

St. Shenouda Society Conference
2023

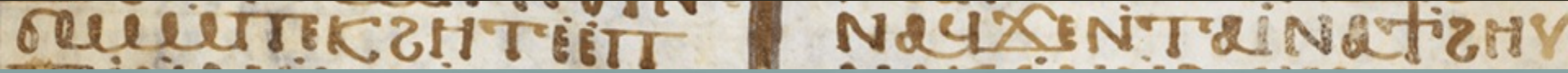
Learning and Teaching Coptic with Coptic Scriptorium's Resources

Caroline T. Schroeder
University of Oklahoma
ctschroeder@ou.edu

*All material licensed **CC-BY 4.0***



**COPTIC
SCRIPTORIUM**

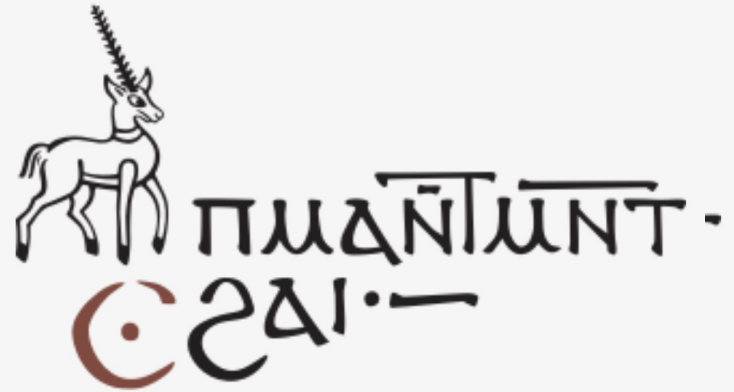


Learning and Teaching Coptic

1. Introduction
2. Reading Coptic Texts with and without translations, linguistic features visible
3. Advanced features of the Coptic Dictionary Online
4. Using Natural Language Processing to analyze Coptic text
5. Searching the Database to Help Identify Puzzling Words or Word Combinations
6. Additional Resources

1. Introduction

copticscriptorium.org



Open source, open access
Interdisciplinary research platform
for Coptic language and literature

- NLP tools
- Digital texts, digital editions
- Searchable database of text and annotations
- Collaborative annotation environment
- coptic-dictionary.org with German partners
- Linked data



Coptic Scriptorium Team

Co-PI

Amir Zeldes, Assoc. Prof. of Linguistics
Georgetown University
amir.zeldes@georgetown.edu

Full team at <https://copticcriptorium.org/about>

Funding

Current: NEH Preservation and Access Humanities Collections Reference and Resources Program (PW-290519-23)

Additionally: NEH Digital Humanities Advancement Grant (HAA-261271-18), NEH Division of Preservation and Access (PW-51672-14); NEH Office of Digital Humanities (HD-51907; HG-229371); DFG; BMBF (German Federal Ministry of Education and Research); Georgetown University; University of Oklahoma; University of the Pacific; Humboldt University; Canisius College

2. Practice Reading Coptic

<https://data.copticscriptorium.org>

Abraham Our Father urn:cts:copticLit:shenoute.abraham

Search ANNIS TEI ANNIS PAULA

Abraham Our Father: YA 518-20	normalized	analytic	diplomatic
Abraham Our Father: YA 525-30	normalized	analytic	diplomatic
Abraham Our Father: ZH fragments 1a-d	normalized	analytic	diplomatic
Abraham Our Father: YA 535-40	normalized	analytic	diplomatic
Abraham Our Father: YA 547-50	normalized	analytic	diplomatic
Abraham Our Father YA 551-54	normalized	analytic	diplomatic
Abraham Our Father XL 93-94	normalized	analytic	diplomatic

PIAΔIΔIANT
Coptic

Home Tools Corpora Documentation News

HOME CORPUS AUTHOR PEOPLE PLACES MSNAME ADVANCED SEARCH Enter URN or que

Corpora

Apophthegmata Patrum
anonymous, unknown - urn:cts:copticLit:ap

Besa Letters
Besa - urn:cts:copticLit:besa

Documentary Papyri
multiple - urn:cts:copticDoc:papyri_info

Dormition of John
unknown - urn:cts:copticLit:misc.dormition_john

Apa Johannes Canons
Johannes the Archimandrite - urn:cts:copticLit:johannes.canons

John of Constantinople Discourse
John IV of Constantinople - urn:cts:copticLit:johnconst.penitence

Pseudo-Chrysostom
Pseudo-Chrysostom, unknown - urn:cts:copticLit:pschrysostom

Life of Aphou


- **Normalized:** spelling regularized, supralinear strokes removed (to enable search); words linked to Coptic Dictionary online
- **Analytic:** can choose to show English translation (where available), part of speech, links to Wikipedia for places/people, etc
- **Diplomatic:** manuscript transcription

2. Practice Reading Coptic

<https://data.copticscriptorium.org>

Abraham Our Father urn:cts:copticLit:shenoute.abraham			
Search ANNIS ↗ TEI ↗ ANNIS ↗ PAULA ↗			
Abraham Our Father: YA 518-20	normalized	analytic	diplomatic
Abraham Our Father: YA 525-30	normalized	analytic	diplomatic
Abraham Our Father: ZH fragments 1a-d	normalized	analytic	diplomatic
Abraham Our Father: YA 535-40	normalized	analytic	diplomatic
Abraham Our Father: YA 547-50	normalized	analytic	diplomatic
Abraham Our Father YA 551-54	normalized	analytic	diplomatic
Abraham Our Father XL 93-94	normalized	analytic	diplomatic

Each “work” may be in multiple documents (*i.e.*, individual sayings in the AP, ms fragments in Shenoute’s writings etc.); shown *in order* of reading

 **COPTIC
SCRIPTORIUM**

[Home](#)

[Tools](#)

[Corpora](#)

[Documentation](#)

[News](#)

[HOME](#)

[CORPUS](#)

[AUTHOR](#)

[PEOPLE](#)

[PLACES](#)

[MSNAME](#)

[ADVANCED SEARCH](#)

Corpora

Apophthegmata Patrum

anonymous, unknown - urn:cts:copticLit:ap

Besa Letters

Besa - urn:cts:copticLit:besa

Documentary Papyri

multiple - urn:cts:copticDoc:papyri_info

Dormition of John

unknown - urn:cts:copticLit:misc.dormition_john

Apa Johannes Canons

Johannes the Archimandrite - urn:cts:copticLit:johannes.canons

John of Constantinople Discourse

John IV of Constantinople - urn:cts:copticLit:johnconst.penitence

Pseudo-Chrysostom

Pseudo-Chrysostom, unknown - urn:cts:copticLit:pschrysostom

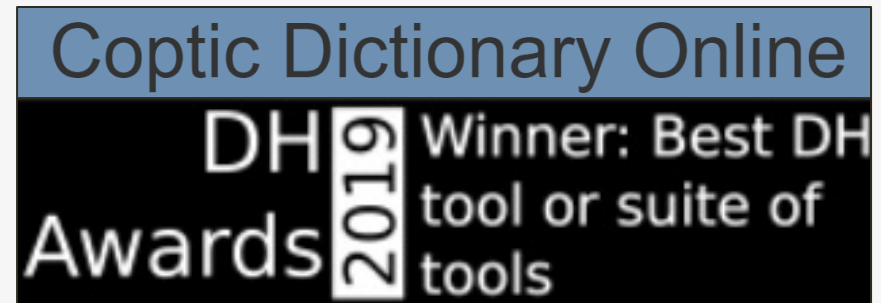
Life of Aphou

- **Normalized:** spelling regularized, supralinear strokes removed (to enable search); words linked to Coptic Dictionary online
- **Analytic:** can choose to show English translation (where available), part of speech, links to Wikipedia for places/people, etc
- **Diplomatic:** manuscript transcription

3. Advanced Features: Coptic Dictionary Online

Multiple partners:

- Coptic Scriptorium (esp. Dr. Amir Zeldes & students at Georgetown U)
- Berlin-Brandenburg Academy
- Göttingen Academy & Göttingen University
- Leipzig and Free Universities
- Göttingen Old Testament Project



<https://coptic-dictionary.org>

4. Using Natural Language Processing to analyze Coptic text

- <https://copticscriptorium.org/tools> -> NLP service
- <https://tools.copticscriptorium.org/coptic-nlp/>

☐ SGML pipeline ☒ Just piped and dashed morphemes

☒ Stretch milestones ⓘ
☒ Tokenize [stk-3.0.0]
 ☒ Automatic
 ☐ From pipes in input
☒ Normalize ⓘ
☒ Tag[flairbert-4.3.0]
☒ Lemmatize
☒ Language of origin
☒ MWE recognition ⓘ
☒ Parse[diabert-UD2.10]
☒ Entity recognition ⓘ

α|q|cωτ̄m_ncω|q_ñσ|π|αγγελος_μ|μ̄nτ-ωβhr_ . ω_τετ̄n|ρωω_ñτωτ̄n_n|ετ|κα-μα'_μ|π|
σατανας_ε|τρε|q|εἰ_ερογν_ñ|q|χωρ̄m_ñ|μ|μα_ετ|ρααβ_ε|π|ρογ̄α_αγω_ñ|τετ̄n|cωωq_μ|πεq|
ρ̄πε_ετε|ñτωτ̄n_πε_μ̄min̄mω|τ̄n_ . _μπε|τ̄n|ρ̄π|meeγε_ñ|τ|αποφασic_μ|π|αποστολος_χε|π|ετ|να|
cωωq'_μ|πε|ρπε_μ|π|νογτε_.

α|q|cωτ̄m_ncω|q_ñσ
π|αγγελος_μ̄nτ-ωβhr_ .
ω_τετ̄n|ρωω_ñτωτ̄n_νετκαμα'_μ|π|σατανας
ετρε|q|εἰ_ερογν_ñ|q|χωρ̄m_ñ|μ|μα_ετογ̄ααβ
μ̄πνογτε_ . αγω_ñτετ̄n|cωωq_μ|πεq|ρ̄πε
ετε|ñτωτ̄n_πε_μ̄min̄mω|τ̄n_ . _μπε|τ̄n|ρ̄π|meeγε
ñταποφασic_μ|π|αποστολος
χε|π|ετ|να|cωωq'_μ|πε|ρπε_μ|π|νογτε_ .

Enter Coptic (Unicode)
text -> output segmented
into words, part of
speech, etc

4. Using Natural Language Processing to analyze Coptic text

- <https://copticscriptorium.org/tools> -> NLP service
- <https://tools.copticscriptorium.org/coptic-nlp/>

☐ SGML pipeline ☒ Just piped and dashed morphemes

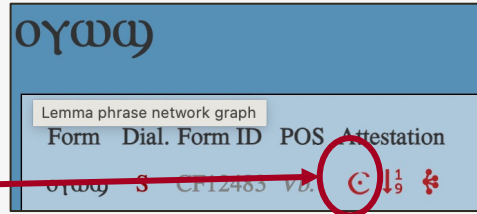

- ☒ Stretch milestones ⓘ
- ☒ Tokenize [stk-3.0.0]
 - ☒ Automatic
 - ☐ From pipes in input
- ☒ Normalize ⓘ
- ☒ Tag[flairbert-4.3.0]
- ☒ Lemmatize
- ☒ Language of origin
- ☒ MWE recognition ⓘ
- ☒ Parse[diabert-UD2.10]
- ☒ Entity recognition ⓘ

α|q|cωτ̄m_ncω|q_ñσ|π|αγγελος_ḿ|ḿñτ-ωβηρ_._ō_τετ̄ñ|ρωωε_ñτωτ̄ñ_n|ετ|κα-μα'_ḿ|π|
σατανας_ε|τρε|q|εῖ_ερογν_ñτ̄q|χωρ̄m_ñ|ḿ|μα_ετ|ογααβ_ḿ|π|νογτε_·_αγω_ñ|τετ̄ñ|cωωq_ḿ|πεq|
ṛπε_ετε|ñτωτ̄ñ_πε_ḿminḿω|τ̄ñ_·_ḿπε|τ̄ñ|ṛ|π|μεεγε_ñ|τ|αποφαςic_ḿ|π|αποστολος_χε|π|ετ|να|
cωωq'_ḿ|πε|ṛπε_ḿ|π|νογτε_.

αϥωτ̄m ncωq ḿσ
ιπαγγελος ḿñτωβηρ .
ō τετ̄ñρωωε ñτωτ̄ñ netκαμα' ḿπσατανας
ετρεqεῖ ερογν ḿqχωρ̄m ḿñμα ετογααβ
ḿπνογτε · αγω ḿτετ̄ñcωωq ḿπεqṛπε
ετεñτωτ̄ñ πε ḿminḿωτ̄ñ · ḿπετ̄ñṛπμεεγε
ñταποφαςic ḿπαποστολος
χεπετ̄ñcωωq' ḿπεṛπε ḿπνογτε .

5. Searching the Database to Help Identify Puzzling Words or Word Combinations

Access database:

- Dictionary (red C) 
- "Search ANNIS" button when reading texts at data.copticscriptorium.org 

- <https://annis.copticscriptorium.org/annis/scriptorium>



5. Searching the Database to Help Identify Puzzling Words or Word Combinations

Searching

- Dictionary headword: lemma=/ογωϣ/
- Normalized word: norm=/ογωϣ/ or norm=/ογ.*ϣ/
- Bound group:
 - norm_group=/ΝΕCΟγωϣ/
 - norm_group=/.*ογεϣογ/
 - norm_group=/ΠΕΤΝαογε.*/

<https://annis.copticscriptorium.org/annis/scriptorium>




6. Additional Resources

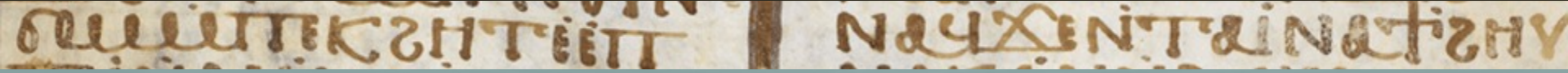
<https://copticscriptorium.org/documentation/>

Documentation / Quick Start

We do have a **wiki for detailed documentation**, but if you want a quick start, here are the most commonly used resources.

Getting started

- **Frequently Asked Questions**
- **Read and browse Coptic texts** in our corpora, including many with aligned translations
- Lookup a word in an **online Coptic Dictionary**, developed with our German partners in the KELLIA project
- ANNIS, our search and visualization platform
 - **ANNIS tutorial and cheat sheet** 
 - **More sample queries for the ANNIS** 
 - A **video of a tutorial** for using ANNIS during our March 2015 workshop
- For an introduction to some of our search functions, see the **recent post** on the Digital Orientalist. 



Thank you!

We appreciate your
feedback!