

Universität
zu Köln

Sustainability strategies for digital humanities systems

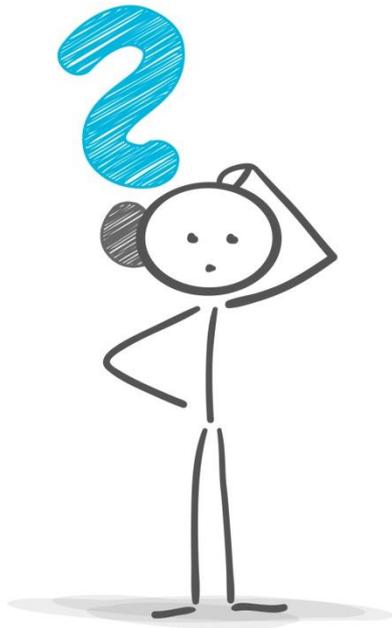


Dr. Brigitte Mathiak

Why is sustainability even an issue in the Digital Humanities?

That's what they say... (in our survey)

What is an important problem with research data? [1]

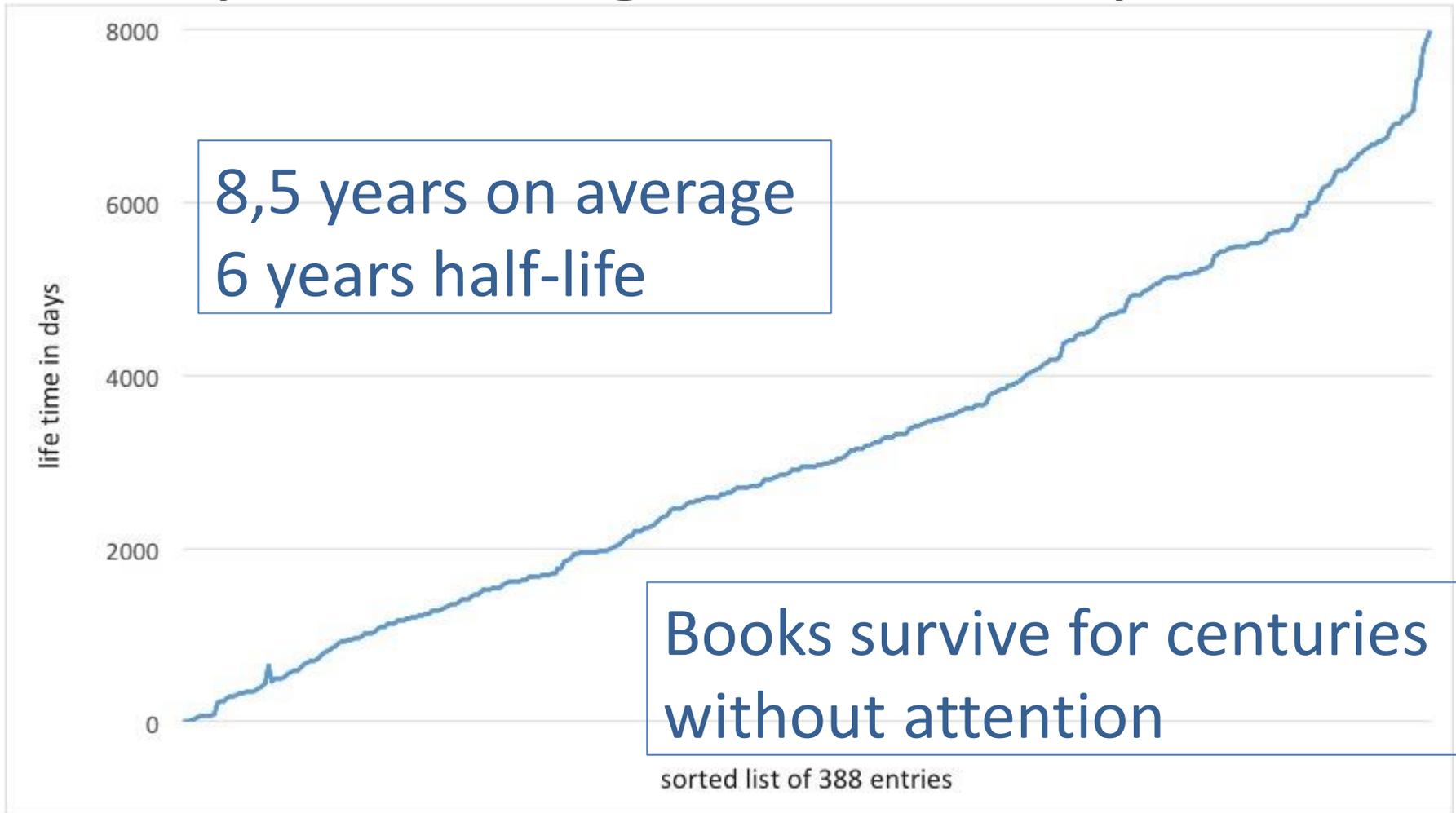


Problems (Top 4)

- 1. Place (68%): No maintenance of the website after the end of research project
- 2. Place (60%): old data formats are not readable with current software
- 3. Place (45%): Data is hard to find
- 4. Place (41%): Data is badly documented and cannot be interpreted adequately

[1] Kronenwett, Simone, and Brigitte Mathiak. 2017. "A Survey on Research Data at the Faculty of Arts and Humanities of the University of Cologne." In *Digital Humanities 2017 Conference Abstracts*, 294 - 298.

Life spans of Digital Scholarly Editions



Aus „Sustainability Strategies for Digital Humanities Systems“. Claes Neuefeind, Brigitte Mathiak, Philip Schildkamp, Unmil Karadkar, Johannes Stigler, Elisabeth Steiner, Gunter Vasold, Fabios Tosques, Arianna Ciula, Brian Maher, Greg Newton, Stewart Arneil, Martin Holmes. Panel auf der ADHO Digital Humanities Conference 2020.

Why look at Digital Scholarly Editions?

- DSE are fairly typical for Humanities projects
- They are (usually) online
- They are (usually) text-based
- They often have a database or XML to store the raw data
- Additionally, they often have interactive components, such as a search function, text statistics, translations, ...
- This combination of data, representation and digital services is very common in Humanities
- It is also particularly difficult to preserve sustainably

What are the strategies?

Overview

- Cheaper maintenance by bundling projects
 - SustainLife, DCH, University of Cologne
 - GAMS, Centre for Information Modelling, University of Graz
- Re-design to preserve
 - Humanities Computing and Media Centre, University of Victoria, Canada
 - Lazarus project, CCeH, University of Cologne
- Sandbox
 - King's Digital Lab, King's College London
- Web Archiving
 - The Internet Archive a.k.a. the Wayback Machine

„Sustainability Strategies for Digital Humanities Systems“. **Claes Neufeind**, **Brigitte Mathiak**, **Philip Schildkamp**, **Unmil Karadkar**, **Johannes Stigler**, **Elisabeth Steiner**, **Gunter Vasold**, **Fabios Tosques**, **Arianna Ciula**, **Brian Maher**, **Greg Newton**, **Stewart Arneil**, **Martin Holmes**. Panel at the ADHO Digital Humanities Conference 2020.

- **Cheaper maintenance by bundling projects**
 - SustainLife, DCH, University of Cologne
 - GAMS, Centre for Information Modelling, University of Graz
- **Re-design to preserve**
 - Humanities Computing and Media Centre, University of Victoria, Canada
 - Lazarus project, CCEH, University of Cologne
- **Sandbox**
 - King's Digital Lab, King's College London
- **Web Archiving**
 - The Internet Archive a.k.a. the Wayback Machine

Overview

- Cheaper maintenance b
 - SustainLife, DCH
 - GAMS, Co
- Re-d

DISCLAIMER!!!
I am not personally affiliated to most of the projects shown here. Any misinformation or out-dated information is due to an honest mistake or misunderstanding on my side and/or represent my personal opinion, not necessarily the opinion of the people actually involved in the project.

– The Wayback Machine



- Goal: Increase sustainability of complex living systems, such as digital editions
- Method: Meta-modelling of the components in TOSCA to streamline maintenance processes, e.g. allowing us to update multiple servers at once, re-start systems automatically, etc.
- Current state: five active prototypes
- Lessons learned:
 - Can be used retroactively on legacy projects
 - It is relatively quick to learn and to apply
 - It does improve maintenance costs significantly
 - But inevitably re-design or shut-down is necessary (security)

The German Early Cinema Database



Attention



This interim version of the German Early Cinema Database is brought to you by the *DCH Cologne*. It should contain all necessary capabilities to ensure data access, but may lack certain features, such as administrative access. If You find a desired feature not working properly, please feel free to contact us at: *info-dch@uni-koeln.de*.



Welcome



If this is your first visit to our site, please make sure to read the following notes. (Clicking on a headline will open or close the chapters.)

- » [About the databases](#)
- » [FAQs](#)
- » [Impressum](#)



The Fairground Cinema Database



Search the Fairground Cinema Database

Information on fairground cinema in Germany and neighbouring countries between 1896 and 1926. Note that this database runs in a different framework for technical reasons.



The Documents Collection



Search Documents

Contains approx. 5,000 texts on early cinema in Germany between 1895 and 1914, from a wide selection of non-film sources, such as specialist journals and general newspapers.



The Film Supply and Programme Databases



Search Films

Information on approx. 45,000 films available on the German market between 1895 and 1920 (irrespective of their country of origin).

Search Programmes

Contains a sample of film programmes from 1905 to 1914, mostly from permanent cinemas, compiled from the newspapers of nine German cities of different regions and sizes (c. 1,200 programmes from c. 100 cinemas, containing c. 3,800 different films).



GAMS, Centre for Information Modelling, University of Graz

- Goal: Get away from stand-alone, custom software
- Method: GAMS is a modular, standards-based software, that projects can use to implement their digital assets, which then will be maintained through the central system
- Current state: over 95,000 digital objects, over 90 projects
- Lessons learned:
 - Projects need to use this from the beginning, or they can start all over; funding pressure is very helpful for this
 - Not all project needs can be accommodated (though most)
 - Projects benefit from features made for others
 - Strong ongoing funding for the centre needed
 - (Complexity is challenging to manage)



Projekte

[Veröffentlichte Projekte](#)[Projekte in Entwicklung](#)

Projekte kategorisiert nach Disziplin, zeitlicher (*Antike, Mittelalter, Neuzeit, Gegenwart*) und räumlicher (*Graz und Steiermark, Österreich, International*) Einordnung; wenn für den Einsatz in der Lehre getagged mit *Didaktik*.



Alexander Rollett Briefedition

HerausgeberInnen: Zentrum für Wissenschaftsgeschichte, Karl-Franzens-Universität Graz

Die Korrespondenz von Alexander Rollett, dem ersten Inhaber einer Professur für Physiologie und Histologie an der 1863 eingerichteten Medizinischen Fakultät der Karl-Franzens-Universität Graz, mit seinem Bruder Emil sowie wissenschaftliche Briefe aus der Scientific Community an Alexander Rollett, erschließen sich über eine Briefedition des Zentrums für Wissenschaftsgeschichte.

Keywords: Wissenschaftsgeschichte, Neuzeit, Graz und Steiermark, International



Aufbruch in die Moderne: Gebrauchsgrafik

HerausgeberIn: Eva Klein

Bereits um 1900 erobert das modern gestaltete Plakat den öffentlichen urbanen Raum und steht damit am Beginn der Geschichte des Grafikdesigns. Eine wissenschaftliche Auseinandersetzung mit den Plakaten und Gebrauchsgrafiken der Steirischen Archive und deren bildwissenschaftliche und kunsthistorische Analyse liefern neue Erkenntnisse zur Steirischen Moderne. In Vergessenheit geratene Gebrauchskünstlerinnen und Gebrauchskünstler werden aufgezeigt, eine Entwicklung festgehalten und in den internationalen Kontext eingegliedert. Graz weist bereits um 1900 rege moderneaffine Tendenzen vor, die mit dem Begriff der Grazer Zeitkunst zusammengefasst werden können.

Keywords: Kunstgeschichte, Neuzeit, Graz und Steiermark



Cantus Network - libri ordinarii of the Salzburg metropolitan province

HerausgeberInnen: Austrian Centre for Digital Humanities and Cultural Heritage, Abteilung Musikwissenschaft, Österreichische Akademie der Wissenschaften



- Goal: Sustain Digital Humanities projects from the University and KISS – Keep it simple
- Method: legacy projects are re-designed into HTML, CSS, Javascript versions only
The idea is that these are comparably stable and can live „on a USB stick“ thereby have very low maintenance
- Current state: 10+ projects
- Lessons learned:
 - Very low is not zero
 - Suitable for legacy projects
 - But much cheaper, if planned from the start
 - The technology stack is surprisingly flexible



Chercher dans des vidéos. Il y a aussi une [liste complète](#) des vidéos.

Lieu

- Burkina Faso
- Californie
- Canada
- France
- île Maurice
- Mali
- Rwanda

Origine

- Burkina Faso
- Cameroun
- Canada
- France
- Haïti
- île Maurice
- Mali

Thème

- cinéma
- culture
- éducation
- éducation post secondaire
- éducation primaire
- éducation secondaire
- francophonie

Age

De :

À :



Great Unsolved Mysteries in Canadian History

[Mysteries](#)[Teachers](#)[About Us](#)[Reviews](#)[News](#)

We need your help!

A team of historians have been trying to solve some historical "cold cases" -- old crimes in which the guilty ones walked, and even more insidious crimes where a whole village may have been complicit. There are other mysteries too, about unusual cases from the Viking age to the Klondike Gold Rush.

Featured Mystery



Quick Access

[Choose your mystery](#)

Site Features

- New - Statues Are Good to Think With
- Check out our Teachers' Guides
- Flash introductions to mysteries
- Web quest style MysteryQuests
- Curriculum Connections
- Newsletter

[Becoming a Historian](#)[Français](#)



Lazarus project, CCeH, University of Cologne

- Goal: Longterm archiving and accessibility of selected legacy projects
- Method: re-designing legacy projects with newer technologies, but adding only new technology, such as search function, no new content
- Current state: 3 projects running again
- Lessons learned:
 - Time lapsed between project end and the beginning of the restoration is very costly (data archeology)
 - Documentation is helpful, but usually not sufficient
 - People are a key factor
 - Never promise you keep something running, unless you have seen it first-hand and up close
 - Costs of restoration higher than original IT costs

Cologne Sanskrit Lexicon

Query:

Dictionaries
 All monier pwg

Input language
 Sanskrit (SLP1) or any other language
 Sanskrit (autodetect input script)
 SLP1 Harvard-Kyoto ISO 15919 Devanagari
 other languages
 English German

Where and how to search
 both lemmas and definitions
 only lemmas only definitions
 exact partial similar



Cologne Sanskrit Lexicon (LAZARUS 2)

Transformation of the
data to TEI-XML

Added search engine
over all lexica

<http://www.sanskrit-lexicon.uni-koeln.de/>

Exact matches (in definitions)

18 results

Matches in monier

अग्नीषोमीयपशु *agnīṣōmīyapaśu* (SLP1: *agnIzomIyapaSu*)
अग्नीषोमीय-पशु *agnīṣōmīya---paśu* (SLP1: *agnIzomIya---paSu*)
m. a victim, generally a sheep or goat, sacred to *agni* and *soma* ([at page 6.1](#))

अजावि *ajāvi* (SLP1: *aj.Avi*)
अजाविर्क *ajāvīka* (SLP1: *aj.Avikā*)
m. pl. (*ajA S3Br.*) *sg.* goats and sheep, small cattle. ([at page 9.3](#))

अजाविक *ajāvika* (SLP1: *aj.Avika*)
अजाविकर्क *ajāvīkā* (SLP1: *aj.Avikā*)
n. sg. goats and sheep, small cattle. ([at page 9.3](#))

अनविध्य *anavīdya* (SLP1: *anaviTya*)
अन-अविध्य *ana-avīdya* (SLP1: *ana-aviTya*)
mfn. (fr. *avi q.v.*), not suited to sheep. ([at page 27.1](#))

अवि *avi* (SLP1: *avi*)
अवि *āvi* (SLP1: *āvi*)
mfn. (√ *av*), favourable, kindly disposed AV. v, 1, 9 ([at page 107.2](#))
m.f. a sheep RV. (mentioned with reference to its wool by strainer) AV. &c ([at page 107.2](#))
m.f. the woollen *soma* strainer RV. ([at page 107.2](#))
m. a protector, lord L. ([at page 107.2](#))

http://dev.cce...agnIzomIyapaSu x +

dev.cceh.uni-koeln.de/dictionary/monier/lemm Suchen

अग्नीषोमीयपशु *agnīṣōmīyapaśu* (*agnIzomIyapaSu*)

अग्नीषोमीयपशु *agnīṣōmīyapaśu* (SLP1: *agnIzomIyapaSu*)
अग्नीषोमीय-पशु *agnīṣōmīya---paśu* (SLP1: *agnIzomIya---paSu*)
m. a victim, generally a sheep or goat, sacred to *agni* and *soma* ([at page 6.1](#))

[click to show TEI-XML code...]

Surrounding entries in Monier-¹ Dictionary

agnīṣōmīyapaśu
SLP1: *agnIzomIyapaSu* lish

अग्नीषोमप्रणयन अग्नीषोमीय अग्नीषोमीयनिर्वाप अग्नीषोमीयपशु

अग्नीषोमीयपश्वनुष्ठान अग्नीषोमीयपुरोडाश अग्नीषोमीययाग

Cite this entry
Impressum



- Goal: How to archive over 100 legacy projects without funding...
- Method: [six options](#) depending on funds and complexity, the low end option is a fully backed-up virtual machine with a placeholder website, another option is a static website and/or webarchives
- Current state: 31 projects at the lowest option, over 100 in total
- Lessons learned:
 - Virtual machines and web archiving have the disadvantage of not being available in web search, counter-measures are needed
 - Very good experiences with explicit Service Level Agreements (SLAs) and a staged approach
 - Newest option is pure data storage in a CKAN instance; both projects as data, and data from projects



The minimalist version (lowest cost option)

This site has been archived as part of [King's Digital Lab](#) (KDL) archiving and sustainability process, following background analysis and consultation with research leads wherever possible.

Project content and data has been stored as a fully backed-up Virtual Machine and can be made available on request (depending on access controls agreed with the Principal Investigator) for a period of at least 2 years from the decommissioning date indicated below.

If you have an interest in this project and would like to support a future phase please contact us by filling in this [form](#).

[Read more >](#)

Project name

Mapping Shakespeare's London

Project principal investigator(s)

Hannah Crawford

Decommission Date

5 March 2018

Archive URL(s)

<http://map.shakespeare.kcl.ac.uk/>

Additional links

[Internet Archive](#)

Overview

Students studying Shakespeare often complain that the playwright's work is remote from the time and place they live in. The aim of this project is to enhance students' understanding of Shakespeare by showing them that the city in which most of his great plays were written is still alive and well on their very own

The Internet Archive a.k.a. the Wayback Machine



- Goal: Archive the Internet
- Method: Websites are crawled and replicated in an index, which make them findable through the Wayback Machine
- Current state: 616 million websites
- Lessons learned:
 - The Internet Archive is very open to archive Digital Humanities content under the provision of both Digital Cultural Heritage and Digital Research
 - However, it is not indexed in Google, so additional effort is needed to make the projects findable to scholars
 - The crawled pages can be slow and/or partially broken, depending on underlying technologies

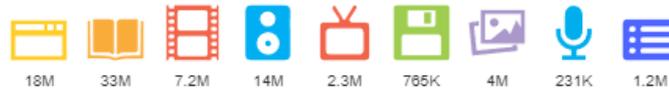
Search the history of over 616 billion web pages on the Internet.



Enter URL or keywords



Internet Archive is a non-profit library of millions of free books, movies, software, music, websites, and more.



18M 33M 7.2M 14M 2.3M 765K 4M 231K 1.2M

Search **GO**

[Advanced Search](#)

Announcements

Reflections as the Internet Archive turns 25

Internet Archive Launches Collaborative, Web-Based Art Resources Preservation and Access Initiative

Brewster Kahle named to the Library of Congress' Copyright Public Modernization Committee

[More announcements](#)

Top Collections at the Archive



KING'S
College
LONDON



HOME ABOUT

Welcome to our map of Shakespeare's London, the first online resource detailing the important ways in which the Early Modern city provided a context for his plays. This site has been produced by a team of students, each asked to research one of six particular locations in his city, and to produce the web content you see here. We hope that it will be useful to other students of Shakespeare and visitors to London. Clicking on any placemark on the Map of London will take you to a list of articles about that particular location.

This is where the map should have been!

Conclusion

- Sustainability for Digital Humanities projects is a challenge and usually requires additional and ongoing funding (e.g. a redesign every 7 years) to keep it fresh
- There are a couple of shoestring options
 - Web archiving
 - Static HTML (with or without Javascript)
- However, not all projects can be transformed in such a way without loss and findability of the resource suffers
- Combination with other measures, such as tombstone pages, archiving of the data and the code mitigate this
- King's Lab has an excellent multi-layered approach
- Also consider joining GAMS-like systems

