features can be elicited, and learnt, from the data. This statistical paradigm cannot be followed for the moment for Gə'əz as there exists no significant corpus for Classical Ethiopic. Additionally, machine learning methods perform best when the number of features to be learnt is limited. This is not the case of Gə'əz, for which we have identified 33 part-of-speech tags that can be accompanied by various features, the number of possible combinations going in several hundred (see § 3 below).

An additional challenge is the absence of an electronic dictionary (lexicon) for Gə'əz. Usually a dictionary is the first digital resource to be developed for a language. Lexicons give important information about the lemma, the root, and morphological features. The *TraCES* project has to build up lexicon and annotated corpus in parallel. This means that before a word in the corpus can be linked to the lexicon, unless it is already present in the initial word-list, the corresponding lemma (with the morphological information, translation, examples) has to be created.

A fully automatic annotation is therefore impossible for Gə'əz at this stage. We adopt a two-stage workflow: (1) at a first stage, texts are manually supplied with detailed linguistic annotation ('deep annotation'). The process is facilitated by a controlled semi-automatic component (batch annotation, see § 3 below); (2) at a second stage, the annotated corpus will be used as training material for a machine learning algorithm. The complete architecture, including the links to the lexicon component, is illustrated in fig. 1.

- 2 See Maegaard et. al 2006 for the definition of a minimum set of resources and tools which are necessary to insert a language on the digital map.
- 3 For further details and a morphosyntactic tagset for the Amharic see Krzyżanowska 2017.

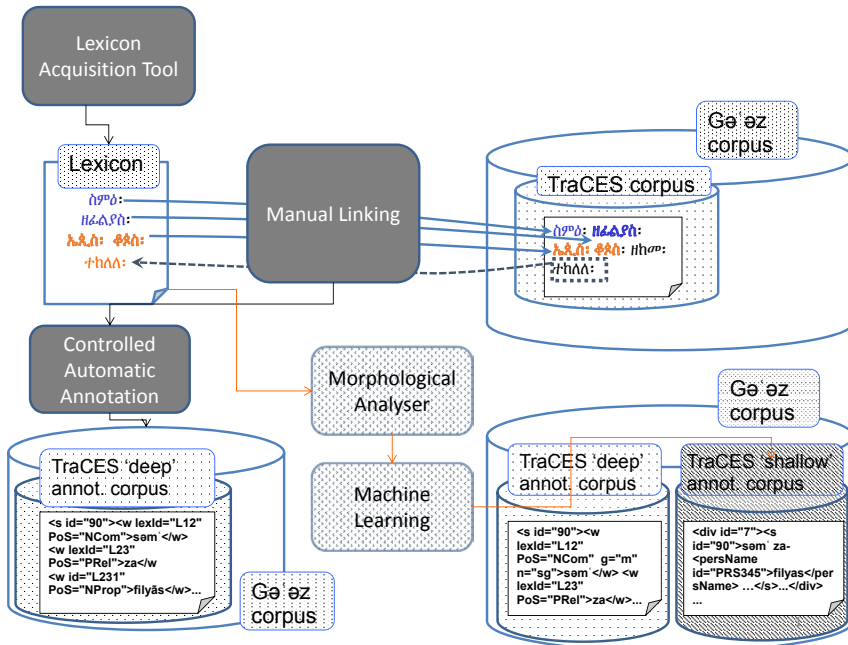


Fig. 1. TraCES modules for linguistic annotation.

### 3. Gə'əz specific requirements for deep annotation and workflow

Over the past years, several language-independent and/or language customizable annotation tools have become available, including, just to name two examples, WebAnno<sup>4</sup> and CorA.<sup>5</sup> It would have been nice to be able to use them, at least as a starting point. However, certain features of Gə'əz—in combination with the high scientific demands of the project—made the use of these or other existing tools impossible. In the following, we explain the reasons and the choices behind the decisions taken for the annotation workflow and the design of the GeTa tool.

As mentioned above, the project part-of-speech tag set is particularly fine grained and consists of 33 different tags (PoS); for many of them additional morphological features must be provided.<sup>6</sup> The PoS are divided into six main categories, of which some have further subdivisions: (1) nominals: nouns (2 subdivisions), pronouns (10 subdivisions), numerals (2 subdivisions); (2) verbs; (3) existentials (affirmative and negative); (4) particles: adverbs (2

4 De Castilho et. al. 2016; <<https://webanno.github.io/webanno/>>.

5 Bollman et. al. 2014; <<https://www.linguistics.rub.de/comphist/resources/cora/>>.

6 For an overview of the tag set and an introduction to the applied annotation principles (in particular to the complex noun annotation), see Hummel and Dickhut 2016.



subdivisions), prepositions, conjunctions, interjections, further particles (9 subdivisions); (5) foreign material; (6) punctuation.

The linguistic annotation is conducted mainly on morphological criteria, but not solely, as morphologically identical forms need to be disambiguated in the context of syntax and semantics. As the examples below show, disambiguation is required at all stages of the annotation: during the process of transliteration, of tokenization, and of assigning the correct PoS tag.

Due to the lack of training material on the one side and the large number of linguistic features on the other, unsupervised machine learning approaches performing automatic tagging were not suitable for our corpus. We opted in a first stage for a semi-automatic workflow as shown in fig. 2. The annotated corpus from this stage will serve as training material for machine learning.

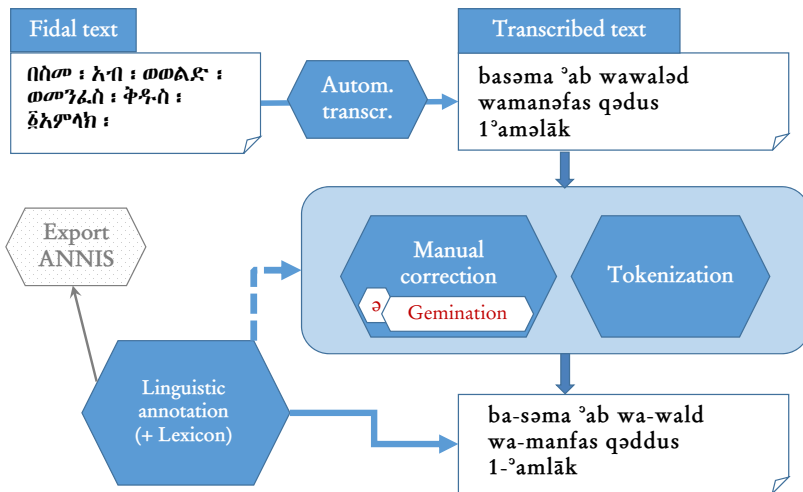


Fig. 2. Annotation workflow.

Because of the syllabic script and the detailed linguistic features to be annotated, any text processing requires a transcription of the Ethiopic text.<sup>7</sup> Therefore, the annotation tool must handle in parallel the text in its original script (*fidal*) and its transcription (respectively, corrected transliteration). Both windows are kept synchronized during all tasks. The transcription is conducted automatically and must be then manually corrected. Corrections concern primarily the presence or absence of the sixth-order vowel (*ə*) be-

7 The graphic unit (GU) ቤታ፡ *betā* consists of two syllables *be* and *tā*, but it has to be tokenized as *bet-ā* ('her house'), which would be impossible on the Ethiopic script.



tween consonants and the gemination of consonants. Both phonetic features are not reflected in the Gə‘əz script, and therefore cannot be implemented automatically. The corrections are performed in a semi-automatic way: having corrected one graphic unit (GU),<sup>8</sup> the annotator can decide whether or not the same correction applies to all other identical occurrences in the text. Sometimes the decision on the correct transliteration can be taken only after a morphological analysis; thus the tool must be able to handle later corrections without losing annotations.

A typical example of manual correction required in connection with a geminated consonant is the form **ደጉቃተሉ ፡** of the verb **ቀተለ ፡**, *qatala*, ‘to make war’. This form is ambiguous, as it can stand for two different verbal moods. It corresponds either to an imperfect indicative (third person, masculine, plural)—in which case the correct transliterated form would be *yəṭqātatalu*, with the reduplication of the second radical (*t*), which is the phonetic feature of the imperfect in Gə‘əz—or to a jussive (third person, masculine, plural), which requires the transliteration without germination: *yəṭqātalu*. This example shows clearly that the correct transliteration can be achieved only after a close analysis taking syntax and semantics into account.

Another classic example illustrates the disambiguation of the epenthetic vowel (ə), again demonstrated with a verb, here with the meaning ‘to save’. The form **ያድኅን ፡** can stand either for an imperfect (third person, masculine, singular) or for a jussive (third person, masculine, singular). In the transliteration, however, the presence or absence of the sixth-order vowel (ə) after the first radical (*d*) differentiates the two verbal moods: with the vowel, *yādəḥən*, the imperfect, without the vowel, *yādhən*, the jussive.

Linguistic annotation also involves an identification of independent tokens (‘tokenization’). We split each complex GU into its smallest analysable units (‘tokens’), to which one can assign a PoS. During this process, the annotator, too, needs to resolve ambiguous forms. Identical GUs may carry different meanings and consequently may be split into a varying number of tokens and assigned different PoS.

For example, **ገብሩ ፡** *gabru* may be translated as ‘they did’; in this case it would be considered a single token, with the PoS ‘verb’ (perfect, third person, masculine, plural). In a different context, the same GU **ገብሩ ፡** *gabru* may carry the meaning ‘his servant’. In this case, it consists of two tokens, each to be assigned a different PoS: *gabr* ‘common noun’ (nominative, pronominal state, masculine by pattern, singular by pattern), and *-u* pronominal suffix (third person, masculine, singular).

8 We define a graphic unit (hereafter GU) as a sequence of characters separated by a word divider ( ፡ ), or by a punctuation sign ( # ); the latter is a GU in its own right.

Based on the tokenized transliteration, we eventually conduct the proper linguistic annotation: assigning the PoS together with its features and values. The annotator is able to check and, if required, to correct or adjust the transliteration and tokenization work done so far. Disambiguation is, however, necessary also during this process. It concerns in particular tokens that can have a different meaning depending on the context.

The token **hᵂᵂ** : *kama* or the token **ᵂᵂ** : *haba* may be annotated as ‘conjunction’ if they precede a verb, or as ‘preposition’ if they precede a noun or a pronoun. The prefix **ᵂ** *la-* occurs in the most cases in the function of a preposition, but it can also function in a final clause as a conjunction if it is attached to a jussive verb form.

Finally, each token is linked to a lemma of the newly established digital lexicon.

#### 4. Underlying data model

The data model of the GeTa tool follows an object-oriented approach. Each object can be located by a unique ID. There are two types of objects:

1. Annotated Objects: GUs, tokens, Gə‘əz characters and transcription letters.
2. Annotation Objects (spans) which are attached to one or more Annotated Objects: morphological annotations, text divisions, editorial annotations. Links between Annotated- and Annotation-Objects are ensured through the IDs. In this way the model enables also the annotation of discontinuous elements (e.g. a Named Entity which does not contain adjacent tokens).

For example the GU-object **ᵂᵂᵂᵂ** : contains the 4 Gə‘əz-character objects **ᵂ**, **ᵂ**, **ᵂ**, **ᵂ** (for synchronization reasons, we consider the word separator **ᵂ** as property attached to the Gə‘əz-character object **ᵂ**). Each of these objects contains the corresponding Transcription-letter objects:

**ᵂ** contains the Transcription-letter objects: *w* and *a*

**ᵂ** contains the Transcription-letter objects: *y* and *ə*

**ᵂ** contains the Transcription-letter objects: *b* and *e*

**ᵂ** contains the Transcription-letter objects: *l* and *o*

During the transliteration and tokenization phase, three Token objects are built: *wa*, *yəbel*, and *o*. Each Token object records the IDs of Transcription-letter objects it contains. Finally, the labels ‘**ᵂᵂᵂᵂ**’ and ‘*wa-yəbel-o*’ are attached to the initial GU object.

Morphological annotation objects are attached to one Token object. They consist of a tag (PoS, e.g. Common Noun) and a list of key-value pairs where the key is the name of the morphological feature (e.g. number). In this way, the tool is robust when adding new morphological features or PoS tags.

As the correspondences between the Gə'əz character and the transcription are unique, the system only stores the labels of the Transcription-letter objects. All other object labels (Token, Gə'əz character, and GU) are dynamically generated throughout a given correspondence table and the IDs, so that the system uses less memory and remains error proof during the transliteration process. In fig. 3 we present the entire data model, hinting also at the other possible annotation levels.

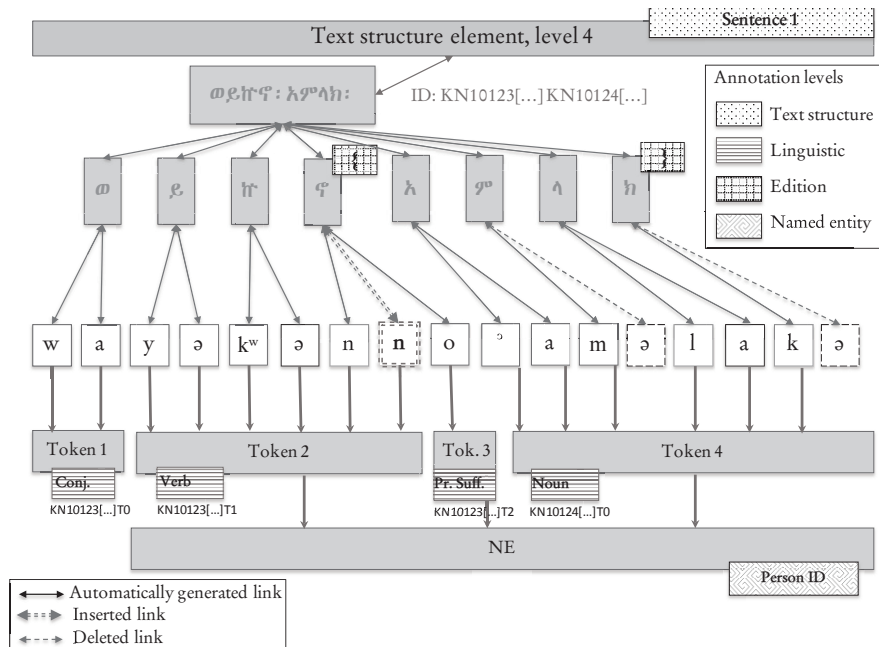


Fig. 3. GeTa data model.

### 5. Interoperability and further work

GeTa is a tailored tool for annotation of Gə'əz texts which enables a deep fine-grained linguistic annotation as well as annotation at other levels. The controlled semi-automatic annotation facilitates the mark-up process but at the same time leaves the full control entirely to the annotator. Units annotated or tokenized automatically are highlighted, so that the user knows anytime if a manual check is necessary. For example automatically generated tokens are displayed in italic, automatically annotated tokens are marked in red.

Corrections to the transcription, as they were described above, can be performed at any moment during the annotation process.

The annotation tool is written in Java 1.8 and is platform independent. The genuine format of the output is JSON. We implemented export functionalities to plain text (TXT) and TEI/XML so that the results can be imported easily to other analytic and visualization applications like Voyant Tools.<sup>9</sup> A special convertor to ANNIS<sup>10</sup> format has been implemented, so that the annotated corpus can be analysed with the powerful mechanism of the ANNIS visualization tool. The corpus will be freely accessible for further research through the ANNIS installation provided by the Hamburger Zentrum für Sprachkorpora.<sup>11</sup> The TEI export will be used for integration with the data available in the project *Beta maṣāḥəft*.<sup>12</sup>

The tool is already able to handle Gəʿəz texts written with the South Arabian alphabet with right-to-left writing direction (early inscriptions). Further work concerns a complete check and adaptation of all functionalities for this alphabet, as well as for unvocalized versions of Gəʿəz texts.

Rules for transliteration, tokenization, and annotation may be extracted from the annotated texts and used for a more advanced automatization of the annotation process.

## References

- Bollmann, M., F. Petran, S. Dipper, and J. Krasselt 2014. ‘CorA: A Web-based Annotation Tool for Historical and Other Nonstandard Language Data’, in K. Zervanou, C. Vertan, A. van den Bosch, and C. Sporleder, eds, *Proceedings of the 8<sup>th</sup> Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH)* (Gothenburg 2014), 86–90, <doi:10.3115/v1/W14-0612>.
- De Castilho, R.E., E. Mújdricza-Maydt, Seid Muhie Yiman, Silvana Hartmann, I. Gurevych, A. Frank, and C. Biemann 2016. ‘A Web-based Tool for the Integrated Annotation of Semantic and Syntactic Structures’, in Erhard W. Hinrichs, Marie Hinrichs, and Thorsten Trippel, eds, *Proceedings of the LT4DH workshop at COLING 2016* (Osaka, 2016), 76–84, <http://aclweb.org/anthology/W16-4011.pdf>.
- Druskat, S. and C. Vertan 2018, ‘Nachnutzbarmachung von Forschungsdaten und Tools am Beispiel altäthiopischer Korpora’, in G. Vogeler, ed., *Kritik der digitalen Vernunft. Abstracts zur Jahrestagung des Verbandes Digital Humanities im deutschsprachigen Raum, 26.02.–02.03.2018 an der Universität zu Köln, veranstaltet vom Cologne Center for eHumanities (CCEH)* (Köln: Universität zu Köln, 2018), 270–273.
- 9 <<https://voyant-tools.org/>>.
- 10 <<http://corpus-tools.org/annis/>>; see Druskat and Vertan 2018.
- 11 <<https://corpora.uni-hamburg.de/hzsk/>>.
- 12 <<https://www.betamasaheft.uni-hamburg.de/>>.

- Hummel, S. and W. Dickhut 2016. 'A Part of Speech Tag Set for Ancient Ethiopic', in A. Bausi and E. Sokolinski, eds, *150 Years after Dillmann's Lexicon: Perspectives and Challenges of Gə'əz Studies*, Supplement to *Aethiopica*, 5 (Wiesbaden: Harrassowitz Verlag, 2016), 17–29.
- Krzyżanowska, M. 2017. 'A Part-of-Speech Tagset for Morphosyntactic Tagging of Amharic', *Aethiopica. International Journal of Ethiopian and Eritrean Studies*, 20 (2017), 210–235.
- Maegaard, B., S. Krauwer, K. Choukri, and L. Damsgaard Jørgensen 2006, 'The BLARK Concept and BLARK for Arabic', in *Proceedings of the 5<sup>th</sup> International Conference on Language Resources and Evaluation (LREC)* (Genoa: LREC, 2006), 773–778 <[http://lrec-conf.org/proceedings/lrec2006/pdf/521\\_pdf.pdf](http://lrec-conf.org/proceedings/lrec2006/pdf/521_pdf.pdf)>.



# The CMCL *Clavis Coptica*. On Producing a Standardized List of (Coptic) Works and Manuscripts

*Tito Orlandi, Accademia Nazionale dei Lincei*

Traditional ways of referring to literary works and manuscripts of the Coptic tradition often lack reliability. A new formalized approach is necessary to create objective and stable repertoires. The paper presents some preliminary conclusions reached during the work on the *Clavis Coptica*.

## **1. Works and their identification**

The necessity to establish a complete and reliable catalogue of all Coptic literary works has long been acknowledged. It was among the initial aims of the CMCL (*Corpus dei Manoscritti Copti Letterari*) project.<sup>1</sup> The digital philology project, launched in the 1970s, offers editions of Coptic texts, catalogues of manuscripts, bibliography, and grammar help for the study of Egyptian Christian culture from the first to the twelfth century CE. Among other contents, the project's rich database currently includes the so-called *Clavis Patrum Coptorum*: a freely retrievable list of the authors and works of the Coptic literature with information on manuscripts, content, and critical problems.<sup>2</sup> This preliminary repertory has already been widely used by the research community to identify Coptic literary works. Yet, the recent development in the discussions on 'Textual Fluidity' (on which more below, §1.2) has not only fueled my interest in the theoretical and methodological aspect of these problems, but also evidenced the necessity to refine the definitions of the works and consequently update the repertory. In the following, I would like to expose synthetically the relevant methodological reflections, as they may also be useful outside the restricted scope of the Coptic *Clavis*.

In the course of my research, it has become clear that the principles according to which texts of the ancient Christian tradition are identified and classified (what is generally referred to as Patristics) derive from centuries-old practice of certain ecclesiastical 'schools', with their particular interests. In this context, the 'school' means any culturally organized identification and presentation of texts and of their respective historical setting, including the attribution, the use in different environments, the reshaping according to different literary genres, etc. The result of this continuous work of systematization, carried on since the very beginning of the Christian literary activity, is a general assessment of the texts guided by religious, and not by properly

1 <<http://cmcl.it>>.

2 <[http://cmcl.it/~cmcl/chiam\\_clavis.html](http://cmcl.it/~cmcl/chiam_clavis.html)>.

historical concerns. The principles and the results therefore clash with modern requirements of historical systematization, calling for a change in paradigm, away from the traditional model.

### ***1.1. The ‘schools’***

I would like to begin by giving a general synthetic survey of the work of the ‘schools’, as I see it, in the three most significant periods of their development, (a) the Late Antiquity, (b) the Renaissance and Enlightenment, and, finally, (c) the mid-nineteenth to early twentieth century.

#### *(a) Late Antiquity*

The late antique period is characterized by the systematization in view of ecclesiastic-liturgical use. This use has many manifestations, each with a different influence on the presentation but also the modification of the texts. The four main manifestations (just to give an example, there are many more) are homiletic, canonical, hagiographic, and polemic (or apologetic). In each case, a text can be submitted to changes to meet the demands, which change with time and depend on ecclesiastic and cultural conditions. Those changes are in turn of different types. Finally (for us), the texts are classified and ordered (typical work of a ‘school’) in a way that tends to conceal the preceding process. Moreover, the texts are arranged in a relative order of importance, which does not correspond to their real meaning and impact in various environments in preceding times and situations. As a notable example, one may mention the monastic literature, such as the Lives of St Basil, Isaiah of Sketis, or Paul of Tamma. Many texts fall into oblivion, others are revised to be useful in new conditions, and still others are produced and spuriously attributed. Even more striking but less important cases include the Lives of St Pachomius and St Anthony, not to speak of the so-called *apocrypha*.

#### *(b) Renaissance and Erudition*

The Renaissance period is characterized by the recovery of Greek patristic texts in the Latin world. They were seen as very authoritative testimonies of a so far lost theological and rhetorical wisdom, therefore were preserved in the same arrangement and textual conditions as they were found. This deference was not without some philological criticism, as it is only to be expected in the period when modern philology was born: the denunciation of the Chrysostomic *spuria* is an obvious example. But, on the whole, the scant quantity and quality of the documents (what survived of the monastic libraries of the eleventh-twelfth centuries and was brought to Europe) led the scholars to adhere to the situation that they represented, without going in search of a more ancient and, in some respect, more genuine situation. In this period of erudite rather than historical studies, the cultural treatment of the texts consisted



above all in submitting them to the recent invention of printing. In fact, our knowledge of the texts and their systematization is still based on the seventeenth and eighteenth-century editions through the collection of Migne.

(c) 1850-1920

Through contacts with the Near Eastern territories and other modern industrial improvements, the quantity and the quality of available documents significantly increased by the mid-nineteenth century. The best example here are the papyrological discoveries. The second half of the nineteenth century was also the period of the scientific classification, and of the birth of Patristics as discipline, inspired by the new sense of history and of a rational critique. The attention was devoted to reconstructing historical (including, as in our case, literary and cultural) phenomena, and their development, through an accurate analysis of the documents, going beyond the earlier assessments and evaluation.

This was at least the theory, and, as such, it would be valid even today. In practice, however, research was (and has since been) carried out under the assumption of a conventional, and an undiscussed idea of *text*, shaped under the influence of the needs of the printing process. This is actually a well-known fact, yet the scholars of Patristics have not drawn any consequences.

Besides, with the advent of printing, diffusion of identical copies became possible, and this imposed the concept of a 'text' as original, invariable, and unique.<sup>3</sup> The accompanying concepts of author, title, layout, chapters, and paragraphs (i.e. internal subdivisions) also assumed a role and a meaning that earlier had been different. There is yet another aspect to keep in mind. The modern history-driven perspective, which was so essential to the progress of the *Altertumswissenschaften*, in the case of Patristics was vitiated by prejudices of religious beliefs and by interests of ecclesiastical confessions. This could happen because, in the organization of the academic studies, the history and literature of Christianity, and of the Churches, was considered of minor or secondary impact, and in any case a prerogative of religious people. As a consequence, patrologists have been inclined to preserve the traditional definition of texts, focusing the discussion on the evaluation of contents—and by doing so charging that same tradition with some modern assumptions that were extraneous to it.

Actually, it was the relatively recent awareness of Christian Oriental literatures, with documents often antedating their Greek and Latin counterparts, and often exhibiting a different character, that has contributed to the new understanding of how a text should be viewed. In fact, I believe that the scholars

3 Cf. Eisenstein 1983, and relative discussions, e.g. De Franceschi 2012.

of oriental traditions enjoy a privileged vision, which they should promote in the face of the traditional (in substance only Greek and Latin) Patrology.

### *1.2. New approaches*

The consequences that can be drawn from the above observations have much in common with some theories dealing with a renewed, post-modern, understanding of the composition and transmission of texts, theories that have recently assumed an identity of their own. In recent times, a special sensibility emerged concerning the difficulty of identifying with certainty and precision *one* literary work, above all in the field of religious studies, particularly the early Jewish and early Christian production. This set of problems has also taken an official name, ‘Textual Fluidity’—and also, less appropriately, ‘New Philology’. In the field of Coptic studies, one could mention in this regard Batovici for the Greek Canon<sup>4</sup> and Lundhaug for the Nag Hammadi manuscripts.<sup>5</sup> This does not mean that I agree with the latter’s proposal about the origin of the Nag Hammadi material. The scholars of Coptic literature have been long conscious of the related problems, not only because the texts were easily and frequently submitted to linguistic rearrangements of many kinds; but also because their transmission has been very tormented, so that often the identification is problematic.

In fact, the ideas proposed by the ‘New Philology’ theory are far from being new, they have existed in philology from the very beginning, but in ‘New Philology’ they assume a sort of operating functionality which previously they had not. When trying to establish a list of the Coptic works, it is not sufficient any more for an identification of a specific work to see the similarity of the text in different manuscripts or its mention or quotation in other works. The same applies to fragments, or to selections transmitted as parts of other works. In fact it may seem reasonable to consider ‘works’ as living creatures, each with an autonomy of its own, which could grow, shrink, change shape, split, yet preserving their original identity.

While I do not want to decline this assumption altogether, it is worth emphasizing that it should be the result of a literary and historical analysis of each work, and not an initial presupposition. I would like to argue that when organizing a reliable repertory (*clavis*), one should avoid acquiescence to a more or less conventional tradition, in particular when it has been showing itself as more and more problematic. We must therefore follow another kind of formal scientific perspective to overcome the errors of the past.

First, it is necessary to avoid a number of confusing issues. One of them is the use of the same term ‘text’ to indicate both an individual work and the

4 Batovici 2016.

5 Lundhaug and Lied 2017.

content of the work in a given document. Another grave error that has been made in the conception and planning of past repertories consists in seeing them as related to, and in a sense accompanying, the historical assessment of the works. Instead, they should be completely independent from the attributed features, like authorship, chronology, or environment. The ordinal number attributed to a work, and consequently the sequence in which they appear, should be fully conventional (in the semiotic sense). It should depend on the documents (manuscripts), in whatever casual order—or better disorder—the author of the repertory considers them. As for the classifications (that can be more than one), these should be left to separate annotations relating to the content (attribution, title, subject, theories), accompanying information on manuscripts and editions, by means of indices—or, in the preferable case of digital treatment, by links that can be retrieved by the users. Good examples here could be the cases of Isaiah of Sketis<sup>6</sup> or of Pachomius<sup>7</sup> and their extremely complicated tradition in Greek and Oriental languages, which cannot be assumed under ‘comprehensive’ numbers. Another example is the case of the Coptic anthology of John Chrysostom *In Hebraeos* (CC 0169, MONB. CR etc.) which is a work by itself and also a collection of excerpta not always matching the Greek text.

### 1.3. *Clavis Coptica*

For the reasons mentioned above, the *Clavis Coptica*—which I have produced, and which is being constantly improved and augmented (and which of course will never be complete)—is conceived for the digital presentation, and especially according to the principles of the relational system.

The starting point is the content of the manuscripts, according to the divisions of what the scribes wanted to present as individual texts, normally on the basis of their models, and only when two or more individual texts are exactly<sup>8</sup> the same they may be subsumed under the same number. The accompanying tables of the related textual units in the sense of the history of literature (i.e. Patrology), authors, titles, literary genres, etc. provide information through cross-references in the relational tables.

The other instruments available for a systematic list of patristic and hagiographic works are the CPG (*Clavis Patrum Graecorum*, with the annexes), BHG (*Bibliotheca Hagiographica Graeca*), BHL (*Bibliotheca Hagiographica Latina*), and the BHO (*Bibliotheca Hagiographica Orientalis*). While it

6 Cf. Suciú 2012.

7 Cf. Veilleux 1980-1982.

8 The value of this ‘exactly’ is uncertain. Of course, much is left to the subjective judgment of the philologist. But due to the great flexibility of the organization that I propose this is not too important.

may seem reasonable to take them as a privileged example, in my opinion it is better to build something new and more formally consistent.

## **2. *The manuscripts and their identification***

Similar observations to those made for the identification of the works are valid (*mutatis mutandis*) for the identification of the manuscripts.

Here we start from a relatively simple classification going back to the Renaissance (e.g. *Laurentianus Plut.* LXXXI, 35; *Parisiensis Graecus* 54), when manuscripts were few and well known, down to the situation of today, especially for the Oriental manuscripts, when shelf marks, or call numbers, are subject to frequent changes, and the real arrangement of the manuscripts is difficult to know, unless visiting the places. It also happens that previously independent fragments are put together under a new number; that pieces move to another collection; that printed catalogues give obsolete or unreliable information on the numbers; that some pieces are dismembered, or put together, etc.

All this leads to the conclusion that if we want to constitute a stable and reliable list of manuscripts we cannot use the current official call numbers, or rather we can refer to them only in a second instance. For a stable reference, we must produce a standard list with its own numbers, from which and to which it will be possible to establish a relation to the shelf marks of libraries, museums, and collections.

At the same time, recent astonishing advances in digital reproduction and storage have resulted in the possibility of forming virtual collections, which may be at the same time much more flexible and to a certain degree more stable than their material counterparts do. Such collections may be manipulated in a countless number of ways, while each item may still keep an unaltered identification number, whichever changes the item or its collection undergo. Therefore, in order to preserve an operative stability, independent of the variations of the items, I maintain that the repertories that aspire to provide standards for the identification of the manuscripts should base on sufficiently international, reliable, and stable collections of digital reproductions.

To obtain this, it is necessary (a) on the theoretical side, to establish a satisfactory ontological definition of the objects, i.e. the manuscripts, which may be made of many different materials; (b) on the operational side, to take advantage in the most rational way of the digital tools available today.

(a) The theory should first establish the basic, minimal entities to which an identity number is assigned. In the case of what are generally called codices, or of fragments thereof,<sup>9</sup> these entities should not be the codices in their

9 In this field, the terminology is of the greatest importance. I have made elsewhere some proposals, see Orlandi 2013.

entirety, even if so preserved in one collection, but single leaves<sup>10</sup> which make them up (or used to make them up). This is necessary because the leaves may have been displaced, extracted, recomposed, and we should be able to show this through proper models. The identity number of these entities is bound in the first instance to a digital reproduction, which directly shows their appearance, and then to the call numbers through which they are (or were) identified in the collections.

(b) The praxis is based on the possibility, well established today, to create, maintain, share, and retrieve archives of digital images at a very low cost.<sup>11</sup> One number identifies every folio (in two images, for recto and verso), and all the numbers with the address of the respective images constitute the first table of a relational database. The other tables list the attributes of the folios (numeration, script, layout, etc.), the codices of which they are or were part, the collections to which they belong or belonged, layout, publications, studies, etc., always by means of tables of entities and tables of relations. In this way, the call numbers themselves are a piece of information added to the entities, and not their primary identification, and may be more than one, following the whimsical history of the manuscripts.

### **References**

- Batovici, D. 2016. ‘The Apostolic Fathers in Codex Sinaiticus and Codex Alexandrinus’, *Biblica* 97/4 (2016), 581–605; <doi:10.2143/BIB.97.4.3186058>.
- De Franceschi, L. 2012. ‘La storia della stampa vista da Elizabeth L. Eisenstein in *Divine art, infernal machine*’, *AIB studi*, 52/2 (2012), 217–225; <doi:10.2426/aibstudi-8338>.
- Eisenstein, E. 1983. *The Printing Revolution in Early Modern Europe* (Cambridge: Cambridge University Press, 1983).
- Lundhaug, H. and L. I. Lied 2017. ‘Studying Snapshots: On Manuscript Culture, Textual Fluidity, and New Philology’, in: H. Lundhaug and L. I. Lied, eds, *Snapshots of Evolving Traditions: Jewish and Christian Manuscript Culture, Textual Fluidity, and New Philology*, *Texte und Untersuchungen zur Geschichte der altchristlichen Literatur*, 175 (Berlin and Boston: De Gruyter, 2017), 1–19; <doi:10.1515/9783110348057-005>.

10 The leaves will be the minimal entity, even if they are part of a bifolium, as it is generally the case. The information on the reconstruction of bifolia, quires, codices, etc. is contained in the relevant tables of the database (cf. below).

11 I mean here the cost of the software; that of the production of the images is responsibility of the authorities of the collections.

- Orlandi, T. 2013. 'A Terminology for the Identification of Coptic Literary Documents', *Journal of Coptic Studies*, 15 (2013), 87–94; <doi:10.2143/JCS.15.0.3005414>.
- Suciu, A. 2012. 'The Louvain Papyrus Fragments from the *Asceticon* of Abba Isaiah', *alinsuciu.com*, 12 May 2012 (<<https://alinsuciu.com/2012/05/12/the-louvain-papyrus-fragments-from-the-asceticon-of-abba-isaiah/>>, last accessed 14 August 2018).
- Veilleux, A. 1980–1982. *Pachomian Koinonia*, I: *The Life of Saint Pachomius and His Disciples*, II: *Pachomian Chronicles and Rules*, III: *Instructions, Letters and Other Writings*, Cistercian Studies, 45, 46, 47 (Kalamazoo, MI: Cistercian Publications Inc., 1980, 1981, 1982).

# Some Remarks about Coptic Colophons and Their Relationship with Manuscripts: Typology, Function, and Structure

*Agostino Soldati, Sapienza Università di Roma*

The brief article offers some remarks about the devotional requirement, the main formal features, and the historical relevance of the colophons often preserved by Coptic manuscripts.

In a seminal contribution on the contrived system of dating by way of a cluster of arithmetical fractions that many Arabic, Persian, and Turkish copyists put in place in the so called *zusammengesetzte Unterschriften*, Gustav Leberecht Flügel observed that, among the areas of Islamic book production, Egypt is the one where the habit of concluding the copy of a manuscript with a dated subscription is most widespread and enduring.<sup>1</sup> The permanence of such a habit might be explained by the existence of another local scribal tradition, which preceded and accompanied the beginnings of Arabic book production, namely the Coptic one. Coptic manuscripts offer us some of the earliest instances of scribal subscriptions within the written cultures of the Christian Orient, which probably even predate those found in the earliest Greek book production.

Before we approach colophons, a feature attested in all manuscript cultures, some clarification on terminology is necessary. Rather than resorting to the supercilious Grecism ‘colophon’, often deplored for its *in vitro* origin dating back to Renaissance proto-typography, one may be tempted to label the ending paratexts of a manuscript with the original ancient term that roughly corresponds to ‘colophon’ in each written tradition. In this sense, it would be natural to use ὑπογραφή, *subscriptio*, and ختم, for the closings of a Greek, Latin, or Arabic manuscript, respectively. Yet, it would not be equally straightforward to find an appropriate synonym for Coptic, based on what we know so far about its technical vocabulary of book production. In the absence of direct attestations of this specific meaning in Coptic, the most likely term for designating ‘colophon’ should perhaps be π-κωρϣ, which often translates σφραγίς, πάεσθαι, or, less probably, τ-τοοβес, expressing rather the material imprint

\* The research leading to these results has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 687567, PI: Paola Buzi, Sapienza Università di Roma), <<http://paths.uniroma1.it>>.

1 Flügel 1855, 357.

of a seal.<sup>2</sup> The Greek loanword τ-χρηγογραφη may also have been used in such sense.

Beyond their more or less elaborate structure, the main feature of Coptic colophons consists in their obvious significance as *Schenkungsurkunden*, namely sorts of documents witnessing a peculiar typology of not altogether material transaction, by which a donee religious institution acquired the property of a book and its donor obtained forgiveness for all the sins committed during his life as well as blessing for the afterlife, thanks to the diuturnal use of his gift for devotional or liturgical purposes. The terms of such a fideistic deal, even including a special intercession and indulgence for a soul expiating in hell all faults committed in life, are eloquently expressed in a passage of Cyril of Alexandria, *De hora mortis*, 188 Amélineau, with a wording definitely frequent in lots of colophons:<sup>3</sup>

εταχεναι τηρου ετετεναγαπη ω πιδος μηαιχρς ογορ νιωηρι ντετκαθολικη  
νεκκλησια εωβενη ετωωπ νηιχωμ νωω εγτ μηωου εβορν επι νηφτ καν ογκογχι  
πε καν ογνιωτ πε ογωε νερφμεγι φανεε πε ενατκην ηενπη νηφτ. τχω δε νμοc  
νωτεν ω πιδος μηαιχρς χερωνι νιβεν εωαφωωπ νογχωμ ντεετηνι εβορν επι  
νηφτ ιχενπηναγ ετογναωω νητηγ ηεντεκκλησια εωωπ πρωμν ετεμμαγ οη  
ηεντογνογ φαγχα νπεφραν επιχωμ ηπωη νσετ ναγ ντεεπροςφορα νζ κωω  
νσοπ ηενογμογ. εωωπ δε οη πρωμν εταφωοπιχωμ αφι εβολ ηενσωνα εωωπ  
αφιρ νογκογχι ννοβι ογορ αγιτη ενικολασιc ιχενπηναγ ετογναωω ηπιχωμ  
ηεντεκκλησια σεναενφ επωωι ηεναηεντ (λβ) ηεννικολασιc ετεφωωπ νηητογ  
σεναηαι ναγ ηεντογνογ.

I said all such things for charity (ἀγάπη) toward you, o Christ-loving laymen (λαός) and sons of the catholic (καθολική) church (ἐκκλησία), for those who buy books for reading and donate them to the house of God, whether (κἄν) they are of small size or big, there shall be an eternal and unceasing memory in the house of God. Thus, I say to you that, o Christ-loving laymen, if any man buys a book and donates it to the house of God, from the moment it is read in the church, if that man is alive, immediately his name is written on the book of life and his offering (προσφορά) will be rendered back to him in blessing multiplied by seven. But if the man who bought the book has left his body (σῶμα), if he committed a little sin and was brought toward the punishments (κόλασις), from the moment the book is read in the church, he will be lifted from hell, from the punishments he will have suffered there, and he will obtain mercy immediately.

In this perspective, the sometimes confusing list of living and deceased persons inserted in the text of colophons may assume a quite clearer relevance to the main goal of these texts. In the earliest instances, namely in the subscriptions in the recently discovered Theban codex of Pseudo-Basilian *Canons* and in the two single leaves kept in Turin (P.Tor. Copt. Inv. Provv. 6266

2 Crum 1939, 398b.

3 Amélineau 1888, 186–187.



and 8548),<sup>4</sup> all likely to be attributed to the seventh century, the colophons appear in their simplest structure. In these *incunabula*, both fundamental elements of the *Schenkungsurkunde* occur: the mention of the church to which each book is offered and the name of its donor, on whose behalf the reader is asked to beseech God. The fact that in one of the Turinese leaves the name of the female donor of the manuscript is withheld does not prevent the reader from beseeching God to have mercy on her and on her relatives, whether dead or alive, in grace of the explicitly referred to divine omniscience.

The classical shape of the Coptic colophon was achieved, in the following century, in the *scriptorium* of Toutōn. There, we record the earliest instances of the normative phrasing that constitutes the bulk and the framework of each Coptic scribal subscription up until the end of the Coptic *Schrifttum*. This model begins with a zealous and polite request of prayer (αριταγαπη --- ψηλη) addressed to any user of the book (ογον νιμ ετναωψ) on behalf of the one who, literally, took care of the book at his own expenses (αφρηπροοψω ἡπειρωωμε ρῆνεφρισε ἡμιν ἡμοσ) and donated it to (αφτααφ ερογν ε-) a religious foundation, whose eponymous saint or angel is asked to intercede (παρακαλει) before God. Eulogies, scriptural direct quotations, Christian *historiolae* paradigmatic for the desired divine intervention enrich that preset canvas in various ways, so that texts often reach a sizeable extent. This main part is always written in a typical sloping uncial.

In most cases, the mention of the copyist and the dating formulas (encompassing in their most complete layout, day, month, indiction, *annus Martyrum* and *annus Hegirae*) are written in a more or less accurate minuscule hand in vulgar Greek (the proper name in nominative precedes a passive aorist ἐγράφη or an odd καλιωγράψατε<sup>5</sup>). In some instances the copyist offers, besides the Greek ὑπογραφή, some information about his work, apart from accustomed expressions of Christian self-effacement (unworthiness to take even his own name, ατῆπσα ἡπραν ἡταγταλοσ εχωσ;<sup>6</sup> acknowledgement of his own unsuitableness, ἡπατηνοι καλωσ αλλα ειχιστω / αλλα ανοκ οσβογι etc.<sup>7</sup>). The main detail is that the text would have been copied in accordance with a certified exemplar of the literary work reproduced (αισραι καταπαντιγραφον ετραρω<sup>8</sup>). This apparently ancillary statement could have been actually included in the text in pursuance of a straightforward *Beschwörung* similar to the one we read in Eusebius' *Historia ecclesiastica* V 20,2. As it is well known, there the author makes the 'one who will transcribe' (μεταγραφόμενος) swear

4 Both republished in van Lantschoot 1929, CV–CVI.

5 About such odd verbal form see Soldati 2017, 26, n. 8.

6 E.g. van Lantschoot 1929, 40–41, n° XXII, 6–8.

7 E.g. van Lantschoot 1929, 216–218, n° CXXI, B11–13.

8 E.g. van Lantschoot 1929, 10–12, n° V, 29–30.

‘that (he) will collate what (he) will have transcribed, ad (he) will amend it according to (a specific) copy, wherefrom (he) will have transcribed (it) carefully’ (ἵνα ἀντιλάβῃς, ὃ μετεγράψω, καὶ κατορθώσης αὐτὸ πρὸς τὸ ἀντίγραφον τοῦτο, ὅθεν μετεγράψω, ἐπιμελῶς).

The salvation ensured to the donor is not the only feature which could closely associate colophons with texts of Christian magic. Another relevant element is the curse, sometimes attached to the colophon, against those who would dare to steal the book. The Coptic wording is akin to the coeval Greek one. As the curses in Greek codices anathematize that ‘the one who has profaned will not be enrolled in the book of life’ (ὁ γοῦν σολήσας μὴ γραφῆ ζωῆς βίβλω),<sup>9</sup> in a similar vein the Coptic copyist echoes, somehow in a more poetic way, ‘might he not take his share from the tree of life’ (νεφι ἡπεμερος εβολ ρῖπωμη ἡπωνῆ).<sup>10</sup>

Beyond their interest for *Urkundenlehre* and religious studies, Coptic colophons, as actual documentary texts, offer plenty of data about medieval Egypt. Moreover, whilst the goods which are the object of many Coptic deeds preserved in papyrus collections are irretrievably lost, preventing us from a full comprehension of the very terms of the transaction they register, colophons are a peculiar documentary category that, in the majority of cases, comes along with the object they sanction the gift of. Usually they disclose to us copious evidence of the cultural, economic and social *milieus* where the books were produced, sold, acquired, preserved and used. Notwithstanding the rich amount of information the colophons offer about the routes manuscripts often embarked on, they also bear witness to the inexorable withdrawal of Coptic culture against the relentless advance of linguistic and cultural Arabization.<sup>11</sup>

## References

- Amélineau, É., 1888. *Monuments pour servir à l'histoire de l'Égypte chrétienne aux IV<sup>e</sup> et V<sup>e</sup> siècles*, Mémoires publiés par les membres de la Mission archéologique française au Caire, 4 (Paris: Ernest Leroux éditeur, 1888).
- Crum, W.E. 1939. *A Coptic Dictionary* (Oxford: At the Clarendon Press, 1939).
- Flügel, G. 1855. ‘Eigenthümlich zusammengesetzte Unterschriften muhammadanischer Manuscripte’, *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, 9 (1855), 357–367.
- 9 MS London, British Library, Add. 19387, f. 235.
- 10 E.g. van Lantschoot 1929, 169–170, n° C, 37–38.
- 11 See the instance of the long Fayyūmic paratext preserved by Vatican City, Biblioteca Apostolica Vaticana, Vat. copt. 68, f. 162v, about which cp. lastly Soldati 2017, 23–24.

- van Lantschoot, A. 1929. *Recueil des colophons des manuscrits chrétiens d'Égypte*, I: *Les colophons coptes des manuscrits sahidiques*, Bibliothèque du Museon, 1 (Louvain: J.-B. Istaš, 1929).
- Soldati, A. 2017. 'A New Piece in the Prosopography Mosaic of the Coptic Scriptorium of Toutōn: Pantouleos, Son of Houmise', *Comparative Oriental Manuscript Studies Bulletin*, 3/1 (2017), 23–32.



# Encoding and Annotation of Ancient Places in Ethiopia

*Solomon Gebreyes Beyene and Pietro Maria Liuzzo,  
Universität Hamburg*

Encoding places in a text and producing annotations offers scholars several ways to investigate new questions and support their research work. In this paper we present the work of annotating a text and the way in which we produced Linked Open Data to share with the Pelagios Commons. Using examples taken from the Chronicles of Christian Ethiopian kings, the paper presents the techniques used to encode the data and the visualizations and insights which can be produced for scholarly research from both the TEI and the RDF.

## ***1. Introduction***

Scholarship on the historical geography of Ethiopia is still in its very early stages. So far, research into the country's ancient and medieval period has largely concentrated on the reconstruction of the political history of kings and churches rather than of the historical geography of the Christian kingdom during its state formation. Very few studies exist that are primarily dedicated to toponyms, definition of regions in given times, or to the location of ancient places. The number of different languages and literary traditions one should master is an additional challenge: at least Gə'əz, Coptic, Arabic, Latin, and other European languages are necessary. This diversity often calls for a heavily collaborative effort. In addition, the style of most of the available sources is very peculiar making the interpretation process problematic. On top of this, the existence of various transcription conventions for Gə'əz makes the simple attempt to answer the question 'where is this place?' and locate it, even on a modern map, a struggle, because one has to try a series of possible variants without any guarantee.

The only available comprehensive work on historical geography is that by Huntingford, completed in 1969 and only published twenty years later.<sup>1</sup> With all the updates made until 1979 and by the editor until 1989, this work has considerable limitations—as already noted by Irvine in his review, it is 'not so much a historical geography as a reasoned topographical commentary established on a chronological basis, an attempt to 'explain Ethiopian topography in its historical setting''<sup>2</sup>—especially in so far as it is based only on a selection of sources, while many more have become available since. The *En-*

1 Huntingford 1989.

2 Irvine 1991, 372.

*cyclopaedia Aethiopica*<sup>3</sup> was a major step forward in the description of most relevant places, but, because of the nature of this work, there is no systematic connection to the toponyms' attestation in the sources, nor to the diachronic level of both toponymy and geolocation.

Scholarly cartography available is also scanty. It includes the historical maps in Huntingford's work, the maps accompanying the entries in the *E Ae*, and those additionally provided in the fifth volume, some scattered maps in print publications, but to this day the colonial Italian publication *Guida dell'Africa Orientale Italiana*<sup>4</sup> is often the tool used to locate historical sites. Thus, in most cases, to 'identify and pin on a map' a place, a region, an administrative unit at a given time is still a big problem.

As a consequence, it is little wonder that resources available online to support research on the historical geography of Ethiopia are inexistent. This is also linked to the lack of relevant primary sources online, as most digital projects in the field have recently focused on the digitization of manuscripts and their cataloguing. Almost no encoded digital online edition of texts belonging to the Ethiopian literary tradition has been produced.

Only in the past year something has been done, in the framework of the long-term project *Beta maṣāḥaft*.<sup>5</sup> Among its efforts, the project aims at building a Gazetteer of places which is based initially on the *index locorum* of the *E Ae* and makes use of annotations in primary sources of toponyms to provide computable and structured data to scholars for further investigation.<sup>6</sup> We have in this way a core base of around 6,000 abstract places, most of the time with at least one label name, the one used as reference in the index of the *E Ae*.

With a Resource Development Grant by the Pelagios Commons,<sup>7</sup> in the summer of 2017 we accomplished a further step towards a reusable resource starting from the text of the *Chronicle of King Galāwdewos*.<sup>8</sup> We developed the methodology and the workflow required for the annotation of historical geographical data in our place records, and we have built the infrastructure required to export the Gazetteer and the annotations to the Pelagios Interconnection Format. In the following pages, we will briefly summarize the en-

3 *E Ae*.

4 Consociazione turistica italiana 1938.

5 Hosted by the Hiob Ludolf Centre for Ethiopian Studies at Universität Hamburg, funded by the Union of the Academies of Sciences in Germany and supervised by the Academy of Sciences in Hamburg; see the project note by Dorothea Reule in this issue.

6 Liuzzo 2017.

7 <<http://commons.pelagios.org/>>, last access 22 May 2018.

8 See the original reports in Liuzzo and Solomon Gebreyes 2017a, 2017b, 2017c .

coding decisions and the annotations model and give some examples of how these annotations are used and visualized.

## 2. *The selected sources*

We have started to annotate ancient places in texts where these are more abundant and can more quickly produce relevant computable outcome. The first text we annotated was the *Chronicle of King Galāwdewos* (CAe 3122, ID: LIT3122Galaw<sup>9</sup>), as a test case to figure out mark-up strategies for Ethiopic texts in *fidal*.

The *Chronicle* is an important source for events involving the Ethiopian kingdom and Christian–Muslim relation in the Horn of Africa in the sixteenth century. Galāwdewos became emperor at the age of eighteen in 1540, and his reign was marked by a successful struggle against ʾAḥmad b. ʾIbrāhīm al-Ġāzī (hence *imām* ʾAḥmad; ID: PRS1522Ahmadb). It was one of the first texts to be encoded in TEI within the *Beta maṣāḥāft* project for its importance.

Other historical texts have been encoded since, including the *Chronicle of Susānyos* (CAe 3951, ID: LIT3951ChronSusenyos<sup>10</sup>) and the *Confessio Claudii* (CAe 1252, ID: LIT1252Confes<sup>11</sup>). We have also acquired annotations produced by Eugenia Sokolinski for the TraCES project<sup>12</sup> on the *Chronicle of ʾAmda Ṣāyon* (CAe 4275, ID: LIT4275ChronAmdS<sup>13</sup>). Finally, we have annotated named entities in the transcriptions of quotations in manuscript descriptions (colophons, additions, etc.). At the time of writing this contribution, we have 3,270 places attestations with identification in the data set, 1,152 of which are in literary works, the others in manuscript text transcriptions or metadata.<sup>14</sup>

In the next section, we will detail how these attestations of toponyms have been marked up in the TEI source data of the *Beta maṣāḥāft* project.

## 3. *Encoding references to places in the source texts*

We encode place attestations in our TEI source files using the `<placeName>` element<sup>15</sup> with an attribute `@ref` which points to an authoritative identifier.

9 Solomon Gebreyes Beyene 2017. Here and in the following, we provide the ID of the named entities in the *Beta maṣāḥāft* database; for the texts, we additionally provide the *Clavis Aethiopica* (CAe) number.

10 Solomon Gebreyes Beyene and Reule 2017.

11 Solomon Gebreyes Beyene 2018.

12 See Bausi 2015 and the TraCES project note in this issue.

13 Pisani et al. 2018.

14 There are also 8,646 such annotations for persons which are not within the scope of this contribution.

15 TEI consortium 2016.

To provide a better understanding of the place names marked up in the Gə‘əz text let us take some more examples from the *Chronicle of King Galāwdewos*. For reasons of military strategy and in order to gather and organize a stronger army, Galāwdewos decided to flee to Šawā, a province where his predecessors were accustomed to camp regularly. The chronicler describes the king’s route in the following way:

*Mār* Galāwdewos then crossed beyond two rivers from the direction of Təgrāy to the direction of the land of Šawā in order to see there those of his flock in righteousness and his innocent people in equity. He reached the country he wished in the month of Ғazirān which is the month of Sane, the beginning of winter months of the Abyssinians. (Chr. Galaw. 21)

In medieval times, the two big rivers were the Takkaze (ID: LOC6850Takkaze) that separates Təgrāy from Gondar and the Blue Nile (ID: LOC1022Abbay) that separates Goḡḡām from Šawā. These are the rivers that are certainly meant here, even if they are not mentioned by name in the text, therefore we add an empty element in the text which points to the best available authority identifier. In the same passage we have places which are named and are thus marked up with a reference to the relative place entity.

We have created a record for each place attested in the text which did not yet have a record in the *Beta maṣāḥaft* Gazetteer but fell within its scope, i.e. Eritrea and Ethiopia in the most inclusive sense.

Below is a sample of the marked-up places taken directly from the text, where we exemplify our use of the attributes @notBefore and @notAfter to encode the relative chronology of each attestation within a text. In this case, we have also made a relative reference explicit, by assigning a date in the Gregorian calendar to dates relative to the literary context:

**ምዕራፍ ፡ ፱ ፡ ወእምዛ ፡ ዓደው ፡ ማር ፡ ገላውዴዎስ ፡ ማዕድተ ፡**  
ref="http://betamasaheft.eu/LOC6850Takkaze"/><placeName  
ref="https://www.wikidata.org/wiki/Q882739"/>><placeName  
ref="http://betamasaheft.eu/LOC1022Abbay" notBefore="1540"  
notAfter="1541">**ክልኢቱ ፡ አፍላግ ፡**</placeName> **መንገለ ፡ ምድረ ፡**  
<placeName ref="http://betamasaheft.eu/LOC5597Sawa"  
notBefore="1540" notAfter="1541">**ሼዋ ፡**</placeName> **እመንገለ ፡ ምድረ ፡**  
<placeName ref="http://sws.geonames.org/444187" notBefore="1540"  
notAfter="1541">**ትግራይ ፡**</placeName> **ከመ ፡ ዶርአይ ፡ መርዒቶ ፡**  
**እለ ፡ ህየ ፡ በጽድቅ ፡ ወይዋሪ ፡ ሕዝቦ ፡ በርጉዕ ፡ ወበጽሐ ፡ ኀበ ፡**<placeName  
ref="http://betamasaheft.eu/LOC5597Sawa" notBefore="1540" no-  
tAfter="1541">**ምድር ፡**</placeName>**ዘፈቀደ ፡ በወርቃ ፡ ሐዘራን ፡ ዘውእቱ ፡**  
**ወርቃ ፡ ሰኔ ፡ ርእሰ ፡ አውራጎ ፡ ክረምቶሙ ፡ ለእልገቡስ ።**



After the king settled in Šawā in the first year of his reign, in 1541, he moved for some time from one locality to another, recruiting armies from the Christian communities, and then finally he stationed in the mountainous region of 'Ifāt (ID: LOC3921Ifat), where he fought the then governor of the region, Naṣraddin (ID: PRS7506Nasraddi), son of *imām* 'Aḥmad (see also below). In the first confrontation, Galāwdewos was defeated, but in the second confrontation, he was victorious. This victory helped him to attract a large number of Christian adherents who joined him.

The range of places involved in this text is not limited to local history, with entities such as the Ottoman empire or Portugal being mentioned. Thus, during the second year of the reign of King Galāwdewos, Portuguese soldiers arrived on the coast of the sea to assist the king and sent a messenger to him to join them as soon as possible. These external entities should also be certainly marked up. For places going beyond the Ethio-Eritrean historical boundaries, we have used whenever possible the Pleiades identifiers. For places which are outside the scope of both our Gazetteer and Pleiades we used Wikidata entities (e.g., the Ottoman empire, Wikidata ID Q12560).

In the example below, the island of Patmos, which is attested in the text as ጥሙስ (Təmus), is marked with a Pleiades ID.

```
<ab>ወእምዝንቲ፡ ኩሉ፡ ለግር፡ ገላውዴዎስ፡ በዓለ፡ ዝንቲ፡ ዜና፡ ኮነ፡ የዓቕቦ፡ ዓቃ
ቢሆ፡ ለኩሉ፡ ከመ፡ ተዓቕቦት፡ አሐቲ፡ ብእሲት፡ ወወልዳ፡ እምነ፡ አርጭ፡ ዓቢይ፡
ዘርእዮ፡ ዮሐንስ፡ በዐለ፡ ራእይ፡ ዘደሲተ፡ በ<placeName ref="https://pleia-
des.stoa.org/places/599872">ጥሙስ #</placeName> ... </ab>
```

After all this *Mār* Galāwdewos, the subject of this story, was protected by the protector of all, just as <was saved> a woman and her child from a great serpent, which John, the seer of the Revelation, saw on the Island of Patmos.

Some of the main problems in the annotation of places in the *Chronicle* were:

- the identification of places referred to very generically as ‘the north border’, ‘the sea’, ‘the mountains’;
- the location of places for which only relative location is provided and no archeological evidence is available;
- the identification of the appropriate entity to refer to for each place;
- the treatment of ethnic groups and respective territories.

#### 4. Encoding information about a Place

Whereas the actual mark-up of places was a relatively quick step, the major associated task was to create a record for those places which were not already present in the *Beta maṣāḥəft* Gazetteer. We followed the good advice of the Pelagios community, adopting their understanding of the concept of place:

**Places** are entirely abstract, conceptual entities. They are objects of thought, speech, or writing, not tangible, mappable points on the earth's surface. They have no spatial or temporal attributes of their own. A **place** can exist in name only in an ancient source, without any material correlate; conversely, an archaeological site can exist as a **place** without an ancient name.<sup>16</sup>

We have also limited the chronological attestation to periods relevant to the project's scope, which have been added to PeriodO,<sup>17</sup> and we have encoded information in our TEI files following the example of the Syriac Gazetteer<sup>18</sup> in the 'Syriaca.org TEI Manual and Schema for Historical Geography'. Records for places contain the multiple names and location related to that conceptual place record, they can have bibliography and a (marked up) description. For those places in the gazetteer for which we have our own coordinates, or a link to an entity in Wikidata which does, we also offer a geoJSON export (see for example Gondar, ID LOC3577Gondar).

Place records in *Beta maṣāḥaft* also contain alignment to other authority files, so that in a reference to a local gazetteer a link to an external resource can also be found. We also record associations with other gazetteers, in `<relation>` elements:

```
<listPlace>
  <place sameAs="https://www.wikidata.org/entity/Q1885762"
    type="monastery"
    subtype="institution">
    <placeName>Dayr as-Suryān</placeName>
    <country ref="https://www.wikidata.org/entity/Q79"/>
    <region ref="https://www.wikidata.org/entity/Q1074945"/>
  </place>
</listPlace>
<listRelation>
  <relation
    name="skos:exactMatch"
    ref="http://www.w3.org/2004/02/skos/core#"
    active="http://betamasaheft.eu/INS0478DS"
    passive="http://syriaca.org/place/360"/>
</listRelation>
```

Institutions records, with coordinates, referring to monasteries and churches in possession of manuscript collections, often with a very long tradition, have been inherited from the *Ethio-SPaRe* project.<sup>19</sup>

16 <<https://pleiades.stoa.org/help/conceptual-overview>>, last access 22 May 2018.

17 <<https://test.perio.do/#/p/Canonical/periodCollections/p03tcss/>>, last access 22 May 2018.

18 Carlson and Michelson forthcoming.

19 <<https://www.aai.uni-hamburg.de/en/ethiostudies/research/ethiospare/>>, last access 22 May 2018.

In the following, we provide five examples of locations where major events took place during the reign of Galāwdewos. They include: (1) the place where the King took refuge during the war and began his political career, (2) the place where he organized his military army, (3) a battlefield where he successfully defeated a strong political contender, (4) the place where the King settled following the end of the turbulent period and (5) the place where he fought Muslim adversaries and was killed by them. We give a brief account of each of these places in connection to Galāwdewos, followed by a brief historical description of these places and a discussion as to where and how these places were mentioned in other relevant historical sources.

### **(1) *Dabra Dāmmo***

The political life of King Galāwdewos began in the northern province of the Christian kingdom, Təgrāy (ID: LOC6569Tegray), at the monastery of Dabra Dāmmo (ID: INS0105DD), which remained relatively inaccessible during the continuous assault of the Muslim army. The monastery served as a shelter for the royal family after King Ləbna Dəngəl (ID: PRS6229LebnaDe) had been defeated by the forces of *imām* 'Aḥmad. It was there that Ləbna Dəngəl died, and his son, Galāwdewos, was crowned as his successor.

The importance of Dabra Dāmmo is believed to go back to the sixth century, when it was, according to the local tradition, the first monastery founded in the Ethiopian highlands by one of the so-called 'Nine Saints' who came to teach Christianity, 'abbā Zamikā'el 'Aragāwi (ID: PRS10581Zamika). It has since enjoyed the reputation of a prominent traditional church educational institution, where several famous religious missionaries and monastic leaders were trained, including, according to their hagiographies, 'Iyasus Mo'a (ID: PRS5633IyasusM) and Takla Haymānot (ID: PRS9151TaklaHa). Thus, it played a pivotal role in the expansion of Christianity in its early years and was also a centre of manuscript production and training in manuscript production. Following the foundation of other monasteries like Ḥayq 'Ēstifanos (ID: INS0327DHE) and Dabra Libānos in Šawā (ID: INS0346DL) it lost its dominance, yet it remained important in the Christian Highlands until it was looted and destroyed by the Turkish forces in 1557.

Due to these facts and its particular geographical setting, surrounded by mountainous cliffs, Dabra Dāmmo has been described, or at least mentioned in passing, in various historical works, such as hagiographies, royal chronicles and travellers' accounts.

Dabra Dāmmo appears in the sixteenth-century hagiography of Zamikā'el 'Aragāwi<sup>20</sup> and in the sixteenth-century hagiographies of the two

20 Guidi 1895.

holy men in the thirteenth and fourteenth centuries—'Iyasus Mo'a and Takla Haymānot—as the most important holy place. Thus, the *Gadla 'Iyasus Mo'a* ('Acts of 'Iyasus Mo'a', CAe 1467, ID: LIT1467Gadlal) narrates that 'Iyasus Mo'a joined the monastery of Dabra Dāmmo at an early age.<sup>21</sup>

The *Chronicle of Zar'a Yā'qob* (reigned 1434–1468; CAe 4646, ID: LIT4646Chronicle<sup>22</sup>) mentions Dabra Dāmmo as a monastery where religious men were anointed as monks.<sup>23</sup>

The European traveller, Pedro Páez, mentions in his 1622 account the monastery of Dabra Dāmmo several times.<sup>24</sup> Here are just two examples:<sup>25</sup>

... she should have put them at Amba Damo, a day's journey from Axum, which is much stronger than Guixen amba.

... on a mountain in the kingdom of Tigre, which they call Amba Damo, which is so secure that one cannot climb up except by means of ropes, where there is a large monastery of monks.<sup>26</sup>

In his book, Manoel Barradas, a Jesuit missionary who wrote about the province of Təgrāy in 1634,<sup>27</sup> describes the geographical feature of Dabra Dāmmo as follows:

It was to this amba, which nature has made impregnable, that the Queen took refuge out of dread for the Moorish king Granha [...] as it seemed to her and to her people that this amba was secure and would keep her so. [...] At the summit of this mountain is a beautiful meadow, all even and almost flat enough, as it would seem, to contain a large city. This meadow is located on a level with the surrounding hills, which on the northern side forms a tip and is what is closest to the amba. [...] On the crest of the mountain there is a church named after Abba Aragavy, who was one of the Nine Priests who came from Rome to Ethiopia to spread the faith, and they hold and venerate him as a saint.<sup>28</sup>

## (2) 'Ifāt

After being crowned in Dabra Dāmmo, Galāwdewos left Təgrāy and arrived in Šawā (ID: LOC5597Sawa); then he stationed at a locality in the sultanate of 'Ifāt, in 1541. He spent some time moving here and there and finally stationed at one of the mountains in 'Ifāt where he fought twice against the son of *imām*

21 Kur 1965, 9, line 20.

22 Reule and Solomon Gebreyes Beyene 2017.

23 Perruchon 1893, 12

24 Páez 2011.

25 It would be very interesting in the future to mark up these sources as well, as attestations of this place name, and analyse the way toponymy varies.

26 Páez 2011, 129, 417.

27 Barradas 1996.

28 Barradas 1996, 158–161.

'Aḥmad, Naṣraddin. Naṣraddin won the first confrontation, but in the second, Galāwdewos was victorious. We do not know anything more precise about the exact place where he stopped and where the battles took place.

'Ifāt was a historical region and sultanate in central Šawā which first flourished around 1285. An integral part of the Rift Valley, the area of 'Ifāt served for millennia as a point of contact between the pastoral or nomadic economic formations of the eastern lowlands and ancient agricultural societies of the Ethiopian highlands. Islam must have been introduced into the region in early times. The growth of the power of 'Ifāt came to a halt in the fourteenth century, when it was incorporated into the Christian kingdom. The strategic importance of 'Ifāt placed it against the Christian state consolidated by rulers of the Solomonic dynasty, whose ambition was to control the caravan trade to the Red Sea. The area of 'Ifāt was reduced to a battleground following the war of the Christian and Muslim in the sixteenth century and followed by the expansion of the Oromo people.

'Ifāt won important value in the study of historical geography both in the local works and traveller accounts. The *Chronicle of 'Amda Šayon* recounted the campaign of the king against the Sultanate of 'Ifāt. The entire chronicle is devoted to the feats of the warrior king against the Sultan of 'Ifāt, Sabraddin (ID: PRS8282sabraddi).<sup>29</sup> It also appears in the *Chronicle of Zar'a Yā'qob*, which reports that the king appointed a governor for the province of 'Ifāt, indicating that it was still part of the Christian kingdom in the fifteenth century:



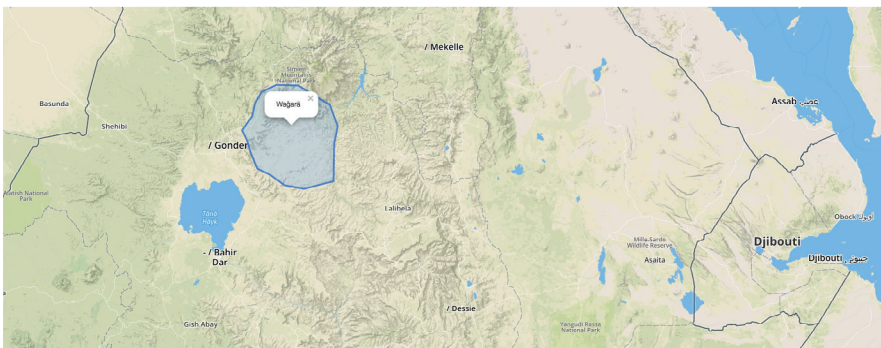
ወሰግሆን ፣ ለኩሎን ፣ አጋቲሆን ፣ ሜሞን ፣ ንጉሥ ፣ አቡሆን ፣ ውስተ ፣ ኩሎ ፣ ምድረ ፣ ኢትዮጵያ ፣ በታሕቴሁ ፣ [...] ወበኢፋት ፣ አሙተ ፣ ጊዮርጊስ ። ('He appointed all of their sisters like them, the King, their father, throughout the land of Ethiopia, under his rule, [...] in 'Ifat, 'Amata Giyorgis').<sup>30</sup>

### (3) *Wağarā*

Having decisively defeated Naşraddin in 'Ifāt in 1542, King Galāwdewos crossed the Blue Nile to Wağarā (ID: LOC6144Wagara), where he confronted *imām* 'Aḥmad himself, and killed him in 1543. Wağarā is the landmark for the victory of King Galāwdewos, which is a district north of Lake Tanā and south-west of Sāmen (ID: LOC5671Semen), a highland region, inhabited mostly by the 'Agaw (ID: ETH1083Agaw).

It was one of the provinces in which Ləbna Dəngəl sought refuge during the Muslim wars of the sixteenth century. After his victory over Ləbna Dəngəl in the mid-1530s, *imām* 'Aḥmad made Wağarā its stronghold, until he was in turn defeated by King Galāwdewos in 1543. Kings Minās (ID: PRS7102Minas) and Śarṣa Dəngəl (ID: PRS8550sardaDe) used Wağarā as a staging ground for repeated campaigns against the Beta 'Ēsrā'el (ID: ETH-1274Betaes) of Sāmen who in 1585 raided and pillaged Wağarā. In the seventeenth century, dissension and rebellion in Wağarā persisted during the reign of Susənyos. Wağarā was also one of the regions where the Portuguese Jesuits proselytized: the Jesuits claimed that there were about 100,000 Catholics in Wağarā in 1630. During the seventeenth and eighteenth centuries, Wağarā was a granary for Gondar. In the mid-twentieth century imperial administrative division, Wağarā was an *awrāggā* administrative unit within Bagemdər, composed of five sub-units (*waradā*).

All chronicles from the fourteenth, fifteenth and later sixteenth, seventeenth and eighteenth centuries mentioned Wağarā in connection to its religious and ethnic composition, agricultural fertility and vastness, which was







appears as one of the provinces of the Christian kingdom which was ruled by the governor Hegano:

ወእምድኅሬሆንሰ፡ አኅዝ፡ ከኩሎ፡ ሢመታተ፡ ኢትዮጵያ፡ ውስተ፡ እደዑ፡ ወሢመ፡ አደክ  
ሽነት፡ ውስተ፡ አህጉር፡ እንዝ፡ ይብል፡ በሸዋ፡ ራቅ፡ ማሰፊ፡ [...] ሄገኖ፡ ወበወጅሂ፡ (And  
then after, the King took over the entire hierarchies of Ethiopia in his hand and ap-  
pointed 'Adakṣənat in the provinces saying that: 'in Šawā rāq māsare [...] and also  
Hegano in Wağ').<sup>31</sup>

### (5) Faṭagār

In the nineteenth year of his reign (1559), King Galāwdewos led a campaign from his royal seat at Wağ in a place called Faṭagār (ID: LOC3061Fataga) to confront the invasion of the region by the Muslim state of 'Adal. He was killed, and his head was cut off and taken as a trophy. Faṭagār was a large historical region (possibly once a Muslim sultanate). It was located in the south-east of Šawā on the northern shore of the 'Awāš river. In the fourteenth and fifteenth centuries, it was gradually integrated into the Christian empire. The region held a strategic position between the Christian Ethiopia and the southern Muslim sultanate. In the beginning of the sixteenth century, Faṭagār came back under the Muslims of 'Adal.

The name Faṭagār appears for the first time in the *Chronicle of 'Amda Šəyon* (1314–1344) where the Muslim ruler claimed to appoint his own representatives in the region of Faṭagār.



31 Perruchon 1893, 14–15.



ወዘንተ ፡ ብሂሎ ፡ ተንሥኦ ፡ ወሐረ ፡ ወበጽሐ ፡ ኅበ ፡ ምድረ ፡ ክርስቲያን ፡ ። ወአውዐየ ፡ ቤ ተ ፡ ክርስቲያን ፡ ወቀተሎሙ ፡ ለክርስቲያን ፡ ወእለ ፡ ተርፋ ፡ ዪወወ ፡ ዕደ ፡ ወአንስተ ፡ ወአ ማብአሙ ፡ ውስተ ፡ ሕጉ ፡ ወእምድኅረዝ ፡ ይቤ ፡ እሠይም ፡ መኳንንተ ፡ ላዕለ ፡ በሓውርተ ፡ ኢትዮጵያ ፡ በከመ ፡ ይሠይም ፡ ንጉሠ ፡ ጽዮን ፡ [...] ፈጠጋር ፡ [...] ('He said so and he departed, he arrived in the land of the Christians, burned down the churches, killed the christians and the survivors he made them prisoners, men and women, and he converted them to his religion. Then he said: 'I will appoint governors on all regions of Ethiopia, as the king of Zion' and he appointed governors: [...] one on Fatagār [...]').<sup>32</sup>

It also appears in the chronicle of the fifteenth-century King Zar'a Yā'qob: the king appointed his own representative there, making it a stronghold of the Christian empire, until it came again under Muslim occupation in the sixteenth century.<sup>33</sup> The Arabic chronicle of *imām* 'Aḥmad's campaigns, *Futūḥ al-Habaša*, written by a contemporary chronicler, mentioned it regarding 'Aḥmad's victories over the Christians during the third decade of the sixteenth century.<sup>34</sup>

### 5. Producing annotations

From the data encoded in TEI, we can already do a lot of computation as we will show in the visualization examples below, but we also use that data to produce geoJSON, KML and RDF to serve even more visualizations and reuses.

There are two kinds of RDF triples produced from the two kinds of information encoded. From all place name attestations, we export annotations in the Pelagios format:<sup>35</sup>

```
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix pelagios: <http://pelagios.github.io/vocab/terms#> .
<http://betamasaheft.eu/api/placeNames/works/LIT1252Confes>
  a pelagios:AnnotatedThing ;
  dcterms:title "Confessio Claudii";
  dcterms:description "A theological treatise believed to be a response of king Galāwdewos to the Jesuit missionaries headed by Oviado.";
  dcterms:source <http://betamasaheft.eu/tei/LIT1252Confes.xml>;
  foaf:homepage<http://betamasaheft.eu/works/LIT1252Confes/main> ;
  dcterms:language "en";
```

32 Marrassini 1993, 52.

33 Perruchon 1893, 15, 30, 47, 67, 71, 93, 112, 137.

34 Šihāb ad-Dīn 2003, 19, 49, 60-61.

35 <<https://github.com/pelagios/pelagios-cookbook/wiki>>, last access 22 May 2018; see Simon et al. 2014.

```

<http://betamasaheft.eu/api/placeNames/works/LIT1252Confes/
annotations/1>
  a oa:Annotation ;
  oa:hasTarget <http://betamasaheft.eu/api/placeNames/works/
LIT1252Confes> ;
  oa:hasBody
    <http://betamasaheft.eu/places/LOC2488Damot> ;
  oa:annotatedAt
    "2018-05-15T16:07:09.022+02:00"^^xsd:date ;

```

Each of the places in the Gazetteer is represented according to the Pelagios Interconnection Format,<sup>36</sup> which is also interoperable with the Syriaca.org places model:<sup>37</sup>

```

@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix lawd: <http://lawd.info/ontology/> .
@prefix pelagios: <http://pelagios.github.io/vocab/terms#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .

<http://betamasaheft.eu/places/LOC2488Damot>
  a lawd:Place ;
  rdfs:label "Dāmōt"@gez;
  dcterms:source <http://betamasaheft.eu/LOC2488Damot.xml> ;
  dcterms:temporal <http://n2t.net/ark:/99152/p03tcsvm7f>;
  lawd:hasName [ lawd:primaryForm "ዳሞት"@"gez ] ;
  lawd:hasName [ lawd:variantForm "Dāmōt"@"gez ] ;
  foaf:primaryTopicOf <http://betamasaheft.eu/places/LOC2488Damot/
main> ;
  dcterms:isPartOf <http://betamasaheft.eu/places/LOC3010Ethiop>;

```

The data is thus shared with Pelagios and becomes available via the Pelagios API, but since the model of the data is the same, direct federated queries can also be run, for example to the Syriaca.org and the *Beta maṣāḥāft* triples stores starting from any SPARQL Endpoint.

## 6. Visualization of the data

Directly from the TEI encoding we can print a summary of all the files containing a reference to a specific entity and co-occurring entities grouped by type, see fig. 1.

In the same way we can directly extract and display with Google charts a graph of the actual attestations of a toponym, see fig. 2.

36 <<https://github.com/pelagios/pelagios-cookbook/wiki/Pelagios-Gazetteer-Interconnection-Format>>, last access 28 May 2018.

37 <<https://github.com/HeardLibrary/semantic-web/blob/master/sparql/syriaca.md>>, last access 22 May 2018; see Michelson 2016.

Paris, Bibliothèque nationale de France, BnF Éthiopien 35	1) pointer only (1550 - 1600), null[placeName]	In the same context of this attestation also the following persons occur	<table border="1"> <thead> <tr> <th>name</th> <th>type</th> </tr> </thead> <tbody> <tr> <td><a href="#">Zašalläse</a></td> <td></td> </tr> </tbody> </table>	name	type	<a href="#">Zašalläse</a>							
name	type												
<a href="#">Zašalläse</a>													
		In the same context of this attestation also the following places occur	<table border="1"> <thead> <tr> <th>name</th> <th>type</th> </tr> </thead> <tbody> <tr> <td><a href="#">Jerusalem</a></td> <td></td> </tr> </tbody> </table>	name	type	<a href="#">Jerusalem</a>							
name	type												
<a href="#">Jerusalem</a>													
	2) ጳጳሳት (1550 - 1600), null[placeName]	In the same context of this attestation also the following persons occur	<table border="1"> <thead> <tr> <th>name</th> <th>type</th> </tr> </thead> <tbody> <tr> <td><a href="#">ዘሥላሴ:</a></td> <td></td> </tr> </tbody> </table>	name	type	<a href="#">ዘሥላሴ:</a>							
name	type												
<a href="#">ዘሥላሴ:</a>													
		In the same context of this attestation also the following places occur	<table border="1"> <thead> <tr> <th>name</th> <th>type</th> </tr> </thead> <tbody> <tr> <td><a href="#">እየሩሳሌም</a></td> <td></td> </tr> <tr> <td><a href="#">እየሩሳሌም</a></td> <td></td> </tr> <tr> <td><a href="#">ሀርተ: ዘዩሳ:</a></td> <td></td> </tr> <tr> <td><a href="#">ደብረ: ቱስላም:</a></td> <td></td> </tr> </tbody> </table>	name	type	<a href="#">እየሩሳሌም</a>		<a href="#">እየሩሳሌም</a>		<a href="#">ሀርተ: ዘዩሳ:</a>		<a href="#">ደብረ: ቱስላም:</a>	
name	type												
<a href="#">እየሩሳሌም</a>													
<a href="#">እየሩሳሌም</a>													
<a href="#">ሀርተ: ዘዩሳ:</a>													
<a href="#">ደብረ: ቱስላም:</a>													
	3) ጳጳሳት (no date), null[placeName]	In the same context of this attestation also the following persons occur	<table border="1"> <thead> <tr> <th>name</th> <th>type</th> </tr> </thead> <tbody> <tr> <td><a href="#">ሐሪ ድጋግል</a></td> <td></td> </tr> <tr> <td><a href="#">ሐሪ: ድጋግል</a></td> <td></td> </tr> <tr> <td><a href="#">ዘርባ: ጦርዳም</a></td> <td></td> </tr> </tbody> </table>	name	type	<a href="#">ሐሪ ድጋግል</a>		<a href="#">ሐሪ: ድጋግል</a>		<a href="#">ዘርባ: ጦርዳም</a>			
name	type												
<a href="#">ሐሪ ድጋግል</a>													
<a href="#">ሐሪ: ድጋግል</a>													
<a href="#">ዘርባ: ጦርዳም</a>													
		In the same context of this attestation also the following places occur	<table border="1"> <thead> <tr> <th>name</th> <th>type</th> </tr> </thead> <tbody> <tr> <td><a href="#">እየሩሳሌም</a></td> <td></td> </tr> </tbody> </table>	name	type	<a href="#">እየሩሳሌም</a>							
name	type												
<a href="#">እየሩሳሌም</a>													
Gwägwäta	1) pointer only (no date), gntoCatedIn[relation]												
Confessio Claudii	1) Damot (no date), null[placeName]												
	2) ጳጳሳት (no date), null[placeName]	In the same context of this attestation also the following persons occur	<table border="1"> <thead> <tr> <th>name</th> <th>type</th> </tr> </thead> <tbody> <tr> <td><a href="#">እየሱስ: ክርስቶስ:</a></td> <td></td> </tr> </tbody> </table>	name	type	<a href="#">እየሱስ: ክርስቶስ:</a>							
name	type												
<a href="#">እየሱስ: ክርስቶስ:</a>													

Fig. 1 Attestations of a place viewed in *Beta mašāḥaft*..

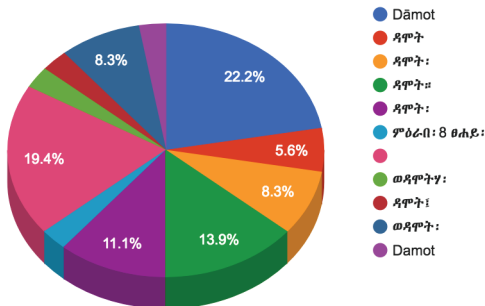


Fig. 2 Distribution of attested place name's forms.

The TEI data converted to geoJSON provides a very practical way to print a map using Leaflet like our homepage map, see fig. 3.

From the TEI we produce also KML which can be visualized with the DARIAH-DE geobrowser and makes use of both the space and time information to allow browsing both dimensions of the data available at the same time, see fig. 4.



Fig. 3 Map of manuscripts repositories based on XML data converted to geoJSON.

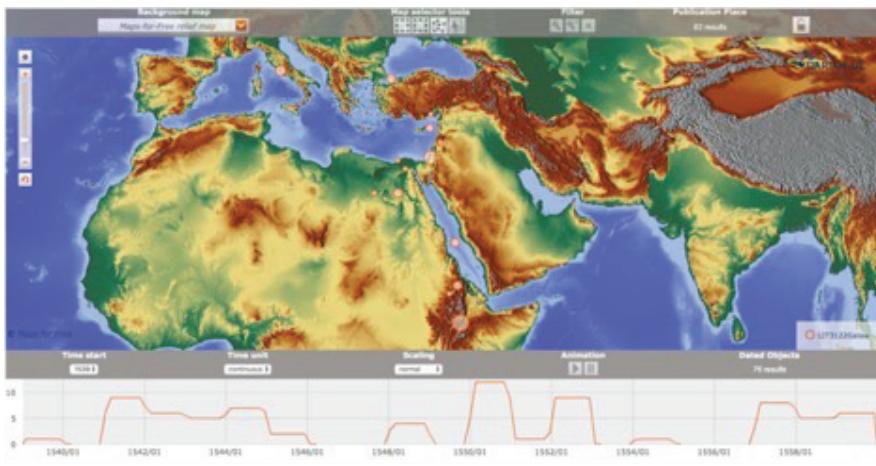


Fig. 4 Dariah-de Geo Browser visualisation of KML data.

From the annotations exposed in RDF, users and other interested parties can build visualizations of several kinds. We describe here only the application of this data as it is published in Pelagios, where it allows one to interrogate the latest dump together with the dumps of all other projects joining in the effort. In Pelagios, the user can search and navigate the available information from the map. One can search for a word, as the text is indexed and see related results.

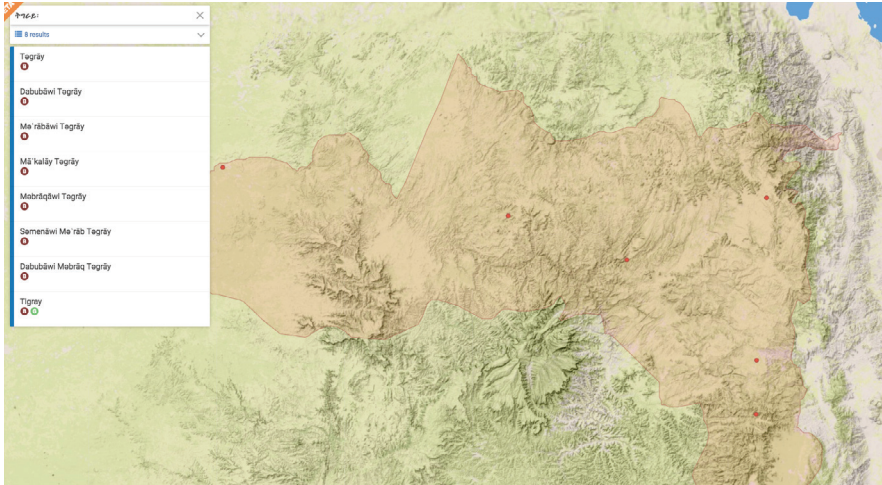


Fig. 5 Visualization of a place from the *Beta masāḥaft* Gazetteer in Peripleo.

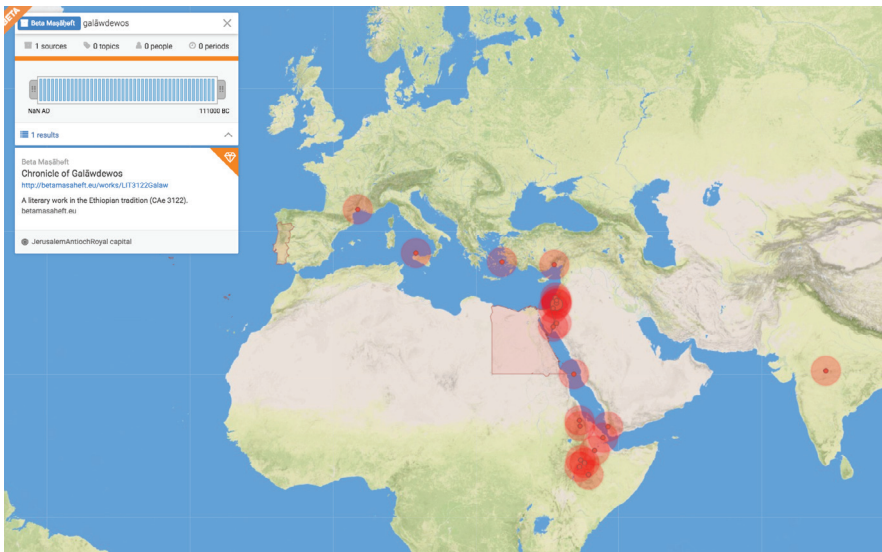


Fig. 6 Places in the *Chronicle of Galāwdewos* viewed in Peripleo.



In fig. 5 we searched a word from the example above and found many places which contain it in the description but since there is an annotation associating this place with the polygon in geonames we can also see that.

We can also search ‘Galāwdewos’ and find the *Chronicle* as annotated resource, as in fig. 6.

Clicking on it, one can see all the annotated places in a map. Clicking on one of these places one can then see how many items link there and for example find out, navigating the triples in this way, that there are also two manuscripts which have been annotated with a reference to that place (fig. 7).

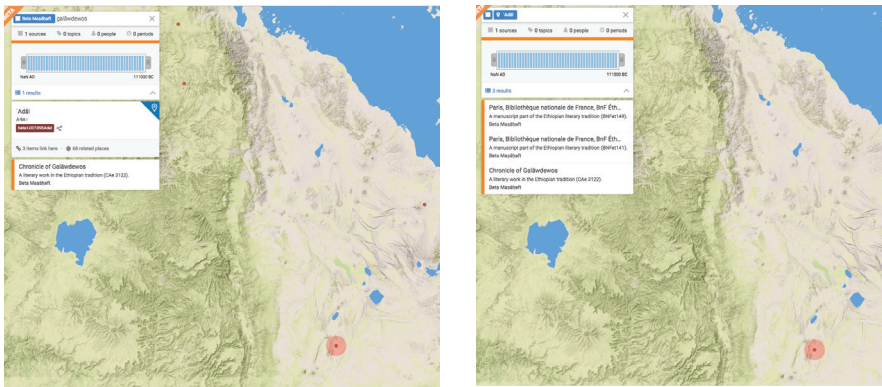


Fig. 7 Navigation of the annotations in Peripleo.

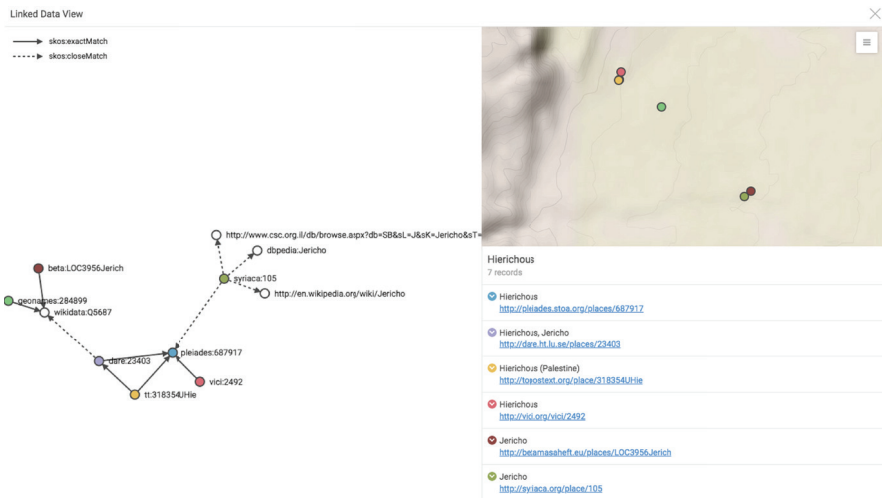


Fig. 8 Peripleo visualization of the relations between place identifiers.

But that is not all. We can filter the annotations by period and see all places in the gazetteer associated to a specific period, or we can see the network of IDs generated from the data of all providers as in the following example. In this way, anyone is able to benefit from the sum of the existing annotations, thus sparing a lot of repetitive authority check work in most cases, and actually having it done much better as a network of declarative associations rather than a flat equivalence (fig. 8).

This last example already shows how data which was provided without any connection to Syriaca.org is joined in Pelagios, and by providing one hook, in this case the equivalence with the Wikidata entity, a whole set of other identifiers becomes accessible, with the resources pointing to them.

While many other projects produce triples and annotations about places, the ways in which to exploit this rich information for comparative and comprehensive studies remains to be explored and is research-question-specific: it is up to the researcher to imagine new questions and produce meaningful queries to support statements which will build towards their answers.

## References

- Barradas, M. 1996. *Tractatus Tres Historico-Geographici (1634): a Seventeenth Century Historical Account of Tigray, Ethiopia*, ed. R. Pankhurst, tr. E. Filleul, Aethiopistische Forschungen, 43 (Wiesbaden: Harrassowitz, 1996).
- Bausi, A. 2015. ‘TraCES: From Translation to Creation: Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages’, in A. Bausi, A. Gori, D. Nosnitsin, and E. Sokolinski, eds, *Essays in Ethiopian Manuscript Studies. Proceedings of the International Conference Manuscripts and Texts, Languages and Contexts: the Transmission of Knowledge in the Horn of Africa, Hamburg, 17–19 July 2014*, Supplement to Aethiopica, 4 (Wiesbaden: Harrassowitz Verlag, 2015), 11–13.
- Carlson, T.A. and D.A. Michelson, eds, forthcoming. *The Syriac Gazetteer*, Syriaca.org: The Syriac Reference Portal (forthcoming).
- Consociazione turistica italiana 1938. *Guida dell’Africa Orientale Italiana* (Milano: Consociazione turistica italiana, 1938 (XVI)).
- EAE. S. Uhlig, ed., *Encyclopaedia Aethiopica*, I: A–C; II: D–Ha; III: Ha–N; in cooperation with A. Bausi, ed., IV: O–X; A. Bausi in cooperation with S. Uhlig, ed., V: Y–Z. *Supplementa. Addenda et Corrigenda. Maps. Index* (Wiesbaden: Harrassowitz Verlag, 2003, 2005, 2007, 2010, 2014).
- Guidi, I. 1895. ‘Il ‘Gadla ’Aragāwi’’, *Memorie della R. Accademia nazionale dei Lincei, Classe di scienze morali, storiche e filologiche*, ser. 5<sup>a</sup>, 2 (1895), 54–96.
- Huntingford, G.W.B. 1989. *The Historical Geography of Ethiopia from the First Century A.D. to 1704*, ed. R. Pankhurst, *Fontes historiae Africanae*, Series varia, 4 (Oxford: Oxford University Press for the British Academy, 1989).

- Irvine, A.K. 1991. Review of G. W. B. Huntingford, *The Historical Geography of Ethiopia from the First Century A.D. to 1704*, *Bulletin of the School of Oriental and African Studies*, 54/2 (1991), 372–373, <doi:10.1017/S0041977X00014968>.
- Kur, S., ed., tr., 1965. *Actes de Iyasus Mo'a, abbé du Convent de St-Etienne de Hayq*, ed., tr. S. Kur, *Corpus Scriptorum Christianorum Orientalium*, 259, 260, *Scriptores Aethiopici*, 49, 50 (Louvain: Secrétariat du CorpusSCO, 1965).
- Liuzzo, P.M. 2017. 'Encoding and Representation Challenges of the Project *Beta maṣāḥāft: Manuscripts of Ethiopia and Eritrea*', in *Proceedings of Balisage: The Markup Conference 2017, Washington, DC, August 1-4, 2017*, Balisage Series on Markup Technologies, 19 (Washington, D.C., 2017).
- and Solomon Gebreyes Beyene 2017a. 'Place Names in the Chronicle of King Galāwdewos: a Prototype Geo-annotated Text of the Ethiopic Tradition | Pelagios Commons' (2017).
- 2017b. 'Report – Ancient Places in the Chronicle of King Galāwdewos' (2017).
- 2017c. 'Final Report: Some Interesting Historical Places in the Chronicle of Galāwdewos | Pelagios Commons' (2017).
- Marrassini, P. 1993. *Lo scettro e la croce. La campagna di 'Amda Šeyon I contro l'Ifāt (1332)*, ed., tr. P. Marrassini, *Studi Africanistici, Serie Etiopica*, 4 (Napoli: Istituto Universitario Orientale, 1993).
- Michelson, D.A. 2016. *Syriaca.org Developing Linked Open Dataset | Project News from Syriaca.org | Vanderbilt University* (2016) <<http://syriaca.org/blog/2016/04/4192016-syriaca-org-developing-linked-open-dataset/>>, accessed 3 April 2018.
- Perruchon, J. 1893. *Les chroniques de Zar'a Yā'eqōb et de Ba'eda Māryām, rois d'Éthiopie de 1434 à 1478 (Texte éthiopien et Traduction)*, ed., tr. J. Perruchon, *Bibliothèque de l'École pratique des hautes études, Section sciences historiques et philologiques*, 93 (Paris: Émile Bouillon, 1893).
- Reule, D. and Solomon Gebreyes Beyene 2017. 'Chronicle of Zar'a Yā'qob', *Beta maṣāḥāft: Manuscripts of Ethiopia and Eritrea, Clavis of Ethiopian Literature* (2017) <<http://betamasaheft.eu/LIT4646Chronicle/>>, accessed 15 May 2018.
- Šihāb ad-Dīn Aḥmad bin 'Abd al-Qāder bin Sālem bin Uṭmān ('Arab Faqīh) 2003. *Futūḥ al Ḥabaša, The Conquest of Abyssinia*, tr. P.L. Stenhouse and R. Pankhurst (Los Angeles, CA: Tsehai Publishers, 2003).
- Simon, R., E. Barker, P. de Soto, and L. Isaksen 2014. 'Pelagios', *ISAW Papers*, 7 (2014).
- Solomon Gebreyes Beyene 2017. 'Chronicle of Galāwdewos', *Beta maṣāḥāft: Manuscripts of Ethiopia and Eritrea, Clavis of Ethiopian Literature* (2017) <<http://betamasaheft.eu/LIT3122Galaw/>>, accessed 15 May 2018.



- 2018. ‘Confessio Claudii’, *Beta maṣāḥəft: Manuscripts of Ethiopia and Eritrea, Clavis of Ethiopian Literature* (2018) <<http://betamasaheft.eu/LIT1252Confes/>>, accessed 15 May 2018.
- and D. Reule 2017. ‘Chronicle of Susənyos’, *Beta maṣāḥəft: Manuscripts of Ethiopia and Eritrea, Clavis of Ethiopian Literature* (2017) <<http://betamasaheft.eu/LIT3951ChronSusenyos/>>, accessed 15 May 2018.
- Páez 2011. *Pedro Páez’s History of Ethiopia, 1622*, eds I. Boavida, H. Penneç, and M. J. Ramos, tr. C. J. Tribe, I, Hakluyt Society, Series III, 23 (London: Ashgate–The Hakluyt Society, 2011).
- Pisani, V., E. Sokolinski, and D. Reule 2018. ‘Tārika ‘Amda Şəyon’, *Beta maṣāḥəft: Manuscripts of Ethiopia and Eritrea, Clavis of Ethiopian Literature* (2018) <<http://betamasaheft.eu/LIT4275ChronAmdS/>>, accessed 15 May 2018.
- TEI consortium 2016. *The TEI Guidelines. Documentation Elements* (2016) <<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/TD.html>>, accessed 24 May 2017.



# Encoding Strategies and the Ethiopic Literary Heritage: The *Physiologus* as a Case Study

Massimo Villa, Universität Hamburg

Producing the Clavis of Ethiopic literature while making the data computer readable for extensive indexing and research purposes is among the aspirations of the project *Beta maṣāḥəft*. Here, we illustrate the challenges faced with and the solutions offered by the project on the example of the *Physiologus*, a literary work, translated into Ethiopic from Greek during the Late Antiquity.

*Beta maṣāḥəft* is a manuscript-centred research environment, and in this initial phase is predominantly focused on codicology. This notwithstanding, one of the project's main spheres of interest is obviously the Ethiopic textual heritage.<sup>1</sup>

Working with texts in a TEI/XML hierarchical structure one faces manifold key issues. A basic terminology-related question which has hitherto received very little scholarly attention in the field of Ethiopian studies is: 'what is a work?'<sup>2</sup> The solution proposed by the team is that, for the purposes of the project, we consider a work any text with an independent circulation.<sup>3</sup> According to this principle we create a new XML-based file for each work. Any work file is associated with a univocal fist-level ID, which is a string consisting of three elements: a fixed sequence LIT that identifies the type of entity, a progressive numerical sequence which is arbitrarily assigned, and an alphabetical part (generally a six-letter sequence) added to help recognize the work. The work ID of the *Wəddāse Māryām* or 'Praise of Mary' is, for instance, LIT2509Weddas. On the contrary, texts with no independent circulation are not considered as works. Being systematically found as sub-units of other works. They can be given their own structure and IDs (e.g. LIT2509Weddas#Monday), and be referred to by pointing to the `@xml:id` anchor inside the file. Any textpart is 'upgraded' to a work as soon as it is found independently.

- 1 *Beta Maṣāḥəft: Die Schriftkultur des christlichen Äthiopiens und Eritreas: Eine multimediale Forschungsumgebung* is a long-term project headed by Prof. Alessandro Bausi, funded within the Academies' Programme, and coordinated by the Union of the German Academies of Sciences and Humanities, under survey of the Akademie der Wissenschaften in Hamburg (2016–2040). The structures and the strategies of encoding illustrated in the present paper are the result of joint efforts by the entire project team (see <https://www.betamasaheft.uni-hamburg.de/en/team/projectteam.html>). See also the project note by Dorothea Reule in this issue.
- 2 See also the contribution by Tito Orlandi in this issue.
- 3 See project guidelines at <http://betamasaheft.eu/Guidelines/?id=definitionWorks>.

To attribute a given text its own appropriate status, i.e. to state whether it is a work or a textpart, has consequences at various levels. First, it encourages researchers to monitor the ways of circulation of the text and to find out systematic correlations among other texts. Secondly, it is interlinked with the creation of a full repertory of Ethiopic texts, i.e. a *Clavis*, which is one of the declared goals of the project. Any *Clavis* requires a numbering; the *Clavis Aethiopica* (CAe) numbering stems from the numerical sequence of the record ID (see above). Consequently, fulfilling the relevant requirements, particularly that of independent circulation, is crucial to the choice to include the text into the *Clavis*.

The creation of work records is currently underway for the following types of textual items: i. texts with an independent circulation; ii. compilations (multiple-text *corpora*) and parts of compilations which can circulate independently at various stages of granularity;<sup>4</sup> iii. each recension of a multiple-recension text plus an extra record for the general abstract work. The record for the general work does not represent the *Urtext* or the genetic ancestor of the textual tradition, yet an ‘architext’ in which the distinctive features of the individual recensions are ideally neutralized, and to which witnesses not properly specified in the existing catalogues are assigned. The relations between the general work and the individual recensions are expressed by means of relation elements (see below).

A case study which proves to be particularly representative of the challenges met by the project team in a multiple-recension tradition is the textual tradition of the Ethiopic *Physiologus*.

4 These include e.g. the hagiographical collection *Gadla Samā’ tāt* (‘Acts of Martyrs’) and the single Acts which are traditionally transmitted within it; the cycle of Marian miracles *Ta’ amməra Māryām* and the individual miracles; a stable hagiographical dossier of a saint and its parts (*Life, Miracles, History* of the translation of his body, etc.). On the *Gadla Samā’ tāt* as a ‘corpus organizer’ see Bausi 2010. Manuscript evidence shows that multiple-text corpora such as the *Acts of the Martyrs* and the *Miracles of Mary* are considered as single units in the local literary and scribal culture. Nonetheless their textual components (the single acts or miracles) can be missing, arranged in a different order, or transmitted separately outside the main corpus. The project records such a documented fluidity in the transmission by treating both the corpus and the single components as independent works. The historical process which led to the development of some multiple-text corpora still lies beyond our knowledge: for instance, a collection such as the *Acts of the Martyrs* appears as one single compilation, with its own traditional label and independent circulation, since the earliest available documentation (thirteenth to fourteenth century). Other compilations, or at least their earliest core, were received as translated works, such as the *Miracles of Mary* and the *Sənkəssār* (‘Synaxarion’), and further enriched with additional original pieces.

**Case study**

The *Physiologus* is a popular literary work, composed in Greek in Alexandria and translated into several Christian oriental languages. The Ethiopic version was produced during the Aksumite Age (fifth–sixth centuries). It counts a variable number of chapters, each providing a legendary description of the natural properties of a species of animal, plant, or mineral, and the explanation of these properties in a moralizing Christian context.

The Ethiopic tradition is somewhat multifaceted (fig. 1). It consists of three recensions: a ‘Homily of the blessed Physiologus’ (*Dərsān za-bəṣu’ Fisālgos*, in short *Phys. α*), a ‘History of the similitudes of the wise Physiologus’ (*Zenā məssāleyāt za-ṭabib Fisə’algos*, in short *Phys. β*), and a ‘History of the wisdom of the wise Physiologus’ (*Zenā ṭəbab za-Fisə’algos ṭabib*, in short *Phys. γ*).<sup>5</sup> Such a scenario entails the creation of four records, one for the general work (LIT1401Physio), and one for each recension: LIT4915PhysA for *Phys. α* (CAe 4915), LIT4916PhysB for *Phys. β* (CAe 4916), and LIT4917PhysC for *Phys. γ* (CAe 4917). The three recensions are disambiguated by the numerical string of their IDs, which is assigned arbitrarily. A printed edition is also extant: it was published by Fritz Hommel in 1877 on the basis of three manuscripts belonging to *Phys. α*. The texts of *Phys. β* and *Phys. γ* are still unpublished.

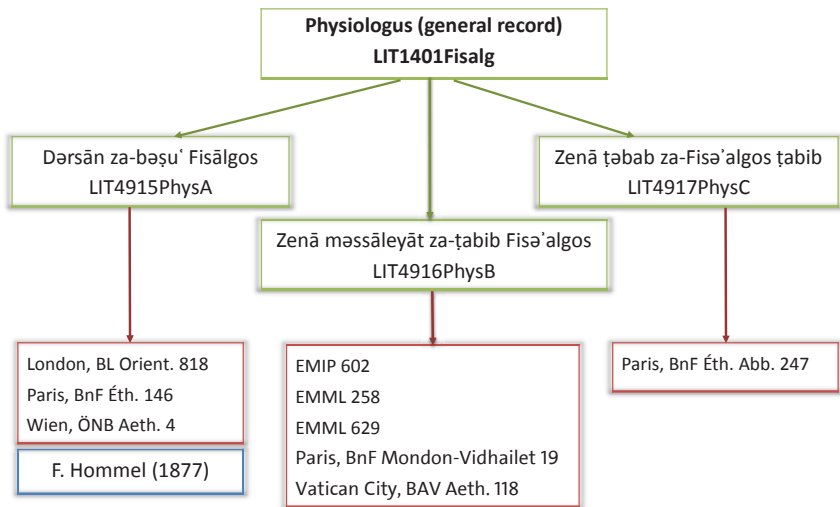


Fig. 1. Diagram illustrating the textual tradition of the Ethiopic *Physiologus*.

5 A detailed picture of the manuscript and textual tradition of the Ethiopic *Physiologus* will be presented in forthcoming publications by the present writer.

Work records are structured through a variety of metadata, in which the relevant pieces of information are properly entered. Some of these components are illustrated hereby.

**Title(s).** The main title of the work is given by a threefold set of elements, which contains different ways of identifying the work: in Ethiopic script (e.g. ‘ድርሳን ፡ ዘብጹዕ ፡ ፊሳልጎስ ፡’ for LIT4915PhysA), in transliteration (‘Dərsān za-bəṣu ‘ Fisālgos’), and in English translation (‘Homily of the blessed Physiologus’). This basic set can be supplemented with additional tags for alternative variant titles, short titles (e.g. ‘Phys. α’), and parallel titles in other traditions.

**Text witnesses.** The list of manuscripts which are actually used in the text edition is explicitly given, while a *recensio* of all encoded witnesses is visible online since the app can retrieve and display all manuscripts whose contents point to the ID of the work in question.<sup>6</sup> The manuscript IDs can be further associated with an `@xml:id` that shortly identifies the witness used in the text edition, and can be identical to the siglum used in the printed edition. For LIT4915PhysA, they are L (= London, BL Orient 818), P (= Paris, BnF Éth. 146), and W (= Wien, ÖNB Aeth. 4).

**Claves.** In addition to the generated CAe number, existing repertoires of claves have also been incorporated.<sup>7</sup> Individual works can be therefore searched for their number in one of the Claves through an appropriate filtered search.

**Keywords.** Keywords help categorize the literary corpus under different typologies. They specify the genre (Bible, Liturgy, Hagiography, etc.) or the age (Aksumite, Gondarine, etc.) of a given work. Keywords allow users to filter their search and automatically extrapolate all works belonging to a certain age or genre. An *ad-hoc* query on the Aksumite texts, for instance, generates a constantly updated repertory of the earlier textual heritage of Ethiopic literature.

**Relations.** Relation elements contain our formal description of the relationships between different entities, in the present case abstract works.<sup>8</sup> Each relation expresses a statement with a subject, a predicate, and an object (technically, a ‘triple’). Relations can or cannot involve items belonging to the same type of entity: in our case we use them to define a relationship between two abstract

6 See guidelines at <<http://betamasaheft.eu/Guidelines/?id=work-teiHeader>>.

7 Among these, the *Clavis Apocryphorum Veteris Testamenti* (CAVT), the *Clavis Apocryphorum Novi Testamenti* (CANT), the *Bibliotheca Hagiographica Orientalis* (BHO), the *Bibliotheca Hagiographica Graeca* (BHG), the Kinefe-Rigb Zelleke’s repertory of hagiographical texts (KRZ), and others.

8 See <<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-relation.html>>.

works. For multiple-recension texts a relation named `saws:isVersionOf` is used.<sup>9</sup> Therefore, a statement like

```
<relation active="LIT4915PhysA"
  name="saws:isVersionOf" passive="LIT1401Physio"/>
```

is a formal way to claim that *Phys. α* is a version of the *Physiologus* as abstract notion of the work. Other relationships among works can also be stated: for multiple-text corpora a relation `saws:formsPartOf` is employed, and for parallel versions in other languages `saws:isVersionInAnotherLanguageOf`.<sup>10</sup> A diagram generated in the website displays all the relations entertained by a given record.

**Text editions.** Being a multiple-recension work, the *Physiologus* is not represented by a unitary text. Each recension is supplied with its own edition or, if existing, more than one. Since any TEI/XML work record points to the abstract notion of a given work, it can host as many text editions as needed. *Phys. α* hosts two editions, the *editio princeps* by Fritz Hommel and a new improved edition accommodating different text-critical choices. Each of them is identified by its own local `@xml:id`. A further challenge has been addressed by the modularity of the work. The *Physiologus* is subdivided into chapters, each obviously identified by a univocal value (e.g. ‘ed2ch3’). However, this reference system, focused on the chapter arrangement and not content-oriented, does not take into consideration that the chapter numbering is variable in the three recensions (*Phys. α* counts 49 chapters, *Phys. β* 57, and *Phys. γ* 58), and a certain subject is attributed different chapter numbers in the distinct recensions. As an example, the chapter on the caladrius (ገራድዮን ፣ *karādyon*, a legendary bird with diagnostic and medical powers) is numbered as 3 in *Phys. α* and as 5 in the remaining recensions. Consequently, a narrative unit with its own ID ‘NAR0014caladrius’ is needed to specify one and the same subject regardless of its numbering.<sup>11</sup>

Once transcribed, the ‘plain text’ can be annotated with tags providing paratextual, text-critical and content-related information. Various scribal phe-

9 Property names are inherited from the ontology developed by the *Sharing Ancient Wisdoms* project; see <<http://www.ancientwisdoms.ac.uk/>>. See also Tupman and Jordanous 2014.

10 Many other relations are used in the *Beta maṣāḥaft* data. They are beyond the scope of this contribution and can be consulted at <<http://betamasaheft.eu/Guidelines/?id=relation>>.

11 The notion of ‘narrative unit’ is indebted to, and further develops, the terminology proposed by Tito Orlandi (2013, 93) for Coptic literary documentation. It is adopted in the project to refer to text portions (paragraphs, chapters, miracles performed by a saint, stanzas, etc.) sharing the same narrative content, no matter if extant in different versions, in multiple recensions of the same work, or even in different works; see <<http://betamasaheft.eu/Guidelines/?id=narrativeUnits>>.



nomena can be recorded, e.g. erasures, omissions, and marginal or interlineal additions. The critical apparatus makes use of the `<xm1:ids>` previously associated with the witnesses. Finally, quotations from the Scriptures and other texts can be marked-up and referenced to the precise verse or line of the source text.

The digital edition is visualized on the website in a three-column structure: the left column outlines the paragraph-based arrangement of the text, the central column displays the text, the right column the critical apparatus. Quotations are clickable and open a pop-up window displaying the source text. Once data from the multiple versions have been entered and properly marked-up, it is desirable to have them displayed together for later comparison. This can be done by enabling a specific function which detects all the existing parallel versions, extrapolates from each the exact portion of text identified by the same narrative unit, and visualizes all relevant portions together (fig. 2). It is important to highlight that the functioning of such a tool arises from a precise encoding strategy, which combines the presence of a relation `saws:isVersionOf` and the reference to a narrative unit. Parallel-version out-

The screenshot displays the digital edition of chapter 4 of the Ethiopic Physiologus. It features a three-column layout:

- Left Column:** Edition metadata including locations like London, Paris, and Vienna, and a 'Compare' button.
- Central Column:** The main text in Amharic script, organized into four sentences. Each sentence has a 'Ugari Alignment' button and a 'Quotations' button. The text includes phrases like 'ጌርካ፡፡ በእንተ፡ ሾሮ፡ ዘበጠ፡ ጸልቃ፡ ዘወጸኦ፡ ግራብ፡'.
- Right Column:** Bibliography sections for 'Editions Bibliography', 'Translation Bibliography', and 'Secondary Bibliography', listing various academic works and their page numbers.

Below the main text, there are two highlighted boxes providing additional context:

- Version Zenā ṭəḥab za-Fis'alǧos ṭəḥib (LIT4917PhysC):** Edition: Paris, Bibliothèque nationale de France, BnF Éthiopian d'Abbadie 247. This version includes a detailed summary of the chapter's content in Amharic.
- Version Zenā massälləyāt za-ṭəḥib Fis'alǧos (LIT4916PhysB):** Edition: Massimo Villa. This version also includes a summary of the chapter's content.

Fig. 2. Parallel versions of chapter 4 of the Ethiopic Physiologus as visualized on the <<http://betamasaheft.eu/>> portal.

puts are particularly desirable to keep track of the discrepancies and the similarities in multiple-recension works (e.g. some *Lives* of most venerated saints) or in works preserved in different oriental traditions.

Finally, text is also linked to the online version of Dillmann's *Lexicon linguae Aethiopicae*. Any individual word is clickable and redirects to the corresponding entry in the *Lexicon*, further complemented by a list of all occurrences of the same lemma in the *Beta maṣāḥaft* corpus.<sup>12</sup> Many opportunities of investigation arise in this way. Search potentialities can be enhanced maximizing the benefits of the data pool at disposal, and reducing the time needed to put them into correlation.

### References

- Bausi, A. 2010. 'A Case for Multiple Text Manuscripts being 'Corpus-Organizers'', *Manuscript Cultures Newsletter*, 2 (2010), 34–36.
- 2016. 'Introduction. 150 Years After Dillmann's Lexicon', in A. Bausi and E. Sokolinski, eds, *150 Years after Dillmann's Lexicon: Perspectives and Challenges of Ga'az Studies*, Supplement to *Aethiopia*, 5 (Wiesbaden: Harrassowitz Verlag, 2016), 3–10.
- Dillmann, C.F.A. 1865. *Lexicon linguae Aethiopicae, cum indice Latino. Adiectum est vocabularium tigre dialecti septentrionalis compilatum a W. Munziger* (Lipsiae: T.O. Weigel, 1865).
- Hommel, F. 1877. *Die aethiopische Uebersetzung des Physiologus nach je einer Londoner, Pariser und Wiener Handschrift* (Leipzig: J. C. Hinrichs'sche Buchhandlung, 1877).
- Orlandi, T. 2013. 'A Terminology for the Identification of Coptic Literary Documents', *Journal of Coptic Studies*, 13 (2013), 87–94.
- Tupman, C. and A. Jordanous 2014. 'Sharing Ancient Wisdoms across the Semantic Web Using TEI and Ontologies', in T. L. Andrews and C. Macé, eds, *Analysis of Ancient and Medieval Texts and Manuscripts: Digital Approaches*, *Lectio: Studies in the Transmission of Texts & Ideas*, 1 (Belgium: Brepols, 2014), 213–228.

12 See Bausi 2016. The tool is developed by the *TraCES* project and is accessible at <<http://betamasaheft.eu/Dillmann/>>.



# Index

*compiled by Eugenia Sokolinski*

- ʿAbbā Garimā Gospels 14  
ʿAbbās 131  
ʿAbbāy, *See* Blue Nile  
ʿAbd al-Rahmān al-ʿArāṣī 91-92  
ʿAbd Allāh Ṣarīf 92  
ʿAbd Allāh ʿUmar b. Ğibrīl al-Aswām al-Sāhītī 91-94  
Abṭālis 85  
Abydos 47  
*Acta Petri* 74-75  
ʿAdal 131-132  
Addis Ababa 91-93, 95  
Afar 30  
Afro-Asiatic 59  
Agaro 7  
Agathonicus 55  
ʿAgaw 130  
Agence nationale de la recherche 33  
ʿAḥmad b. ʿIbrāhīm al-Ġāzī 123, 125, 127-128, 130-131, 133  
Akademie der Wissenschaften in Hamburg 13, 143  
Aksum, Aksumite 14, 60, 128, 145-146  
Alexander of Alexandria 53  
Alexandria 53, 116, 145  
ʿAmata Giyorgis 130  
ʿAmda Šəyon 60-61, 63, 123, 129, 131-132  
Amharic 30, 98  
*Anaphora of Jesus Christ* 80-81  
*Anaphora of Mary* 81  
Ancient Egyptian 39, 43  
Ancient Greek 43  
ANNotation of Information Structure (ANNIS) 63-64, 104  
Anthony 108  
Aphrahat 24  
*Apostolic Canons* 85  
al-ʿArabāt al-Madfūnah, *See* Abydos  
Arabia 31  
Arabic 30, 43, 46-47, 59-60, 84-85, 90-91, 98, 115, 118, 121, 133  
script 30, 89-90  
Arcadia 83  
Asswan 44  
Athanasius 51, 53, 84  
Atom format 43  
ʿAwāš 131-132  
Baʿada Māryām 131  
Bagemdār 130  
al-Balyanā, *See* Abydos  
Barradas, Manoel 128  
*Barrington Atlas of the Greek and Roman World* 43-44  
Basel 81  
Basil of Caesarea 52-53, 108  
Beta ʿĪsrāʿel 130  
Beta mašāḥḥ: Manuscripts of Ethiopia and Eritrea 5-8, 13-16, 22, 31-32, 34, 37, 60-62, 82, 85, 104, 122-126, 134-135, 137, 143, 147, 149  
Bible 15, 74, 79, 80, 146, 148  
Bible Society, London 82  
Biblioteca Apostolica Vaticana 23, 53, 90, 118  
*Bibliotheca Hagiographica Graeca* (BHG) 14, 54, 111, 146  
*Bibliotheca Hagiographica Latina* (BHL) 54, 111  
*Bibliotheca Hagiographica Orientalis* (BHO) 14, 54, 111, 146  
Bibliothèque nationale de France 18  
Blue Nile (ʿAbbāy) 124, 130-131  
Bodmer Papyri 74  
Bohairic 53  
*Book of Mysteries (Mašḥafa mašṭir)* 60, 81  
British Museum 83  
*Burda* 95  
Caesarea 52-54

- Cairo 72, 74  
*Canons* (Pseudo-Basil) 74, 116  
 Cappadocia 53  
 Carpius 60  
 Cerulli, Enrico 81-82, 90-91  
 Christianity 14, 29, 39, 49, 60, 107, 109-110, 117-118, 127, 145  
 Christian Orient 5, 60, 80, 97, 109, 115, 145  
*Chronicle of 'Amda Šəyon* 60-61, 63, 123, 129, 132  
*Chronicle of Galāwdewos* 6, 21-22, 122-125, 137-138  
*Chronicle of Susānyos* 123  
*Chronicle of Zar'a Yā'qob* 128-129, 131, 133  
 Chrysostom, *See* John Chrysostom  
*Clavis Aethiopica* (CAe) 14, 17, 54, 61, 123, 144  
*Clavis Apocryphorum Novi Testamenti* (CANT) 14, 54, 146  
*Clavis Apocryphorum Veteris Testamenti* (CAVT) 14, 54, 146  
*Clavis Coptica* (CC) 41, 52, 69, 107, 111  
*Clavis Patrum Copticorum*, *See* *Clavis Coptica* (CC)  
*Clavis Patrum Graecorum* (CPG) 54, 69, 111  
*Colloquium cum Palladio* (Cyril) 83  
 Comparative Oriental Manuscript Studies (COMSt) 8-9  
*Confessio Claudii* 123, 133  
 Constantinople 54  
 Copenhagen 5  
 Coptic 5-7, 9, 39-47, 50-56, 69-71, 74, 76, 107, 110-111, 115-118, 121, 147  
 Coptic Egypt 45, 118  
 Coptic Museum, Cairo 74-76  
 Copto-Arabic 60  
 Coptos 74  
 CorA 99  
 Corpus dei Manoscritti Copti Letterari (CMCL) 5, 9, 40-41, 52, 54-55, 71, 107  
 Council of Ephesus 83  
 Cushitic 30, 59  
 Cyril of Alexandria 51, 83-84, 116  
 Dabra Dāmmo 127-128  
 Dabra Libānos (Šawā) 127, 131  
 DARIAH-DE 136  
 Dayr as-Suryān 126  
*De anima et corpore* (Melito) 53  
*De cella* (Agathonicus) 55  
*Dəgg<sup>va</sup>* 85  
*De hora mortis* (Cyril) 116  
 Deir el-Shelwit 48  
*De Melchisedech* (Cyril) 83  
 Demetrius of Antioch 51  
*Dərsān za-bāšu 'Fisālgos* 145-147  
*De recta fide* (Cyril) 83  
 Digital Atlas of Roman Empire (DARE) 44  
 Dillmann, C. F. A. 7, 61, 79, 82  
 Dionysias 46  
 Dishna Papers 74  
 Djibouti 29  
 Dynastic Egypt 49  
 Egypt 39-40, 42-46, 48-49, 118  
 ElasticSearch 8  
*Encōmion* (Pisenthius) 74  
*Encyclopaedia Aethiopica* (EAE) 9, 13-14, 23, 82, 91, 121  
 English 41, 46, 52, 82, 146  
 Ephrem the Syrian 71-73  
*Epistle of Eusebius to Carpius* 60  
*Epistula Apostolorum* 80  
 Erho, Ted 15  
 Eritrea 29, 124  
 Ethiopia 5, 7, 13-15, 19, 24, 29-30, 33, 36, 38, 60, 80, 89, 121-124, 128-129, 132-133  
 Ethiopian Manuscripts Archives (EMA) 5, 7-8, 33-38  
 Ethiopic (Gə'əz) 5-7, 9, 13-15, 18, 21, 59-62, 79-82, 84-86, 97-104, 121, 123-124, 143-146, 148  
   script (*fidal*) 100-103, 146  
 Ethio-SPaRe: Cultural Heritage of Christian Ethiopia. Salvation, Preservation, Research 9, 126  
 Europe 15-16, 89, 108, 121  
 European Research Council 5, 29, 39, 59, 69, 79, 89, 97, 115  
 Eusebius 54, 60, 117  
 eXist-db 13, 61

- Expositio decalogi* (John Chrysostom) 85  
 Faṭagār 131-133  
*Faṭḥ al-Raḥmānī* 31  
*Fawātiḥ al-falāḥ wa-bawāriḥ al-ṣalāḥ*  
*ft ḡikr mawlid al-nāṭiq bi-l-naḡāḥ*  
 (Kabīr Ḥamza b. Kabīr Maḥmūd) 96  
 Fayyūm 48, 71, 118  
 al-Fayyūmī, Naṣīr al-Dīn Muḥammad 95  
 First Cataract 44  
 Flickr 43  
 Flügel, G. L. 115  
*Futūḥ al-Ḥabaša* 133  
 Gabala 54  
*Gadla 'Iyasus Mo'a* 128  
*Gadla Lālibalā* 81  
*Gadla Samā'īāt* 144  
 Galāwdewos 21, 34, 122-125, 127-133, 137-138  
 Gə'əz, *See* Ethiopic (Gə'əz)  
 geoJSON 126, 133, 136  
 Geonames 44  
 GeTa Tool 62-64, 97, 99, 102-103  
 GitHub 13, 46  
 Goḡḡām 124  
 Golden Gospel 24  
 Gondar 124, 126, 130, 146  
*Gospel of Matthew* 60  
*Gospels* 14, 24, 60  
 Graf, Georg 84  
 Greek 14, 41-43, 46-47, 51-55, 59-60,  
 69-70, 76, 83, 85, 108-111, 115-118,  
 143, 145  
 Gurage 7  
 HaCohen, Ran 15  
*ḥadīṭ* 91, 93  
 Hagarāy 35-36  
 Ḥamara Noḥ 24  
 Hamburg 5, 13, 59  
 Hamburger Zentrum für Sprachkorpora 64, 104  
 Harar 31, 92, 95  
 Harari 90-91  
 Hāšim 31  
*Haymānota 'abaw* 82, 84  
*Hāymānotu la-Yā'qob 'Elberād'i* 84  
 Ḥayq 'Istifanos 127  
 Hebrew 80, 98  
 Hegano 132  
 Hiob Ludolf Centre for Ethiopian Studies  
 5, 13, 32, 59, 122  
*Historia ecclesiastica* (Eusebius) 117  
*History of the Patriarchs of Alexandria* 53  
*History of the similitudes of the wise*  
*Physiologus, See Zenā massāleyāt*  
*za-ṭabib Fisā'algos*  
*History of the wisdom of the wise Physio-*  
*logus, See Zenā ṭabab za-Fisā'algos ṭabib*  
*Homily of the blessed Physiologus,*  
*See Dərsān za-bəsu 'Fisālgos*  
 Hommel, Fritz 145, 147  
 Horizon 2020 39, 69, 115  
 Horn of Africa 29, 31, 89-90, 123  
 Huntingford, G. W. B. 121  
 'Ifāt 125, 128-130  
*iḡāza* 93  
 India Office, London 81-83  
*In Hebraeos* (John Chrysostom) 111  
*In Ioseph patriarcham* (Ephrem) 71-73  
 Institut de Recherche et d'Histoire des  
 Textes (IRHT) 33  
 Institute of Ethiopian Studies, Addis  
 Ababa 91-93, 95  
*Interpretation of the Decalogue, See*  
*Expositio decalogi*  
 Isaac of Kalamon 71-73  
 Isaiah of Sketis 108, 111  
 Isaias 74-75  
 Islam 7, 9, 29-30, 60, 89-90, 93, 115,  
 129, 131-132  
 IslHornAfr: Islam in the Horn of Africa:  
 A Comparative Literary Approach  
 5-8, 29, 31-32, 89-90, 94, 96  
 Italian 52, 90, 122  
*I'tirāf al-'ābā'* 84  
 'Iyasus Mo'a 127-128  
 Jacob Baradaeus 84  
 Java 62, 104  
 Jerusalem 80  
 Jesuits 128, 130, 133  
 Jesus Christ 80-81, 83, 116  
 John Chrysostom 53-54, 85, 108, 111  
 JSON 43, 62, 104, 126, 133, 136

- Judaism 60, 110  
*Judith, Book* 80  
 Kabīr Ḥamza b. Kabīr Maḥmūd b. Kabīr 'Alī al-Ḥarallī al-Awsī al-Ḥanafī 96  
 Kamāl from Agaro 7  
*Kəbra nagašt* 60-61  
*Kidān za-nagh* (Matins) 80  
 Kidron 80  
*Kings, Books of* 80  
*Kitāb al-farā'id* 89-95  
 KML: Keyhole Markup Language 43, 133, 136  
 Krapf, Johann Ludwig 81  
 Landesforschungsförderung Hamburg 5  
 Late Antiquity 14, 39, 42, 46, 48-49, 59-60, 108, 143, 145  
 Latin 43, 69, 79, 110, 115, 121  
   script 62  
 Ləbna Dəngəl 35-36, 38, 127, 130-131  
*Letter to the Galatians* 55  
*Lexicon linguae Aethiopicae* (Dillmann) 61, 79-86, 149  
*Life of Anthony* 108  
*Life of Isaiah of Sketis* 108, 111  
*Life of Pachomius* 108  
*Life of Paul of Tamma* 108  
*Life of Secundus the Silent* 60  
*Life of St Basil* 108  
 Linked Open Data 8, 44, 121  
 Location Referenced External Apparatus 22  
 London 82, 83, 84, 86  
 Ludolf, Hiob 81  
 Luxor 46  
*Māḥbara mə'əmanān* 86  
 Makāna Šəllāse 35-36  
 Manqabad 46  
 Marḥa Krəstos 131  
 Marina 83  
 Mary 34, 81  
*Maṣḥafa dəgg<sup>w</sup>a* 82  
*Maṣḥafa məṣṭir*; *See Book of Mysteries*  
*Maṣḥafa ʾomār* 84  
*Mawāšə't* 86  
*Mawlid šaraf al-'ālamīn* 31, 95  
 Mediterranean 43-44  
 Melchisedek 83  
 Melito 53  
 Memnoneia (Djeme) 43  
 Memphis 48  
 Middle Ages 33, 42, 46, 48, 118, 121, 124  
 Minās 130  
*Missal* 81  
 Monastery of Saint Phoibammon 46  
 Monastery of the Archangel Michael 71, 73  
 Mount of Olives 80  
 Naḡ Hammadi 110  
 Naṣīr al-Dīn Muḥammad al-Fayyūmī,  
   *See al-Fayyūmī*  
 Naṣraddin b. 'Aḥmad 125, 129-130  
 Near East 60, 109, *See also* Christian Orient  
 Nestorius 51  
 New Testament 80  
 New York 71  
 Nile Delta 48  
 Nine Saints 127-128  
 Nubia 48  
 Old Testament 74  
*On the Nativity* (Demetrius) 51  
*On the Right Faith* (Cyril), *See De recta fide*  
*On the Trinity* (Athanasius) 84  
 Oromo 30, 129  
 Ottoman Empire 125, 127  
 Oxford 86  
 Pachomius 108, 111  
 Paez, Pedro 128  
 Palladius 83  
*Paralipomena Ieremiae* (4 Baruch) 71-73  
 Paris 5, 33, 72  
 PATHs: Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature. Literary Texts in their Geographical Context. Production, Copying, Usage, Dissemination and Storage 5-9, 39-40, 42, 44, 46-47, 50-55, 69-70, 73, 75-76  
 Patmos 125  
 Patristics 107-110  
 Patrology, *See* Patristics  
 Pauline Epistles 55  
 Paul of Tamma 108

- Pelagios Commons 43, 121-122, 125, 133-134, 137, 139
- Pelagios Interconnection Format 122, 134
- PeriodO 50, 126
- Peripleo 137-138
- Persian 115
- Phantoou 73
- Physiologus* 143-148
- Pierpont Morgan Library, New York 69-73
- Pisenthuis of Coptos 74
- Plato 51
- Pleiades 22, 40, 43-44, 47-48, 50, 125
- Portugal 125
- PostgreSQL 8
- Pseudo-Basil 74, 116
- Ptolemaic Egypt 45, 49
- Qerällos* 82-84
- Quod deus non est auctor malorum* (Basil) 52
- RDF: Resource Description Framework 6, 43-44, 121, 133, 137
- Red Sea 44, 129
- Renaissance 108, 112
- Republic* (Plato) 51
- Rift Valley 129
- Roman Egypt 45, 49
- Rome 5, 39, 72, 128
- Ruby on Rails 8
- Sabraddin of 'Ifāt 129
- Šāfi'ī 94
- Sahara 44
- Sahidic 52
- Samuel, Book of* 80
- Sapienza Università di Roma 5, 39, 69, 115
- Šarša Dəngəl 130-131
- Šawā 124-125, 127-129, 131-132
- Sawāsəw* 84
- Schriftkultur des christlichen Äthopiens und Eritreas: Eine multimediale Forschungsumgebung. *See* Beta mašāḥəft
- Səmen 130
- Semitic 30, 59, 82, 98
- Sənkassār* 144
- Senodos* 82, 84-85
- Septuagint* 80, *See also* Bible
- Seventh Framework Programme (EU) 29, 59, 79, 89, 97
- Severian of Gabala 54
- Sharing Ancient WisdomS (SAWS) 22, 24, 147
- SNAP: Standards for Networking Ancient Prosopographies 24
- Solomonic Dynasty 129
- Somali 30
- Somalia 29
- Somaliland 29
- South Arabian script 104
- South Semitic 59
- Stuttgart 80
- Susənyos 123, 130
- La syntaxe du codex. Essai de codicologie structurale* (Andrist et al.) 69-70, 73, 75-76
- Syriac 6, 9, 53, 126
- Syriaca.org 5-6, 8, 22, 126, 134, 139
- Ṭabiba ṭabibān* 86
- Ṭaḥmīs al-Fayyūmī 'alā al-Burda* 95
- Takkaze 124
- Takla Haymānot 127-128
- Tamben 35-36
- Tanā 130
- Tansə'a Krəstos 35-36
- Ṭayyib al-Wanāḡī al-Šadrī 91-93
- Ta'amməra Māryām* 144
- Teaching of the 318 Nicaean Fathers* 85
- Ṭəgrāy 124, 127, 128
- TEI: Text Encoding Initiative 6-7, 13, 15-17, 22, 33-35, 37, 41, 61-62, 104, 121, 123, 126, 133-134, 136, 143, 147
- Testament of Our Lord in Galilee 80
- Testament of Our Lord Jesus Christ, See Testamentum Domini*
- Testamentum Domini* 60, 80-81
- Thebes 43, 46-48
- Theodosius II 83
- Toutōn 117
- TraCES: From Translation to Creation: Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages 5, 8, 59, 61-64, 79, 97-99, 123, 149
- Transmission of Knowledge in the Red Sea Area 5



- Trinity 84  
 Trismegistos 40, 47-48  
 Tübingen 81, 83-86  
 Turin 116-117  
 Turkish 115, *See also* Ottoman Empire  
 TXT 61, 104  
 Union of the German Academies of Sciences and Humanities 5, 13, 143  
 Universitätsbibliothek Tübingen 83  
*'Unwān al-šarīf* 95  
 URI: Universal Resource Identifier 6, 15, 43-44  
 URN: Uniform Resource Name 47  
 USA 15  
 Valley of the Kings 46  
 Valley of the Queens 46  
 Vanderbilt University 5  
 Vatican 53, 90, 118, *See also* Biblioteca Apostolica Vaticana  
 VIAF: Virtual International Authority File 7, 40, 52  
 Vici.org 44  
 VisColl 20  
*Vita Moysis Archimandritae* 55  
*Vita Pisentii* 55  
*Vita Samuelis Archimandritae* (Isaac of Kalamon) 71-73  
 Voyant Tools 104  
 Waḡ 61, 131-132  
 Waḡarā 130  
 Wagner, Ewald 91  
 Waldəbba 34  
*waqf* 7  
 WebAnno 99  
*Wəddāse Māryām* 18, 143  
 Wikidata 23, 44, 47, 125-126, 139  
 Wikipedia 44, 47  
 Wright, William 82  
 Württembergische Landesbibliothek, Stuttgart 80  
 XML: Extensible Markup Language 6-8, 15-18, 21, 33, 35-37, 41, 43, 60-62, 104, 136, 143, 147  
 XPath 38  
 XQuery 38  
 XSLT 61  
 Yāgbā Şayon 23  
 Yəkunno 'Amlāk 103, 131  
 Yemen 31  
 Zagwe 131  
 Zamikā'el 'Aragāwi 127-128  
 Zar'a Yā'qob 128-129, 131, 133  
 Zebdār people 131  
*Zenā məssāleyāt za-ṭabib Fisə'algos* 145, 147, *See also* *Physiologus*  
*Zenā ṭəbab za-Fisə'algos ṭabib* 145, 147, *See also* *Physiologus*  
 Zion 133  
 Zotero 32