

So Miyagawa
Kyoto University and Georg-August-University of Göttingen

Digitization of Coptic Manuscripts and Digital Humanities: Tools and Methods for Coptic Studies¹

Abstract:

In this paper, we focus on Coptic manuscripts and their digitization. After providing a brief history of Coptic language manuscripts, we discuss aspects of their digitization before moving on to digital editions, archives, corpora and catalogues, specifically. We also report on the various global advances in web platforms and digital tools and their application in the digitization of Coptic manuscripts. We conclude by elaborating on some of the future directions for Coptic manuscript digitization, emphasizing avenues for collaboration, strengthened interoperability, and improving the user experience.

Keywords: Digital Humanities, Coptic Manuscripts, Digital Archive, Digital Edition, Digital Corpus

¹ This article is a modified and extended English version of Miyagawa (2019e), a journal article originally written in Japanese. Miyagawa (2019e) appeared in a Japanese journal called *Chūgoku* 21 [『中国21』], volume 51, pp. 39–76. This research was supported by German Research Foundation (DFG: Deutsche Forschungsgemeinschaft)’s Collaborative Research Centre (SFB: Sonderforschungsbereich) 1136 “Education and Religion in the Mediterranean and Surrounding Cultures from Antiquity to the Middle Ages and Classical Islam” (Bildung und Religion in Kulturen des Mittelmeerraums und seiner Umwelt von der Antike bis zum Mittelalter und zum Klassischen Islam). In addition, some of the topics in this paper are included in several installments of “Digital Humanities in Europe and the Middle East,” a series of articles by the author in the monthly e-newsletter *Digital Humanities Monthly* [『人文情報学月報』], edited by the International Institute for Digital Humanities located in Tokyo (Miyagawa 2018b, 2018c, 2018d, 2018e, 2018f, 2019a, 2019b, 2019c, 2019d).

1. Introduction

When Napoleon Bonaparte invaded Egypt in 1798, he ushered in a century and a half of systematic removal of Egyptian manuscripts and other treasures by Western imperial powers. Generations of European archaeologists and excavators, supported by the economic and political might of their nations, discovered treasures which had been lying hidden in Egyptian soil for centuries, and transported half of them back to Europe for study and display. European collectors also made direct contact with institutions such as St. Catherine's Monastery on Mount Sinai, as well as with local buyers and archivists, in order to buy up troves of Egyptian manuscripts. These joined the archaeological finds in European libraries, museums, art galleries, and private collections, where a process of digitization has currently been underway for some years². This accelerating digitization of Coptic manuscripts in Europe is understood, for the purpose of this paper, to be the scholarly activity designed to make them accessible to the open public.

This paper is organized as follows: Chapter 1 presents an outline of the digitization of Egyptian manuscripts in Europe, focusing on manuscripts written in the Coptic language. Thereafter, Chapter 2 discusses digital catalogues, and Chapters 3, 4, and 5 address digital archives, digital editions, and digital corpora, respectively, with the last also offering a summary of the paper³.

² See Louis (2008) for an account of how the manuscripts of the White Monastery found their ways, page by page and fragment by fragment, from the library and cache of the White Monastery near Sohag to various libraries and museums in Europe and North America. The history and current holdings of these manuscripts could be regarded problematic in ways that Edward Said has criticized as "Orientalism" (Said, 1978).

³ The author of this paper is an expert in philology and linguistics of Ancient Egyptian-Coptic, Koine Greek, and Old Nubian. In addition, since the author works in Germany, this paper is biased towards available German sources. There are many excellent projects in Europe, such as Europeana (<https://fontane-nb.dariah.eu/index.html>, last viewed on March 30, 2019) and DARIAH-EU's digital edition of Theodor Fontane's notebooks (<https://fontane-nb.dariah.eu/index.html>, last viewed on March 30, 2019). However, these do not mainly address Egypt, so this paper only describes in detail that part of the Europeana Collections which is related to Coptic manuscripts. In addition, it is impossible to cover all existing projects in this essay. Rather, I will focus on the projects related to Coptic manuscripts, which I am familiar with and use in my research.

The Egyptian language has the world's longest scribal history, dating back approximately 5000 years. According to Dreyer (2011:128), the writing on ivory tags from the Abydos U-j tomb was radiocarbon-dated to 3350-3150 BCE. The Egyptian language is the sole member of the Egyptian branch in the Afro-Asiatic language family⁴, which stretches from North and Central Africa to West Asia.

Egyptian has been used for a wide range of texts, from sacred through administrative to popular, and is written in Hieroglyphic, Hieratic, Demotic, and Coptic scripts. Of these, the so called "Egyptian scripts" (Hieroglyphic, Hieratic, and Demotic) are unique to Egypt, while the Coptic script consists of the Greek script with a few phonetic letters added from the Demotic to represent sounds not found in Greek. The Egyptian writing system consists mainly of phonograms for one to three consonants, logograms for specific words, and determinatives for word categories. The Coptic alphabet, on the other hand, is a phonemic alphabet in which each letter corresponds to a single vowel or consonant phoneme except for a few bi-phonemic letters. The diachronic stages of the language are Old Egyptian, Middle Egyptian, Late Egyptian, Demotic Egyptian, and Coptic Egyptian, usually referred to simply as Coptic⁵. The Coptic language is still used in Coptic Christian liturgy and chanting in the Coptic Orthodox Church and Coptic Catholic Church.

Table 1 shows the scripts and dates of these linguistic stages, as well as the larger groupings (Earlier Egyptian and Later Egyptian) based on their grammars.

⁴ This was proposed by Greenberg (1963) and is now accepted by many linguists. Prior to Greenberg's (1963) study, this language family was called "Semitic-Hamitic." In addition to Egyptian, the Afro-Asiatic languages include Semitic, to which Arabic and Hebrew belong; Cushitic, to which Somali belongs; Berber, to which Silha belongs; Chadic, to which Hausa belongs; and Omotic, proposed by Fleming (1969), accepted in recent years and about which there are various theories (Frajzyngier & Shay, 2012, p. 7). The details of grouping and branching vary among scholars.

⁵ The word "Coptic" comes from the Arabic word *qubṭ*, but it is generally believed that it is derived from the Greek *αἴγυπτος*, which means "Egypt" and which originated in *ḥwt-kʿ-ptḥ* "house of the *ka* of Ptah," an alias of Memphis in Ancient Egyptian.

Language		Script	Period
Pre-Old Egyptian		Proto-Egyptian script	32 nd – 27 th century B.C.
Earlier Egyptian	Old Egyptian	Hieroglyphs Hieratic script	27 th –21 st centuries B.C., and 7 th century B.C. (archaic revival)
	Middle Egyptian		23 rd century B.C. – 44 th century A.D.
Later Egyptian	Late Egyptian		14 th – 7 th century B.C.
	Demotic	Demotic script Old Coptic script	8 th century B.C. – 5 th century A.D.
	Coptic	Coptic script	3 rd century – 16 th century A.D., 19 th century – 21 st century A.D. ⁶

Table 1: Based on historical classification by Kammerzell (2000, p. 97), with script entries added by Miyagawa, Yoshino and Nagai (2018)

In 334 B.C., Alexander III⁷, King of Macedonia, otherwise known as Alexander the Great, began his campaign against Achaemenid Persia. Having conquered Asia Minor, he moved on to the other parts of the Persian empire, including Egypt. Alexander’s death, on June 10, 323 B.C., launched a series of civil wars among his would-be successors competing for supremacy, during which General Ptolemy, founder of the Ptolemaic dynasty, became the governor of Egypt. By 305 or 304 B.C., Ptolemy had begun styling himself King Ptolemy I Soter⁸ of Egypt. Ptolemaic Egypt survived until 30 B.C., when it fell to the forces of Rome. The allied armies

⁶ Kammerzell (2000, p. 97) divides the period of Coptic into three: Standard Coptic (3rd–12th centuries A.D.), Late Coptic (11th–16th centuries A.D.), and Neo-Coptic (19th–20th centuries A.D.). Since Kammerzell’s (2000) paper was written during the 20th century, Coptic is described as a language up to the 20th century. Coptic is, however, still used by Coptic Christians in the 21st century, so Miyagawa, Yoshino & Nagai (2018) modified “20th century” into “21st century”, as per Table 1.

⁷ Ancient Greek: Ἀλέξανδρος Γ’. Alexandros the Great (Ancient Greek: Ἀλέξανδρος ὁ Μέγας). Date of birth and death: 356–323 BC.

⁸ Ancient Greek: Πτολεμαῖος Α’ Σωτήρ. Date of birth and death: 367/366 BC–283/282 BC.

of the last Ptolemaic queen, Cleopatra VII Philopator⁹, and Marcus Antonius of Rome were defeated by Octavian (later Augustus, the first Roman Emperor) at the Battle of Actium in 31 B.C.

After the division of Roman-occupied territories on the death of Theodosius I in 395, Egypt was part of the Eastern Roman, or Byzantine, Empire¹⁰ for almost three centuries: There was a brief occupation by the troops of Sassanid Persia from 618/619 to 628, but it was not until 641 that it was wrested from Byzantine control by the expanding Islamic Caliphate¹¹ which was making rapid progress through the Arabian Peninsula.

During the centuries in which Egypt was ruled by the Ptolemaic, Roman, and Byzantine empires, many linguistic changes occurred. Greek became the language used in politics, economics, trade, religion, and learning, having established itself throughout the eastern Mediterranean, and retained its dominance even under Roman and Byzantine rule. It was only after the Islamic conquest of the 7th century that Arabic became the language of administration. Until then, Greek tended to be the language of the upper classes, while Coptic remained the language of the people. Because of the importance and widespread use of the language, many Greek words were absorbed into Coptic: not just content words such as verbs and nouns, but also function words such as infinitives and prepositions¹².

The written language, too, underwent changes. Several attempts were made to write Egyptian using Greek letters during the Ptolemaic and early Roman periods. This variant of Egyptian written with Greek letters plus several phonetic letters derived from Demotic is called Old Coptic¹³. Although the Egyptian script continued to be used until the 5th century A.D., Greek and other scripts of popular origin became widespread and standardized from the 2nd century A.D. onwards, mainly due to the spread

⁹ Ancient Greek: Κλεοπάτρα Ζ' Φιλοπάτωρ. Date of birth and death: 69 BC–30 BC.

¹⁰ For a history of this period, see Kaibara, 2015.

¹¹ Under the second legitimate caliph Umar (‘Umar ibn al-Khaṭṭāb) and led by General Amr ibn al-‘Ās.

¹² See Miyagawa, 2018a, pp. 278-279, p. 281.

¹³ Satzinger, 1991, and Satzinger, 1975, are two of the references to Old Coptic. Although it is called “Old” Coptic, its grammar suggests that it is a Greek version of Egyptian, a popular script. For this reason, Kammerzell classifies Old Coptic as an Egyptian folk script (Kammerzell 2000, p. 97). Miyagawa, Yoshino, & Nagai (2019) believe that Old Coptic mainly refers to the script, not the language, and translate it as *Ko-Koputo-Moji* meaning “Old Coptic script”.

of Christianity. The Coptic script is a combination of the ancient Greek and a few letters derived from the demotic script.

There were also regional differences in the language used. For example, Coptic was used more in Upper Egypt than in Lower Egypt (the Nile Delta), as the latter was home to Alexandria, the center of Hellenistic culture. Hence, a large number of Coptic writings on papyrus, ostraca, parchment, and paper have been found in Upper Egypt, where an arid climate has aided their survival. Many of these are of great importance in the fields of religion and Biblical studies. The Nag Hammadi Library, for example, is famous for the large number of manuscripts apparently written by Gnostics, whose existence could previously only be inferred by the refutations of Christian apologists such as Irenaeus.

Another important cache of documents was the Manichaean literature arguably found at Madinat Madi in Fayoum¹⁴, such as the *Kephalaia*¹⁵, which provide insight into the teachings of Manichaeism from a Manichaean perspective.

Also, it was in Egypt that the Christian monastic system flourished under Paulus of Thebes, Antony, and other reclusive monks and nuns, who have left documents of enormous interest in the Coptic language: the letters of Antony in the 3rd and 4th centuries¹⁶, the writings of Pachomius in the 3rd and 4th centuries¹⁷, and the letters and sermons of Shenoute in the 4th and 5th centuries.

¹⁴ In terms of dialect, it was written in the Lycopolitan dialect (Semi-Akhmimic dialect). Since this dialect was not used in Fayoum, we can assume that these documents were probably written in the south and brought to Fayoum in some way.

¹⁵ As regards the Berliner *Kephalaia*, which is housed in Berlin and was partly transferred to the Soviet Union and Poland as a result of the Battle of Berlin at the end period of World War II, the Coptic text and German translations are Polotsky & Böhlig (1940), Böhlig (1966) and numerous publications by Wolf-Peter Funk, for example, Funk (2020, 2018, 2000, 1999). An English translation is by Gardner (1995). See also Gardner, BeDuhn & Dilley (2014) for the Dublin *Kephalaia* in the Chester Beatty Library, Dublin, Ireland, which has been the subject of a recent publication project.

¹⁶ Samuel Rubenson (1995) has argued from this letter that Antony was not an “uneducated monk” as depicted by Athanasius in his biography of Antony.

¹⁷ For a transcription and translation of a Coptic text of Pachomius’ works, see “The Instruction of Apa Pachomius the Archimandrite” in Budge, 1913, pp. 352-382. Louis-Théophile Lefort has also published more Coptic texts of Pachomius’ works (Lefort, 1956a) and French translations (Lefort 1956b).

The Coptic language continued to produce new literature until the 14th century, when the Triadon¹⁸, the last known literature composed in Coptic, was created. Since then, however, the language has survived in various ways. It has continued to be copied – especially liturgical texts, chants, and the Bible – and in recent years digitization has been promoted by both scholars and believers alike. Iqladiyus Labib and others have been trying to revive Coptic as a spoken language since the 20th century, and there is an ongoing Coptic revival movement¹⁹.

There are several dialects within Coptic. In late Antiquity, Sahidic became a standard dialect. Then, after the Patriarchate moved to Cairo, Sahidic was replaced by Bohairic, which is still used in Coptic Christian liturgy today. The Fayoumic dialect is found mainly in Fayoum, and the Oxyrhynchite dialect (also called “Mesokemic”, or “Middle Egyptian”) is spoken in Central Egypt. The Lycopolitan (“Semi-Akhmimic”) dialect is often used in Manichaean and some Gnostic literature. Finally, there is the Akhmimic dialect, a dialect used in the south around between Akhmim and Aswan, which is a relatively old dialect attested to in Biblical translations and apocrypha²⁰.

The digitized Coptic documents discussed in this paper mainly date to late Antiquity, from the latter period of the Roman Empire to the Early Middle Ages, and were written in ink on a variety of materials, including papyrus, parchment, ostraca (pottery shards), wood, and, in later periods, paper. The paper considers inscriptions, as well as texts written in ink.

The digitization process encompasses several stages, from the digitization of metadata, images, and texts through visualization of annotations, statistical analysis data, and so on. The following chapter presents a digital catalogue of literature metadata.

¹⁸ See von Lemm, 1903, for the Coptic text and Nagel, 1983, for the German translation. English translations can be found in MacCoull, 1997, and Alcock, 2013a, 2013b.

¹⁹ Basta (1991) is the entry for Iqladiyus Labib in *The Coptic Encyclopedia*.

²⁰ See Kasser, 1991, and Funk, 1988.

2. Digital Catalogue (Metadata)

2.1 CMCL: Corpus dei Manoscritti Copti Letterari²¹ (Italy)

The history of Digital Humanities may be said to originate in Italy, with the *Index Thomisticus* produced in 1951 by the Italian Jesuit Father Roberto Busa, S.J.²² This work is an index of the writings of Thomas Aquinas, O.P., the most influential philosopher in the medieval philosophy of the Catholic Church. In effect, Fr. Busa created a machine-searchable textual corpus of Aquinas' writings, including his most famous work, the *Theological Compendium*. Fr. Busa had been planning the work since 1946, but it only came into being after he met Thomas J. Watson, the founder of IBM, in 1949, and was introduced to the then-nascent computer technology²³. According to the website of the University of Tübingen, the *Index Thomisticus*, which took over 30 years to complete, contains the 118 texts of the complete works of Thomas Aquinas himself, and 61 related texts by other authors²⁴. Roberto Busa pioneered the use of computers to archive materials for research in the humanities.

If the history of Digital Humanities as a whole can thus be said to have started in Italy, so did the history of Coptic Digital Humanities in particular. The *Corpus dei Manoscritti Copti Letterari* (CMCL) (Orlandi 1990:397) project was begun at the Università degli Studi di Roma "La Sapienza"²⁵ in 1968. Led by Tito Orlandi²⁶, a professor at the university, it aimed to bring together the Coptic literature which had been scattered across pages and smaller fragments throughout museums, libraries, and research institutions around the world. Thereafter, the gathered literature would be reconstructed as codices, and the catalogue would be

²¹ <http://www.cmcl.it/>, last accessed on March 30th, 2019.

²² See Busa, 1951. In addition, the work of Burdick *et al.* (2012, p. 123) has a commentary on Father Roberto Busa, S.J.

²³ For more on this encounter, see Masoner, 2018.

²⁴ <https://uni-tuebingen.de/fakultaeten/philosophische-fakultaet/fachbereiche/neuphilologie/seminar-fuer-sprachwissenschaft/arbeitsbereiche/allg-sprachwissenschaft-computerlinguistik/ressourcen/corpora/index-thomisticus-baumbank/>, last accessed on March 30th, 2019.

²⁵ Today, this university is called the Sapienza Università di Roma (Sapienza University of Rome).

²⁶ Besides being a Coptic philologist, Orlandi has contributed numerous articles on the application of information technology to the humanities in informatics journals. For a list of his publications in the field of information technology, see <http://www.cmcl.it/~orlandi/pubinf.html>, last accessed on April 28th, 2019.

computerized. The CMCL is not only a database of reconstructed codex metadata, however; it also contains photographs and reprints of some, albeit somewhat inconsistent, references. A grammar of Coptic by Orlandi and information on the authors of Coptic documents are also available in English or Italian. This project is ongoing, and has had a significant impact on Coptic Studies.

The first stage of the reconstruction process for the CMCL project, given the fragmentation of so much Coptic literature, is to ascertain where it was disassembled. Before the CMCL, the attempts at reconstruction made by various bibliographers were themselves scattered across many articles and reports, gathering and checking, which was a daunting task in itself. Although a subscription was previously necessary to use the CMCL²⁷, most of the data is now available free of charge at PATHs under the Creative Commons²⁸ (CC) Attribution-NonCommercial-Share Alike 4.0 International License (CC BY-NC-SA 4.0) (see Section 2.2. for further details). Reconstructions can now be checked instantly by any interested parties; hence, it is in widespread use among Coptic philologists.

Stephen Emmel, for example, made extensive use of the CMCL for his *Shenoute's Literary Corpus* (2004), which attempts to reconstruct the entire literary codex of Shenoute, the most prolific of all Coptic authors, and now constitutes the basis of research in this area. Prior to Emmel's work, Shenoute's writings had been so scattered that only partial reconstructions had ever been possible. The CMCL has since reimplemented the results of Emmel's research; thus, his achievements are now reflected in the very database on which he drew. Therefore, the CMCL project has become an exemplar of the circulation of knowledge and resulting development of the scholarly community.

²⁷ The author had an individual subscription before coming to the Georg-August-University of Göttingen, thereafter using the subscription of the Seminar for Egyptology and Coptic Studies of the University.

²⁸ Creative Commons, <https://creativecommons.org/>, last accessed on March 31st, 2019.

In the past, if there was no mention of rights such as copyright, one would have to apply to the rights holder for secondary use every time. However, with the license established by Creative Commons, it is easy to see at a glance the extent to which secondary use is allowed and under what conditions, by merely writing a short notice, for example, CC BY-NC-SA 4.0. Globally, more and more projects are displaying Creative Commons licenses, such as PATHs and the Digital Papyrus Collection of the Heidelberg University Library.

2.2 PATHs: Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature²⁹ (Italy)

The PATHs project, funded by a grant from the European Research Council and led by Professor Paola Buzi of the Sapienza University of Rome, provides detailed information on the geographical mappings of Coptic manuscripts and on their material structure, that is, how the pages are folded and what the quires physically are³⁰, together with other codicological details, some of which are visualized. In February of 2019, the results were released to the public in the form of an atlas. Exploring the atlas and clicking on a given area will reveal which documents are believed to have been created or found there. The metadata contained in the atlas is based on the CMCL data, but further detail has been added. Alternatively, searches can be made using the Trismegistos Number (TM Number), which will be described in the next section. The PATHs project is available to the public under the same CC license as the CMCL.

²⁹ <https://atlas.paths-erc.eu/>, last accessed on March 30th, 2019.

³⁰ A piece of parchment/papyrus/paper with two pages on one side, or four pages on both sides, was folded in the middle, and then four or more sheets of parchment/papyrus/paper folded in the same way were placed on top of each other to form a single quire. One quire, or multiple quires, were then bound together to form a codex. According to the modern English definition of the quire, a huge sheet of paper is first prepared and then folded in multiple turns to form a quire, but the modern quire is different from the quire of late Antiquity.


PATHs **Atlas** Places Manuscripts Works Authors Titles Colophons Collections How to cite "PATHs" [Get more](#)

View record

Places [Show all](#) [Search](#) [Advanced search](#) [Saved queries](#)

paths.places.112
<http://paths.uniroma1.it/atlas/places/112> [Copy to clipboard](#)

PATHs ID	112
Site name	Monastery of Shenoute at Atripe
Trismegistos GeolD	3121
Pleiades Id	756530
Region	Northern Upper Egypt
Area	Sohag
Nome	Panopolites
Province	Thebais



Links

524 referenced items in Manuscripts

Fig. 1 Screenshot of the Monastery of Shenoute at Atripe (also known as the White Monastery) from the PATHs atlas³¹.

PATHs **Atlas** Places Manuscripts Works Authors Titles Colophons Collections How to cite "PATHs" [Get more](#)

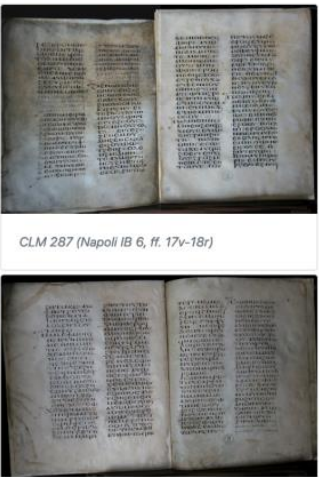
paths.manuscripts.287

<http://paths.uniroma1.it/atlas/manuscripts/287> [Copy to clipboard](#)

Manuscript identifiers

Coptic Literary Manuscript (CLM) ID	287
CMCL	MONB.BA
TM	108394
LDAB	108394
Alias	Besas' Codex A / Zoega CCIV
Codex stratigraphy	Single codicological unit, incomplete and dismembered.
Modern history	Dismembered in 5 shelfmarks, 3 of them - namely Napoli IB 6, Wien, K. 965 and London, Or. 8810 - containing different sections of the original codicological unit. Napoli IB 6 from the Borgia collection. London, Or. 8810 from the Curzon collection (cf. Lavton 1987: 87).

Attached images



CLM 287 (Napoli IB 6, ff. 17v-18r)

CLM 287 (Napoli IB 6, ff. 30v-31r)

Fig. 2 MONB.BA manuscript containing the writings of Besa, the successor of Shenoute, originated in the White Monastery. Two photographs of the manuscript are shown³².

³¹ <https://atlas.paths-erc.eu/places/112/>, last accessed on March 30th, 2019.

³² <https://atlas.paths-erc.eu/manuscripts/287/>, last accessed on March 30th, 2019.

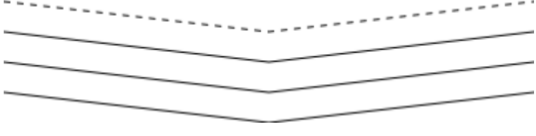
Quire no.	9
Quire layout	1x/1x/1x/1x* <div style="text-align: right; color: red;">1x*/1x/1x/1x</div> <div style="text-align: right;">Total: 12 pages</div> 
Parchment quire typology	F like facing like
Quire notes	Wien, K. 965.1-3 and 965.4-6. Ancient pagination ρκε-ρλ + <ρλδ-ρλδ> + ρλε-ρμ. Ancient quire-signature θ on first and last page.

Fig. 3 Composition of the ninth quire of the same MONB.BA codex by PATHs³³. In the above figure, one sheet of four pages is missing in the area marked by the dotted line.

2.3 Trismegistos³⁴

Trismegistos Texts is a metadata database of documents dating from the 8th century B.C. to the 8th century A.D., mainly from the Mediterranean world, currently under development by Mark Depauw of the Catholic University of Leuven, Belgium. The database was initiated in 2004, when Depauw, then studying at the University of Cologne, won the Sofja Kovalevskaja Award of the Alexander von Humboldt-Stiftung³⁵. At present, the Catholic University of Leuven is the primary operator of the project. The scope of the database is enormous, as it covers not only

³³ <https://atlas.paths-erc.eu/manuscripts/287/>, last accessed on March 30th, 2019.

³⁴ <https://www.trismegistos.org/>, last accessed on March 30th, 2019. For a more detailed explanation of Trismegistos, see Miyagawa, 2019b. The name “Trismegistos” comes from Hermes Trismegistos, who was a combination of the Greek god Hermes, the Egyptian god Thoth, and the alchemist Hermes, in the Hellenistic culture of Egypt.

³⁵ https://www.trismegistos.org/about_history.php/, last accessed on March 30th, 2019.

Greek, Egyptian, and Latin literature, mainly of Egyptian origin, but also literature excavated outside Egypt in the Celtic languages, Ge'ez, and Central Asian languages. While Trismegistos Texts is the main database, there are also databases of names of people, places, ancient documents, current holdings, authors, etc., and an application that visualizes these databases in a network diagram. The data can be easily used by other services through APIs³⁶.

The central Trismegistos database records basic information, such as the date of the literature, the region where it is supposed to have originated, the institution where it is currently stored, and the language, material, and genre in which it was written³⁷. Metadata such as the name of the work (if it is a work), publication information (if an edition of the work has been published), etc., is also recorded. Since this is a metadata database, there are no photographs or actual texts. An important recent development is the use of Trismegistos Numbers (TM Numbers), which are ID numbers assigned to each document by Trismegistos Texts. Although this system of identification is currently only standard for documents in Digital Humanities projects related to the ancient and late Antique Mediterranean world, such as Greek and Coptic, it is hoped that in the future it will be used to link the results of various projects. Hence, it will be possible to accumulate a large amount of information related to a particular document at one time.

Two other Trismegistos databases, Trismegistos Places and Trismegistos Names, allow users to compare the correspondence between the Egyptian names of people and places written in Hieroglyphic, Hieratic, Demotic, Coptic, Greek, and Latin scripts, as well as the frequency of mentions in documents from any period or region. Trismegistos Archives and Trismegistos Collections, meanwhile, use numbers, pie charts, and bar graphs to show which regions currently have collections and archives, which periods, languages, and materials they cover, and so on. Trismegistos Authors gives all data related to an author registered in Trismegistos Texts, while Trismegistos Editors is a database of editors of published text editions.

Lastly, Trismegistos Networks enables users to draw a network diagram using network theory to visually check which names are closely

³⁶ <https://www.trismegistos.org/dataservices/>, last accessed on April 28th, 2019.

³⁷ Metadata is not the images or text of the main literature, but the surrounding or background information of the literature, such as the material, age, storage location, and language of the text.

connected, using data from Trismegistos Places and People. Figure 4 shows an example of Trismegistos Networks.

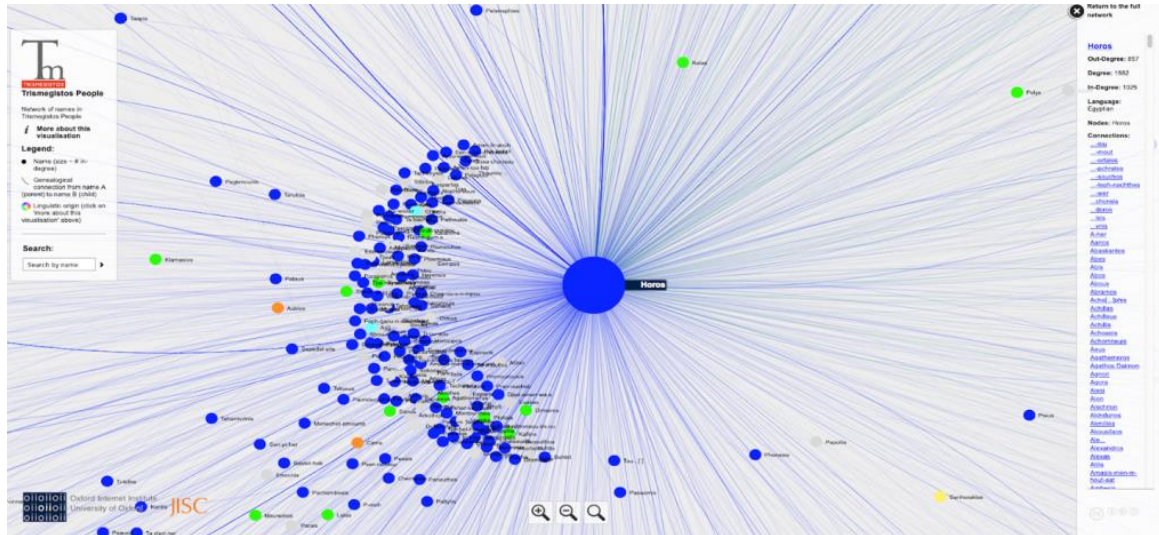


Fig. 4 “Shenoute”, searched with Trismegistos People using Trismegistos Networks and focusing on Horos³⁸

One difference between Trismegistos and CMCL/PAThS concerns the language, region, and date of the metadata. CMCL/PAThS collects only Coptic literary manuscripts; that is, it includes literary and religious literature but excludes documentary literature, such as administrative and economic documents. Trismegistos Texts, in contrast, collects documents from the 8th century B.C. to the 8th century A.D., mainly of Egyptian origin, in various languages, of both documentary and literary genres. The other major difference is that Trismegistos is based on units of collections held in museums and libraries, whereas the CMCL/PAThS is based on reconstructed codices, which is very important in the case of documents from the White Monastery Library, which were scattered throughout the world.

³⁸ https://www.trismegistos.org/network/6_2015_03_13/?search=Shenoute#Horos/, from Miyagawa (2019b)

3. Digital Archives

In contrast to Trismegistos and CMCL/PAThs, which are catalogues of metadata and peripheral information on the documents, digital archives mainly contain images of the documents themselves. Since the theme of the current paper is the digitization of manuscripts, I will use the term “digital archive” to refer to databases of media (here, images of manuscripts)³⁹ with their metadata. Although digital archives often include the metadata of the manuscripts, such data is often limited to collections of one institute or a narrow genre. In contrast, Trismegistos and CMCL/PAThs address all the related manuscripts, regardless of which collection they are found in. Therefore, digital archive databases usually have far fewer entries than Trismegistos and CMCL/PAThs. Nevertheless, they are very valuable tools, as they provide researchers with images of the most needed literature. Some of the digital archives described below have adopted the IIIF⁴⁰, an international framework for the mutual use of images that has gained attention in recent years and is expected to promote both mutual and secondary use. Digital archives such as BerlPap⁴¹, on the other hand, read and display external textual data on Papyri.info⁴² through APIs⁴³.

3.1 DigiVatLib⁴⁴ (Vatican Library)

The Vatican City is a city-state surrounded by the far larger city of Rome. It is home to the Pope, the supreme leader of the Roman Catholic Church and the Vatican’s head of state. Among the country’s most important buildings are the Vatican Palace, St. Peter’s Basilica, the Vatican

³⁹ Since the main focus of this paper is the digitization of manuscripts, “media” here refers to images of manuscripts.

⁴⁰ Shortened for “International Image Interoperability Framework.” See <https://iiif.io/>, last accessed on May 9th, 2021.

⁴¹ In this paper, BerlPap is included in the digital archive rather than the digital corpus since BerlPap brings the data of digital texts from outside, such as from <http://papyri.info/>.

⁴² <http://papyri.info/>, last accessed on May 9th, 2021.

⁴³ Shortened for Application Programming Interface, a system which connects one application with others.

⁴⁴ <https://digi.vatlib.it/>, last accessed on March 31st, 2019. The Vatican Library and DigiVatLib are also discussed in Miyagawa, 2019c.

Museums, and, not least, the Vatican Library⁴⁵, which houses valuable documents from around the world. One of the Library's holdings is the Borgia Collection, amassed by Cardinal Borgia in the 18th century, which includes Coptic documents. Of these, non-biblical documents such as Shenoute's have been transferred to the "Vittorio Emanuele III" National Library in Naples⁴⁶; however, specific manuscripts, mainly the Biblical manuscripts translated into Coptic, remain in the Vatican. The Vatican Library, in cooperation with NTT Data in Japan and others, has been working to digitize these manuscripts and has opened its digital archive, DigiVatLib, to the public⁴⁷. The archive contains not only historical manuscripts in Western languages, but also Coptic manuscripts, Japanese Christian documents, and many other valuable items. The images of the documents available on DigiVatLib are based on the International Image Interoperability Framework (IIIF)⁴⁸ and can therefore be viewed and annotated by external IIIF-compliant viewers⁴⁹ by inputting the URL of the JSON file⁵⁰, called the "IIIF Manifest".

3.2 BodmerLab⁵¹

The IIIF is being actively implemented in the digital archives of many museums and libraries that deal with Egyptian and Greek manuscripts, among them BodmerLab, the online digital archive of the Bodmer Collection of the Bodmer⁵² Foundation, whose beta version has

⁴⁵ Italian: Biblioteca Apostolica Vaticana, which translates to "Vatican Apostolic Library"; since "Vatican Library" is more common in English, however, I will use this term.

⁴⁶ Italian: Biblioteca Nazionale di Napoli "Vittorio Emanuele III"; Buzi (2009) is a catalogue of this Borgia collection transferred to the National Library of Naples. Miyagawa (2018a) is a Coptic text, Japanese translation, and grammatical commentary on a fragment of a manuscript of the writings of Abbot Besa, one of the late antique Egyptian monastic leaders and successor of Shenoute, in the Borgia Collection of the "Vittorio Emanuele III" National Library in Naples.

⁴⁷ For the digital archiving project of the Vatican Library by NTT Data, see <http://www.nttdata.com/jp/ja/services/sp/dataforthefuture/>, last accessed on March 30th, 2019.

⁴⁸ <https://iiif.io/>, last accessed on March 31st, 2019.

⁴⁹ For a detailed description of this method, see Miyagawa, 2019c.

⁵⁰ JSON stands for JavaScript Object Notation, which is a light-weight format for data interchange.

⁵¹ <https://bodmerlab.unige.ch/>, last accessed on March 31st, 2019.

⁵² The name of this foundation was named after its founder; yet although Martin Bodmer was born in 1899 in the German-speaking city of Zurich, the Bodmer Foundation is located in the French-speaking city of Cologne, Switzerland.

been open to the public since October of 2018⁵³. The Bodmer Collection is famous among Coptic and Biblical scholars for its very old Greek and Coptic Biblical papyri. Papyrus Bodmer 6, for example, one of the oldest Coptic Bible translations, is a translation of the Old Testament book of Proverbs into the ancient Proto-Sahidic dialect⁵⁴ (also known as the P dialect, Paleo-Theban dialect, etc.), written around the 3rd century. Figure 5 shows Papyrus Bodmer 6 opened with Mirador, a IIIF viewer⁵⁵.

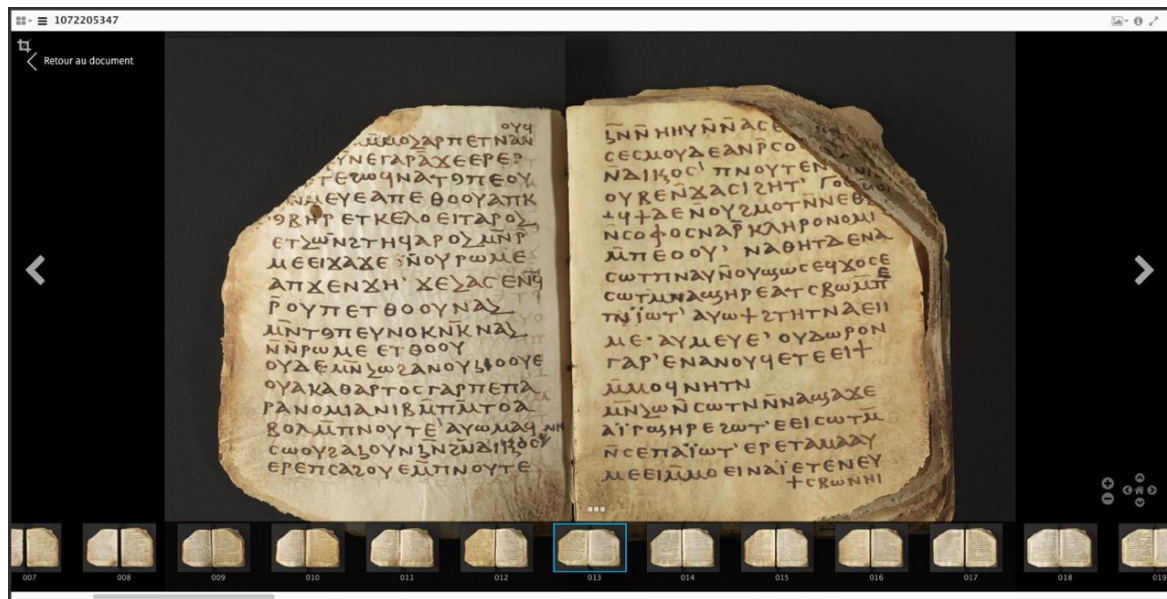


Fig. 5 Screenshot of the IIIF image of Papyrus Bodmer 6, p. 13, opened in Mirador⁵⁶.

3.3 Gallica⁵⁷ (National Library of France)

The Bibliothèque Nationale de France (National Library of France) also houses a large number of Coptic manuscripts, including those of

⁵³ For a more detailed description of BodmerLab, see Miyagawa, 2018e.

⁵⁴ For more information on this dialect, see Kasser, 1994, 1982.

⁵⁵ In addition to Mirador, IIIF-compatible viewers include Universal Viewer (<https://universalviewer.io/>, last accessed on July 29th, 2021) and IIIF Curation Viewer (<http://codh.rois.ac.jp/software/iiif-curation-viewer/>, last accessed on March 31st, 2019).

⁵⁶ <https://bodmerlab.unige.ch/fr/constellations/papyri/mirador/1072205347?page=013>, last accessed on March 31st, 2019.

⁵⁷ <https://gallica.bnf.fr/>, last accessed on March 31st, 2019. Note that a more detailed description of Gallica and IIIF is given in Miyagawa (2019d).

Shenoute⁵⁸. Originally built by Charles V as a royal library in 14th-century Paris, it was renamed the Bibliothèque Nationale de France during the French Revolution. Having been known as the Imperial National Library during the 19th century, it returned to its current name in 1994. The Bibliothèque Nationale stores its documents in seven buildings, mostly located in Paris. It operates the *Gallica* digital archive, which is one of the largest in the world. One can⁵⁹ acquire an IIIF Manifest and, by using it, display the image on one's own IIIF viewer installed on their own website, such as Mirador, Universal Viewer, or IIIF Curation Viewer⁶⁰.

3.4 Biblissima⁶¹

Biblissima is another important digital archive developed in France. Its IIIF Collections - Manuscripts & Rare Books⁶² is a collection of IIIF images from digital archives that provide IIIFs, such as Gallica, British Library, Bodleian Library, Europeana Regia⁶³, and e-codex⁶⁴. As of March 31st, 2019, there were only 12 Coptic documents on Biblissima imported from Gallica through IIIF, but it is expected that this number will increase with the development of Gallica and other digital archives.

3.5 Bodleian Library

Britain has a long history of purchasing, or otherwise acquiring, valuable documents from all over the world as a result of its vast empire and economic and political might. In addition to the many collections in London, the Ashmolean Museum, Bodleian Library, and Clarendon Press in Oxford, the Fitzwilliam Museum in Cambridge, the John Rylands Library and Manchester Museum in Manchester, and the Garstang Museum in Liverpool all have valuable collections, many in some form of digital archive. The Digital Bodleian⁶⁵, the digital archive of the Bodleian Library

⁵⁸ See Louis (2008), for the story of how many of the manuscripts found in the library of the White Monastery came to the Bibliothèque Nationale de France.

⁵⁹ This method is described in Nagasaki (2016).

⁶⁰ <https://gallica.bnf.fr/accueil/fr/content/accueil-fr?mode=desktop/>, last accessed on March 31st, 2019.

⁶¹ <http://beta.biblissima.fr/>, last accessed on March 31st, 2019. A more detailed description of Biblissima is given in Miyagawa (2019d).

⁶² <https://iiif.biblissima.fr/collections/>, last accessed on March 31st, 2019.

⁶³ <https://pro.europeana.eu/project/europeana-regia>, last accessed on July 29th, 2021.

⁶⁴ <https://www.e-codex.eu/>, last accessed on July 29th, 2021.

⁶⁵ <https://digital.bodleian.ox.ac.uk/>, last accessed on March 31st, 2019.

at the University of Oxford, was one of the first to adopt the IIIF and has become a model of IIIF-compliant digital archives. Nonetheless, although the Bodleian Library holds Coptic manuscripts, no IIIF images of this part of its collection were available on Digital Bodleian as of March 27th, 2019⁶⁶.

3.6 British Library

The British Library⁶⁷ in London, established in 1973 as an independent library of the British Museum, has a large collection of Coptic and Old Nubian manuscripts, such as those found at Edfu in Upper Egypt and those found in the remains of the White Monastery Library, especially the manuscripts of the texts of Abbot Besa. Other famous manuscripts often belong to the Greek Bible, such as the Codex Sinaiticus⁶⁸ and Codex Alexandrinus. When the Library separated from the Museum, Coptic manuscripts were placed under the jurisdiction of the former, while pre-Coptic Egyptian manuscripts written in Hieroglyphs, Hieratic, and Demotic scripts were placed under the jurisdiction of the latter. The British Library is also taking an advanced approach, and has opened its digital archives to the public through both old and new viewers. The new viewer is based on the Universal Viewer, which supports IIIF⁶⁹. However, images of few Coptic documents are available, and most still use the old viewer.

3.7 BerlPap⁷⁰

The Egyptian Museum and Papyrus Collection⁷¹ is the institution that manages the Egyptian collection of the Neues Museum on the Museum Island, one of many museums in Berlin. All the museums on the

⁶⁶ In Digital Bodleian, a search for “Coptic” yielded one hit, but it was in Arabic and not a Coptic bilingual reference:

<https://digital.bodleian.ox.ac.uk/inquire/Discover/Search/#/?p=c+1,t+Coptic,r+srs+O,rsp+10,fa+,so+ox%3A%5Easc,scids+,pid+,vi+/>, accessed on March 31st 2019.

⁶⁷ <https://www.bl.uk/>, last accessed on March 31st, 2019. The British Library’s catalogue of Coptic literature in printed form is Layton (1987).

⁶⁸ The digital edition of the Sinai Codex is available at <http://codexsinaiticus.org/> (last accessed on March 31st, 2019).

⁶⁹ “There’s a new viewer for digitised items in the British Library’s collections” (December 7th, 2016) <https://blogs.bl.uk/digital-scholarship/2016/12/new-viewer-digitised-collections-british-library.htm/> (last accessed on March 31st, 2019).

⁷⁰ <https://berlpap.smb.museum/>, last accessed on March 31st, 2019.

⁷¹ German: *Ägyptisches Museum und Papyrussammlung*.

Island – the Altes Museum, the Neues Museum, the Pergamon Museum, the Bode Museum, and others – have different institutions managing each collection. For example, the Islamic collection in the Pergamon Museum is managed by the Museum of Islamic Art⁷². Recently, the Egyptian Museum and the Papyrus Collection launched a digital archive service called BerlPap, which makes its papyrus manuscripts available online.

Papyrus is a type of paper made by stretching, combining, and drying the fibers of the papyrus plant, a member of the *Apiaceae* family which is mainly found along the banks of the Nile River. Until parchment became popular during the Roman Empire, papyrus was the main recording medium, alongside ostraca and inscriptions. Due to its simplicity, it was used for a wide range of purposes, including religious, administrative, economic, and literary documents, and for multiple languages, including Egyptian (mainly Hieratic, Demotic, and Coptic scripts, with some written Hieroglyphs such as the Book of the Dead), Greek, Latin, Old Nubian, Meroitic, Pahlavi, Aramaic, and Arabic.

BerlPap uses photographs of papyrus documents on their own, as well as texts and translations from Papyri.info, and mainly takes metadata of manuscripts from Trismegistos, as explained in the Digital Editions section. Papyri.info, also discussed in the Digital Editions section, is a website for reprinting papyrus documents. Users can download reprints and metadata in TEI XML format from BerlPap. As of March 31st, 2019, the website only contained bilingual documents in Greek and Coptic. The number of Coptic manuscripts on this website will likely increase in the future. The main source are the Greek papyri excavated in Egypt. Their images can be viewed in the DFG Viewer⁷³.

3.8 Heidelberg University Library Digital Papyrus Collection⁷⁴ and Europeana Collections

The Coptic and Greek papyri held by the Heidelberg University Library are also famous among scholars, in particular the collection of Coptic magical papyri. The library maintains one of the most advanced

⁷² German: *Museum für Islamische Kunst*.

⁷³ See the following slides by Sebastian Mayer presented at the KIM-Workshop at the University Library Mannheim, titled “IIIF im Kontext des DFG-Viewers”: https://wiki.dnb.de/download/attachments/132748423/2018-04-11_KIMWS18_Meyer-IIIF%20DFG-Viewer.pptx?version=1&modificationDate=1523606567000&api=v2/, last accessed on March 31st, 2019.

⁷⁴ German: Heidelberg Papyrussammlung - digital. https://www.ub.uni-heidelberg.de/helios/digi/hd_papyrus.html/, last accessed on March 31st, 2019.

digital archives of papyrus literature in Europe, linked to various web archives and databases through the BerlPap API and Papyri.info API. Papyrus images are licensed under CC BY-SA 4.0⁷⁵ and IIF, and can therefore be browsed externally⁷⁶. The Europeana Collections also use the Library's IIF manifest. Fig. 6 shows an example of the Europeana Collection's secondary use of the image of P. Heid. Inv. Kopt. 193 using the IIF manifest of Heidelberg University Library.

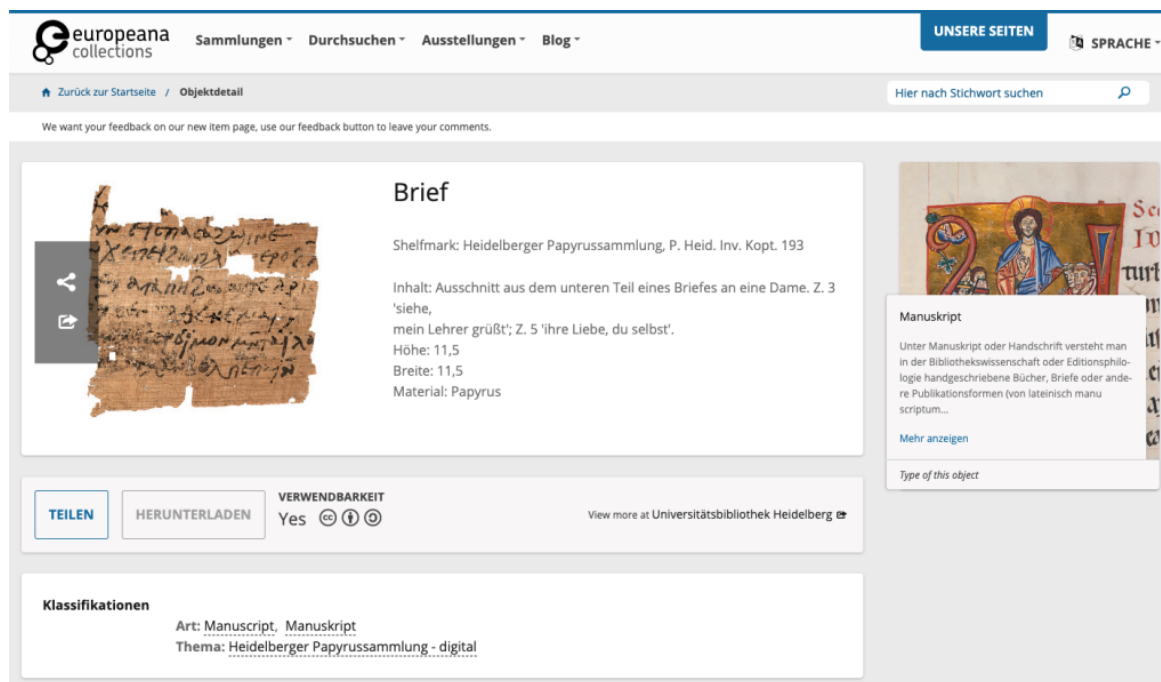


Fig. 6 Secondary use of images from P. Heid. Inv. Kopt. 193 in Europeana Collections⁷⁷ through IIF.

⁷⁵ For more information on the different types of Creative Commons licenses, see slide 23 of the slides made by Francis Bond named “The Great Game: Sherlock in Popular Culture Sherlock Holmes and Herlock Sholmes” (<http://compling.hss.ntu.edu.sg/courses/hg8011/pdf/hg8011-09-game.pdf/>, last accessed on March 31st, 2019).

⁷⁶ <https://www.europeana.eu/portal/en/search?q=what%3A%22Heidelberger+Papyrussammlung+-+digital%22&view=grid/>, last accessed on March 31st, 2019.

⁷⁷ http://www.europeana.eu/portal/de/record/o7932/diglit_p_kopt_193.html?q=what%3A%22Heidelberger+Papyrussammlung+-+digital%22#dcId=1554038967569&p=80/, last accessed on March 31st, 2019.

3.9 DVCTVS⁷⁸

The Palau Ribes Collection was assembled by the collector Palau Ribes, and is housed in a Jesuit institution in the Catalan city of Barcelona. Among the important Coptic and Greek papyri that it contains are the complete Coptic Sahidic dialect translations of the Gospels of Mark, Luke, and John. The University of Pompeu Fabra⁷⁹, a relatively new university in Barcelona (established 1990), maintains an online digital archive of the documents in these collections called DVCTVS. Users can download the image data of the manuscripts as they are, with a T.M. number and linked to the entry of the literature in question in Trismegistos⁸⁰. As of April 28th, 2019, DVCTVS did not support IIIF.

3.10 Cleo⁸¹

Cleo is a digital archive created by a Dutch team led by Heleen Wilbrink, with support from the SIDN Fund and the Google Cloud Startup Program. It provides access to the ancient Egyptian collections of the National Museum of Antiquities in Leiden⁸², the Metropolitan Museum of Art in New York, and the Brooklyn Museum, including papyrus documents. A special feature of the site is that it uses artificial intelligence technology for its query search engine, making it easier and more efficient for users to search.

3.11 Other collections

In addition, the Leipzig Library has digitally published the Ebers Papyrus, a medical papyrus written in the ancient Egyptian Hieratic script, as an IIIF image with annotated translation⁸³. This single papyrus could serve as a model of annotated IIIF documents for Digital Humanities in

⁷⁸ <http://dvctvs.upf.edu/>, last accessed on March 31st, 2019.

⁷⁹ Catalan: *Universitat Pompeu Fabra*.

⁸⁰ See, for example,

<http://dvctvs.upf.edu/catalogo/ductus.php?operacion=introduce&ver=1&nume=340/> (last accessed on March 31st, 2019). This page is a part of a Coptic Sahidic dialect translation of *The Gospel According to John*, in P.PalauRib.inv.183.

⁸¹ <https://cleo.aincient.org/pages/en/>, last accessed on March 31st, 2019.

⁸² Dutch: *Rijksmuseum van Oudheden*.

⁸³ <https://papyrusebers.de/>, last accessed on March 31st, 2019.

Egyptology. The Austrian National Library⁸⁴ also houses important Coptic manuscripts, including some from the White Monastery Library, and has published images of some of these. On December 5th, 2018, it was announced that an online digital archive would be opened by the Chester Beatty Library in Dublin, Ireland, whose holdings include the oldest manuscript fragments of the New Testament, Coptic Manichaean manuscripts, and Biblical manuscripts⁸⁵. Today, the Chester Beatty Digital Collection offers many images of various precious Coptic manuscripts available online with the IIF Manifests⁸⁶.

⁸⁴ German: *Österreichische Nationalbibliothek*. <https://www.onb.ac.at/>, last accessed on March 31st, 2019.

⁸⁵ <https://viewer.cbl.ie/viewer/>, last accessed on April 11th, 2021.

⁸⁶ <https://viewer.cbl.ie/viewer/index/>, last accessed on May 9th, 2021. See also Miyagawa, 2020.

4. Digital Editions⁸⁷

Thus far, the present paper has described the development of digital editions, which mainly provide images and metadata of literature. Although such data is already very useful, the “new” type of philologists, who use computerized methods to analyze the literature, as well as linguists who perform statistical analysis of corpora, require texts that are encoded in Unicode, digitized, and tagged with philological or linguistic information. Digital editions can be used to digitize texts, tag them with page, column, line information, other ornamental characters, parts of speech, or lemmata, and make them available digitally, mainly on the Web, as philologists have previously done in paper editions. The Text Encoding Initiative (TEI) is central to this work. The TEI⁸⁸ is an international organization that aims to develop standards for the markup of texts for the Humanities.

The file format used is the Extensive Markup Language (XML), which itself has a wide range of uses; for example, Microsoft Word’s .docx extension, Excel’s .xlsx, PowerPoint’s .pptx, etc. are all types of XML files. The tags in XML themselves are infinitely extensible, and can be customized by the user so that even the most complex documents can be marked up for machine readability. In the Humanities, XML can be used to mark up the textual features and linguistic information of texts by tagging. However, if each researcher marks up the texts in their own manner, it is difficult for them to show each other the files or develop each other’s work together. The TEI was consequently established in 1987 in order to promote collaboration among researchers and projects by standardizing text mark-up. The latest version of TEI is called TEI P5⁸⁹. Today, TEI XML is the standard format for digitizing texts in the Digital Humanities⁹⁰.

⁸⁷ There are currently a variety of digital edition projects in Europe, North America, and Israel. For a list of such projects, see the Catalogue of Digital Editions compiled by Greta Franzini (<https://dig-ed-cat.acdh.oeaw.ac.at/>, last accessed on March 31st, 2019).

⁸⁸ <https://tei-c.org/>, last accessed on March 31st, 2019.

⁸⁹ For guidelines on P5, see <https://tei-c.org/guidelines/p5/> (last accessed on March 31st, 2019).

⁹⁰ The Niedersächsische Staats- und Universitätsbibliothek Göttingen is developing a TEI XML-based digital edition of Theodor Fontane’s notebooks as a project of DARIAH-EU, a European digital infrastructure project. <https://fontane-nb.dariah.eu/test/index.html/>, last accessed on March 31st, 2019.

The following important Coptic digital editions are further described below: the New Testament Virtual Manuscript Room, a digital edition of the New Testament manuscripts; the Digital Edition of Coptic Old Testament, a digital edition project of the Coptic Old Testament; and Papyri.info, a user-participatory digital edition of documentary papyri.

4.1 New Testament Virtual Manuscript Room⁹¹

The Institute for the New Testament Textual Studies at the University of Münster⁹², founded by the New Testament scholar Kurt Aland, is famous for its editions of the Nestle-Aland edition of the Greek New Testament⁹³, which constitutes the basis for many modern Bible translations. It also edits the *Editio Critica Maior* (ECM), which encompasses all the textual variants that could not be included in the Nestle-Aland edition⁹⁴. In order to contribute to the ECM, the Virtual Manuscript Room (VMR) is used to carry out the digital revision process: reprinting the manuscript, creating a diplomatic edition of it, and outputting it to HTML and Epidoc, a subset of TEI XML, viewing and adjusting images, publishing it online as a digital edition, visualizing variants, etc. (Griffitts 2017). Epidoc was originally designed for the markup of Greek inscriptions, but it can also be used for documents in other languages and other media, such as papyrus, parchment, and paper. The VMR is designed to foster collaboration among scholars; hence, each user logs in to the portal and all changes made to the file are stored in Git for version control.

The VMR is the brainchild of Troy Griffitts of the University of Birmingham's Institute for Textual Studies and Electronic Editing, who has been developing the CrossWire Bible Society's Bible study software, and of Ulrich Schmid⁹⁵ of the Wuppertal/Bethel Theological College⁹⁶. The two researchers/developers are currently working on the Digital Edition of the Coptic Old Testament project at the Göttingen Academy of Sciences and Humanities. The New Testament Virtual Manuscript Room (NT-VMR) website, published by the Institute for New Testament Texts

⁹¹ <http://ntvmr.uni-muenster.de/de>, last accessed on March 31st, 2019. Note that the Virtual Manuscript Room is also described in detail in Miyagawa, 2018c.

⁹² German: Institut für neutestamentliche Textforschung (INTF).

⁹³ The current latest edition is the 28th edition (Nestle *et al.* 2012).

⁹⁴ See http://egora.uni-muenster.de/intf/projekte/ecm_en.shtml/ (last accessed on March 31st, 2019).

⁹⁵ He is also a prominent researcher on Marchion; see Schmid, 2012.

⁹⁶ German: *Kirchliche Hochschule Wuppertal/Bethel*.

at the University of Münster, uses this application to view diplomatic editions of Greek, Coptic, Latin, and Syriac manuscripts of the New Testament. Dissimilarities between manuscripts can be visualized by using CollateX, as shown in Figure 7.

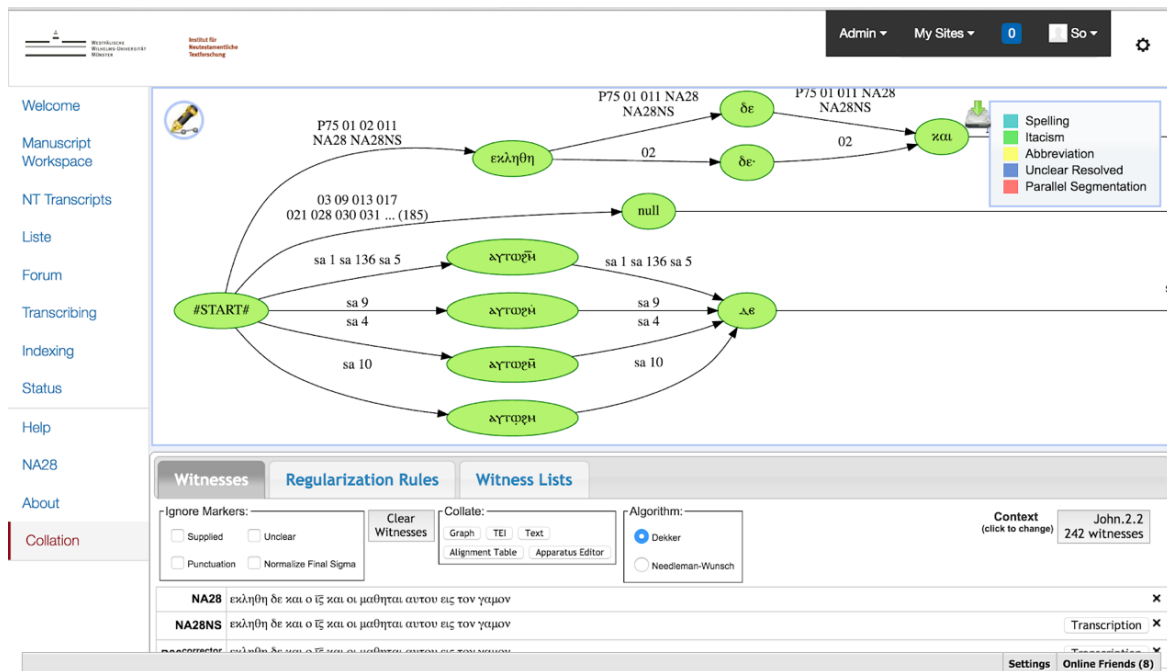


Fig. 7 Visualization of similarities and dissimilarities between manuscripts using NT-VMR with CollateX⁹⁷.

4.2 Digital Edition of the Old Testament in Coptic⁹⁸

While the NT-VMR is nearing completion, another project currently using the VMR is Göttingen Academy of Sciences and Humanities’⁹⁹ Digital Edition of the Coptic Old Testament project, which began in 2015. The project is led by Coptacists Heike Behlmer and Frank Feder, who have been working on the digital edition of the Coptic Old Testament¹⁰⁰. They aim to collect manuscripts of the Coptic Old Testament that have not yet been

⁹⁷ http://ntvmr.uni-muenster.de/de_DE/collation/, last accessed on March 31st, 2019.

⁹⁸ <http://coptot.manuscriptroom.com/>, last accessed on March 31st, 2019.

⁹⁹ German: *Akademie der Wissenschaften zu Göttingen*.

¹⁰⁰ See Behlmer & Feder, 2017.

fully edited¹⁰¹, compile a diplomatic edition using VMR, and, eventually, publish a critical edition. Although only part of the diplomatic editions of all the manuscripts of the Coptic Old Testament are currently available to the broader public, diplomatic editions of numerous manuscripts, visualization of variants, as well as the critical edition will be available in the future. To a certain extent, this project was a continuation of the *Biblia Coptica*¹⁰², a non-digital cataloguing project of Coptic Bible manuscripts under Karlheinz Schüssler in Austria, aiming to develop a project on the Coptic Old Testament already being run by Peter Nagel, an emeritus professor at the Martin Luther University Halle-Wittenberg.

4.3 The Canons of Apa Joannes the Archimandrite¹⁰³

The Canons of Apa Joannes the Archimandrite is a digital edition of the writings of Coptic liturgical scholar Diliانا Atanassova, who studied under Schüssler and is now based in Göttingen, under the auspices of the Austrian Science Fund¹⁰⁴. Johannes, commonly called “Apa Johannes” to distinguish him from the other Johannes, was the abbot of the so-called “White Monastic Federation” of the White, Red, and Women’s Monasteries founded by Shenoute’s uncle, Pcol (Pjol/Pgol). This project is a digital rendering of the diplomatic editions¹⁰⁵ and images of manuscripts of Johannes’ letters and sermons, compiled by the monks for the discipline and practice of the monastery. This digital edition was created using VMR in cooperation with the “Digital Edition of the Coptic Old Testament” project, and is available online through the servers of the Göttingen Academy of Sciences and Humanities, Germany. Figure 8 shows the web page using the VMR.

¹⁰¹ See Toda, 2016, pp. 172-179.

¹⁰² The results have been published by Harrassowitz Verlag in the “Biblia Coptica: Die koptische Bibeltexte” series (Schüssler 1995, 1996, 1998, 2000, 2003, 2004, 2006, 2007, 2009, 2010, 2011, 2012, 2015).

¹⁰³ <http://coptot.manuscriptroom.com/de/web/apa-johannes/>, last accessed on March 31st, 2019. Grant number FWF Project P22641-G19.

¹⁰⁴ German: *Fonds zur Förderung der wissenschaftlichen Forschung* (FWF).

¹⁰⁵ This is an edition in which the text is transcribed exactly as it appears in the manuscript. In many cases, the layout is close to that of the manuscript. On the other hand, an edition in which the text is revised by comparing the differences between various manuscripts is, aptly, called a critical edition.

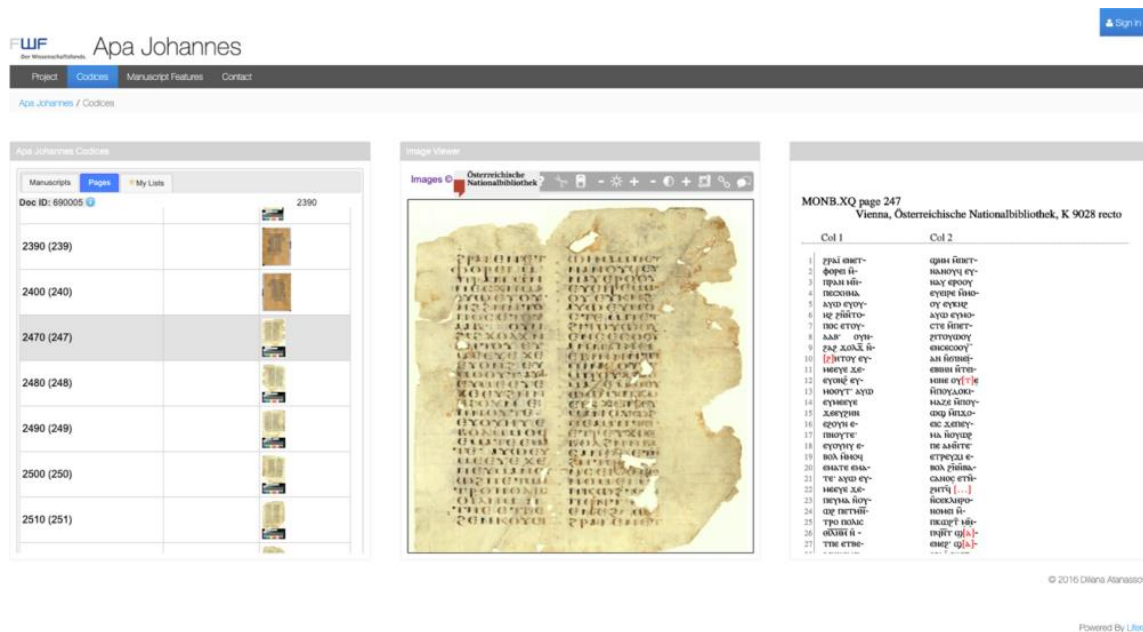


Fig. 8 Page from the MONB.XQ codex, p. 247, Austrian National Library K 9028 recto¹⁰⁶.

4.4 Papyri.info¹⁰⁷

Papyri.info is a website developed by the Duke Collaboratory for Classics Computing at Duke University and the Institute for the Study of the Ancient World at New York University in the United States, in collaboration with various researchers and institutions in Europe¹⁰⁸. It focuses on diplomatic editions of documentary papyri added through the site's online editor using a mark-up schema for papyrology called Leiden+. However, the data itself is recorded in Epidoc, a subset of TEI XML which can be downloaded by clicking on the link labeled XML on the page of each diplomatic edition. The digital diplomatic editions of Coptic papyri and ostraca on the website is supervised by Alain Delattre, a scholar of Coptic papyrology at the Free University of Brussels, and by other renowned experts. BerlPap, mentioned in Section 3.6, uses the Papyri.info API to make secondary use of the textual data of Papyri.info.

¹⁰⁶ <http://coptot.manuscriptroom.com/de/web/apa-johannes/>, last accessed on March 31st, 2019.

¹⁰⁷ <http://papyri.info/>, last accessed on March 31st, 2019.

¹⁰⁸ It is based on data from the Duke Databank of Documentary Papyri; see <https://papyri.info/docs/ddbdp/>, last accessed on March 10th, 2021.

5. Digital corpora and tools

The corpus discussed in this chapter is a Unicode-encoded, digitized text that has been tagged with linguistic information such as parts of speech and lemmata to enable linguistic analysis. In the sections below, I discuss the tools that can be used to mark up the corpus for textual re-use, such as quotations and allusions.

5.1 Coptic SCRIPTORIUM and KELLIA

Coptic SCRIPTORIUM¹⁰⁹, officially titled the Sahidic Corpus Research: Internet Platform for Interdisciplinary Multilayer Methods, is a project which aims to make available a linguistically and philologically analyzed multi-layered corpus of texts from Coptic literature online, enabling researchers to easily conduct linguistic analysis. It is currently directed by Amir Zeldes of Georgetown University, Washington, DC and Caroline Schroeder of the University of Oklahoma, and is funded by the National Endowment for the Humanities.

When the project was initiated in 2013, Zeldes was a doctoral student at Humboldt University Berlin, working on “KOMeT: Corpus linguistics methods for ePhilology with TEI XML”¹¹⁰, a project supported by the German Ministry of Research and Education to create a corpus of Besa’s letters in ANNIS, the prototype of the Coptic SCRIPTORIUM. Schroeder, meanwhile, was a visiting researcher at the University of Göttingen. Many people from Germany, especially Göttingen, are still working on the Coptic SCRIPTORIUM, including Alin Suciu, Diliانا Atanassova, and Heike Behlmer from the Digital Edition of the Coptic Old Testament project.

Data has been added to the SCRIPTORIUM, partly due to the work of the Koptische/Coptic Electronic Language and Literature International Alliance (KELLIA), a project which aims to create collaborative platforms and standards to promote collaborations among Digital Humanities projects related to Coptic Studies.

The Coptic SCRIPTORIUM displays its corpus on the web through ANNIS, a corpus visualization tool devised by developers at the Humboldt

¹⁰⁹ See Schroeder & Zeldes, 2016. The website is <http://copticSCRIPTORIUM.org/>, last accessed on March 31st, 2019. The Coptic SCRIPTORIUM is partially described in Miyagawa, 2018b.

¹¹⁰ <https://korpling.german.hu-berlin.de/komet/>, last accessed on March 31st, 2019.

University of Berlin, including Amir Zeldes. ANNIS allows a corpus to be displayed as any number of freely configurable layers. The diplomatic edition layer, for example, reproduces all diacritical marks and punctuation; the normalized text layer standardizes spelling and symbols; the part-of-speech layer tags parts of speech; the lemma layer displays lemmata; the morphological layer divides words into morphemes; and the loanword layer shows whether a word is of Egyptian, Greek, or Semitic origin.

There is also a translation-into-English layer, a line number layer, a column number layer, and a page number layer. Thus, ANNIS enables Coptic SCRIPTORIUM to display both the linguistic and philological information of the corpus. Clicking on a lemma reveals its meaning because the lemmata are linked to the lexical database of Coptic Dictionary Online¹¹¹, which was developed by the KELLIA project in collaboration with the *Thesaurus Linguae Aegyptiae*¹¹² project run by the Berlin-Brandenburg Academy of Sciences and Humanities.

Syntactic information is also tagged by Universal Dependencies¹¹³, and for certain corpora syntactic trees can also be viewed in ANNIS. Figure 9 depicts a corpus from the Coptic SCRIPTORIUM opened in ANNIS to show the annotation layers and syntactic tree. Tagging is automatic, and the Coptic NLP Service¹¹⁴, currently available as a product of the KELLIA project, automatically tags the input Coptic text for morphological analysis, part-of-speech information, lemma information, syntactic information, and whether it is a loanword or not, in SGML format. Tools such as the tokenizer and lemmatizer are available on GitHub¹¹⁵. After automatic tagging has been carried out, the text is checked manually.

Currently, the SCRIPTORIUM contains a multi-layered corpus of Coptic New Testament and Old Testament texts, as well as monastic literature by writers such as Shenoute, Besa, and Apa Johannes, which were written in the Sahidic dialect of Coptic, literary texts such as the *Apophthegmata Patrum*, which were translated into the Sahidic dialect, and more.

¹¹¹ <https://corpling.uis.georgetown.edu/coptic-dictionary/>, last accessed on March 31st, 2019.

¹¹² <http://aaew.bbaw.de/tla/index.html/>, last accessed on March 31st, 2019.

¹¹³ <https://universaldependencies.org/>, last accessed on March 31st, 2019.

¹¹⁴ <https://corpling.uis.georgetown.edu/coptic-nlp/>, last accessed on March 31st, 2019.

¹¹⁵ <https://github.com/CopticScriptorium/>, last accessed on March 31st, 2019.

The screenshot shows the ANNIS interface with the search query `pos="VBD"`. The search results are displayed in a table format, showing the first 10 of 120 matches. The first match is from the `apophthegmata.patrum` corpus, specifically from the `norm_group` `ΔCΠΑΘΓΕ`. The search results are displayed in a table format, showing the first 10 of 120 matches. The first match is from the `apophthegmata.patrum` corpus, specifically from the `norm_group` `ΔCΠΑΘΓΕ`.

norm_group	ΔCΠΑΘΓΕ	.	ΔΥΘ	ΠΕΧΑ	Ε
norm	Δ	·	ΔΥΘ	ΠΕΧΑ	Ε
pos	APST	PPERS	V	PUNCT	CONJ
lemma	Δ	NTOC	ΠΑΘΓΕ	ΔΥΘ	ΠΕΧΕ
func	aux	nsbj	root	punct	cc
orig_group	ΔCΠΑΘΓΕ	.	ΔΥΘ	ΠΕΧΑ	Ε
orig	Δ	·	ΔΥΘ	ΠΕΧΑ	Ε
translation	When she saw him, she was amazed. And she said to him, "W				
p	p				
pb_xml_id	EG15				
cb_n	1				
lb_n	25				26

Fig. 9 Screenshot of *Apophthegmata Patrum* opened in ANNIS at the Coptic SCRIPTORIUM¹¹⁶.

5.2 eTRAP and the German Research Foundation Collaborative Research Centre 1136¹¹⁷

From 2015 to 2020, a project was underway in Göttingen to analyze intertextuality, such as quotations and allusions in Coptic literature, using digital tools. It is based on subproject B 05, “Scriptural Interpretation and Educational Traditions in the Coptic-speaking Egyptian Christianity of Late Antiquity: Shenoute, Canon 6”¹¹⁸, which is itself part of the Collaborative Research Centre (SFB) 1136 “Education and Religion in the Mediterranean and Surrounding Cultures from Antiquity to the Middle

¹¹⁶ From Miyagawa et al., 2018, p. 141.

¹¹⁷ The topic of this section, TRACER, is also discussed in Miyagawa, 2018d, 2018f.

¹¹⁸ German: *B 05 Schriftauslegung und Bildungstraditionen im koptischsprachigen ägyptischen Christentum der Spätantike*: Shenute, Kanon 6.

Ages and Classical Islam” project¹¹⁹, funded by the German Research Foundation (DFG). The authors aim to automatically find quotations, allusions, and other text re-uses between the Coptic monastic writings of Shenoute, Besa and the Coptic Bible by using recent applications: the TRACER¹²⁰ software developed in the eTRAP project, TRAViz software, and the Coptic SCRIPTORIUM tools (Coptic NLP Service). TRACER consists of more than 700 algorithms to find text re-uses such as quotations, while TRAViz¹²¹ software is used for visualization in HTML, JavaScript, and CSS by highlighting similarities or displaying them as textual alignments, as shown at the top of Figure 10. As a research associate of the eTRAP team, I am currently working on digital editions of Shenoute’s *Canon 6* and Besa’s letters and sermons on VMR, and running TRACER on the transcriptions comparing the text of the digital edition of the Coptic Bible provided by the Digital Edition of the Coptic Old Testament project to find many previously undiscovered quotations¹²². Miyagawa et al. (2018) advocate for tagging corpora that can be displayed in the ANNIS of the Coptic SCRIPTORIUM with information on quotations and allusions found by TRACER.

¹¹⁹ German: Sonderforschungsbereich 1136 “Bildung und Religion in Kulturen des Mittelmeerraums und seiner Umwelt von der Antike bis zum Mittelalter und zum Klassischen Islam.”

¹²⁰ Büchler *et al.*, 2014.

¹²¹ Jänicke *et al.*, 2015.

¹²² For example, TRACER found 14 quotations that previous researchers had failed to find in the writings of Shenoute’s *Canon 6* from the Coptic translation of the Old Testament Psalms (Miyagawa *et al.* forthcoming).



Fig. 10 Two Coptic texts (the Psalms in the Coptic Sahidic dialect and “Is It Not Written” from Shenoute’s *Canon 6*), automatically morphologically analyzed by the Coptic NLP Service, TRACER analysis, and TRAViz visualization. The upper screen shows the differences between the two texts using text alignment, and the lower screen shows the parallel text. Areas of high similarity are highlighted in blue in the parallel text (bottom).

This project machine-trained *OCROpy*, an optical character recognition (OCR)¹²³ software based on neural network models, to recognize Coptic letters and digitize the text of existing printed editions. The text was then used as the basis for transcriptions of photographs of the manuscripts in the VMR.

¹²³ For more on OCR for Coptic, see Lincke *et al.*, 2020, and Miyagawa *et al.*, 2019.

6. Conclusion

This paper has focused on the digitization of Coptic manuscripts, from digital catalogues through archives, editions, and corpora. Some have been licensed by the Creative Commons, forming the basis of open-linked data using TM numbers, geographic Pleiades numbers, or PAThs numbers, which constitute I.D.s for each codex. In the future, all these databases will be interlinked to enable easier data collection. Although the catalogues are fairly well developed and archives are being established, many manuscripts are still not accessible online. Digital editions are available only for documentary papyri, the Coptic Bible, and some of the Church Fathers and accounts of early monastic leaders, which, together, represent only a fraction of works in Coptic. It would take a great deal of time, money, and labor to reproduce all the Coptic manuscripts digitally and make the images and texts available online. However, a number of projects for the digitization of Coptic literature are now underway, especially in Europe and the United States. In addition to assigning I.D.s to each item in the codex, museums, and libraries, I.D.s for each work will also be necessary. The Canonical Text Service¹²⁴ has been developed for the Digital Humanities of Classics (Latin and Classical Greek), mainly at the University of Leipzig, with I.D.s assigned to each work. It is expected that such a service will also be needed for Coptic literature¹²⁵. The KELLIA project brought together multiple ongoing projects, a model which, from now on, will become increasingly important in the management of the digitization of Coptic manuscripts. Museums, libraries, and other organizations must work together. Several digital infrastructure projects are underway, such as DARIAH-DE¹²⁶ and CLARIN-D¹²⁷ in Germany¹²⁸, while research promotion organizations such as the German Ministry of Research and Education (BMBF) and German Research Foundation (DFG) are encouraging the promotion of Linked Open Data (LOD), a platform which

¹²⁴ See Tiepmar & Heyer, 2017.

¹²⁵ See Almas & Schroeder, 2016. Recently, Coptic SCRIPTORIUM is developing a canonical text service for the Coptic literature. See <http://data.copticscriptorium.org/>, last accessed on July 30, 2021.

¹²⁶ <https://de.dariah.eu/>, last accessed on March 31st, 2019.

¹²⁷ <https://www.clarin-d.net/>, last accessed on March 31st, 2019.

¹²⁸ Recently, the National Research Data Infrastructure (NFDI; Nationalen Forschungsdateninfrastruktur) of German Research Foundation (DFG; Deutsche Forschungsgemeinschaft) has been launched to promote digital infrastructure development for academic research projects in Germany.

has been introduced to link data openly using technologies such as RDF¹²⁹ across countries. Although I have only presented examples from Germany here, the number of such infrastructure projects is increasing world-wide, and the movement to encourage openness of data is spreading rapidly around the world. In this global context, digital catalogues, digital archives, digital editions, digital corpora, and other linguistic and geographic databases related to Coptic manuscripts are expected to be increasingly interlinked, promoting networking among projects and enhancing user convenience. It is hoped that this will further accelerate the digitization of Coptic manuscripts.

¹²⁹ RDF stands for “Resource Description Framework”.

References:

- Alcock, A., 2013a. "Triadon: English version of a 14th century religious poem written in Coptic and Arabic. Part One."
<https://suciualin.files.wordpress.com/2013/02/triadon.pdf>. Last accessed on March 31st, 2019.
- Alcock, A., 2013b. "Triadon: English version of a 14th century religious poem written in Coptic and Arabic. Part Two."
<https://suciualin.files.wordpress.com/2013/04/triadon-part-two.pdf>. Last accessed on March 31st, 2019.
- Almas, B. & Schroeder, C. T., 2016. "Applying the canonical text services model to the Coptic SCRIPTORIUM." *Data Science Journal* 15(3): 1-13. DOI: 10.5334/dsj-2016-013.
- Basta, M., 1991. "Iqladiyus Labib (1873-1918)". In Atiya, A. S. (ed.): *The Coptic Encyclopedia*, vol. 4. New York, Macmillan, 1302.
- Behlmer, H. & Feder, F., 2017. "The Complete Digital Edition and Translation of the Coptic Sahidic Old Testament. A New Research Project at the Göttingen Academy of Sciences and Humanities." *Early Christianity* 8: 97–107.
- Böhlig, A., 1966. *Manichäische Handschriften der Staatlichen Museen Berlin, Bd. 1: Kephalaia: Zweite Hälfte, (Lieferung 11/12, Seite 244–291)*. Stuttgart, Kohlhammer.
- Büchler, M., Burns, P. R., Müller, M., Franzini, E. & Franzini, G., 2014. "Towards a Historical Text Re-Use Detection." In Biemann, C. & Mehler, A. (eds.): *Text Mining, Theory and Applications of Natural Language Processing*. Basel, Springer International Publishing, pp. 221–238.
- Budge, E. A. W., 1913. *Coptic Apocrypha in the Dialect of Upper Egypt*. Oxford, Oxford University Press.
- Burdick, A., Drucker, J., Lunenfeld, P., Presner, T. & Schnapp, J., 2012. *Digital Humanities*. Cambridge, MIT Press.
- Busa, R. S. J., 1951. *Sancti Thomae Aquinatis Hymnorum Ritualium Vara Specimina Concordantiarum: A First Example of Word Index Automatically Compiled and Printed by IBM Punched Card Machines*. Vol. N. 7. Serie II. Milan, Fratelli Bocca.
- Buzi, P., 2009. *Catalogo Dei Manoscritti Copti Borgiani Conservati Presso La Biblioteca Nazionale "Vittorio Emanuele III" Di Napoli: Con Un Profilo Scientifico Di Stefano Borgia E Georg Zoega, E Una Breve Storia Della Formazione Della Collezione Borgiana*. Roma, Accademia Nazionale dei Lincei.

- Dreyer, G., 2011. "Tomb Uj: A Royal Burial of Dynasty 0 at Abydos." In Teeter, E. (ed.): *Before the Pyramids: The Origins of Egyptian Civilization*, Oriental Institute Museum Publications 33. Chicago, The Oriental Institute of the University of Chicago, pp. 131–138.
- Emmel, S., 2004. *Shenoute's literary corpus*. Two volumes. Leuven, Peeters.
- Fleming, H. C., 1969. "The Classification of West Cushitic within Hamito-Semitic." In McCall, D. F. et al. (ed.), *Eastern African History*. New York, Washington, London: Frederick A. Praeger, pp. 3–27.
- Frajzyngier, Z. & Shay, E., 2012. Introduction. In Frajzyngier, Z. & Shay, E. (eds): *The Afroasiatic Languages*. Cambridge, Cambridge University Press, pp. 1–17.
- Funk, W.-P., 2020. *Kephalaia I: Supplementa. Manichäische Handschriften der Staatlichen Museen zu Berlin*, Bd. 1,5. Stuttgart, Kohlhammer Verlag.
- Funk, W.-P., 2018. *Kephalaia I: Zweite Hälfte, Lieferung 17/18. Manichäische Handschriften der Staatlichen Museen zu Berlin*, Bd. 1. Stuttgart, Kohlhammer Verlag.
- Funk, W.-P., 2000. *Kephalaia I: Zweite Hälfte, Lieferung 15/16. Manichäische Handschriften der Staatlichen Museen zu Berlin*, Bd. 1. Stuttgart, Kohlhammer Verlag.
- Funk, W.-P., 1999. *Kephalaia I: Zweite Hälfte, Lieferung 13/14. Manichäische Handschriften der Staatlichen Museen zu Berlin*, Bd. 1. Stuttgart, Kohlhammer Verlag.
- Funk, W.-P., 1988. "Dialects Wanting Homes: A Numerical Approach to the Early Varieties of Coptic." In Fisiak, J. (ed.), *Historical Dialectology, Regional and Social. Trends in Linguistics. Studies and Monographs 37*, Berlin, Mouton de Gruyter. pp. 149–192.
- Gardner, I. (ed.), 1995. *The Kephalaia of the Teacher: The edited Coptic Manichaean texts in translation with commentary*. Leiden, Brill.
- Gardner, I. BeDuhn, J. D. & Dilley, P., 2014. *Mani at the Court of the Persian Kings: Studies on the Chester Beatty Kephalaia Codex*. Leiden, Brill.
- Greenberg, J. H., 1963. *The Languages of Africa. International Journal of American Linguistics*, Vol. 29, No. 1, Part II. Bloomington, Indiana University.
- Griffitts, T., 2017. *Software for the Collaborative Editing of the Greek New Testament*. Ph.D. Thesis, University of Birmingham.
- Jänicke, S., Geßner, A., Franzini, G., Terras, M., Mahony, S. & Scheuermann, G., 2015 "TRAViz: A Visualization for Variant Graphs."

- Digital Scholarship in the Humanities* 30 (suppl. 1): i83–i99. DOI: 10.1093/llc/fqvo49.
- Kaibara, A. [貝原哲生], 2015. *Religious Conflict and Its Development in 6th -7th Century Egypt: Based on a Regional Perspective*. [6–7 世紀エジプトにおける宗教的対立とその展開—地域的視点に立脚して—]. PhD thesis, Osaka City University.
- Kim, B.-Y. [金甫榮], 2019. “Archives and Digital Humanities: The Key to Successful Collaboration” [アーカイブズと人文情報学—コラボレーションを成功に導く鍵となるのは—]. *Digital Humanities Monthly* [人文情報学月報] 92(1).
- Kammerzell, F., 2000. “Egyptian possessive constructions: a diachronic typological perspective.” *STUF – Language Typology and Universals* 53(1): 97–108.
- Kasser, R., 1994. “Approche de la langue copte proto-saïdique.” In Berger, C., Clerc, G. & Grimal, N. (eds.), *Hommages à Jean Leclant*, Bibliothèque d’Étude 104, vol. IV. Le Caire, Institut Français d’Archéologie Orientale, pp. 152–162.
- Kasser, R., 1991. “Dialects.” In Atiya, A. S. (ed.), *Coptic Encyclopedia*, vol. 8. New York, Macmillan, pp. 87–97.
- Kasser, R., 1982. “Le dialecte protosaidique de Thebes.” *Archiv für Papyrusforschung* 28: 67–81.
- Layton, B., 1987. *Catalogue of Coptic Literary Manuscripts in the British Library Acquired Since the Year 1906*. London, British Library.
- Lefort, L.-T., 1956a. *Œuvres de s. Pachôme et de ses disciples*, CSCO 159. Louvain, L. Durbecq.
- Lefort, L.-T., 1956b. *Œuvres de s. Pachôme et de ses disciples*, CSCO 160. Louvain, L. Durbecq.
- von Lemm, O., 1903. *Das Triadon*. St. Petersburg, Commissionaires de l’Académie Impériale des Sciences.
- Lincke, E.-S., Bulert, K. & Büchler, M., 2019. “Optical Character Recognition for Coptic fonts: A multi-source approach for scholarly editions.” In *Proceedings of the 3rd International Conference on Digital Access to Textual Cultural Heritage*, pp. 87–91.
- Louis, C., 2008. “The Fate of the White Monastery Library.” In Gabra, G. & Takla, H. N. (eds.), *Christianity and Monasticism in Upper Egypt: Volume 1, Akhmim and Sohag*. Cairo, The American University in Cairo Press, pp. 83–90.
- MacCoull, L. S. B., 1997. “The *Triadon*: An English Translation.” *Greek Orthodox Theological Review*, 42: 83–148.

- Masoner, A., 2018. “Ein Jesuitenpater als Computerpionier”. science ORF.at. <https://science.orf.at/stories/2904212/>. Last accessed on March 30th, 2019.
- Miyagawa, S. [宮川創], 2018a. Linguistic materials and grammatical annotations in the Said dialect of Coptic: fragments of texts by Basa in the collection of the National Library of Vittorio Emanuele III, Naples [コプト語サイド方言の言語資料と文法注釈 –ナポリ・国立ヴィットーリオ・エマヌエーレ3 世図書館蔵・ベーサによるテキストの断片–]. *Journal of Kijutsuken: Descriptive Linguistics Study Groupe* [言語記述論集] 10: 271-320.
- Miyagawa, S. [宮川創], 2018b. A Cross-section of Digital Humanities Related to Late Antiquity in Sweden and the United States [スウェーデンおよびアメリカ合衆国における古代末期関連のデジタル・ヒューマニティーズの一断面]. *Current Trends of Digital Humanities in Europe and the Middle East* [欧州・中東デジタル・ヒューマニティーズ動向] 3. *Digital Humanities Monthly* [人文情報学月報] 83(1).
- Miyagawa, S. [宮川創], 2018c. New Trends in Digital Biblical Manuscript Studies: Virtual Manuscript Room [デジタル聖書写本学の新潮流–Virtual Manuscript Room–]. *Current Trends of Digital Humanities in Europe and the Middle East* [欧州・中東デジタル・ヒューマニティーズ動向] 6. *Digital Humanities Monthly* [人文情報学月報] 86(1).
- Miyagawa, S. [宮川創], 2018d. Research on Intertextuality and Text Re-Use in Digital Humanities [デジタル・ヒューマニティーズにおけるテキスト・リユースと間テキスト性の研究]. *Current Trends of Digital Humanities in Europe and the Middle East* [欧州・中東デジタル・ヒューマニティーズ動向] 7. *Digital Humanities Monthly* [人文情報学月報] 87(1).
- Miyagawa, S. [宮川創], 2018e. Bodmer Collection Opens Online Database of Manuscripts / University of Hamburg to Open Exzellenzcluster (German Research Foundation) for Manuscript Studies [ボドマー・コレクションが写本のオンライン・データベースを公開／ハンブルク大学が写本学のエクスツェレンツクラスター（ドイツ研究振興協会）を開設へ]. *Current Trends of Digital Humanities in Europe and the Middle East* [欧州・中東デジタル・ヒューマニティーズ動向] 8. *Digital Humanities Monthly* [人文情報学月報], 88(1).
- Miyagawa, S. [宮川創], 2018f. The History and Development of Intertextuality Studies and Re-use of Texts [テキスト・リユースと間テ

- クスト性研究の歴史と発展]. *Current Trends of Digital Humanities in Europe and the Middle East* [欧州・中東デジタル・ヒューマニティーズ動向] 9. *Digital Humanities Monthly* [人文情報学月報], 89(1).
- Miyagawa, S. [宮川創], 2019a. Digital Archive of Papyrus Literature in German-speaking Countries [ドイツ語圏のパピルス文献デジタル・アーカイブ]. *Current Trends of Digital Humanities in Europe and the Middle East* [欧州・中東デジタル・ヒューマニティーズ動向] 10. *Digital Humanities Monthly* [人文情報学月報] 90(1).
- Miyagawa, S. [宮川創], 2019b. Trismegistos: a set of web databases of metadata of Egyptian, Greek, Latin, Syriac and other languages' texts and related names of people and places from the 8th century B.C. to the 8th century A.D., and services of Linked Open Data [Trismegistos : 紀元前8世紀から紀元後8世紀までのエジプト語・ギリシア語・ラテン語・シリア語などの文献のメタデータや関連する人名・地名などのウェブ・データベース群, および, Linked Open Data のサービス]. *Current Trends of Digital Humanities in Europe and the Middle East* [欧州・中東デジタル・ヒューマニティーズ動向] 11. *Digital Humanities Monthly* [人文情報学月報], 91(1).
- Miyagawa, S. [宮川創], 2019c. Digital Archiving of Coptic Documents Corresponding to IIF (1): Vatican Library. [IIF に対応したコプト語文献のデジタルアーカイブ (1) : バチカン図書館]. *Current Trends of Digital Humanities in Europe and the Middle East* [欧州・中東デジタル・ヒューマニティーズ動向] 12. *Digital Humanities Monthly* [人文情報学月報] 92.
- Miyagawa, S. [宮川創], 2019d. Digital Archiving of Coptic Documents Corresponding to IIF (2): Bibliothèque Nationale de France and Biblissima. [IIF に対応したコプト語文献のデジタルアーカイブ (2) : フランス国立図書館と Biblissima]. *Current Trends of Digital Humanities in Europe and the Middle East* [欧州・中東デジタル・ヒューマニティーズ動向] 13. *Digital Humanities Monthly* [人文情報学月報] 93(1).
- Miyagawa, S. [宮川創], 2019e. Current Trends and Future of Digitization of Coptic Manuscripts in Europe [欧州におけるコプト語文献のデジタル化の現状と未来]. *Chūgoku* 21 [中国21] 51: 39–76.
- Miyagawa, S. [宮川創], 2020. Development of Chester Beatty's Digital Collection. Focusing on Egyptian papyrus documents. [Chester Beatty

- Digital Collections の発展：エジプト出土パピルス文献を中心に]. *Digital Humanities Monthly* [人文情報学月報] 105.
- Miyagawa, S., Yoshino, H. & Nagai, M. [宮川創・吉野宏志・永井正勝], 2018. Hieroglyphs and the Egyptian Language [ヒエログリフと古代エジプト語]. Introduction to Ancient Egyptian Hieroglyphs: Reading the Rosetta Stone [古代エジプト語のヒエログリフ：ロゼッタストーン読解] 1. *Hituzigusa: Hituzi Shobo's Web Magazine* [未草：ひつじ書房のウェブマガジン]. <http://www.hituzi.co.jp/hituzigusa/2018/08/31/hieroglyph-1/>. Last accessed on March 31st, 2019.
- Miyagawa, S., Yoshino, H. & Nagai, M. [宮川創・吉野宏志・永井正勝], 2019. Phonograms of Hieroglyphs: Review of One Consonantal Letters and Easy Divine Names [ヒエログリフの表音文字：一子音文字のおさらいとやさしい神名]. Introduction to Ancient Egyptian Hieroglyphs: Reading the Rosetta Stone [古代エジプト語のヒエログリフ：ロゼッタストーン読解] 6. *Hituzigusa: Hituzi Shobo's Web Magazine* [未草：ひつじ書房のウェブマガジン]. <http://www.hituzi.co.jp/hituzigusa/2019/01/21/hieroglyph-6/>. Last accessed on March 31st, 2019.
- Miyagawa, S., Büchler, M. & Behlmer, H., forthcoming. “Computational Analysis of Text Reuse/Intertextuality: The Example of Shenoute Canon 6.” In Takla, H. N., Emmel, S. & Mikhail, M. S. A. (eds.), *Proceedings of the 11th International Congress of Coptic Studies*. Leuven, Peeters.
- Miyagawa, S., Bulert, K., Büchler, M. & Behlmer, H., 2019. “Optical Character Recognition of Typeset Coptic Text with Neural Networks.” *Digital Scholarship in the Humanities* 34(suppl. 1): i135–i141. DOI: 10.1093/llc/fqz023.
- Miyagawa, S., Zeldes, A., Büchler, M., Behlmer, H. & Griffiths, T., 2018. “Building Linguistically and Intertextually Tagged Coptic Corpora with Open Source Tools.” In Suzuki, C. (ed.), *Proceedings of the 8th Conference of Japanese Association for Digital Humanities*. Tokyo, Center for Open Data in the Humanities, pp. 139–41.
- Nagasaki, K., 2016. Introduction to IIIF, An International Standard for Image Sharing in ‘Digital Archives’ (continued) [「デジタルアーカイブ」における画像共有のための国際規格IIIFについてのご紹介（続）]. digitalnagasaki's blog [digitalnagasakiのブログ], <http://digitalnagasaki.hatenablog.com/?page=1465181142>. Last accessed on March 31st, 2019.

- Nagasaki, K. & Shimoda, M. [永崎研宣・下田正弘], 2018. Advanced Utilization of Digital Archives Cultivated by Openness: Through the Operation of IIF Manifests for Buddhist Studies [オープン化が拓くデジタルアーカイブの高度利活用 : IIIF Manifests for Buddhist Studies の運用を通じて]. In *Proceedings of Jinmonkon 2018* [じんもんこん 2018 論文集], pp. 389–394.
- Nagel, P., 1983. *Das Triadon: ein sahidisches Lehrgedicht des 14. Jahrhunderts*. Halle (Saale), Abt. Wissenschaftspublizistik der Martin-Luther-Universität Halle-Wittenberg.
- Nestle, E., Nestle, E., Aland, B., Aland, K., Karavidopoulos, J., Martini, C. M. & Metzger, B. M., 2012. *Novum Testamentum Graece*. 28th ed. Stuttgart, Deutsche Bibelgesellschaft.
- Orlandi, T., 1990. “The Corpus dei Manoscritti Copti Letterari.” *Computers and the Humanities*, 24: 397–405.
- Polotsky, H. J. & Böhlig, A., 1940. *Kephalaia: Erste Hälfte, Lieferung 1–10*. Manichäische Handschriften der Staatlichen Museen Berlin, Bd. 1. Stuttgart, Kohlhammer.
- Rubenson, S., 1995. *The Letters of St. Antony: Monasticism and the Making of a Saint*. Minneapolis, Fortress Press.
- Said, E. W., 1978. *Orientalism: Western Conceptions of the Orient*. New York, Vintage Books.
- Satzinger, H., 1991. “Old Coptic.” In Atiya, A. S. (ed.), *The Coptic encyclopedia*, volume 8. New York, Macmillan, pp. 169b–175b.
- Satzinger, H., 1975. “The Old Coptic Schmidt Papyrus.” *Journal of the American Research Center in Egypt* 12: 37–50.
- Schmid, U., 2012. *Marcion und sein Apostolos: Rekonstruktion und historische Einordnung der marcionitischen Paulusbriefausgabe*. Berlin, Walter de Gruyter.
- Schroeder, C. & Zeldes, A., 2016. “Raiders of the Lost Corpus.” *Digital Humanities Quarterly* 10 (2): 1–13.
- Schüssler, K.-H., 2015. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten, sa 185 – 260*. Biblia Coptica, Bd. 2, Lfg. 2. Bearbeitet von Frank Feder und Hans Förster. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 2012. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten, sa 121 – 184*. Biblia Coptica, Bd. 2, Lfg. 1. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 2011. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten, sa 721–780*. Biblia Coptica, Bd. 4, Lfg. 4. Wiesbaden, Harrassowitz Verlag.

- Schüssler, K.-H., 2010. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten*, sa 673-720. Biblia Coptica, Bd. 4, Lfg. 3. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 2009. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten*, sa 621-672. Biblia Coptica, Bd. 4, Lfg. 2. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 2007. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten*, sa 586-620. Biblia Coptica, Bd. 4, Lfg. 1. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 2006. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten*, sa 561-585. Biblia Coptica, Bd. 3, Lfg. 4. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 2004. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten*, sa 541-560. Biblia Coptica, Bd. 3, Lfg. 3. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 2003. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten*, sa 521-540. Biblia Coptica, Bd. 3, Lfg. 2. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 2001. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten*, sa 500-520. Biblia Coptica, Bd. 3, Lfg. 1. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 2000. *Das sahidische Alte und Neue Testament: vollständiges Verzeichnis mit Standorten*, sa 93-120. Biblia Coptica, Bd. 1, Lfg. 4. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 1998. *Das sahidische Alte und Neue Testament: sa 49-9*. Biblia Coptica, Bd. 1, Lfg. 3, 2. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 1996. *Das sahidische Alte und Neue Testament: sa 21-48*. Biblia Coptica, Bd. 1, Lfg. 2. Wiesbaden, Harrassowitz Verlag.
- Schüssler, K.-H., 1995. *Das sahidische Alte und Neue Testament: sa 1-20*. Biblia Coptica, Bd. 1, Lfg. 1. Wiesbaden, Harrassowitz Verlag.
- Tiepmar, J. & Heyer, G., 2017. "An Overview of Canonical Text Services." *Linguistics and Literature Studies* 5(2): 132-148.
- Toda, S. [戸田聡], 2016. Early Christianity and Bible Translations [初期キリスト教と聖書翻訳]. *Bulletin of the Graduate School of Letters, Hokkaido University* [北海道大学文学研究科紀要] 150:159-199.
- Toda, S. [戸田聡], 1995. Antonios the Uneducated Monk? - A Trend in the Study of Early Monasticism [無学な修道者アントニオス? — 初期修道制研究の一動向]. *Orient* [オリエント] 38(2): 162-174.