

4-25-2024

Preserving the Uncrawlable: Serving the Server

Andrew McDonnell
University of Kentucky, mcdonnell@uky.edu

Follow this and additional works at: https://uknowledge.uky.edu/libraries_present



Part of the [Archival Science Commons](#), [Byzantine and Modern Greek Commons](#), and the [Digital Humanities Commons](#)

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Repository Citation

McDonnell, Andrew, "Preserving the Uncrawlable: Serving the Server" (2024). *Library Presentations*. 273.
https://uknowledge.uky.edu/libraries_present/273

This Presentation is brought to you for free and open access by the University of Kentucky Libraries at UKnowledge. It has been accepted for inclusion in Library Presentations by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu, rs_kbnotifs-acl@uky.edu.

Preserving the Uncrawlable: Serving the Server

tiny.cc/ukweb2024



Andrew McDonnell, Digital Archivist
University of Kentucky Libraries
Special Collections Research Center
April 25, 2024



University of Kentucky Libraries Web Archiving

Established in 2018

<https://archive-it.org/organizations/915>

The screenshot shows the Archive-IT website interface. At the top right, there are social media icons for Facebook, Twitter, and LinkedIn, along with a 'Login' button. Below this is a navigation bar with 'HOME', 'EXPLORE', 'LEARN MORE', and 'CONTACT US' links. A tagline reads: 'The leading web archiving service for collecting and accessing cultural heritage on the web. Built at the Internet Archive.' A small classical building icon is also present.

The main content area features a green header for 'University of Kentucky Libraries'. On the left is the 'UK Libraries' logo. To the right, the text includes: 'University of Kentucky Libraries', 'Archive-It Partner Since: Jun, 2018', 'Organization Type: Colleges & Universities', and 'Organization URI: <http://libraries.uky.edu/>'. A detailed description follows, explaining the library's role in selecting, preserving, and providing access to archived websites and social media. It also includes a link to a web archiving guide and a form for reporting harmful content.

Below the description is a search section titled 'Narrow Your Results'. It includes a 'Subject' dropdown menu, a 'Sort By: Count | (A-Z)' option, and a search input field with a 'Search' button and a 'Clear' button. Below the search field, there are several colored buttons representing different filters or categories.

Suda On Line: Byzantine Lexicography

Founded in 1998

Still live:

www.cs.uky.edu/~raphael/sol/sol-html



Suda On Line: Byzantine Lexicography

This site has moved to a temporary new home.



sfoa.org

[Register](#) [About SOL](#) [Features](#) [Editors](#) [Rights](#) [Help](#)

Published by the
Stoa Consortium

Senior Editor

[David Whitehead](#)

Technical Director

[Raphael Finkel](#)

Managing Editors

William Hutton

Catharine Roth

Patrick Rourke

Elizabeth Vandiver

David Whitehead

Founders

Raphael Finkel

William Hutton

Patrick Rourke

Ross Scaife †

Elizabeth Vandiver

Programmers Emeriti

Huar En Ng

Mukund Chandak

Shahid Saleem Mohammed

Kamal Shah

Help

[Search Tips](#)

[Display Options](#)

[About BetaCode](#)

[Headword List](#)

[Search the Greek](#)

[Abbreviations](#)

Σάρδ;

Σασσιανός;

Σαῖ;

Σαββαριμάχος;

Σαδικαί;

Σαῖ γ' ἔθρα;

Σαῖτιοί;

Σαυτηριανός;

Σαῖται;

Σαυτί;

Σαυτί;

Σαυτί;

Σαυτί;

[Entire list of entries](#)

Login ID: Password:

The Suda on Line

For Attic phrase in Plato let them seek,

I poach in Suidas for unlicens'd Greek.

— Alexander Pope, *The Dunciad* 4.227-8

Pope's 'Suidas' is not a man but a work, The *Suda* (or **Stronghold**): a massive 10th century Byzantine Greek historical encyclopedia of the ancient Mediterranean world, covering the whole of Greek and Roman antiquity and also including Biblical and Christian material.

Preserved in several medieval manuscripts, it has been edited and published several times since the end of the 14th century in traditional hard-copy scholarly editions, most recently that of Ada Adler (Teubner, 5 volumes: 1928-1938, reprinted 1971). The Suda On Line (SOL) project, begun in 1998 as part of the Stoa Consortium, opens up this stronghold of information by means of a freely accessible, keyword-searchable database, with English translations, notes, bibliography, and links to other electronic resources. With contributions (as Translators and/or Editors) from more than two hundred people worldwide, the SOL reached the landmark of all entries being translated and "vetted" (edited) to a usable standard on July 21, 2014. But more can, and will, be done.

A fuller history of the project may be found [here](#).

Using this website

Users of our website have three options.

1. Without registration you can gain immediate access to the [entire list of entries](#). It may be sufficient if you already know which entry or entries (by Adler number) you want to consult. You may also undertake simple searching here:

Either



Suda On Line: Byzantine Lexicography

Founded in 1998

https://wayback.archive-it.org/11384*/https://www.cs.uky.edu/~raphael/sol/sol-html/

Looks Great!

You are viewing an archived web page collected at the request of [University of Kentucky Libraries](#) using [Archive-It](#). This page was captured on 21:47:56 Sep 21, 2019, and is part of the [University of Kentucky Faculty, Staff, and Student Work Web Archives](#) collection. The information on this web page may be out of date. See [All versions](#) of this archived page. Found 0 archived media items out of 0 total on this page. [Metadata](#) [Enable QA](#)

 Suda On Line: Byzantine Lexicography
This site has moved to a temporary new home.  [stoa.org](#)

[Register](#) [About SOL](#) [Features](#) [Editors](#) [Rights](#) [Help](#)

Published by the
Stoa Consortium

Senior Editor
[David Whitehead](#)

Technical Director
[Raphael Finkel](#)

Managing Editors
William Hutton
Catharine Roth
Patrick Rourke
Elizabeth Vandiver
David Whitehead

Founders
Raphael Finkel
William Hutton
Patrick Rourke
Ross Scaife †
Elizabeth Vandiver

Programmers Emeriti
Huar En Ng
Mukund Chandak
Shahid Saleem Mohammed
Kamal Shah

[Entire list of entries](#)

Login ID: Password: [Login](#)

The Suda on Line

For Attic phrase in Plato let them seek,
I poach in Suidas for unlicens'd Greek.
— Alexander Pope, *The Dunciad* [4.227-8](#)

Pope's 'Suidas' is not a man but a work, The *Suda* (or **Stronghold**): a massive 10th century Byzantine Greek historical encyclopedia of the ancient Mediterranean world, covering the whole of Greek and Roman antiquity and also including Biblical and Christian material.

Preserved in several medieval manuscripts, it has been edited and published several times since the end of the 14th century in traditional hard-copy scholarly editions, most recently that of Ada Adler (Teubner, 5 volumes: 1928-1938, reprinted 1971). The Suda On Line (SOL) project, begun in 1998 as part of the Stoa Consortium, opens up this stronghold of information by means of a freely accessible, keyword-searchable database, with English translations, notes, bibliography, and links to other electronic resources. With contributions (as Translators and/or Editors) from more than two hundred people worldwide, the SOL reached the landmark of all entries being

House of lies.

Suda On Line: Byzantine Lexicography

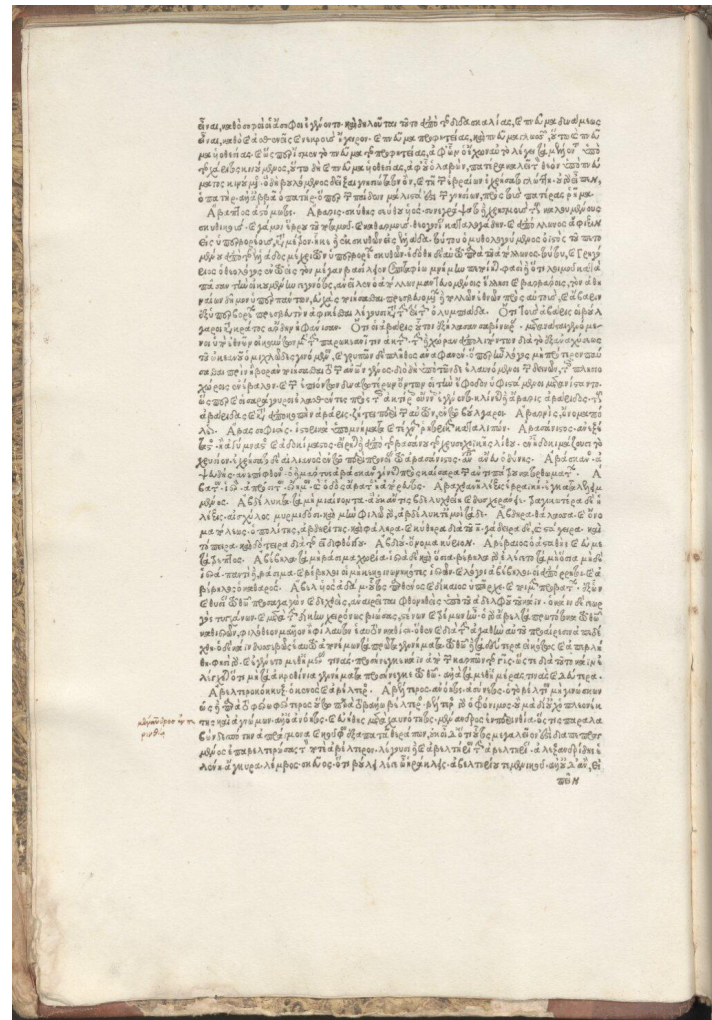
Founded in 1998

- What and Why?(Appraisal)
- How? (Capture/Ingest)
- Larger philosophical questions
(Context/Respect des fonds)

Suda Off Line: Byzantine Lexicography

1499 edition housed by the Bavarian State
Library and shared online by the Munich
Digitization Center (MDC)

https://daten.digital-sammlungen.de/bsb0005075/6/image_16



Suda On Line: Byzantine Lexicography

Founded in 1998

- Translation of 10th century encyclopedia written in Greek
 - Crowd-sourced project, pre-dating Wikipedia
 - Product of the Stoa Consortium for Electronic Publication in the Humanities
 - Only English translation of the Suda, in any format
-

Suda On Line: Byzantine Lexicography

Founded in 1998

Still live:

www.cs.uky.edu/~raphael/sol/sol-html



Suda On Line: Byzantine Lexicography

This site has moved to a temporary new home.



sfoa.org

[Register](#) [About SOL](#) [Features](#) [Editors](#) [Rights](#) [Help](#)

Published by the
Stoa Consortium

Senior Editor

[David Whitehead](#)

Technical Director

[Raphael Finkel](#)

Managing Editors

William Hutton

Catharine Roth

Patrick Rourke

Elizabeth Vandiver

David Whitehead

Founders

Raphael Finkel

William Hutton

Patrick Rourke

Ross Scaife †

Elizabeth Vandiver

Programmers Emeriti

Huar En Ng

Mukund Chandak

Shahid Saleem Mohammed

Kamal Shah

Help

[Search Tips](#)

[Display Options](#)

[About BetaCode](#)

[Headword List](#)

[Search the Greek](#)

[Abbreviations](#)

Σάρδ;

Σασσιανός;

Σαῖ;

Σαβαστιανός;

Σαδικαί;

Σαῖ γ' ἔθρα;

Σαῖτις;

Σαυτηριανός;

Σαῖται;

Σαυτί;

Σαυτί;

Σαυτί;

Σαυτί;

[Entire list of entries](#)

Login ID: Password:

The Suda on Line

For Attic phrase in Plato let them seek,

I poach in Suidas for unlicens'd Greek.

— Alexander Pope, *The Dunciad* 4.227-8

Pope's 'Suidas' is not a man but a work, The *Suda* (or **Stronghold**): a massive 10th century Byzantine Greek historical encyclopedia of the ancient Mediterranean world, covering the whole of Greek and Roman antiquity and also including Biblical and Christian material.

Preserved in several medieval manuscripts, it has been edited and published several times since the end of the 14th century in traditional hard-copy scholarly editions, most recently that of Ada Adler (Teubner, 5 volumes: 1928-1938, reprinted 1971). The Suda On Line (SOL) project, begun in 1998 as part of the Stoa Consortium, opens up this stronghold of information by means of a freely accessible, keyword-searchable database, with English translations, notes, bibliography, and links to other electronic resources. With contributions (as Translators and/or Editors) from more than two hundred people worldwide, the SOL reached the landmark of all entries being translated and "vetted" (edited) to a usable standard on July 21, 2014. But more can, and will, be done.

A fuller history of the project may be found [here](#).

Using this website

Users of our website have three options.

1. Without registration you can gain immediate access to the [entire list of entries](#). It may be sufficient if you already know which entry or entries (by Adler number) you want to consult. You may also undertake simple searching here:

Either

Why Won't you Just Crawl?

...



Server-Side Operations

Using this website

Users of our website have three options.

1. Without registration you can gain immediate access to the [entire list of entries](#). It may be sufficient (or a large number) you want to consult. You may also undertake simple searching here:

or find **Either**

Choose a subject area

- aetiology
- agriculture
- architecture
- art history
- athletics
- biography
- botany
- children
- Christianity
- chronology
- clothing
- comedy
- constitution
- daily life
- definition
- dialects, grammar, and etymology
- dreams
- economics
- epic

100 per page

2. Registration as a Guest user allows you to access these but would rather be a registered user. See the [instructions](#).
3. Registration as an author allows you to upload appropriate qualifications, if approved by the Managing Editor. See the [register as an editor](#).

Citing the Suda On Line

Detailed instructions for citing

All our entries, including the advanced search features of how we apply the

advanced search features, including the Greek login ID `guest` and the password `sol` in the

appropriate qualifications, if approved by the Managing Editor. See the [register as an editor](#).

published work or in CVs, may be found [here](#).

is covered by a Creative Commons 'Attribution-NonCommercial-ShareAlike' license. See the [site](#).

Source identifications in the Suda on Line

Preserving
function, not
just form.

You are viewing an archived web page collected at the request of [University of Kentucky Libr](#)
[University of Kentucky Faculty, Staff, and Student Work Web Archives](#) collection. The inform
archived media items c

Suda On Line

The search string is empty
Search file invoked

Idea 1:
Preserve the Whole Server
...

The Importance of the Donor

Passive vs. Active
Web Archiving

A bird in hand.

OVF/OVA file

- Open Virtualization Format/Appliance
- Preserve the full disk image as a virtual machine
- Document content and instructions
- Easy!

A bird in hand.

OVF/OVA file

- Open Virtualization Format
- Preserve the full disk image as a virtual machine, import in VirtualBox
- Document content and instructions
- ~~-Easy!~~

Possible?

—

A bird in hand.

OVF/OVA file

- Open Virtualization Format
- Preserve the full disk image as a virtual machine
- Document content and instructions
- ~~-Easy!~~
- ~~Possible?~~
- Possible!

The .OVA

It's something.

Pros

- The full server is preserved
- It works

Cons

- Offline storage limits access
- Long-term sustainability?
- Requires full access to server
- Requires all website components share one server (no external resources/API integration)

Idea 2:
ReproZip-Web
...

ReproZip-Web

They're
presenting here
tomorrow!

—

ReproZip-Web

Capture server-side actions

Also capture external website elements

Package & reproduce sites far more complex than Suda On Line

—

ReproZip-Web

Pros:

- Potentially ideal for efficiently capturing DH projects
- Smaller output than server image
- Lower barrier to serving access than VM

—

ReproZip-Web

Challenges:

- Technical complexity

- Requires access/permissions to servers

- Widespread acceptance of .rpz standard/sustainability

ReproZip-Web

That being said:

Oui.

.RPZ

—



A Larger Question

...

What is a website when it's no longer online?



The Offline Suda On Line

To what extent is
SOL a tool?

To what extent a
text?

—

The Offline
Suda On Line

Aboutness &
Context.

—

The Future of the
Suda On Line

The conversation
continues...

—

References

- Boss, Katherine, Ilya Kreymer, Vicky Rampin, and Rémi Rampin. “Server-Side Web Archiving Using Repozip-Web”. *iPRES 2023 Proposal Submission*. https://drive.google.com/file/d/1D9h_qc7jd4KP58DxlsFhzNxr7ckiPtPy/view . Accessed 2024 April 14.
- Mahoney, Anne. “Tachypaedia Byzantina: The Suda On Line as Collaborative Encyclopedia”. *digital humanities quarterly*. 3:1, 2009, <https://www.digitalhumanities.org/dhq/vol/003/1/000025/000025.html> . Accessed 2024 April 14.
- Suda On Line: Byzantine Lexicography*. <https://www.cs.uky.edu/~raphael/sol/sol-html/index.html> . Accessed 2024 April 14.
- Suda*, 1499, National Library of Munich, <https://daten.digitale-sammlungen.de/0005/bsb00050756/images/index.html?fp=193.174.98.30&id=00050756&seite=1> . Accessed 2024 April 14.
- Theimer, Kate. “Archives in Context and as Context.” *Journal of Digital Humanities*. June 2012, <https://journalofdigitalhumanities.org/1-2/archives-in-context-and-as-context-by-kate-theimer/> . Accessed 2024 April 14.
- University of Kentucky Web Archives, <https://archive-it.org/organizations/915> . Accessed 2024 April 14.

Thank you!

Questions?

Andrew McDonnell
Digital Archivist
University of Kentucky Libraries
Special Collections Research Center
mcdonnell@uky.edu

Presentation Link:
tiny.cc/ukweb2024