



Πολιτισμική Πληροφορική

Έργα ΔΗ

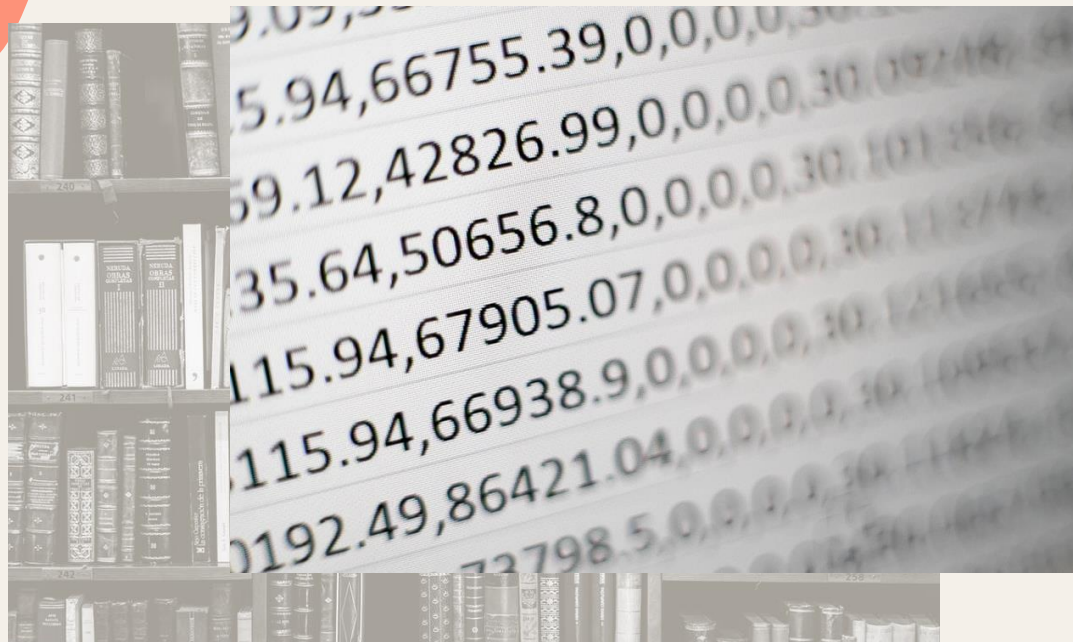
kkarpou@panteion.gr - @kkarpou

Κύκλος ζωής σε ένα έργο DH



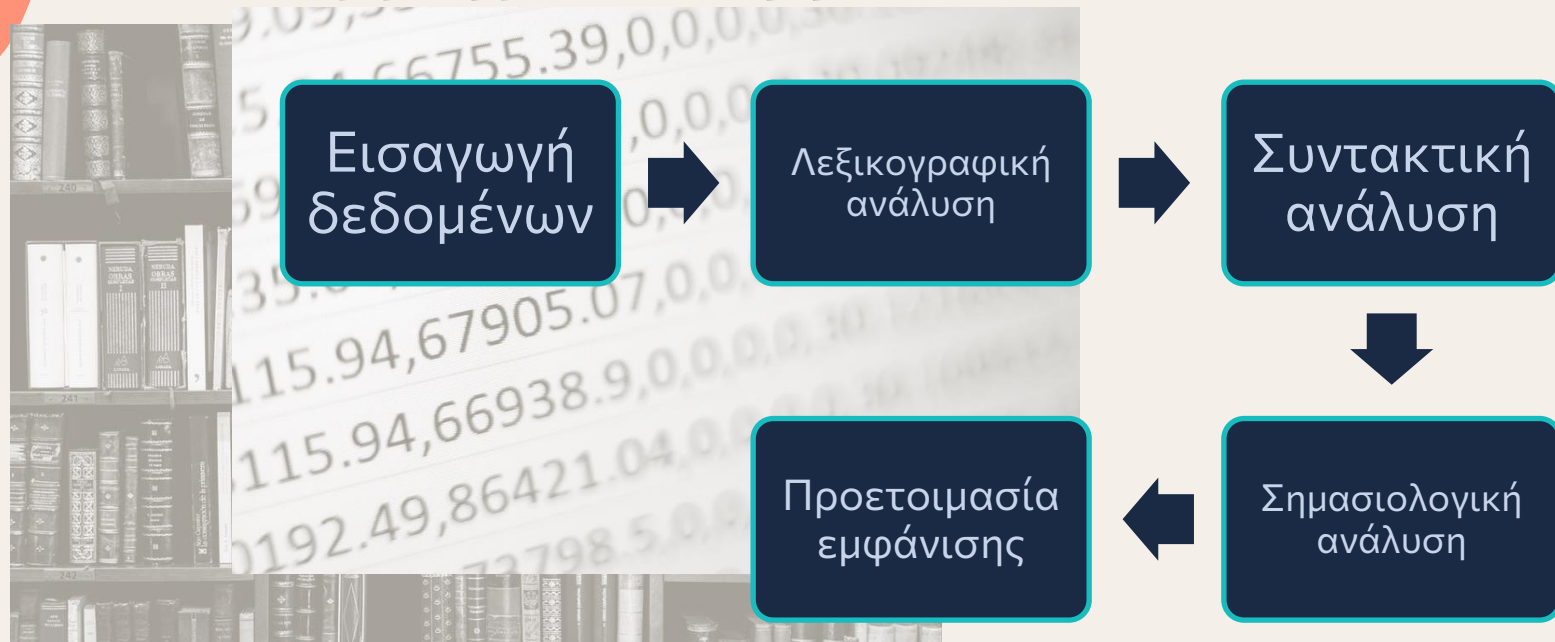
Πρωτογενή τεκμήρια

Κύκλος ζωής σε ένα έργο DH



Δομημένα δεδομένα

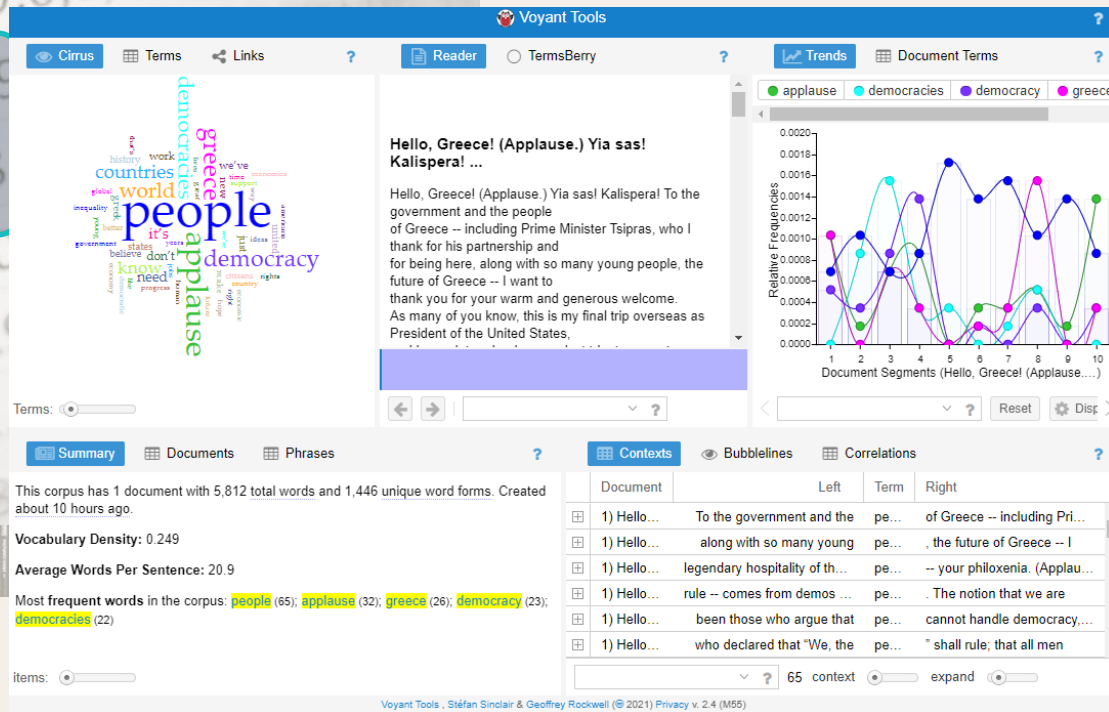
Κύκλος ζωής σε ένα έργο DH



Υπηρεσίες επεξεργασίας

Κύκλος ζωής σε ένα έργο DH

Εισαγωγή
δεδομένων



Παρουσίαση αποτελεσμάτων

The project as a basic unit

- “Digital Humanities projects typically involve **multiple circles of researchers**, from faculty and staff to students and community partners. A project’s complexity and scale generally implies the **involvement of multiple strata of personnel** from **within** and **across** institutions of learning.”
 - http://jeffreyschnapp.com/2013/01/17/short-guide-to-the-digital_humanities/

Είδη έργων DH

1949: Father Roberto Busa begins working with IBM on Index Thomisticus

1966: The journal Computers and the Humanities founded

1973: Computer Applications and Quantitative Methods in Archaeology (CAA)

1977: Association for Literary and Linguistic Computing (ALLC)

1978: Founding of the Association for Computers and The Humanities (ACH)

1985: Perseus Digital Library Project begun at Harvard

1986: Brown University Women Writers Project developed, bringing together early modern women's studies and electronic text encoding

1987: Text Encoding Initiative (TEI) developed to set **digital scholarship** standards

Είδη έργων DH

2004: A **Companion to the Digital Humanities** is printed

2005: Alliance of Digital Humanities Organizations formed from ALLC & ACH

2006: The US's National Endowment for the Humanities (NEH) Office of Digital Humanities is established; two pivotal reports comes out in US & Europe

2008: HathiTrust, Google Books, and Internet Archive established

2009: Digital Humanities Manifesto 2.0 released

2012: **Debates in Digital Humanities** is printed

2016: A New Companion to the Digital Humanities is printed

2016: Debates in Digital Humanities 2016 is printed

2019: Debates in Digital Humanities 2019 is printed

Είδη έργων DH

- Collection-based projects (συλλογή πολιτιστικών τεκμηρίων)

CORPUS THOMISTICUM
INDEX THOMISTICUS
by Roberto Busa SJ and associates
web edition by Eduardo Bernot and Enrique Alarcón
English version

Search:

[concordances](#) [terms](#) [works](#) [options](#) [new search](#)

Instructions:

You may type in any number of terms. By default, it is assumed that you want to search for all of the given terms within each of the textual units contained in the whole Index Thomisticus database, and that the given terms may appear in any order whatsoever with the only restriction that they be contiguous. However, you may override these settings by pressing the **options** button and modifying the defaults, or by using the advanced expression syntax given below. Having typed in the desired terms or expression, you can always press any of the following buttons:

concordances To read occurrences of the given terms (or expression) in their respective context.

terms To refine word selection for each given term by checking or unchecking individual forms, and to examine statistical information regarding their occurrence in the works that are selected.

works To examine statistical information concerning the occurrences of the given terms or expression in all listed works. Notice: You can press this button at any time to select which works (individually or by category) will be included in the search.

UNIVERSITY OF RICHMOND

DAILY DISPATCH

RICHMOND, VA FUNDING BY THE INSTITUTE OF MUSEUM AND LIBRARY SERVICES 1860-1865

THE RISE AND FALL OF A CAPITAL

On the eve of the Civil War, Richmond, Virginia's capital since 1780, bustling with the commerce, politics, and enterprise of an active southern city. The Dispatch, one of the city's earliest newspapers, reported on local and world events, creating snapshots of daily life. Both the city and newspaper would experience major changes by early 1865, and both would be left in ruins by April of that year. Using the transcriptions of Richmond's non-partisan newspaper and other contextual documents as guides, discover how dramatically the city, and thus the people and fortunes of the south, changed during the course of the war.

Search the Collections

[Search by Word](#)

[Search by Phrase](#)

[Boolean Search](#)

DARTMOUTH DANTE PROJECT

"Nel mezzo del cammin di nostra vita..."

The Dartmouth Dante Project (DDP) is a searchable full-text database containing more than seventy commentaries on Dante's Divine Comedy - the Commedia.

Use the links on the right to [search](#) the text of the poem and the commentaries, to read more [about](#) the project, to see a [list](#) of the commentaries in the database, and to learn how to [use](#) the database.

Dartmouth Dante Project

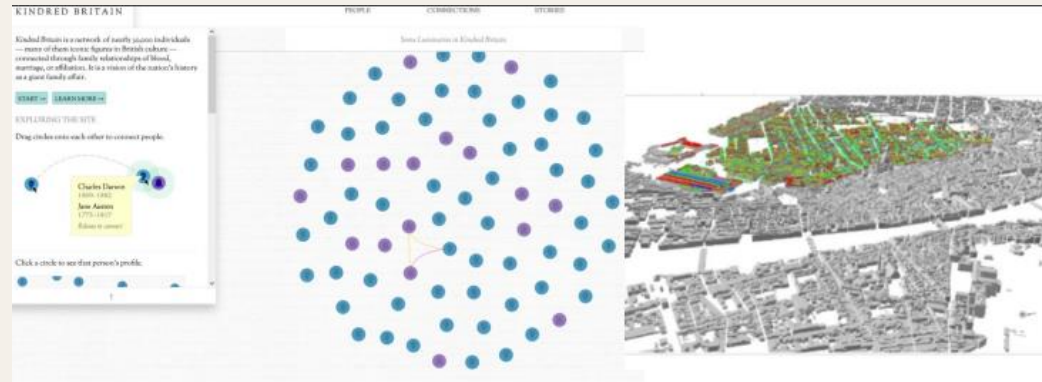
[New Search](#)
[About the DDP](#)
[List of Commentaries](#)
[Contact the DDP](#)
[Using the DDP](#)
[Dante Lab: Next Gen Reader](#)

Top of page | Copyright © 2021 Trustees of Dartmouth College

Contact Us

Είδη έργων DH

- Analysis-based projects (ανάλυση και αποτίμηση)



The DECIMA WebGIS

The heart of DECIMA is a free-to-use WebGIS Online engine. More detailed explanations of the map, and how to work with them, can be found in the User Guide.

[Open the DECIMA WebGIS](#)

[WebGIS User Guide](#)

Mining the Dispatch



Chart Type

course ☒ topic proportion
of articles % per point space

Includes articles with 0% or more of their words in this topic:

▲ increase threshold to 50%

Exemplary Articles

Wed, Sep 21, 1864

FOUR THOUSAND FIVE HUNDRED DOLLARS REWARD.

—I will pay four thousand five hundred dollars for the apprehension and delivery to me the la

EXPLORING MEDIEVAL MARY MAGDALENE

INTRODUCTION

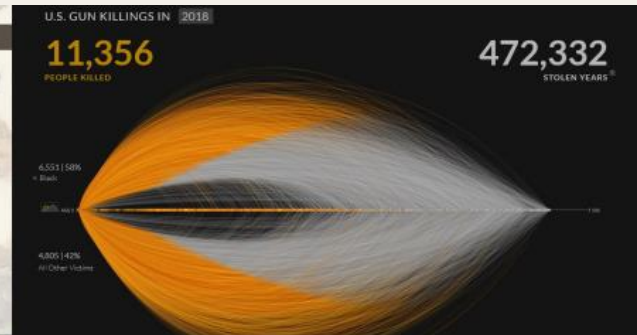
USER GUIDELINES

INTRODUCTION



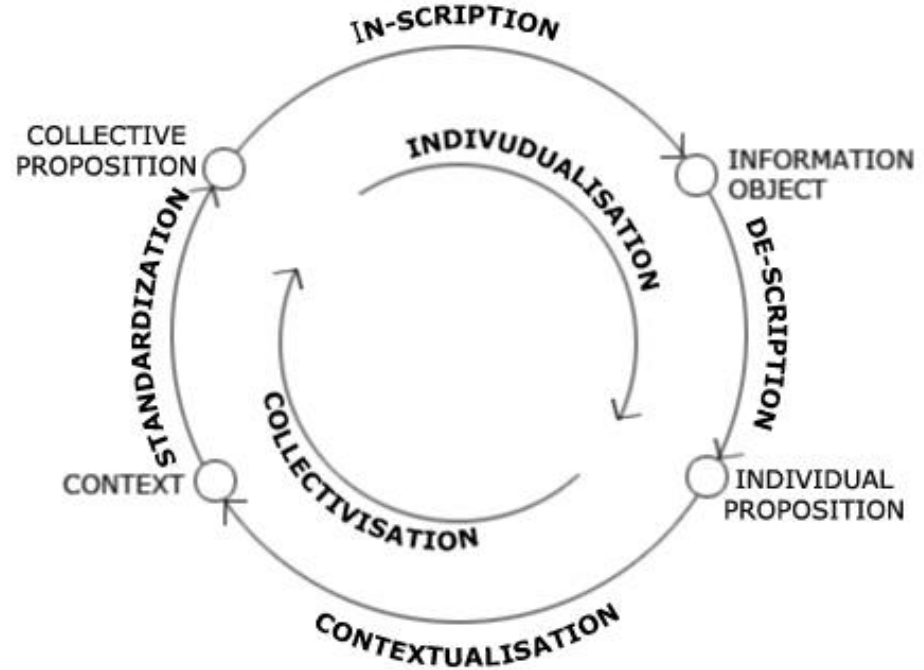
Είδη έργων DH

- Experience-based projects (δημιουργία εμπειρίας χρήστη)



Information Cosmopolitics

- Tabak, E. Information Cosmopolitics: An Actor-Network Theory Approach to Information Practices. Elsevier-Chandos, Oxford (2015).



Information Cosmopolitics

- In-scription
 - Όταν ένα “information object” αφήνει ένα ίχνος, το οποίο μπορούμε **να αναζητήσουμε, να καταγράψουμε ή να επεξεργαστούμε**
 - Information object: οτιδήποτε (και οποιοσδήποτε) μπορεί να παράγει πληροφορία
 - Μπορούμε να **απαντήσουμε ερωτήσεις**, να αναζητήσουμε σχέσεις ή να αιτιολογήσουμε τη μορφή του
- De-scription
 - Κάθε διαδικασία καταγραφής **μεταβάλλει την πρωτογενή πληροφορία**
 - Υποκειμενικότητα;
 - Transforming, αντί για informing

Information Cosmopolitics

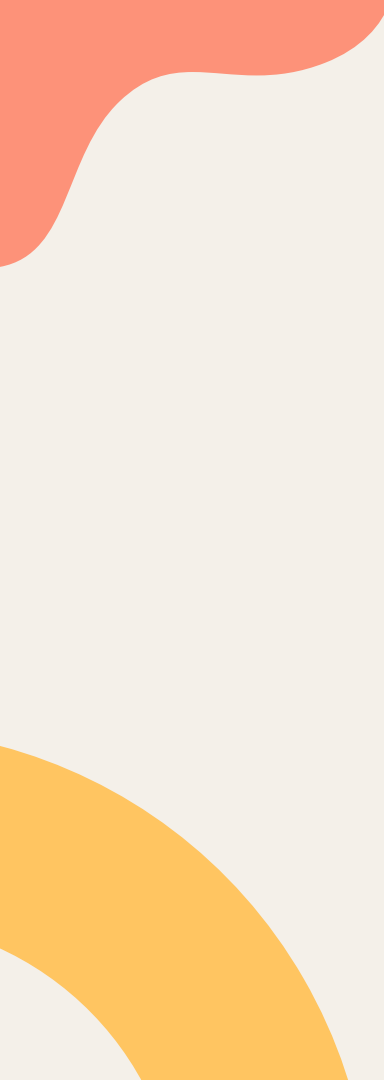
- Contextualisation
 - **Πλαίσιο**, περιβάλλον, συμφραζόμενα
 - Κοινή γλώσσα περιγραφής για ετερογενείς οντότητες (άτομα, αντικείμενα, τοπωνύμια, κλπ.)
 - Το πλαίσιο μπορεί να αναδείξει συντελεστές που **δεν είναι εμφανείς στο τεκμήριο**
 - Π.χ. ο ζωγράφος για έναν πίνακα, ο στρατιωτικός που οδήγησε το στράτευμα στον πόλεμο, κλπ.
- Standardization
 - Τα τεκμήρια κουβαλούν διαφορετικές πληροφορίες, σχέσεις, πλαίσιο
 - Εδώ διαμορφώνουμε μια **ολιστική πρόταση** σχετικά με τα τεκμήρια
 - Η πρόταση εμπεριέχει την υποκειμενικότητα κι έτσι δεν είναι ορθή ή λανθασμένη, αλλά αναλυτική ή εύγλωττη

Διαχείριση ενός έργου DH

- Οι σύγχρονες μεθοδολογίες δε διαφέρουν από την κλασική διαχείριση έργων (project management)
- Ο ρόλος του **Product Owner** είναι να βρίσκεται ανάμεσα στην ομάδα που υλοποιεί και τον «πελάτη»
 - Ορίζει τις επιμέρους ενέργειες (tasks), τις προτεραιότητες, το χρονοδιάγραμμα
 - Μικρές ομάδες από συγκεκριμένα tasks → sprint
- Ο **Scrum master** διαχειρίζεται την επικοινωνία μέσα στην ομάδα, αναζητά τις απαντήσεις στις ερωτήσεις τους και όσες πληροφορίες λείπουν, ελέγχει την πρόοδο των εργασιών
- Σε ένα ερευνητικό έργο, οι ρόλοι μεταβάλλονται
 - Μοντέλο ARDev

Διαχείριση ενός έργου DH

- Στην πραγματικότητα, σε ένα έργο DH οι ρόλοι και οι ομάδες αντιστοιχούν στα επιμέρους στοιχεία
- **D-role:** τεχνικά ζητήματα αποθήκευσης και ανάκτησης, web server, interfaces, ψηφιακά εργαλεία
- **H-role:** περιεχόμενο, γνώση/εμπειρία, βασικές προσεγγίσεις humanities
- **DH-role:** ισορροπία ανάμεσα στους δύο
 - Κίνδυνος να μην επωφεληθούμε από την τεχνολογία ή να μην απαντήσουμε στα θεμελιώδη ερωτήματα
 - Παραγωγή νέας γνώσης, πειραμάτων, μεθοδολογιών μέσα από interactive εφαρμογές, οπτικοποίηση, ανάλυση ή και κώδικα



Έργα DH

ArchNet

<http://archnet.org/>

“The Aga Khan Documentation Center at MIT (AKDC@MIT) and the Aga Khan Trust for Culture are pleased to invite you to explore Archnet Next, a new iteration of the largest online library focused on the built environments of Muslim societies.”



The Story

The Aga Khan Documentation Center at MIT (AKDC@MIT) and the Aga Khan Trust for Culture are pleased to invite you to explore ArchNet Next, a new iteration of the largest online library focused on the built environments of Muslim societies. ArchNet Next is the first major reimagining of the site since 2013. ArchNet Next has been completely redesigned to be more responsive to users in all corners of the world and on all sorts of devices in order to remain the essential reference on the architecture, planning, and landscape design in Muslim societies. Explore and enjoy.

Featured



Aga Khan Museum: Publications

[Learn more](#)



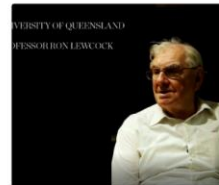
Bing and Harrington Balkan Archive (BHBA)

[Learn more](#)



Architecture and Urban Development of the Deccan Sultanates

[Learn more](#)



Eastern Architectures Secular and Religious: From Pre-historic Times to the Early Twentieth Century

[Learn more](#)



Historic Photos of Architecture from The Metropolitan Museum of Art, Department of Islamic Art

[Learn more](#)



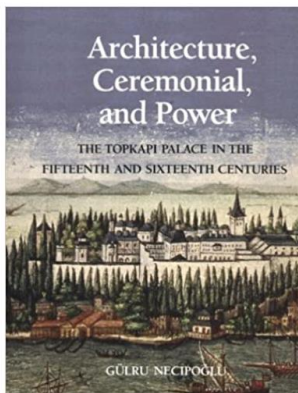
A3-ArchNet Collaborative for the Documentation of Africa's Built Heritage

[Learn more](#)

Feedback

ArchNet

<http://archnet.org/>



Architecture, Ceremonial, and Power: The Topkapi Palace in the Fifteenth and Sixteenth Centuries

Type: book Year: 1992

[Download](#)

Today the Topkapi Palace in Istanbul seems a haphazard aggregate of modest buildings no longer capable of conveying imperial power. Yet it is at once the most celebrated of all Islamic palaces and the least understood. Gülru Necipoglu brings together largely unpublished sources, both written and visual, along with information derived from the architectural remains to uncover the processes through which the meaning of the palace was once produced, before it came to represent a stereotyped microcosm of oriental despotism imbued with the exotic otherness of the East. She relocates the Topkapi in its historical context, a context that included not only the circumstances of its patronage, but the complex interaction of cultural practices, ideologies, and social codes of recognition.

Long out of print, this publication is now available from MIT Press Direct as a free e-book [here](#)

The MIT Press

Citation

Necipoglu, Gülru. *Architecture, Ceremonial, and Power: The Topkapi Palace in the Fifteenth and Sixteenth Centuries*. Cambridge: The MIT Press, 1992

Associated Sites

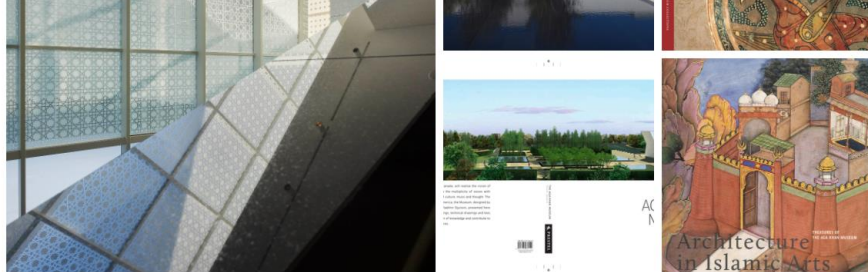


Topkapi Sarayı
Istanbul, Turkey

[Feedback](#)

ArchNet

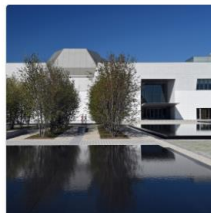
<http://archnet.org/>



Aga Khan Museum: Publications

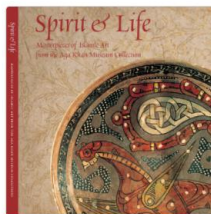
Publications of the Aga Khan Museum.

Sites



Aga Khan Museum
Toronto, Canada

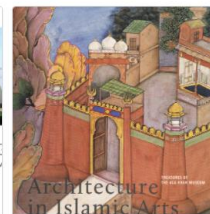
Documents



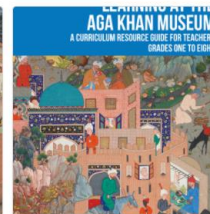
Spirit and Life
catalog



The Aga Khan Museum, Toronto
book



Architecture in Islamic Arts:
Treasures of the Aga Khan Museum
catalog



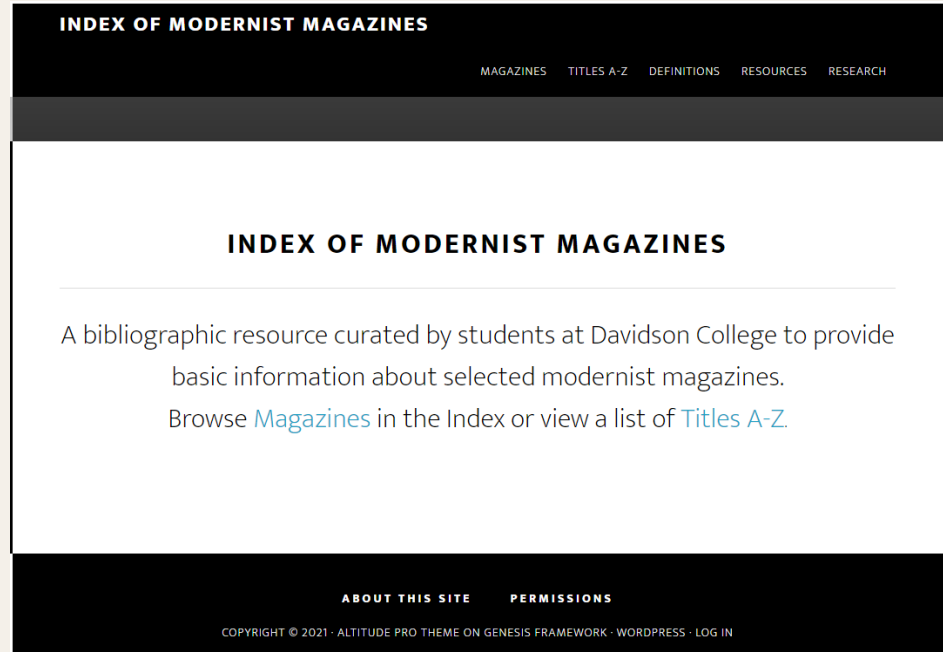
Learning at the Aga Khan Museum:
A Curriculum Resource Guide for
Teachers Grades One to Eight
teaching guide

Feedback

History of Modernist Magazines

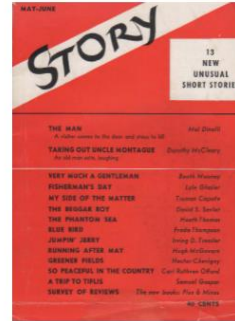
<https://modernistmagazines.org/>

“A bibliographic resource curated by students at Davidson College to provide basic information about selected modernist magazines.”



History of Modernist Magazines

<https://modernistmagazines.org/>



▲ Facts

Title:

Story: The Magazine of the Short Story

Subtitle varied:

The only magazine devoted solely to the short story

(Apr./May 1931 – Apr. 1933)

Devoted solely to the short story (June 1933 – Jan. 1937)

The magazine of the short story (Feb. 1937 – 1964)

Date of Publication:

April 1931 – Summer 1948

1960 – 1967

Place(s) of Publication:

Vienna, Austria

Majorca, Spain

New York, NY

Frequency of Publication:

Bi-monthly (frequency varied from quarterly to monthly)

Circulation:

600 copies in 1933, a figure that climbed to 21,000 copies by the late 1930s

Publisher:

Story Magazine, Inc., New York, NY (June 1933 – Sept. 1934;
Sept. 1935 – Summer 1948; 1960-1967)

Random House Magazine, Inc., New York, NY (Nov. 1934 –

History of Modernist Magazines

<https://modernistmagazines.org/>

The *Index of Modernist Magazines* provides factual information about the [little magazines](#) that fueled and broadcast [modernism](#). You can access and learn more about many of these magazines in digital archives such as:

The Modernist Journals Project

A joint project of Brown University and the University of Tulsa, The Modernist Journals Project (MJP) contains searchable and downloadable digital editions of over 30 modernist magazines, along with scholarly introductions and other helpful tools and information. The MJP's [December 1910 Collection](#) is a great resource for exploring beyond little magazines.

The Modernist Magazines Project

The UK equivalent of the MJP, this website offers sample issues of magazines, an index, and other helpful information, along with related book projects.

The Blue Mountain Project

This Princeton-based digital archive includes many European and American avant-garde magazines. It emphasizes international and experimental literary and artistic journals.

The International Dada Archive

Begun in 1979 by the University of Iowa libraries, the International Dada Archive provides images, resources, and information about the Dada movement.

Newsstand: 1925

- This virtual newsstand from the summer of 1925 gives you access to American popular and pulp magazines of the era.

Linked Jazz

<https://linkedjazz.org/>

“Our goals are:
To uncover meaningful connections between documents and data related to the personal and professional lives of jazz artists, and
To develop broadly applicable tools and methods for working with Linked Open Data.”

LINKED JAZZ

[HOME](#)[NEWS](#)[DATA](#)[TOOLS](#)[PUBLICATIONS & PRESENTATIONS](#)[BLOG](#)[NETWORK VISUALIZATION](#)[ABOUT](#)[CONTACT US](#)

Linked Jazz Now Part of Semantic Lab at Pratt

Linked Jazz joins our other LOD projects under the umbrella of the Semantic Lab at Pratt. As we continue to expand the Linked Jazz semantic work in new directions, you can follow our progress at semlab.io.

Revealing the Relationships of the Jazz Community

Linked Jazz is a research project investigating the application of Linked Open Data technologies to digital cultural heritage materials. Our goals are:

- To uncover meaningful connections between documents and data related to the personal and professional lives of jazz artists, and
- To develop broadly applicable tools and methods for working with Linked Open Data.

The project draws on jazz history materials in digital format to expose relationships between musicians and reveal their community network. New modes of connecting cultural data have the potential to open up new and unprecedented avenues of research and community engagement.



LINKED JAZZ
API



NETWORK
VISUALIZATION TOOL

Linked Jazz

<https://linkedjazz.org/>

Data Sources

Oral History Transcripts

We've utilized over 50 transcripts of oral history interviews in our project. The original interviews come from the Hamilton College Jazz Archive, Rutgers Institute for Jazz Studies Archives, Smithsonian Jazz Oral Histories, UCLA's Central Avenue Sounds Series, and the University of Michigan's Nathaniel C. Standifer Video Archive of Oral History.

Hamilton College:

Benny Powell, Benny Waters, Bill Berry, Billy Taylor, Bob Haggart, Buddy DeFranco, Buddy Tate, Buster Williams, Charles Davis, Charles McPherson, Clark Terry, David Murray, Doc Cheatham, Ed Shaughnessy, Gerald Wiggins, Harold Ousley, Herbie Hancock, Jane Jarvis, Jimmy Lewis, Jimmy Owens, Joe Williams, Leslie Johnson, Lionel Hampton, Marian McPartland, Milt Hinton, Mona Hinton, Oscar Peterson, Phil Woods, Red Holloway, Ron Carter, Roswell Rudd, Slide Hampton, Stanley Kay, Vi Redd

Rutgers:

Mary Lou Williams

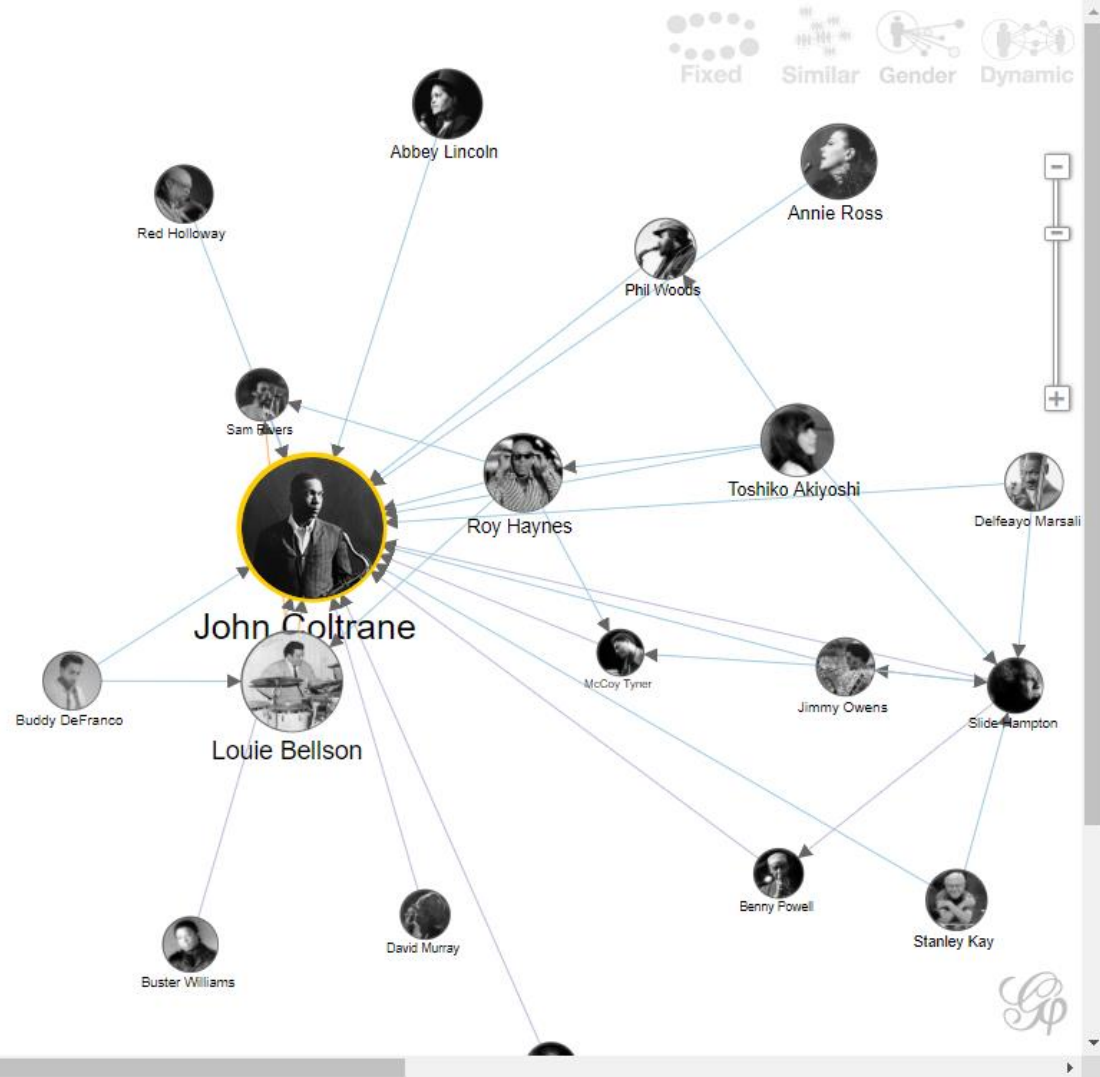
Smithsonian:

Abbey Lincoln, Annie Ross, Buddy DeFranco, Danny Barker, Dave Brubeck, Delfeayo Marsalis, Jimmy Scott, John Levy, Louie Bellson, Nancy Wilson, Roy Haynes, Toshiko Akiyoshi

UCLA:

Linked Jazz

<https://linkedjazz.org/>



Million Song Dataset

<http://millionsongdataset.com>

“The Million Song Dataset is a freely-available collection of audio features and metadata for a million contemporary popular music tracks.

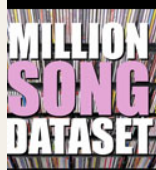
Its purposes are:

- To encourage research on algorithms that scale to commercial sizes

- To provide a reference dataset for evaluating research

- As a shortcut alternative to creating a large dataset with APIs (e.g. The Echo Nest's)

- To help new researchers get started in the MIR field”



Million Song Dataset

[Home](#)[Getting the dataset](#)[Code](#)[Tutorial](#)[Tasks / Demos](#)[More data](#)[Forum](#)

Welcome!

The **Million Song Dataset** is a freely-available collection of audio features and metadata for a million contemporary popular music tracks.

Its purposes are:

- To encourage research on algorithms that scale to commercial sizes
- To provide a reference dataset for evaluating research
- As a shortcut alternative to creating a large dataset with APIs (e.g. The Echo Nest's)
- To help new researchers get started in the MIR field

The core of the dataset is the feature analysis and metadata for one million songs, provided by [The Echo Nest](#). The dataset does not include any audio, only the derived features. Note, however, that sample audio can be fetched from services like [7digital](#), using [code](#) we provide.

The Million Song Dataset is also a cluster of complementary datasets contributed by the community:

- [SecondHandSongs dataset](#) -> cover songs
- [musiXmatch dataset](#) -> lyrics
- [Last.fm dataset](#) -> song-level tags and similarity
- [Taste Profile subset](#) -> user data
- [thisismyjam-to-MSD mapping](#) -> more user data
- [tagtraum genre annotations](#) -> genre labels
- [Top MAGD dataset](#) -> more genre labels

The Million Song Dataset started as a collaborative project between [The Echo Nest](#) and [LabROSA](#). It was supported in part by the NSF.

How to get started

To get a sense of the dataset, you can look at this [description](#) of one of the million songs.

Million Song Dataset

<http://millionsongdataset.com>

An Example Track Description

Below is a list of all the fields associated with each track in the database. This is simply an annotated version of the output of the example code [display_song.py](#). For the fields that include a large amount of numerical data, we indicate only the shape of the data array. Since most of these fields are taken directly from the Echo Nest Analyze API, more details can be found at [the Echo Nest Analyze API documentation](#).

A more technically-oriented list of these fields is given on the [field list](#) page.

This example data is shown for the track whose track_id is TRAXLZU12903D05F94 - namely, "Never Gonna Give You Up" by Rick Astley.

artist_mbid: db92a151-1ac2-438b-bc43-b82e149ddd50
the musicbrainz.org ID for this artists is db9...
artist_mbtags: shape = (4,)
this artist received 4 tags on musicbrainz.org
artist_mbtags_count: shape = (4,)
raw tag count of the 4 tags this artist received on musicbrainz.org
artist_name: Rick Astley
artist name
artist_playmeid: 1338
the ID of that artist on the service playme.com
artist_terms: shape = (12,)
this artist has 12 terms (tags) from The Echo Nest
artist_terms_freq: shape = (12,)
frequency of the 12 terms from The Echo Nest (number between 0 and 1)
artist_terms_weight: shape = (12,)
weight of the 12 terms from The Echo Nest (number between 0 and 1)
audio_md5: bf53f8113508a466cd2d3fda18b06368
hash code of the audio used for the analysis by The Echo Nest
bars_confidence: shape = (99,)
confidence value (between 0 and 1) associated with each bar by The Echo Nest
bars_start: shape = (99,)

Million Song Dataset

<http://millionsongdataset.com>

Year recognition

This is a supervised task easy to set up with the dataset. From audio features, probably "segments_timbre", and possibly some other information like "energy", "danceability", etc, try to predict the year or decade when this song was released.

A simplified dataset for that task is available on the [UCI Machine Learning Repository](#). We plan to include preliminary results in our ISMIR '11 submission.

Some code has been created to get you started, see [YearPrediction](#) folder. You also want to get (or recreate using track_metadata.db) the list of all tracks for which we have the year information (515576 tracks): [tracks_per_year.txt](#). So everyone reports comparable results, we provide a train/test split of the 28,223 artists that have at least one song with year info.

NEW SPLIT AS OF Feb 17, 2011, we now make sure no test artist is in the 10K songs subsets.

The 2,822 test artists authored 49,436 tracks with year information, about 10% of the whole set. We split according to artists, and not according to tracks, to avoid the producer effect. See [split_train_test.py](#) for details, and [train](#) and [test](#) the actual split.

Year recognition or year prediction? is it more important to know the actual year of the song, or which year this song would best fit in? I tend towards the latter (TBM).

Imputation of missing data

Imputation of missing data in time series is [well-known](#). Recently, we have studied imputation of beat-aligned chroma features using The Echo Nest data. The Million Song Dataset can easily be used to further experiment with this task. Code and ICASSP '11 paper available [here](#) (yes, we plug our own work in this case).

And since we are already doing it: imputation was first investigated as a mean to evaluate the result of our clustering of beat-chroma patterns in a large dataset, see our [ISMIR '10](#) paper and code for preliminary work. So, anyone knows how to cluster a million songs?

Robots Reading Vogue

<http://dh.library.yale.edu/projects/vogue/>

“Few magazines can boast being continuously published for over a century, familiar and interesting to almost everyone, full of iconic pictures — and also completely digitized and marked up as both text and images. [...] Students, librarians and faculty are excited about the possibilities of working with Vogue to explore questions in fields from gender studies to computer science.”



Few magazines can boast being continuously published for over a century, familiar and interesting to almost everyone, full of iconic pictures — and also completely digitized and marked up as both text and images. What can you do with over 2,700 covers, 400,000 pages, 6 TB of data? Students, librarians and faculty are excited about the possibilities of working with Vogue to explore questions in fields from gender studies to computer science. We highlight some early experiments below:

Word Vectors



Using temporal word embeddings to study the shifting notion of beauty in Vogue, by Sydney Bowen '21.

Explore »

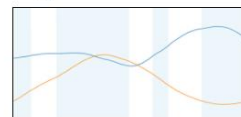
Cover Averages



Visual continuity and change across the decades.

Explore »

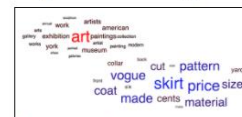
n-gram Search



Search and compare word usage in 400,000+ pages within ads, articles, or all texts.

Explore »

Topic Modeling



Self-organizing themes determined by word cooccurrence.

Explore »

Advertisements



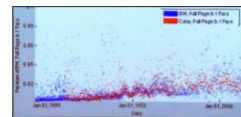
Sort ads by frequency, date, and industry.

Statistics



View figures on circulation, ratio of articles to advertisements, price per issue, and

Student Work



See projects conducted by Yale students using this data.

Colormetric Space



Hue, Saturation and Lightness.

Robots Reading Vogue

<http://dh.library.yale.edu/projects/vogue/>

"Art"

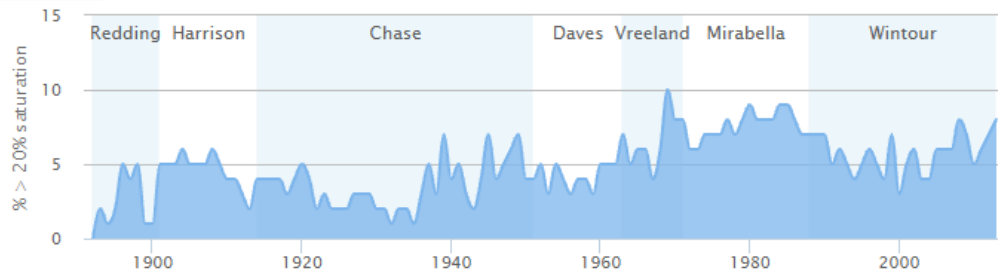
Art Words



Art Phrases



Art over Time



1995 Articles

- Michael Brenson. "Features: Spare Essentials" 1 Oct. 1995: 322
- David Colman. "Vogue Arts: Art Notes: Bodice Ripper" 1 Oct. 1995: 190
- Michael Fitzgerald. "Vogue Arts: Art: Really Rembrandt?" 1 Oct. 1995: 185
- M.G. Lord. "Features: The Rose" 1 Oct. 1995: 328
- David Colman. "Vogue Arts: Art: Short Takes" 1 Sep. 1995: 374

Robots Reading Vogue

<http://dh.library.yale.edu/projects/vogue/>

FabricSpace : Clustering

Recent work on [Word Embedding](#), a method of modeling text in vector space, has opened up new ways of visualizing the semantic relationship between terms in a large corpus.

In particular, Ben Schmidt has done work on [food terms in 19th century newspapers](#), and Ryan Heuser has investigated [cultural-semantic debates in the 18th century](#), each using Word Embedding Models as a significant part of their toolset.

One way to begin with Word Embedding is to gather a set of terms to explore. Unlike some other machine learning methods, Word Embedding allows for more control over the terms that are important to you:

1. You can use your own domain expertise to select words you know will be significant in the corpus
2. You can leverage the ability of Word Embedding Models to show *similar* terms to those you think are meaningful.

This computationally-assisted, but human-mediated, form of constructing lists of meaningful words has definite advantages over techniques such as standard [topic modeling](#). You're deeply involved in the choice of words that are meaningful to your investigation, rather than just hoping that those words turn up somewhere in a cluster or topic.

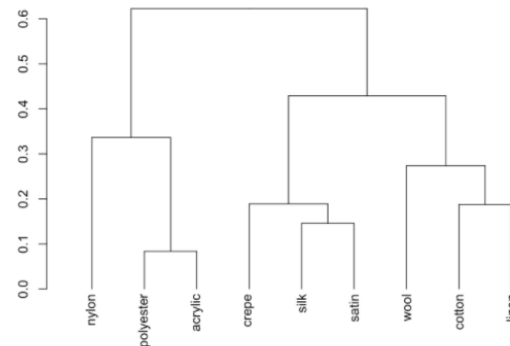
Ben Schmidt [suggests](#) one way to think about this potential of Word Embedding Models as:

...a supervised form of topic modeling: it lets you assemble a list of words that typically appear in similar contexts. But instead of letting the algorithm choose a fixed number of topics, WEMs let you choose how expansive you want the explored space to be.

Words and Arbitrary Gradients

There are a number of interesting things (similarity, analogy) you can do with Word Embedding Models. Schmidt and Heuser, after having assembled terms that are meaningful to them, go on to do sophisticated analyses of these words' spatial relationships in imaginary vector space, as a proxy for semantic meaning. Schmidt, for example, maps hundreds of foods against the axes of "meatness" and "vegetableness." These axes are constructed dynamically by other words in the newspapers themselves, and as humorous as the example is, it's a great expression of Word Embedding's possibilities to map terms to a gradient between any two points. With the Vogue archive, we eventually hope to be able to map fashion terms against similar spectra, for example formal→casual and high→low.

Cluster Dendrogram of materials in Vogue, 1892-2013



Rossica Society

<https://ufdc.ufl.edu/ros1>

"The Rossica Society is a world-wide society devoted to all aspects of Russian philately, from the pre-stamp days of Imperial Russia to current post-Soviet philately. The purpose of the Society is to unite all philatelist throughout the world with an interest in Russian philately and provide them a means of broadening their philatelic knowledge through inter- communication via the award winning Rossica Journal."

Archive of

The Rossica Society

of Russian Philately

HOME

SEARCH OPTIONS

BRIEF VIEW

THUMBNAIL VIEW

TABLE VIEW

PRINT

All Items

1 - 20 of 101 matching titles

NEXT

LAST

NARROW RESULTS BY:

Language

English (137)

Russian (37)

Creator

Rossica Society of Russian Philately (121)

Canadian Society of Russian Philately (60)

Union of Soviet Socialist Republic (50)

Russia, Empire of (20)

Union of Soviet Socialist Republics (8)

Nagorno-Karabakh (4)

Russian Soviet Federated Socialist Republic (4)

Rossica Society (3)

Nagorno-Karabakh (Azerbaijan) - Nagorno-Karabakh, Armenian Republic of (2)

Ukraine, Republic (1992-) (2)

Show More >>

Publisher

Canadian Society of Russian Philately (60)

K.L. Wilson (43)

Rossica Society (3)

Rossica Society of Russian Philately (3)

Rossica Society of Russian Philately (2)

Svazidat (1)

Subject: Topic

Stamp collections (237)

Periodicals (76)

Postage-stamps (76)

Stamp collecting (76)

Antarctica (16)

Discovery and exploration (14)

Mail (5)

Armenia (4)

Azerbaijan (4)

Kremlin (Moscow, Russia) (4)

Show More >>

1



10 years of Soviet Antarctic exploration

Publication Date: 1987

Creator: Union of Soviet Socialist Republic

Format: Artifact

Source Institution: University of Florida

2



1923 Aviation cover posted by the Soviet Philatelic Association to England.

Publication Date: 1923

Creator: Russian Soviet Federated Socialist Republic

Format: Artifact

Source Institution: University of Florida

3



20 years of Soviet Antarctic exploration: the trek to the South Pole

Publication Date: 1978

Creator: Union of Soviet Socialist Republic

Format: Artifact

Source Institution: University of Florida

4



20th Anniversary of the Space Age - two stamps on first day cover (FDC) and stamp issues and souvenir sheet in the set, 1977.

Publication Date: 1977

Creator: Union of Soviet Socialist Republic

Format: Artifact

Source Institution: University of Florida

Rossica Society

<https://ufdc.ufl.edu/ros1>

1



[AH] Wallace Williams, SoMar meets Ron Hill SoMAR in Athens, Greece 2016

Publication Date: 2016
Creator: Wallace Williams, Digital, Photographer, founder Society of Olympic Marathon Runners
Publisher: Virgin Islands Council on the Arts Project
The Virgin Island Space (St. Croix, Virgin Islands USA)
Format: pdf text and color image
Subjects: Caribbean Area
Wallace Williams
St. Croix, Virgin Islands USA
Virgin Islands Track and Field Federation
Society of Olympic Marathon Runners
SOMar
Virgin Islands Pace Runners
Caribbean—Runners—Running—Marathon—Road Running
—Cross-Country—Long Distance Running—Track and Field
—Athletics—Endurance Events—Sports
St. Croix, Virgin Islands USA—St. Thomas—St. John
Campbellsville, Kentucky USA—Taylor County
IAAF—World Athletics (world governing body for track and field)—road running—cross-country
Ron Hill
Source Institution: The Virgin Island Space

2



157. Comparing medical oncologists' and surgeons' treatment recommendations in early stage HR+, HER2- breast cancer (BC) patients (pts): A subanalysis of the Multidisciplinary Application of Genomics in Clinical Practice (MAGIC) survey

Links: ([publisher version](#) | [internal citation](#))
Publication Date: 2014-11-30
Creator: Rouzier, R. (author)
Markopoulos, C. (author)
Aapro, M. (author)
De Laurentiis, M. (author)
Linderholm, B. (author)
Neven, P. (author)
Ducrocq, C. (author)

Publisher version
Check access

Rossica Society

<https://ufdc.ufl.edu/ros1>

DESCRIPTION	PAGE TURNER	THUMBNAILS	PAGE IMAGES	SEARCH	PRINT	SEND	+ADD	SHARE
--------------------	--------------------	-------------------	--------------------	---------------	--------------	-------------	-------------	--------------

TABLE OF CONTENTS		◀ FIRST ◀ PREVIOUS Go To: Page 8 NEXT ▶ LAST ▶	
--------------------------	--	--	--

Click on image below to switch to zoomable version

SelfieCity

<http://selfiecity.net/>

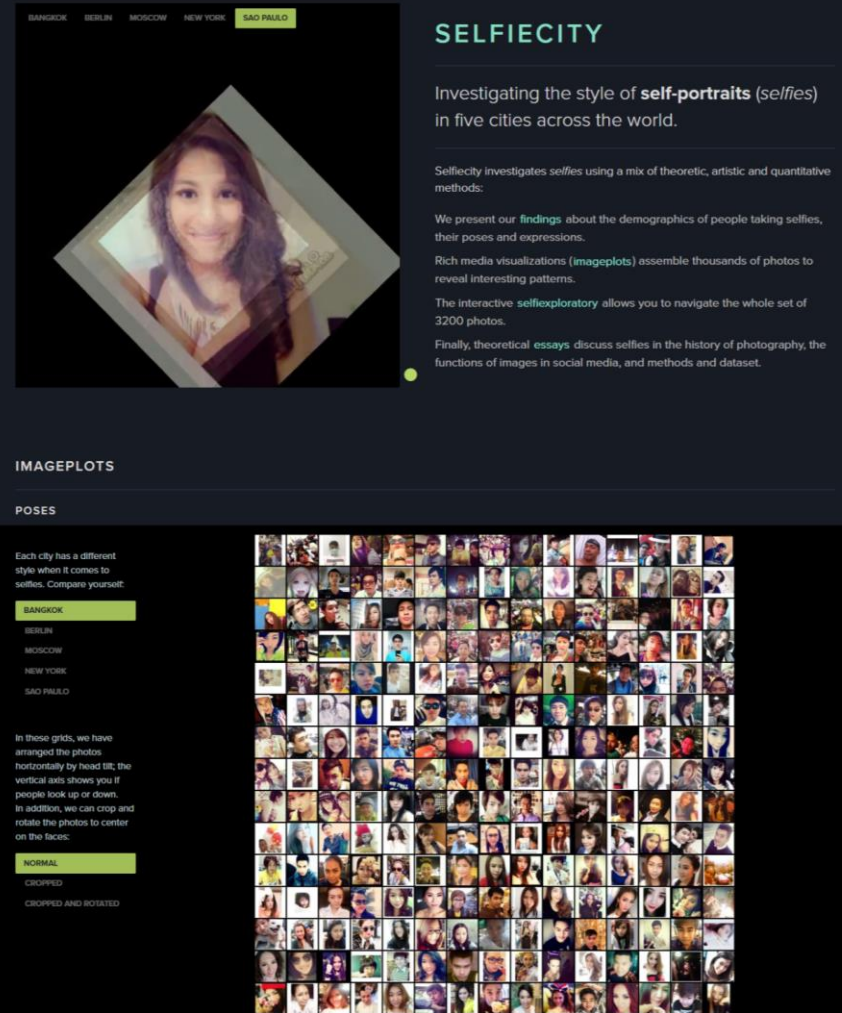
“Selfiecity investigates selfies using a mix of theoretic, artistic and quantitative methods:

We present our findings about the demographics of people taking selfies, their poses and expressions.

Rich media visualizations (imageplots) assemble thousands of photos to reveal interesting patterns.

The interactive selfieexploratory allows you to navigate the whole set of 3200 photos.

Finally, theoretical essays discuss selfies in the history of photography, the functions of images in social media, and methods and dataset.”



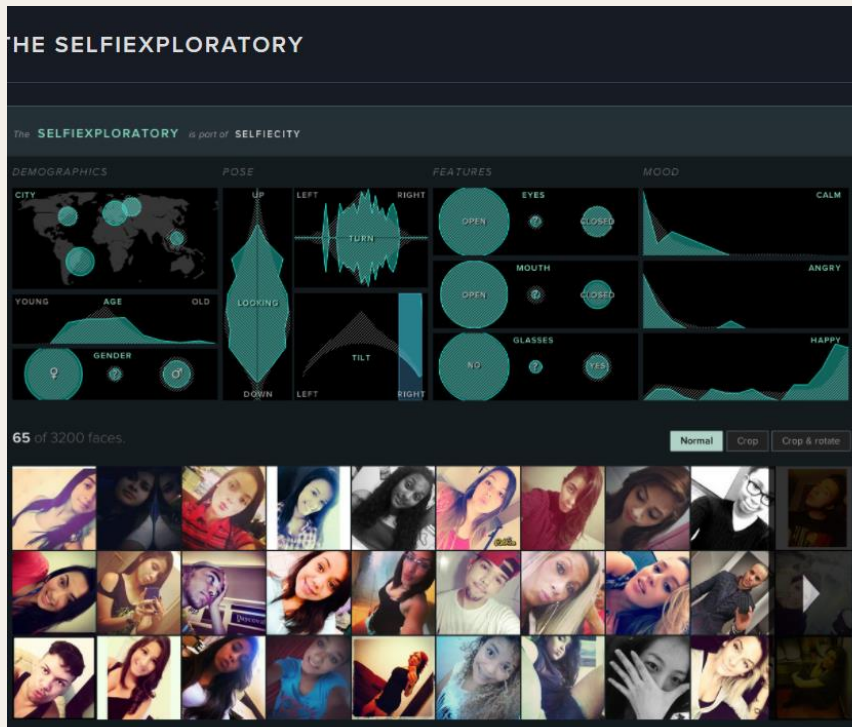
SelfieCity

<http://selfiecity.net/>



SelfieCity

<http://selfiecity.net/>



Experiment with all the data
we collected.

*Do angry people tilt their heads more strongly?
And what is a characteristic mood for people in
Moscow? Find out!*

→ LAUNCH

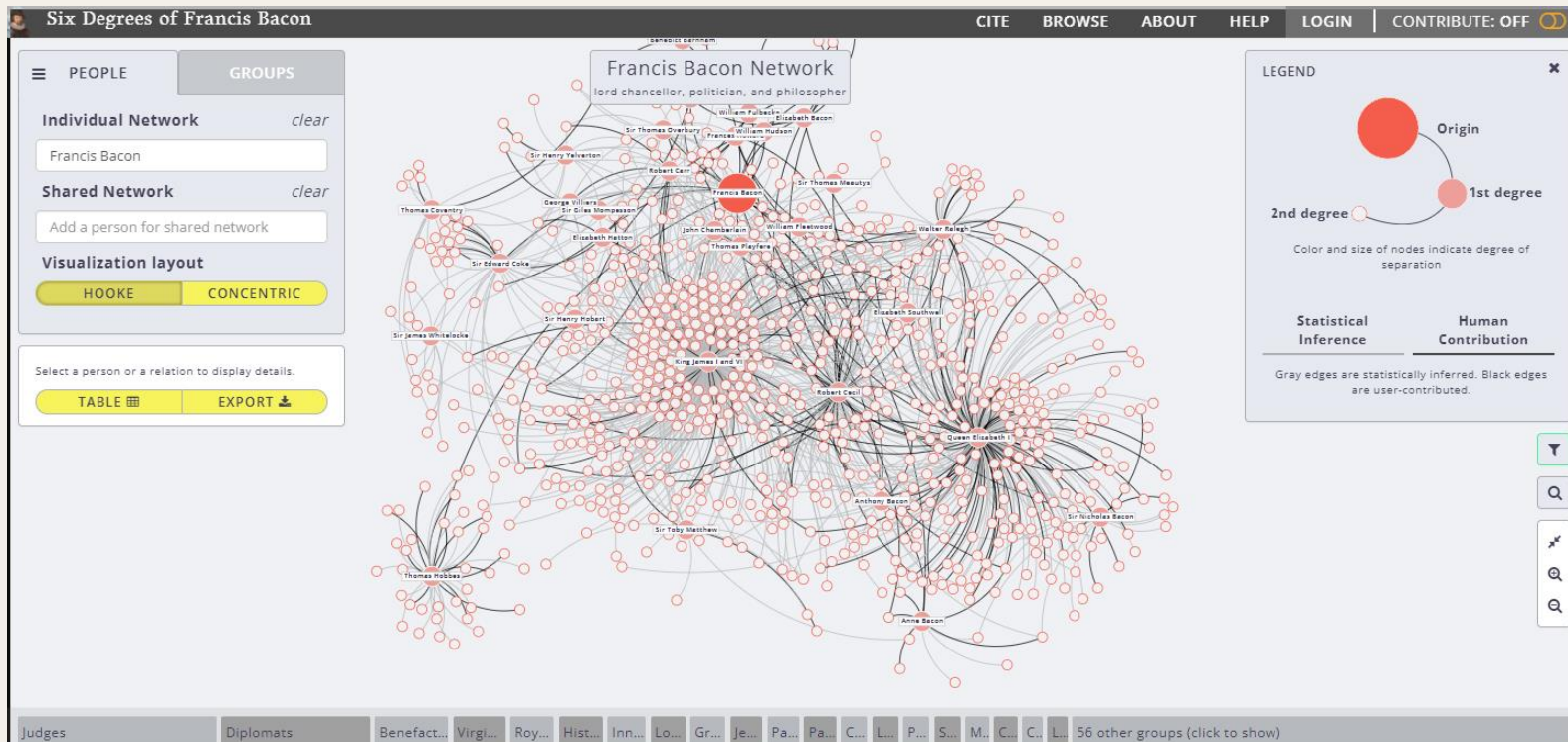
Six degrees of Francis Bacon

<http://www.sixdegreesoffrancisbacon.com/>

“Six Degrees of Francis Bacon is a digital reconstruction of the early modern social network that scholars and students from all over the world can collaboratively expand, revise, curate, and critique. Unlike published prose, Six Degrees is extensible, collaborative, and interoperable: extensible in that people and associations can always be added, modified, developed, or removed; collaborative in that it synthesizes the work of many scholars; interoperable in that new work on the network is put into immediate relation to previously studied relationships. This website is hosted by Carnegie Mellon University Libraries, and data is available for download both on this site and as part of the Folger Shakespeare Library's digital collections.”

Six degrees of Francis Bacon

<http://www.sixdegreesoffrancisbacon.com/>



Six degrees of Francis Bacon

<http://www.sixdegreesoffrancisbacon.com/>

RECENTLY APPROVED SUBMISSIONS

New People

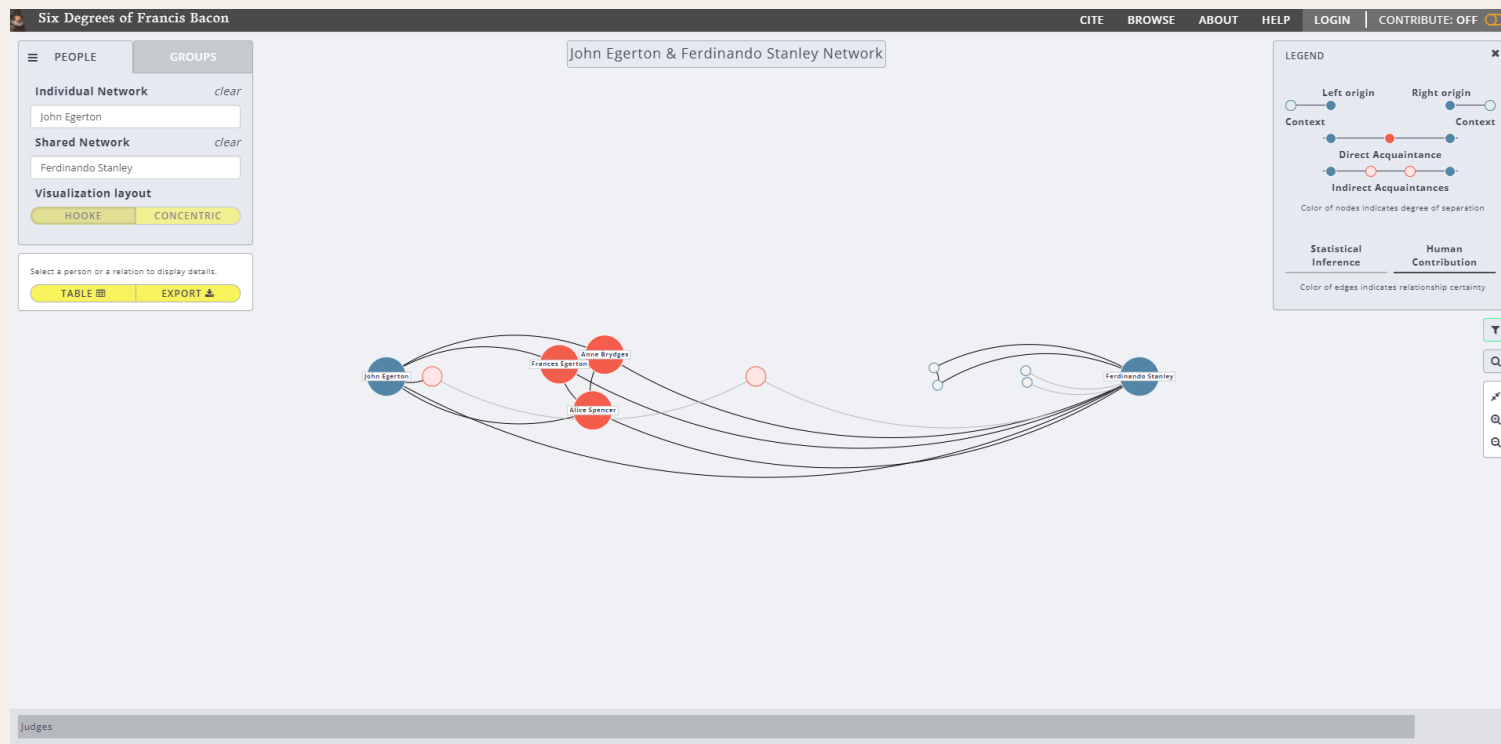
Name	Historical Significance	Birth	Death	Gender	Created By	References
Mary Fillis		CA 1577	AF/IN 1597	female	ZariaH	
Mabel Dacre		CA 1490	CA 1533	female	ZariaH	
Alice Sutton		CA 1483	IN 1554	female	ZariaH	
Susan Shore		CA 1577	IN 1648	female	ZariaH	
Mary Parker		AF/IN 1496	AF/IN 1606	female	ZariaH	

New Relationships

Source Name	Target Name	Start Date	End Date	Created By	References
Richard Stonley	William Paulet	AF/IN 1520	BF/IN 1572	cnwarren	Heal, F. Stonley, Richard (1520/21–1600), administrator and diarist. Oxford Dictionary of National Biography. Retrieved 15 Jun. 2021, from https://www.oxforddnb-com.proxy.library.cmu.edu/view/10.109.9780198614128-e-109639 .
Sir William Petre	Richard Stonley	AF/IN 1520	BF/IN 1572	cnwarren	Heal, F. Stonley, Richard (1520/21–1600), administrator and diarist. Oxford Dictionary of National Biography. Retrieved 15 Jun. 2021

Six degrees of Francis Bacon

<http://www.sixdegreesoffrancisbacon.com/>



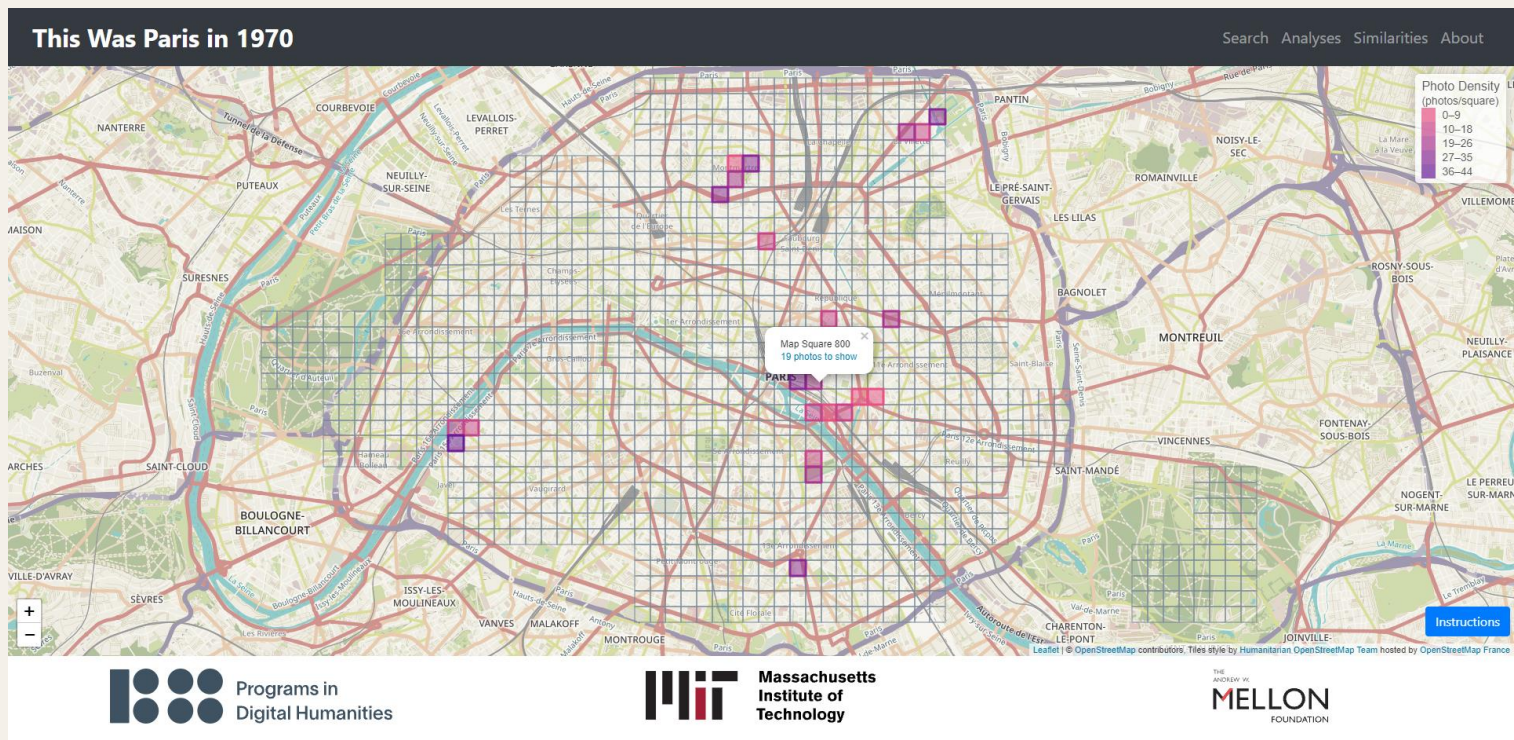
This Was Paris in 1970

<https://paris1970.dhmit.xyz/>

“Welcome to the map view of Paris in 1970! The images in this archive are from a photography competition in 1970's Paris hosted to capture Paris before rapid industrial change altered the city permanently.”

This Was Paris in 1970

<https://paris1970.dhmit.xyz/>



This Was Paris in 1970

<https://paris1970.dhmit.xyz/>

This Was Paris in 1970

[Search](#) [Analyses](#) [Similarities](#) [About](#)

Map Square 800



This Was Paris in 1970

<https://paris1970.dhmit.xyz/>

This Was Paris in 1970

[Search](#) [Analyses](#) [Similarities](#) [About](#)

[Back to Map Square](#)

Map Square 800, Photo 19



Photographer

Unknown (Number: 10085)

Photographer Caption

RUE CLOCHE PERCÉE (de la rue rivioli)

Visual Analysis

None selected

% Whitespace

19%

Objects Detected

None

Average Amount of Detail

113

Is It a Portrait

Yes

Similar Photos (% Similarity)

- [Map Square 955, Photo 39](#) (51%)
- [Map Square 237, Photo 8](#) (51%)
- [Map Square 6, Photo 22](#) (53%)
- [Map Square 237, Photo 2](#) (55%)
- [Map Square 236, Photo 1](#) (56%)
- [Map Square 123, Photo 20](#) (56%)
- [Map Square 6, Photo 2](#) (57%)

Women Writers Project

<https://www.wwp.northeastern.edu/>

“The Women Writers Project is a long-term research project devoted to early modern women's writing and electronic text encoding. Our goal is to bring texts by pre-Victorian women writers out of the archive and make them accessible to a wide audience of teachers, students, scholars, and the general reader.”

WOMEN WRITERS PROJECT

HOME

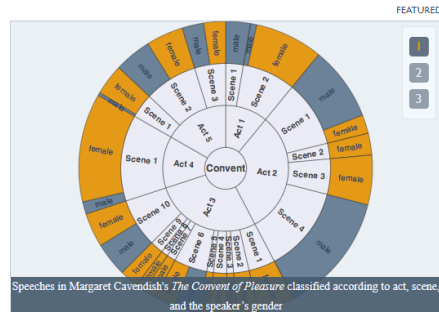
WOMEN WRITERS ONLINE

EDUCATION & OUTREACH

RESEARCH & PUBLICATIONS

ABOUT

The Women Writers Project is a long-term research project devoted to early modern women's writing and electronic text encoding. Our goal is to bring texts by pre-Victorian women writers out of the archive and make them accessible to a wide audience of teachers, students, scholars, and the general reader. We support research on women's writing, text encoding, and the role of electronic texts in teaching and scholarship.



ANNOUNCEMENTS

September 2, 2021

Call for Participation: Word Vectors for the Thoughtful Humanist »

June 15, 2021

Five New Texts Added to Women Writers Online »

June 17, 2020

Eight New Texts Added to Women Writers Online »

March 2, 2020

Women Writers Online is Free for the Month of March »

FEATURED QUOTATION

No! instead of aspiring to be inferior men, let us content ourselves with becoming noble women...but let not sex be carried into every thing.

Lucy Aikin
Epistles on Women, 1810

UPCOMING EVENTS

MAY 16 Word Vectors for the Thoughtful Humanist: Intensive, Teaching-focused »

BLOG POSTS

May 25, 2021
Representing Race in the Early Modern Archive »

April 21, 2021
Experiments in Tokenization for Word Embedding Models »

Feb 1, 2021

Women Writers Project

<https://www.wwp.northeastern.edu/>

WWP PUBLICATIONS AND PRESENTATIONS

WOMEN WRITERS ONLINE AND RELATED RESOURCES

- *Women Writers Online*, a digital collection of early women's writing in English
- *Women Writers in Context*, a collection of scholarly essays and explorations of women's writing
- *WWP Curricular Materials*, including teaching materials and a searchable database of syllabi
- *Women Writers in Review*, a collection of eighteenth- and nineteenth-century reviews, publication notices, literary histories, and other texts responding to works by early women writers

TOOLS

[Tweet](#)

REPORTS ON GRANT-FUNDED WORK

- *Taking TEI Further*, 2011–2016, funded by the National Endowment for the Humanities: [Final report and white paper](#)
- *Cultures of Reception*, 2010–2015, funded by the National Endowment for the Humanities: [Final report](#)
- *Encoding Names for Contextual Exploration in Digital Thematic Research Collections*, 2008–2009, funded by NEH: [Final report and white paper](#)
- *Seminars in Humanities Text Encoding with TEI*, 2007–2009, funded by NEH: [Final report](#)
- *Creating a Guide for Encoding Early Printed Books*, 2003–2006, funded by NEH: [Final report](#)
- *A Textbase of Pre-Victorian Women's Writing in English*, 1997–2000, funded by NEH: [Final report](#)
- *Renaissance Women Online*, 1997–1999, funded by the Andrew W. Mellon Foundation: [Initial report](#) and [final report](#)

Women Writers Project

<https://www.wwp.northeastern.edu/>

TEXTS INCLUDED IN WOMEN WRITERS ONLINE, SORTED BY TITLE

View list sorted by author or date, or view only recently added texts.

Academia: or, the Humours of the University of Oxford, 1691 — D'Anvers, Alicia (Clarke)

The Accession Speech and Prayer, 1558-01-23 — Elizabeth I

Account of the American Baptist Mission to the Burman Empire, 1827 — Judson, Ann H.

Address Delivered With Applause at the Federal-Street Theatre, 1802 — Gannett, Deborah (Sampson)

An Address to the People of Wapping and its Environs, c. 1770 — Cooper, Mrs.

Admirable Events (Epistle Dedicatory), 1639 — Du Verger, Susanne, trans. (from Jean-Pierre Camus)

The Adulateur, 1773 — Warren, Mercy (Otis)

Adventures of Eovaaï, Princess of Ijaveo, 1736 — Haywood, Eliza (Fowler)

The Agreeable Caledonian, 1728 — Haywood, Eliza (Fowler)

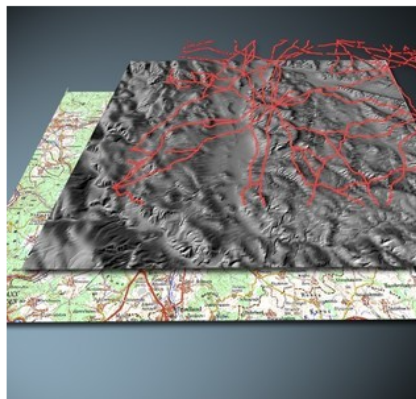
Albina, a Tragedy, 1813 — Cowley, Hannah (Parkhouse)

TOOLS

[Tweet](#)

Viabundus

<https://fgho.eu/en/projects/viabundus>



The project contributes to the study of the premodern economy and society by creating a digital infrastructure for individual research projects. The digital map of premodern overland trade routes in northern Europe will be accessible online as an Open Access Resource (date to be announced). Printed atlases like Bruns/Weczercka's 'Hanseatic Trade Roads' (Hansische Handelsstraßen, 1962-1968) and research literature on trade routes provide the basic data for this digital research tool, supplemented by archival research. Next to digitized road systems, the map and the accompanying database will also cover spatiotemporal data, such as fairs and tolls etc.

Read up on the project and the progress we made on the [Project-Blog](#) and on [Twitter](#).

The Viabundus project is carried out by the IHLF (Institute for Historical Regional Studies) at the University of Göttingen in cooperation with the FGHO. Work on the area of today's Lower Saxony is funded by [Pro*Niedersachsen](#) and is set to run until the end of 2020. It covers the area of today's Lower Saxony. You can find further information on the project [here](#).

Viabundus

<https://fgho.eu/en/projects/viabundus>

Opening remarks by Dr. Angela Huang, director of the Research Centre for Hanse and Baltic History:



Σύνδεσμοι

- Tabak, E. “A Hybrid Model for Managing DH Projects,” <http://www.digitalhumanities.org/dhq/vol/11/1/000284/000284.html>
- Bezerra, D. R., Dias-Neto, A. C. and da Silva Barreto, R. “ARDev: a methodology based on scrum principles to support research management on software technologies,” Proceedings of 24th Annual International Conference on Computer Science and Software Engineering, Riverton, NJ, November 2014, <https://dl.acm.org/doi/10.5555/2735522.2735584>
- Svensson, P. “Humanities Computing as Digital Humanities,” Digital Humanities Quarterly, 3(3). Available at: <http://www.digitalhumanities.org/dhq/vol/3/3/000065/000065.html>
- Tabak, E. “Jumping between context and users: A difficulty in tracing information practices,” Journal of the Association for Information Science and Technology, 65(11): 2223–32.

Ερωτήσεις αυτοαξιολόγησης

- Ποιοι μπορεί να είναι οι ρόλοι των συμμετεχόντων σε ένα έργο DH, ποια δεδομένα διαχειρίζεται ο καθένας και για ποιες διεργασίες είναι υπεύθυνος;
- Πώς θα μπορούσαμε να ορίσουμε τα κριτήρια αποδοχής μιας διεργασίας;
 - Π.χ. της ψηφιοποίησης των εξωφύλλων ενός περιοδικού
- Αναφέρετε πλεονεκτήματα και μειονεκτήματα της συνεργασίας διαφορετικών φορέων σε ένα έργο DH
 - Π.χ. δημόσιος/ιδιωτικός τομέας, εταιρία/πανεπιστήμιο/ερευνητές, κλπ.
- Αν έχουμε τα ίδια δεδομένα και τεχνολογία στη διάθεσή μας, πώς μπορούμε να διαφοροποιήσουμε το τελικό αποτέλεσμα και με ποια κριτήρια;