

Towards the Preservation of Coptic Cultural Heritage via Digitization by Enquiring Coptic Inscriptions in Cairo Museums

Argyro Kontogianni^{1*}, Theodore Ganetsos², Evangelos Papakitsos³ & Hind Salah El-Din⁴

¹Department of Industrial Design and Production Engineering, Faculty of Engineering, University of West Attica, Thivon 250, 122 41 Egaleo, Athens, Greece. ²Laboratory for Materials in Cultural Heritage, Department of Materials Engineering, Faculty of Technology, University of Novi Sad, Bulevar cara Lazara 1, 21000 Novi Sad, Serbia. ³School of Physics, Faculty of Science, University of the Witwatersrand, Private Bag 3, Wits 2050, Johannesburg, South Africa. ⁴Egyptology Department, Faculty of Archaeology, Cairo University, 12613 Al Orman, Giza, Greater Cairo, Egypt.
Corresponding Author (Argyro Kontogianni) Email: akontogianni@uniwa.gr*



DOI: Under Assignment

Copyright © 2026 Argyro Kontogianni et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Article Received: 20 July 2026

Article Accepted: 25 September 2026

Article Published: 30 September 2026

ABSTRACT

Egypt presents a long-standing tapestry of cultures, and its language constitutes an integral part of this heritage. The Coptic language, which marks the final stage in the evolution of Ancient Egyptian, appeared in Egypt as a complete, autonomous writing and spoken system around the 3rd century AD, despite its earlier origins. Thus, Coptic cultural heritage represents a unique bridge between ancient Pharaonic traditions, Hellenistic culture, and early Christian spirituality. The term "Coptic" derives ultimately from the Greek name for Egypt (Aigyptios) through Arabic usage (Qibt). Thus, Coptic refers to the native Egyptian population and language across Late Antiquity. Today, it refers to the indigenous Christian population of Egypt and their distinct cultural, linguistic, and artistic traditions spanning over two millennia. Coptic inscriptions are found on a wide range of materials, including ivory, linen, wood and papyrus. The fragility of these materials, combined with the fact that many artifacts are fractured, complicates recognition of the content of these inscriptions. Furthermore, the different eras of these artifacts across various historical periods -alongside regional dialectal variations -further complicates their interpretation. The present study focuses on Coptic artifacts bearing inscriptions housed in the museums of Cairo -specifically the Egyptian Museum and the Coptic Museum- and explores how the content of these inscriptions can be digitized by a semi-automated software tool for the computer assisted translation, thereby contributing to the preservation of cultural heritage.

Keywords: Coptic Art; Cultural Heritage Preservation; Cultural Heritage Digitization; Cairo Museums; Egyptian Museum; Coptic Museum; Digital Technology; Language Software Engineering; Cultural Informatics Applications; Software Application; Lexical Databases.

1.0. Introduction

Egypt is a country with a long and diverse cultural history and likewise other ancient civilizations it has its own language and script systems. Coptic is the last stage of this linguistic evolution, after Hieroglyphic, Hieratic and Demotic. Following Alexander the Great's conquest of Egypt in 332 BCE and the establishment of the Ptolemaic Kingdom, Greek became the official language of government and administration. While ordinary Egyptians continued speaking Egyptian, educated Egyptians and administrators routinely used Greek. During this Hellenistic period, Egyptian scribes began experimenting with using the Greek uncial alphabet to write down Egyptian words [1],[2]. This early phase -known to scholars as Old Coptic- was initially used to transcribe pagan texts, magical spells, and native names to ensure accurate pronunciation, since older Egyptian scripts did not explicitly record vowel sounds. So, in order to overcome the barrier of the pronunciation, as the Greek alphabet was fully phonemic and included clear vowel sounds, developers of the script adopted 24 standard Greek letters (and a numerical sign derived from a defunct Greek letter) while adding 7 more letters borrowed directly from Demotic script to represent specific consonantal sounds that did not exist in the Greek language (such as the "sh", "f", "kh" or "h" sounds) [3], creating an alphabet of 32 letters (Fig. 1). Of course, every dialect varies in their specific letter counts and creates its own alphabetic format.

The development of Coptic as a major written and literary language cannot be reduced to the rise of Christianity or to the activity traditionally attributed to St. Mark. The Christianization of Egypt and the development of Coptic literature were important historical processes, but the relationship between them was complex, and Coptic textual

production also reflects other religious and social contexts. The Coptic tradition preserves sacred scriptures, monastic writings, non-canonical texts, administrative papyri, and other literary and documentary texts. Coptic script is also found on wood, ostraca, linen, bones, stone, and other materials [4],[5]. The present study is focused on searching for Coptic inscriptions on artifacts other than ancient books in Cairo Museums, especially at Egyptian Museum and Coptic Museum, in general research for digitization and preservation of Cultural Heritage.

Α (Alpha)	Greek	Π (Pi)	Greek
Β (Beta)	Greek	Ρ (Rho)	Greek
Γ (Gamma)	Greek	Σ (Sigma)	Greek
Δ (Delta)	Greek	Τ (Tau)	Greek
Ε (Epsilon)	Greek	Υ (Upsilon)	Greek
Ζ (Zeta)	Greek	Φ (Phi)	Greek
Η (Eta)	Greek	Χ (Chi)	Greek
Θ (Theta)	Greek	Ψ (Psi)	Greek
Ι (Iota)	Greek	Ω (Omega)	Greek
Κ (Kappa)	Greek	ϣ (Shai)	Demotic
Λ (Lambda)	Greek	ϝ (Fai)	Demotic
Μ (Mu)	Greek	ϙ (Khai)	Demotic
Ν (Nu)	Greek	Ϡ (Hori)	Demotic
Ξ (Xi)	Greek	ϡ (Janja)	Demotic
Ο (Omicron)	Greek	ϣ (Chima)	Demotic
		ϛ (Ti)	Demotic

Figure 1. Coptic alphabet and its letter origin; Ϛ (So) used exclusively for number 6

(Source: A. Kontogianni, adapted from standard Coptic).

Digitizing and preserving Coptic heritage—encompassing ancient manuscripts, non-canonical literature, monastic architecture, unique liturgical music, and artwork—is an urgent priority for the relevant historians, linguists, and cultural institutions [6],[7]. Thus, recent systematic research on museum digital transformation shows that digital technologies are increasingly being applied to the preservation, documentation, accessibility, education, and presentation of cultural heritage collections [8].

In this study we are trying to approach Coptic artifacts at two Cairo Museums, the Egyptian Museum and Coptic Cairo Museum, and introduce a human-supervised software application designed to translate rather short Coptic inscriptions into Greek and English. Built for offline use not only to desktops, but also on portable devices like tablets, the lightweight system processes both digital texts and fragmented inscriptions on physical artifacts such as wood, linen, and stone. Adapted from a framework originally created for other ancient languages [9],[10], the tool integrates three core components: an interactive interface, a Coptic–Greek/English digital dictionary, and a dual-strategy search engine utilizing both linear and weighted linear searches. As the first Greek-focused translation tool for Coptic, its novelty lies in addressing material-based artifacts rather than standard corpora, while offering insights into optimizing ancient language databases for computational linguistics and cultural heritage digitization.

1.1. Study Objectives

The study focuses on the following objectives:

- 1) Identify and document Coptic inscriptions on artifacts preserved in Cairo museums, with particular emphasis on the Egyptian Museum and the Coptic Museum, including inscriptions found on materials such as wood, linen, stone, ostraca, and other non-manuscript artifacts.
- 2) Investigate the potential of digital technologies for the preservation of Coptic cultural heritage, addressing the fragility, fragmentation, and deterioration of physical artifacts and their inscriptions through non-destructive digital documentation.
- 3) Introduce a semi-automated, human-supervised software tool for the digitization and computer-assisted translation of short Coptic inscriptions into Greek and English, particularly inscriptions encountered directly on physical artifacts.
- 4) Evaluate the applicability and functionality of the proposed software through preliminary field testing in a museum environment, focusing on essential tasks such as Coptic character input, lexical searching, translation retrieval, correction of input, and automated session recording.
- 5) Establish a systematic digital workflow for the preservation, linguistic analysis and translation of fragile Coptic inscriptions, while examine how to optimize lexical retrieval through the comparison of conventional and frequency-weighted search strategies to facilitate efficient access to Coptic linguistic data for scholars.
- 6) To establish a human-in-the-loop approach to the digital analysis of fragmented and non-standard Coptic inscriptions, enabling scholars to interactively record, examine, translate, and potentially reconstruct damaged textual material without requiring a fully digitized corpus in advance.

2.0. Context and Methods

2.1. Coptic artifacts at Cairo Museums

2.1.1. The Coptic Museum

The Coptic Museum in Old Cairo houses the world's largest collection of Coptic Christian artifacts and manuscripts [11]. It is located within the historical site of the Roman Fortress of Babylon; the building itself is a masterpiece, featuring intricate mashrabiya (wooden lattice) screens, carved wooden ceilings, and recycled architectural elements from historic Coptic churches and homes. Marcus Simaika Pasha, a prominent Coptic figure and scholar, founded it in 1908 to gather and preserve Coptic antiquities scattered across Egyptian churches, desert monasteries, and private collections [12]. The museum documents the transition of Egyptian culture from ancient pharaonic and Greco-Roman traditions to Christianity and Islam, and holds approximately 16,000 artifacts spanning from the 3rd to the 19th centuries, including textiles, stonework, icons, metalwork, and illuminated manuscripts [13].

From 1,200 artifacts exhibited at 27 galleries, we can find script on derived items. The museum's manuscript library houses one of the world's most significant collections of early Christian and Gnostic texts. Among its most notable holdings is the Nag Hammadi Library [14], consisting of 13 leather-bound papyrus codices containing 52 ancient Coptic treatises and more than 1,000 digitized pages or leaves, including the renowned *Gospel of Thomas*.

Another important manuscript is the Coptic Psalter, a complete 4th–5th century book of Psalms written on parchment in the Sahidic dialect. The collection also includes hundreds of illuminated liturgical codices, gospels, lectionaries, and prayer books featuring parallel columns in Coptic and Arabic [15].

Text-bearing objects are also found among the museum’s collections of icons, textiles, and decorative arts [16]. Many illuminated icons, wooden altar screens (mashrabiya) and other carved woodworks feature Coptic and Arabic inscriptions identifying saints or recording dedicatory prayers. Similarly, a few linen and wool textiles are decorated with woven inscriptions or monogrammed names, illustrating the integration of written language into both religious and everyday artistic traditions (Fig. 2).



Figure 2. Coptic Tapestry Wall Hanging, 5th–6th century AD; Linen and wool, preserved at the Coptic Museum in Cairo, Egypt; Epigraph: (“IC”) (“XC”) “ΦΟΙΒΑΜΜΩΝ” “C[Y]HYTOC” (Source: A. Kontogianni).



Figure 3. Ostraca exhibited at Coptic Museum, spanning the 6th–8th centuries CE (Source: A. Kontogianni).

In addition to manuscripts and text-bearing objects, the museum preserves hundreds of ostraca (Fig. 3), i.e., broken pottery fragments and limestone flakes that served as inexpensive writing materials in antiquity [17]–[20]. These

contain a wide range of everyday texts, including personal letters, legal documents, tax receipts, school exercises, and medical recipes written in Coptic and Greek.

The museum also contains numerous examples of inscribed stonework and funerary monuments. These include carved gravestones and funerary stelae bearing Greek and Coptic epitaphs commemorating Christians and monks, as well as architectural elements such as lintels, capitals, and wall panels engraved with donor names, biblical verses, and religious dedications.

2.1.2. The Egyptian Museum

The Egyptian Museum in Cairo (often referred to as the Museum of Egyptian Antiquities or EMC), located on the northern side of Tahrir Square, is the oldest archaeological museum in the Middle East and houses the world's largest collection of Pharaonic antiquities [21]. Designed by French architect Marcel Dourgnon and opened in 1902, the building is a neoclassical landmark featuring a salmon-pink facade, towering glass skylights, and double-height central halls [22]. The museum houses over 120,000 to 170,000 artifacts spanning more than 5,000 years of Egyptian history, from the Predynastic period to the Greco-Roman era. The Egyptian Museum's collection is overwhelmingly devoted to Pharaonic Egypt, so the vast majority of inscribed objects are in hieroglyphic, hieratic, Demotic, and Greek (Fig. 4) [23]. Coptic inscriptions are present, but they represent only a small fraction of the museum's inscribed material. However, there is a small but significant number of Coptic inscriptions and objects in the Egyptian Museum's collection, including stone inscriptions and other inscribed artifacts. These are only published items, not the museum's entire holdings, so they should not be interpreted as a complete inventory [24].



(a)



(b)

Figure 4. Samples of inscribed ostraca, (pottery and limestone fragments), 5th–8th century AD: (a) papyrus from Qasr Ibrim, 6th–8th century AD, and other inscribed artifacts at the Coptic Museum in Cairo, Egypt; (b) papyrus preserved at the Egyptian Museum, Inventory No: 15942 (Source: A. Kontogianni).

2.2. The Egyptian Museum collections management database

Since most artifacts are not exhibited in the museum, searching the museum's collections management database for further information was essential. There we could find some interesting information such as:

- **Registration and Inventory Identifiers**, like *JE Number (Journal d'Entrée)*, the primary chronological accession number assigned to objects upon arrival at the Cairo Museum or *CG Number (Catalogue Général)*, the primary master catalog index number assigned to published artifacts, etc.;
- **Physical Medium and Material Typology**, queries filtering by material use strict technical terms like stone, clay, metals, etc.;
- **Chronology and Historical Period**, searches can be scoped by historical era or dynasty such as Predynastic, Pharaonic Eras, Late Antiquity, etc.;
- **Object Type and Function**, objects are categorized by their physical form and intended function, like epigraphy and texts -inscriptions, ostraca, papyrus rolls, codices, or funerary goods, as coffins, canopic jars, shabtis, mummy masks, amulets, daily life & utility, such as jewelry, tools, utensils, cosmetics, textiles, furniture, etc.;
- **Provenance and Site of Origin**, queries can be filtered by excavation location or original site, *Lower Egypt* (Saqqara, Giza, Tanis, Memphis, Alexandria), *Upper Egypt* (Thebes, Luxor/Karnak, Abydos, Elephantine, Philae, Edfu), *Oases and Deserts* (Fayum, Sinai, Western Desert), etc.

Searching the Egyptian's Museum collections management database, we also found total 3,261 artifacts. From them, 119 had labeled officially "inscribed": the majority of artifacts (53) were of limestone, there are also 11 artifacts of marble, 1 artifact of clay, 11 of sandstone, 1 artifact of alabaster, 1 artifact of linen, 6 artifacts of wood (unspecified), 3 artifacts of bone, 11 artifacts of unspecified material, 1 artifact of mica, 3 artifacts of silver, 1 artifact of reed, 3 artifacts of pottery, 4 artifacts of papyrus, 5 artifacts of terracotta, and 4 artifacts of plaster. Furthermore, there are 283 manuscripts from papyrus, paper and parchment.

2.3. Digital Repositories and Tools for Coptic Studies

As Coptic culture bridges the gap between Pharaonic Egypt and the Mediterranean Christian world, preserving its legacy requires tackling the physical decay of artifacts, alongside the fragmentation of texts scattered across global collections. Except for physical conservation and chemical restorations, scientists through digital humanities try to conserve them via digital documentation and linguistic preservation, while systematic research confirms the rapid expansion of digital technologies in cultural heritage conservation and identifies continuing challenges related to technological integration, sustainability, and interdisciplinary application [25]. All these artifacts serve as essential primary evidence for tracking the historical development of the Coptic language and offer valuable insights into daily life in Egypt, starting from the early Roman period. The ecosystem of online tools supporting Coptic (spanning dialects such as Sahidic and Bohairic) has expanded significantly in recent years. Designed to support university scholars, language students, and church practitioners alike, these digital assets fall into the four main categories presented next.

2.3.1. Academic Textual Corpora and Primary Source Databases

These digital platforms enable advanced searching, linguistic evaluation, and direct reading of Coptic primary sources:

- **Coptic SCRIPTORIUM** [26]: A premier open-access platform for computational Coptic research. It hosts deeply annotated Sahidic and Bohairic manuscripts integrated with Natural Language Processing (NLP) tools for searching syntax, parts of speech, and vocabulary.
- **Digital Edition of the Coptic Old Testament** [27]: Led by the Göttingen Academy, this long-term endeavor aims to assemble a virtual library and critical online version of the Coptic biblical text by linking dispersed manuscript fragments.
- **PAThs (Archaeological Atlas of Coptic Literature)** [28]: A comprehensive spatial and textual database cataloging Coptic literary traditions according to geography, authorship, and codicological history.
- **Koptoo (Koptische Ostraka Online)** [29]: Maintained by LMU Munich, this portal indexes non-literary documentary records, such as administrative receipts and private letters inscribed on ceramic sherds (ostraca).
- **Brussels Coptic Database (BCD)** [30]: Managed by the Université Libre de Bruxelles, this inventory contains metadata for over 100,000 documentary texts.
- **Coptic Magical Papyri** [31]: Hosted by the University of Würzburg, this initiative curates and analyzes digital records of Coptic protective spells, healing rituals, and esoteric texts.

2.3.2. Digital Lexica and Dictionaries

Modern reference portals have replaced traditional print volumes for searching Coptic roots and definitions:

- **Coptic Dictionary Online** [32]: A searchable dictionary produced by the KELLIA initiative, allowing cross-language queries between Coptic, English, German, and French.
- **Crum's Coptic Dictionary Online** [33]: Digitized versions of Walter Ewing Crum's seminal print dictionary remain accessible via interactive online interfaces.
- **Database and Dictionary of Greek Loanwords in Coptic (DDGLC)** [34]: A specialized lexical tool designed to trace Greek vocabulary imported into Coptic texts.

2.3.3. Ecclesial and Liturgical Libraries

Digital initiatives serving Coptic Orthodox worship communities host digitized hymnody, prayer books, and theological collections:

- **OneCopticLibrary** [35]: A multilingual platform (covering Coptic, English, Arabic, French, and German) combining liturgical readings, rites, and musical transcriptions with phonetics.
- **Coptic Treasures Project** [36]: A digital repository storing thousands of scanned books, academic studies, and recordings on church history, Coptic grammar, and theology.

2.3.4. Instructional Tools, Typefaces, and Keyboards

Applications designed to assist with reading, typing, and analyzing Coptic scripts:

- **Marcion** [37]: An open-source desktop environment for Coptic, Latin, and Greek research. It features an offline edition of Crum's dictionary and the complete Nag Hammadi codices.
- **Antinoou Font** [38]: A Unicode-compliant typeface commissioned by the International Association for Coptic Studies to standardise Coptic script rendering across operating systems.
- **CopticforAll** [39]: An interactive educational website providing self-paced lessons, translated hymns, and language modules for learners.

These resources inform the present project, offering the opportunity to examine established corpora, lexical data, metadata management practices, and the conventions guiding the design of the lexical database and the project's search workflow [40]. Coptic SCRIPTORIUM is used as a source of corpus data for compiling frequency information for the experimental weighted-search component described in Section 4. The present application does not reproduce these repositories; rather, it draws selectively on their data and methodological practices for a specialized, offline tool intended for short inscriptions on physical artifacts.

3.0. Results and Discussion

3.1. Developing a Software Tool for Coptic-Greek

Notwithstanding the wide array of Coptic digital initiatives outlined above, existing tools remain ill-suited for the specialized requirements of the framework proposed in this study. Although recent work has also demonstrated the growing potential of computational approaches to Coptic translation, including the development of parallel Coptic–English and Coptic–Arabic corpora and neural machine-translation systems, highlighting both the progress of Coptic language technology and the continuing need for specialized resources [41], the vast majority of current platforms are engineered exclusively for desktop workstation environments or computational corpus processing or depend entirely on continuous internet connectivity. Additionally, many of these resources present notable usability hurdles, lack support for Greek-language outputs, and fail to incorporate specialized functionalities necessary for offline field analysis of non-papyrus material culture.

To bridge this gap, field specialists require an offline, dedicated utility tailored to detect and decipher Coptic inscriptions directly from physical objects, serving as an indispensable tool for analyzing brief epigraphic texts and conducting field linguistics. Accordingly, the application introduced in this work was designed to streamline direct linguistic analysis for Coptic scholars. Featuring an intuitive, user-centric interface, the software removes technical barriers, enabling domain experts without specialized computational training to engage effectively with this material.

Engineering software for Coptic language introduces formidable computational hurdles. Chief among these is the traditional reliance on *scriptura continua* (the complete absence of word boundaries) alongside frequent diacritics, non-standard punctuation, and complex scribal abbreviations. These features necessitate a robust text

normalization preprocessing phase prior to any downstream algorithmic processing. Furthermore, extensive dialectal diversity generates significant orthographic variance: a single lexical item may be present with disparate spellings, while a single surface form can simultaneously resolve to multiple dictionary lemmas. In order to overcome these obstacles, we propose a semi-automated approach, which suits better to physical inscriptions on material artifacts, as, in contrast to literary manuscripts, epigraphic inscriptions are typically weathered, fragmented, or syntactically non-standard. Additionally, such physical artifacts rarely feature parallel Greek translations, despite the deep historical, cultural, and lexical interdependencies that bind the two linguistic traditions [42].

The adapted software framework incorporates several key architectural and User Interface (UI) design choices:

- **Virtual Keyboard Input:** An interactive, on-screen Coptic keyboard enables sign-by-sign input via touch or click, bypassing non-intuitive key-mapping combinations on physical keyboards.
- **Integrated Query Feedback:** Executing a search instantly retrieves the target lemma, phonetic transliteration, and associated lexical attributes.
- **Transparent Lexical Storage:** The underlying dictionary relies on accessible spreadsheet formats, simplifying database updates, error corrections, and custom additions by non-technical scholars.
- **Memory-Time Optimization:** The retrieval algorithm minimizes dynamic time complexity at the cost of static spatial organization. Lemmata are partitioned into a two-dimensional Cartesian space, ordered firstly by character count and secondly alphabetically. This structural constraint bounds the search candidate subset to fewer than 100 entries, i.e., the theoretical threshold where average linear search efficiency equals that of a binary search.
- **Constant-Space Linear Search:** Within this constrained candidate set, an in-place linear search operates with an average time complexity of $O(n)$ and a space complexity of $O(1)$, where n represents the number of candidate dictionary entries considered during the search, $O(n)$ denotes linear time complexity, and $O(1)$ denotes constant space complexity. This approach yields 100% retrieval accuracy for correctly spelled dictionary lemmata.
- **Implementation Platform:** The system was engineered in Visual C# within Microsoft Visual Studio to take advantage of its robust native UI framework. Structurally, the software integrates three primary modules modeled on the predecessor platform: (i) a digital Coptic–Greek dictionary, (ii) a specialized search engine, and (iii) the graphical user interface.

3.2. Lexical Database Architecture and Content Expansion

At the core of the system lies a single, unified spreadsheet that functions as the primary **lexical database**, structured for easy updating and data addition. Vocabulary entries are categorized first by word-length (spanning eight tiers: one to seven letters, and more than seven) and secondarily sorted alphabetically. The matrix uses four core fields: Coptic lexical entry, Greek translation, English translation, and scholarly commentary (including citations, Sahidic or other dialectal variants, and epigraphic context). Compiled from major printed and digital

Coptic lexica, all words have been translated into Greek. The repository is actively being scaled to contain roughly 12,000 headwords verified by expert lexicographers, with explicit inclusion of dialectal variants and alternative spellings to optimize direct word-level matching without requiring full morphological parsing.

Data retrieval is tested using two distinct search algorithms: **Linear Search**, which operates over a Cartesian dictionary ordered strictly by word-length and alphabetization and **Weighted Linear Search**, which sorts dictionary entries in descending order, based on historical occurrence frequency, applying the Pareto Principle (where 20% of vocabulary accounts for 80% of searches). Frequency counts are tracked dynamically via a fifth database column, updated from corpus occurrences and re-sorted after each session. While both algorithms share identical formal time complexity, weighted search offers superior average-case retrieval speed for large datasets. In preliminary field testing on brief, highly fragmented inscriptions, performance differences between the two methods remained negligible due to small query sizes, categorized indexing, and rapid processing speeds relative to human input.

The third component of the software, the *interactive interface* (Fig. 5), emphasizes operational simplicity for field research. Users tap character keys on the left-panel virtual Coptic keyboard to construct the observed string, which is displayed in the query header. By clicking the “Search” button, if the queried string is matched, the corresponding Greek translation, English translation, and academic comments render in the results panel. If the term is missing or misspelled, an unsuccessful query notification appears. Other buttons allow users to edit input errors (“Correction”), reset the interface (“Clear”), or consult system documentation (“Help”). Finally, if the user selects “Exit”, it automatically exports all search outputs from the active session into a standard plaintext (.txt) log file for subsequent analysis.



(a)



(b)

Figure 5. (a) The user interface [38], (b) tested at Coptic Liturgical Codex (14th–18th century AD), housed in the Coptic Museum in Cairo (Source: A. Kontogianni).

3.3. Preliminary Field Testing

We have to point out that the software is under construction, so throughout its development, the software underwent initial functional and on-site testing led by the research team. Because both the application and its underlying vocabulary database are continually growing, this assessment represents a preliminary proof-of-concept demonstration rather than a rigorous statistical evaluation. Field testing is centered on essential

user tasks: entering Coptic characters via an on-screen keyboard, executing search requests, displaying Greek and English translations, editing typed input, resetting the display area, and generating automated session exports. In these initial tests, the software reliably executed all core routines for terms already stored in its lexicon. Crucially, the platform functions entirely offline, offering a significant practical benefit when working inside museums or remote archaeological sites, as demonstrated in Figure 5.

3.4. Discussion

As art is inherently bound to the historical era in which it was created, language serves as its most direct mirror. Furthermore, Greek and Coptic are bound together not only in the foundational matrix of the Coptic language, but also because these two scripts frequently coexist on the same artifact. At the same time, all the artifacts we discovered across the two museums of Cairo (the Egyptian and the Coptic) highlight the fragmentary nature of the Coptic language and the fragility of the materials bearing its script. To leverage this connection, the software tool presented here is designed to aid scholars in recording damaged inscriptions on physical artifacts and potentially reconstructing missing text by identifying epigraphic equivalents across related sources. Additionally, by using the embedded digital lexicon in a reverse-lookup mode, modern iconographers can produce authentic Neo-Coptic artwork, featuring grammatically correct and contextually appropriate devotional text.

As noted, the system conducts database queries through two distinct mechanisms: a standard sequential search through an alphabetical dictionary, and a frequency-weighted sequential search. The necessary frequency distributions are being compiled using datasets from the Coptic SCRIPTORIUM initiative, i.e., an open-access digital project established by Caroline T. Schroeder and Amir Zeldes [43] for reading, analyzing, and querying Coptic texts. This platform encompasses a wide range of documents, such as magical spells, biblical books (Ruth and Mark), apocryphal works like the Assumption of John, and hagiographic biographies. Implementing this comparative search, the design yields notable benefits: it enhances search speed and precision, while offering an experimental setup to evaluate how digital humanities repositories should be organized to optimize performance and stability.

Computational methods for decoding ancient writing systems rely heavily on specialized processing algorithms and dedicated lexical resources [44]. A prominent example of the former is the pioneering unsupervised model created by Snyder et al. [45] to decipher Ugaritic, i.e., a long-lost Northwestern Semitic language recorded in a 14th-century BCE cuneiform script. Their algorithmic approach functions autonomously by drawing on an extensive Hebrew dictionary, exploiting the close linguistic connection between Ugaritic and Hebrew. Although this automated method successfully decodes 90.53% of words with known cognates, it highlights that even advanced statistical systems depend on scholarly review to resolve remaining textual irregularities.

In contrast to fully automated statistical systems, our approach prioritizes a computer-assisted, human-in-the-loop framework for analyzing ancient texts. Instead of autonomous decipherment, the system operates as an interactive lexical environment that supports language acquisition, translation, and self-directed study. Moreover, unlike unsupervised models that demand a fully digitized corpus upfront, our tool is functional immediately, allowing epigraphers to digitize materials dynamically as they examine physical artifacts firsthand [42].

4.0. Conclusion

4.1. Commentary

In the two museums of Cairo, the Coptic heritage is immense. Nevertheless, while the artifacts are either exhibited or catalogued, there is no clear overview of those bearing inscriptions, making the search process almost chaotic. So, this significant Coptic epigraphic material remains untapped by scholars. Coptic heritage is not limited to written text; it includes distinct art, architecture, and musical traditions, everyday letters, tax receipts, land leases, and medical recipes. Digitizing these documents can facilitate broader access to museum collections and support the preservation, interpretation, and educational use of cultural heritage resources [46] and it can also provide an unprecedented insight into the real daily lives of ordinary Egyptians in late antiquity. This inspired the creation of the present digital tool, which establishes a systematic workflow for preserving Coptic heritage. Given that Coptic inscriptions often survive on highly delicate substrates prone to decay, a human-centered digital process is critical. Facilitating non-destructive digitization, translation, structural analysis, and contextual interpretation helps protect these fragile artifacts from damage caused by manual handling. The developed application provides a practical environment in which short Coptic inscriptions can be searched, translated, and documented through an integrated lexical database. The comparison between Linear Search and Weighted Linear Search also establishes a basis for examining alternative retrieval strategies within a Coptic lexical environment, but also gives the opportunity to reach broader conclusions about how lexical databases can be more functional. Although the preliminary tests were not intended as a statistical performance evaluation, the application successfully supported the principal functions required for the examination of inscriptional material, including character input, lexical retrieval and translation display.

Finally, this software forms part of a wider suite of language-processing applications developed at the University of West Attica, Athens, dedicated to the study of ancient languages/scripts. This broader research program focuses on deciphering ancient scripts, safeguarding vulnerable cultural heritage, and generating digital textual replicas for both archival preservation and the continuation of traditional artistic practices in the modern era.

4.2. Future Suggestions

The findings of the present study provide a basis for further development and application of the proposed approach. Future work could focus on expanding the lexical resources, improving the search and retrieval methods, and further evaluating the software in real-world cultural heritage environments. The following directions are therefore proposed:

- 1) Expansion of the Coptic lexical database to incorporate additional newly identified inscriptional vocabulary.
- 2) Further evaluation of the Linear Search and Weighted Linear Search algorithms using larger datasets and more extensive Coptic corpora to assess their retrieval performance under different search conditions.
- 3) Integration of the Coptic lexical database with additional corpus-processing tools to facilitate the analysis and processing of larger Coptic text corpora.

- 4) Extended field testing with Coptic scholars, epigraphers, and museum professionals to evaluate the usability, accuracy, and practical applicability of the software in diverse cultural heritage environments.
- 5) Integration of the software application into a network for processing ancient (such as Linear A and Linear B), under the auspices of the University of West Attica.

Declarations

Source of Funding

The field research in Cairo was funded by the Erasmus+ Programme of the University of West Attica, Greece (Prot. No.: E+/ICM_SMS5_2023_23/4/2025).

Competing Interests Statement

The authors have declared that no competing financial, professional or personal interests exist.

Consent for publication

All authors consented to the publication of this research work.

Authors' Contributions

All authors contributed to the manuscript.

Informed Consent

Not Applicable.

Availability of data and material

The data presented in this study are not yet available, until the end of the project.

Institutional Review Board Statement

Not Applicable.

Ethical Approval

Not Applicable.

Acknowledgment

The authors would like to express their thankfulness to the following persons: Professor Dr. Harby Ezzeldin Ahmed for his invaluable assistance in this research; Special thanks are also extended to the Director of the Coptic Museum, Mrs. Jihan Atef, the Director of the Egyptian Museum in Cairo, Dr. Ali Abdelhalim, and the staff of the two museums for their support during our field research; The staff of the Erasmus+ Office of the University of West Attica, for their assistance with the funding formalities; Emer. Professor Elias Sideras-Haddad of the University of the Witwatersrand, South Africa, for facilitating the research herein.

Declaration of Artificial Intelligence

No artificial intelligence tools were used for this study.

References

- [1] Bagnall, R.S. (1996). *Egypt in Late Antiquity*. Princeton University Press, Princeton, USA. <https://doi.org/10.1515/9781400821167>.
- [2] Salah El-Din, H. (2007). *Coptic script: A study in nascence and evolution* [in Arabic-unpublished]. MA thesis, Cairo University, Cairo, Egypt.
- [3] Kontogianni, A., Ganetsos, T., Zacharis, N., Kousoulis, P., & Papakitsos, E.C. (2021). A detailed study about Egyptian-Coptic and software engineering. *Archaeology*, 9: 28–33. <https://doi.org/10.5923/j.archaeology.20210901.06>.
- [4] Atiya, A.S. (1991). *The Coptic Encyclopedia*. Macmillan Publishing Company, New York, NY.
- [5] Gabra, G. (2010). *Coptic Monasteries: Egypt's Monastic Art and Architecture*. The American University in Cairo Press, Cairo, Egypt.
- [6] Immerzeel, M., van der Vliet, J. (2004). *Coptic Studies on the Threshold of a New Millennium*. Peeters Publishers, Leuven, Belgium.
- [7] Gabra, G., & Takla, H.N. (2008). *Christianity and Monasticism in Upper Egypt*, (Vol 2). American University in Cairo Press, Cairo, Egypt. <https://doi.org/10.5743/cairo/9789774163111.001.0001>.
- [8] Li, J., Zheng, X., Watanabe, I., & Ochiai, Y. (2024). A systematic review of digital transformation technologies in museum exhibition. *Computers in Human Behavior*, 161: 108407. <https://doi.org/10.1016/j.chb.2024.108407>.
- [9] Kontogianni, A. (2014). *Research Software: Development of a Linear B Lexicon* [in Greek]. Postgraduate dissertation, Inter-disciplinary Joint Postgraduate Course “Technoglossia” (MSc), National & Capodistrian University of Athens and National Technical University of Athens, Athens, Greece.
- [10] Mavridaki, A. (2025). *Development of calculation methods and tools that will contribute in digitization, studying and interpreting the Linear A script*. Doctoral Thesis, Department of Informatics & Computer Engineering, University of West Attica, Athens, Greece.
- [11] Gabra, G., Alcock, A. (1993). *Cairo: The Coptic Museum & Old Churches*. Egyptian International Publishing Company, Longman: Harlow, England.
- [12] Simaika Pasha, M.H. (1938). *A Brief Guide to the Coptic Museum and to the Principal Ancient Coptic Churches of Cairo*. The Government Press: Cairo, Egypt.
- [13] Capuani, M., Meinardus, O.F.A. (2002). *Christian Egypt: Coptic Art and Monuments Through Two Millennia*. Liturgical Press: Collegeville, MN.
- [14] Robinson, J.M. (1996). *The Nag Hammadi Library in English*, 4th rev. ed. E.J. Brill, Leiden, Netherlands.
- [15] Gabra, G. (2006). *The treasures of Coptic art in the Coptic Museum and churches of Old Cairo*. American University in Cairo Press: Cairo, Egypt.

- [16] Thomas, T.K. (2007). Coptic and Byzantine Textiles Found in Egypt: Corpora, Collections, and Scholarly Perspectives. In Bagnall, R.S. (Ed.), *Egypt in the Byzantine World, 300–700*, Pages 137–162, Cambridge University Press.
- [17] Crum, W.E., & Brightman, F.E. (2021). *Coptic Ostraca: From the Collections of the Egypt Exploration Fund, the Cairo Museum and Others*. Creative Media Partners: London, England.
- [18] Salah-Eldin, H. (2016). Two Biblical Ostraca from the Coptic Museum in Cairo. In *Life in Egypt during the Coptic Period*, Pages 237–244, Bibliotheca Alexandrina.
- [19] Salah-Eldin, H. (2013). A New Light on Coptic Cryptography. *Abgadiyat*, 8(1): 59–66. <https://doi.org/10.1163/22138609-90000009>.
- [20] Salah El-Din, H. (2018). An Ostrakon Related to Archangel Michael's Cult in Upper Egypt. *Scrinium*, 14: 369–372. <https://doi.org/10.1163/18177565-00141p25>.
- [21] Egyptian Museum Cairo (EMC). <https://egyptianmuseumcairo.eg/> (Accessed: July 31, 2026).
- [22] El-Shahawy, A., Al-Miṣrī, M. (2005). *The Egyptian Museum in Cairo*. American University in Cairo Press: Cairo, Egypt.
- [23] Hawass, Z. (2010). *Inside the Egyptian Museum with Zahi Hawass*. American University in Cairo Press: Cairo, Egypt.
- [24] Tiradritti, F., & De Luca, A. (1999). *Egyptian Treasures from the Egyptian Museum in Cairo*. Harry N. Abrams: New York.
- [25] Li, W., Xie, Q., Ao, J., Lin, H., Ji, S., Yang, M., & Sun, J. (2025). Systematic review: a scientometric analysis of the status, trends and challenges in the application of digital technology to cultural heritage conservation (2019–2024). *NPJ Heritage Science*, 13: 90. <https://doi.org/10.1038/s40494-025-01636-8>.
- [26] Coptic Scriptorium. <https://copticcriptorium.org/> (Accessed: August 14, 2026).
- [27] Digital Edition of the Coptic Old Testament. <https://coptot.manuscriptroom.com/> (Accessed: August 14, 2026).
- [28] PAThs (Archaeological Atlas of Coptic Literature). <http://paths.uniroma1.it/> (Accessed: August 16, 2026).
- [29] Koptoo (Koptische Ostraka Online). <https://www.en.ag.geschichte.uni-muenchen.de/index.html> (Accessed: August 12, 2026).
- [30] Brussels Coptic Database (BCD). <https://mnamon.sns.it/index.php?page=Risorse&id=5&lang=en> (Accessed: July 31, 2026).
- [31] Coptic Magical Papyri. <https://www.coptic-magic.phil.uni-wuerzburg.de/> (Accessed: August 12, 2026).
- [32] Coptic Dictionary Online. <https://corpling.uis.georgetown.edu/coptic-dictionary/> (Accessed: July 31, 2026).

- [33] Crum's Coptic Dictionary Online. <https://www.bintel.org.uk/crum.html> (Accessed: July 31, 2026).
- [34] DDGLC (Database and Dictionary of Greek Loanwords in Coptic). <https://www.gwdg.de/> (Accessed: August 2, 2026).
- [35] OneCopticLibrary. <https://onecopticlibrary.com/> (Accessed: August 2, 2026).
- [36] Coptic Treasures Project. <https://ml.coptic-treasures.com/> (Accessed: August 2, 2026).
- [37] Marcion. <http://marcion.sourceforge.net/> (Accessed: August 2, 2026).
- [38] Antinoou Font. <http://www.evertype.com/fonts/coptic/> (Accessed: August 2, 2026).
- [39] CopticforAll. <https://copticforall.com/> (Accessed: August 14, 2026).
- [40] Nappi, M.L., Buono, M., Chivăran, C., & Giusto, R.M. (2024). Models and tools for the digital organisation of knowledge: accessible and adaptive narratives for cultural heritage. *Heritage Science*, 12: 112. <https://doi.org/10.1186/s40494-024-01219-z>.
- [41] Saeed, M., Mohamed, A., Mohamed, M., Shehata, S., & Abdul-Mageed, M. (2024). From Nile Sands to Digital Hands: Machine Translation of Coptic Texts. In *Proceedings of the Second Arabic Natural Language Processing Conference*, Pages 298–308, Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.arabicnlp-1.25>.
- [42] Kontogianni, A., Kantaros, A., Ganetsos, T., Kousoulis, P., Mouzala, M.G., & Papakitsos, E. (2026). Digitization and Preservation of Cultural Heritage: Translating the Coptic Scripts on Artifacts. *Applied Sciences*, 16(14): 7147. <https://doi.org/10.3390/app16147147>.
- [43] Schroeder, C.T., & Zeldes, A. (2016). Raiders of the Lost Corpus. *Digital Humanities Quarterly*, 10(2): 1–15. <https://doi.org/10.63744/k66m3teq42p9>.
- [44] Sommerschild, T., Assael, Y., Pavlopoulos, J., Stefanak, V., Senior, A., Dyer, C., Bodel, J., Prag, J., Androutsopoulos, I., & de Freitas, N. (2023). Machine Learning for Ancient Languages: A Survey. *Computational Linguistics*, 49(3): 703–747. https://doi.org/10.1162/coli_a_00481.
- [45] Snyder, B., Barzilay, R., & Knight, K. (2010). A Statistical Model for Lost Language Decipherment. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Pages 1048–1057, Association for Computational Linguistics. <https://aclanthology.org/P10-1107>.
- [46] Yap, J.Q.H., Kamble, Z., Kuah, A.T.H., & Tolkach, D. (2024). The impact of digitalisation and digitisation in museums on memory-making. *Current Issues in Tourism*, 27(16): 2538–2560. <https://doi.org/10.1080/13683500.2024.2317912>.