

General notes on tool collection



- selection of tools useful for DH
- online/web & offline/desktop applications
- (mainly) open & proprietary software
- 'grown' and open collection (non-exhaustive and exemplary selection)
- different tools – different purposes

Tool collection

Code editors and IDEs

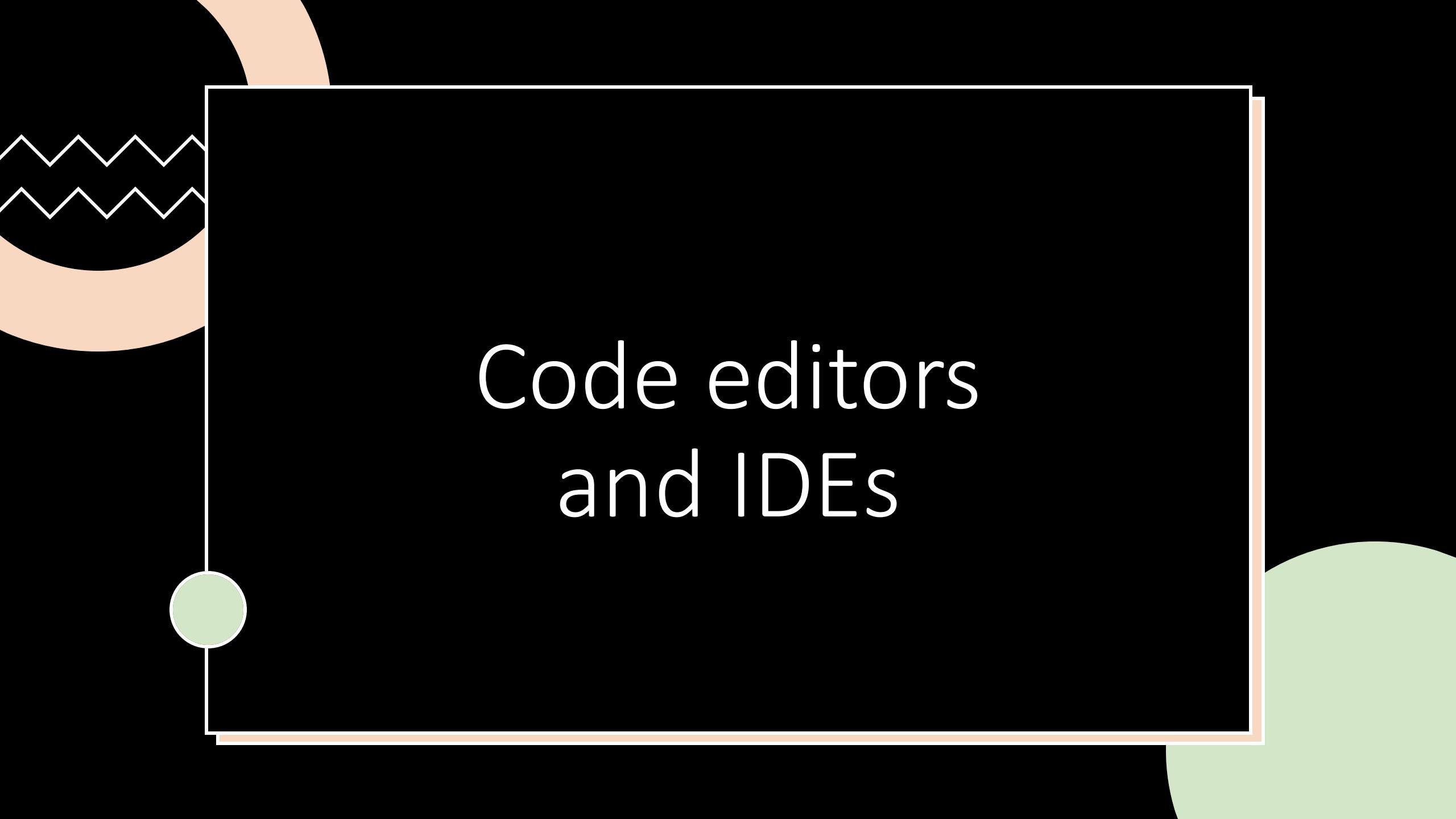
Automated Text Recognition

Annotation

Data Collection & Organisation

Data Analysis & Visualisation

GIS, 3D Visualisation, and more...



Code editors and IDEs

Visual Studio Code <https://code.visualstudio.com/>

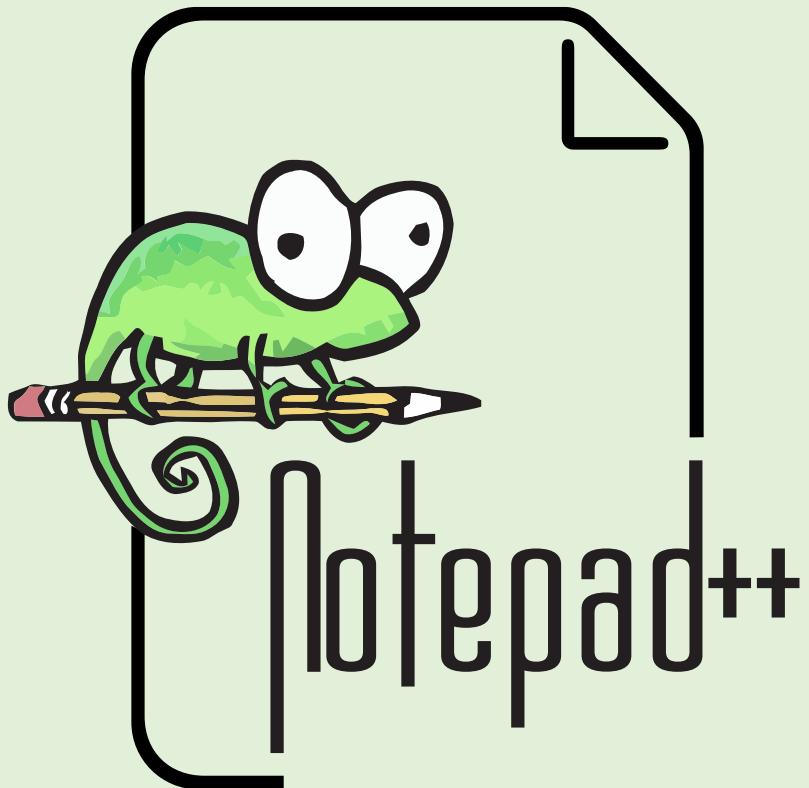
- Open Source Code Editor
- Extensions for various (programming) languages



The screenshot shows the Visual Studio Code interface. On the left, there's a sidebar with icons for file operations like copy, paste, search, and refresh. The main area has a dark theme. At the top, it says "EXTENSIONS: MARK...". Below that is a list of installed extensions:

- Python by Microsoft (42.7M)
- C/C++ by Microsoft (23.5M)
- Jupyter by Microsoft (22.7M)
- ESLint by Dirk Baeumer (16.5M)
- Prettier - ... by Prettier (15.3M)
- Pylance by Microsoft (15.3M)
- Live Server by Microsoft (15M)

On the right, there are three code editors showing JavaScript files: "serviceWorker.js", "App.js", and "# App.css". The "serviceWorker.js" file contains code related to service workers, including registration and navigation logic. The "App.js" and "# App.css" files show basic boilerplate code.



<https://notepad-plus-plus.org>

- Free and open-source text/code editor
- "Notepad replacement", but can do much more!
- Only for Windows

The screenshot shows the Notepad++ interface with a file named "hands-on_1.md" open. The file contains a series of numbered steps and notes for configuring Git. The code block below is a transcription of the visible text in the Notepad++ window.

```
1  # Hands-on instructions
2
3  ## 1. Configuring git and creating a new repository from the command line
4
5  + Open a terminal (UNIX shell / bash is assumed).
6
7  + Let's make sure that the `git` command is available: `$ git --version`
8
9  + One-time Git setup (set author info, text editor, line endings):
10   + `$ git config --global user.name "<your name>"`  

11   + `$ git config --global user.email "<username>@users.noreply.github.com"`  

12   + `$ git config --global core.editor "nano -u"`  

13   + `$ git config --global core.autocrlf input` for Mac/Linux, `true` for Windows  

14   + Check settings: `git config --global --list` (or `$ cat ~/.gitconfig`)
15
16  + Notes:
17    + The email you set here will potentially be visible to everyone on the internet, which is why you might want to use the scheme above.  

18    + By default, the text editor is used for writing commit messages, but you can also use `git commit -m '<message>'`. Set your preferred edi
19    + These (and all other config settings) can also be set on a per-repository level by omitting the `--global` flag. Per-repository settings
20    + For more information on line endings (and Git config in general), see <https://www.git-scm.com/book/en/v2/Customizing-Git-Git-Configurations>
21
22  + Let's create a brand-new repository from scratch:
23    + Using the `cd` command, navigate to the parent directory of where you want to create your new repository (e.g. `~/Documents` if you want
24    + Create a new folder there and `cd` into it: `$ mkdir my_first_repo`, `$ cd my_first_repo`
25    + Initialize a new repository: `$ git init`
26
27  + Take a quick look at the newly created `.git` directory:
28    + List contents of current directory ('-a' option shows hidden items): `$ ls -a`
29    + List contents of `.git`: `$ ls .git`
30    + View `.git/config` file: `$ cat .git/config`
31
32
33  ## 2. Working with files and commits from the command line
34
35  + We continue with the fresh repository created in the first hands-on session.
36
37  + Create and save a new text file with some text (traditionally "hello world") in it: `$ echo 'hello world' > hello.txt` (or use your favorite
38
39  + `$ git status` should list the newly created file under "Untracked files".
```

Code editors or IDEs?

- IDE = Integrated Development Environment
- More tools for debugging and compiling
- Usually for a specific language

More info: <https://blog.jetbrains.com/webstorm/2024/03/ides-vs-code-editors>



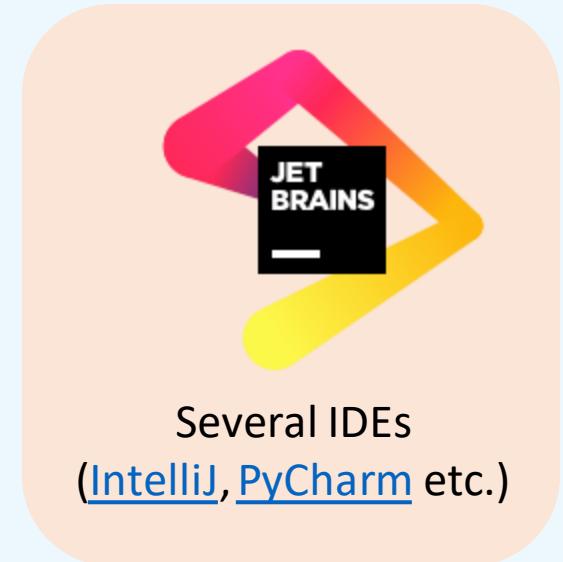
[Visual Studio](#)
(Win, Mac, Linux)



[Xcode](#)
(Mac only)



eclipse
[Eclipse IDE](#) for Java
(Win, Mac, Linux)



GitHub Student Developer Pack

Learn to ship software like a pro. There's no substitute for hands-on experience. But for most students, real world tools can be cost-prohibitive. That's why we created the GitHub Student Developer Pack with some of our partners and friends.

[Sign up for Student Developer Pack](#)

Love the pack? Spread the word

<https://education.github.com/pack>

Experiences

Discover the best ways to use pack offers with Experiences. Experiences are curated bundles of pack partner products, GitHub tools, and other resources that are designed for you learn new skills and make the most out of the Student Developer Pack and your journey in Global Campus.



Tools

AI Prompting & Technical Writing

This Experience offers the tools, top tips and resources you'll need to get started becoming a technical wizard with words.

Offers in this bundle



[Learn more about AI Prompting & Technical Writing >](#)



Tools

Security and Monitoring

Explore industry standard tools that keep your code safe, your permissions secure and your infrastructure monitored so that you can confidently scale your brilliant ideas.

Offers in this bundle



[Learn more about Security and Monitoring >](#)



Primer

Primer: Copilot

What is Copilot and how do you get started using it? Use this Experience as a Primer on Copilot. We share tutorials, guides and templates to learn how to use AI in your development process.

Offers in this bundle



[Learn more about Primer: Copilot >](#)

Automated Text Recognition (ATR)



<https://readcoop.eu/de/transkribus/>

- Layout and text recognition
- HTR = Handwritten Text Recognition, can be used for handwritten and printed text
- Online and local version available (Expert Client recommended)

The screenshot shows a historical document from 1706. The text is handwritten in German. A transcription table below the document lists numbered items corresponding to the text:

5-1	Ankunfft aller Hoch= und nidrigen Stands=Personen.
6-1	Den 30. Januari 1706.
6-2	Cärntner=Thor. Herr Häckel / vom
6-3	Kriegs=Commissariat, kombt auß Jta=
6-4	lien / log. beym wilden Mann.
6-5	Den 31. dito.
6-6	Cärntner=Thor. Herr Graff von Leng=
6-7	heim / kombt von Grätz / log. auf der
6-8	Mehl=Gruben.

A green circle with the number 6 is placed over the word "Mehl=Gruben". In the bottom right corner, there is a small illustration of a bird and the word "Transkribus".



Transkribus ScanTent read&search Über Ressourcen Pläne & Preise

0

Anmelden

App

Sprache: Deutsch



Handschrift Druckschrift

KI-Modell: The German Giant I

Ziehen Sie hier ein Bild rein
oder

Datei auswählen...

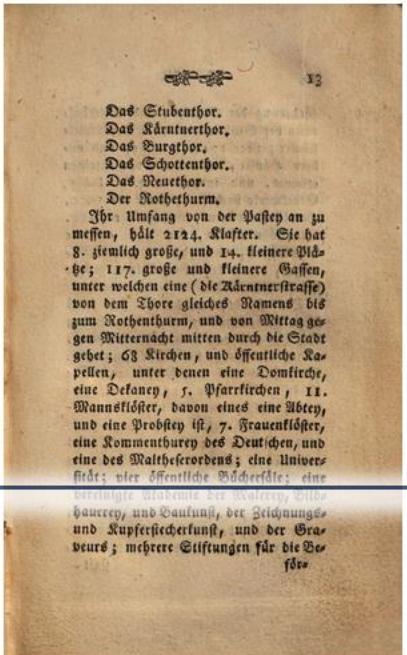
PNG oder JPG bis zu 10 Mb

Mit dem Hochladen eines Bildes akzeptieren Sie unsere Allgemeinen Geschäftsbedingungen und unsere Datenschutzrichtlinie.



Sprache: **Deutsch****Handschrift** Druckschrift

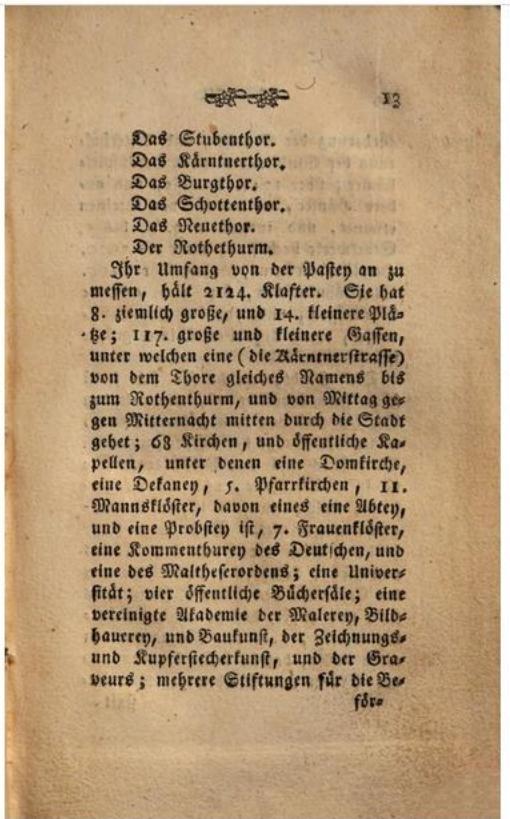
KI-Modell: The German Giant I ⓘ

**Transkribieren..**

Sprache: Deutsch

Handschrift Druckschrift

KI-Modell: The German Giant I ⓘ

+
-

73

Das Stubenthor.
Das Kärtnerthor.
Das Burghor.
Das Schottenthor.
Das Neuethor.
Der Rotheturm.
Ihr Umfang von der Pastey an zu messen, hält 2124. Klafter. Sie hat 8. ziemlich große, und 14. kleinere Plätze; 117. große und kleinere Gassen, unter welchen eine (die Kärtnerstrasse) von dem Thore gleiches Namens bis zum Rotheturm, und von Mittag gegen Mitternacht mitten durch die Stadt geht; 63 Kirchen, und öffentliche Kapellen, unter denen eine Domkirche, eine Dekaney, 5. Pfarrkirchen, 11. Mannsklöster, davon eines eine Abtey, und eine Probstey ist, 7. Frauenkläster, eine Kommissarien des Deutschen, und eine des Malteserordens; eine Universität; vier öffentliche Büchersäle; eine vereinigte Akademie der Maleren, Bildhauern, und Baukunst, der Zeichnungs- und Kupferstecherkunst, und der Graueurs; mehrere Stiftungen für die Bes

Ⓐ Anderes Bild
hochladen

PDF

DOC

>In die Zwischenablage
kopieren





Wiener Zeitung

paragraph

mit 200., s. andere mit 300., und 1. dets mit 500. Ducaten abgeführt worden. Und zumal die dermalige Pachtung des sogenannten Lotto di Genova in andere Hände gekommen, so wird hiermit zu wissen gesucht, daß alldiejenige, welche etwann noch ein oder andere Gewinne an sich zu hängen haben, sich in Zeit von 6. Monaten gegen Probierung ihrer Originalloose behörig melden sollen, wo im widrigsten nach Verstießung dieser Zeit, nach Vor schrift der Allerhöchsten Verordnung derselben loose für null und nichtig angesehen, und ohne einzigen Ausnahm oder Ver bes halt für ungültig und kastlos erklärt werden sollen.

Liste deren Verstorbenen zu Wien in und vor der Stadt

Den 11. Mart. In der Stadt.
Der Wohlebehag. Dr. Joh. Bapt. Meiss, Chur pfalz. Hofräht, und R. R. Hofkriegs-rähts agent, im Regensburgerhof, alt 69. J.
Den Hrn. Martin. Panholzer, F. R. Kammer thürhüt. f. R. Elsab., im fl. Burgis. H. in der Vogler gat, alt 5. J.
Fr. Maria Eva Ranft. Witwe, im Seelauffl. H. im tief. Grab., alt 2. J.
Jacob Kiezer, Herrschaffl. Vertier, im Walleris. H. am Kohlmarkt, alt 47. J.
Köhrl. Wallin, led., im Auferschmidis. H. im tief. Grab., alt 38. J.
Cathar. Haberleiter, ledig., im Goberis. H. in der Wolseit, alt 80. J.

Vor der Stadt.

Dem Joh. Rich. Schuf. Stadt-musico, f. R. Churf., bezw geld. 1. am Spitalb., s.v.J.

list

wen verlissatt.

In der Stadt.

Dem Hrn. Domin. Ant. Jäger, Burgeis. Handelsmann, f. Jr. Anna Bar. im Wahingeris. H. am alt. Baurenmarkt, alt 77. J.

Vor der Stadt.

Anna Mayrin, ledig., bezw weiss. Stern in der Josephs., alt 61. J.

Dem Adam Weir, arm. im Grandsait., außer Mariahilf, alt 80. J.

Dem Paul Kober, Endl., 15. Garb., im Pinapflis. H. auf der Landstr., alt 12. J.

Franz Wagner, Schmalztrag., im Wegentreteris. H. in der Leopoldsh., alt 68. J.

Sophia s. Person darunter 1. Kind.

Den 13. Mart. In der Stadt.

Dem Hochadel geb. Hrn. Reich. von Seitzinger, Heer R. R. Rejst. Hofräht, und Referendar, bez dem Hofkriegs-räht, f. gr. Carlina, geb. v. Bludenzhofen, welche sich zerstalten, ist beim gr. Rö. in der Schulerstr. abgabom R.R. Staats und Landgerichte beschaut worden, alt 51. J.

Der Wohl. Dr. August. Auregus, neistl. Priest. im Abrechekburgis. H. im Berg-hof, 76. J.

Dem Kor. Beisser, Burgeis. Schuhmacher. W. Eva, im Schelleris. H. im Körbergs., alt 66. J.

Dem Joh. Christian Sch. — in, Burgeis. Handelsmann, f. R. Joh. Christian, im Kernis. H. am Salz gr., alt 1. J.

Dem Math. Crombeder, gen. R. R. Krab., f. W. Eva, bezw gold. A. S. E. in der Himmel storts gat, alt 69. J.

Dem Joh. Mich. Frank, Herrschaffl. Heobud., f. W. Magdal., im Coffe-federis. H. auf der Scheitnbaust., alt 26. J.

Gerd. Krepliu, led. M., im Eilis. H. in der Gub gerstr., alt 29. J.

Vor der Stadt.

Der Wohl. adelige. Dr. Joh. a. Dorelli, gen. R. R. Oberkraut-müts, 15. Altingeris. H. in der Josephs., alt 50. J.

heading

catch-word



Field Model

ID: 56638

Private Model

German Newspapers 18th Century - WrDiarium M3

by NinaClaudia.Rastinger@oeaw.ac.at Nov 16, 2023

Trained paragraph, heading, header, list, main-title, imprint, count-date, catch-word, figure, separator-double, separator-single, signature
Tags

% Accuracy 81.81%

Trained on handwritten

Edit Show Description

Transkribus Scholarship Programme

The Transkribus Scholarship Programme is designed to encourage and support future scholars as this is one of our core values. With this programme we want to facilitate the research in historical documents.

The programme is open to students and teachers from any programme or discipline. Through this global programme you can get free credits to use for the Handwritten Text Recognition technology in Transkribus. Transkribus Scholarships are granted based on projects, this means that you need to apply with a specific project that you plan to spend the credits of your Scholarship on.



<https://readcoop.eu/de/transkribus/scholarship/>



eScriptorium

eScriptorium: A Digital Text Production Pipeline for Print and Handwritten Texts using machine learning techniques.



Automatic Transcription

Apply OCR/HTR to images of printed and handwritten documents using shared open models.



Manual transcription

Make use of an ergonomic user interface leveraging modern browser technology to edit segmentations and transcriptions.



Train Models

Create new models or finetune existing ones to improve automatic recognition.



Import/Export

Import and export models and texts transcriptions in a variety of formats. Access data through a full REST API.

<https://escriptorium.readthedocs.io/en/latest/>

OCR4all

Optical Character Recognition (and more) for everyone

Setup guide, user guide, developer documentation and more.

Get Started

Fully free and open-source

OCR4all is and will stay completely free and open-source. No subscriptions, paywalled features or private code.

OCR-D compatible

All future versions of OCR4all are built to be fully compatible with the OCR-D ecosystem



<https://www.ocr4all.org/>

Flexible applicable

From the high-quality processing of challenging manuscripts to the mass full-text recognition of printings

Designed with usability in mind

Create complex OCR workflows through the UI without the need of interacting with code or command line interfaces.

Powerful layout and text annotation included

Manually annotate, correct or compare layout and text elements using the powerful LAREX editor.

Easy cross-platform deployment

Docker and a single command are all it takes to get OCR4all up and running, regardless of your operating system.

Tutorials

- Transkribus Tutorial Playlist on
Youtube: <https://www.youtube.com/playlist?list=PL7UbQtd4qlhIMP1KfdjGW3C-KXTxw4KYb>
- Transkribus Helpcenter: <https://help.transkribus.org/getting-started>
- ForTEXT Tutorial for OCR4all: <https://fortext.net/tools/tools/ocr4all>



Annotation

TEI

<oxygene/>[®]
XML Editor

Text Encoding Initiative

W.H.Auden

77 St Mark's Place
New York City 3
New York 10003
U.S.A.

Air-Mail

Die Baronin
Stella Musulin

Fridau
3200 Obergrafendorf
N.Ö.
Austria



amp-transcript_0004.xml X

TEI

```
157      </surface>
158    </facsimile>
159  <text type="letter" hand="#typed">
160    <body>
161      <div type="transcription" xml:id="transcription_amp_0004" xml:lang="de">
162        <div type="envelope">
163          <pb facs="https://iiif.acdh.oeaw.ac.at/amp/amp_0019/" type="env">
164            <ab>
              <rs type="person" ref="#amp_person_1">W.H.Auden</rs><lb/>
              <address>
                <rs type="place" ref="#amp_place_id_29">77 St Mark's P
                  New York City 3<lb/>
                  New York 10003<lb/>
                  U.S.A.</rs>
              </address>
            </ab>
            <ab>
              <hi rend="underline">Air-Mail</hi>
            </ab>
            <ab xml:lang="de">
              <rs type="person" ref="#amp_person_2">Die Baronin<lb/>
              <space quantity="3" unit="char"/>Stella Musulin</rs><lb/>
              <address>
                <space quantity="5" unit="char"/><rs type="place" ref=
                <space quantity="6" unit="char"/>32<del>o</del>OO Ober
                <space quantity="8" unit="char"/>N.<unclear>Ö</unclear
                <space quantity="9" unit="char"/><foreign xml:lang="en">
              </address>
            </ab>
            <pb facs="https://iiif.acdh.oeaw.ac.at/amp/amp_0020/" type="env">
```

<https://amp.acdh.oeaw.ac.at/>

Semantic Annotation without the pointy brackets

Work on texts and images. Identify and mark named entities. Use your data in other tools or connect to other data on the Web.
Without the need to learn code.



Shared Collections

Upload texts and images. Build your own collections. Invite others to collaborate.



Collaborative Annotation

Create notes. Tag. Discuss. Recogito tracks provenance and maintains version history.



Semantics Included

Tag places, people and events and reveal the relations between them.



Open Data

Export your data into other tools. Make the results of your work more open, visible and re-usable.



<https://catma.de/>

= Computer Assisted Text Markup and Analysis

CATMA 6.0.0

Edgar Allan Poe

THE TELL-TALE HEART-2

Edgar Allan Poe

TELL-TALE HEART

TRUE! —nervous —very, very dreadfully nervous I had been and am; but why will you say that I am mad? The disease had sharpened my senses —not destroyed —not dulled them. Above all was the sense of hearing acute. I heard all things in the heaven and in the earth. I heard many things in hell. How, then, am I mad? Harken! and observe how healthily —how calmly I can tell you the whole story.

It is impossible to say how first the idea entered my brain; but once conceived, it haunted me day and night. Object there was none. Passion there was none. I loved the old man. He had never wronged me. He had never given me insult. For his gold I had no desire. I think it was his eye! yes, it was this! He had the eye of a vulture —a pale blue eye, with a film over it. Whenever it fell upon me, my blood ran cold; and so by degrees —very gradually—I made up my mind to take the life of the old man, and thus rid myself of the eye forever.

Collection currently being edited
Tell-Tale Heart Jan

Tagsets

Tagsets	Tags
character	opponent hero(ine) hero
narrator	attitude questionable morals questionable factual claims
style	address to the reader repetition exclamation parallelism

Visualizations

KWIC

DISTRIBUTION

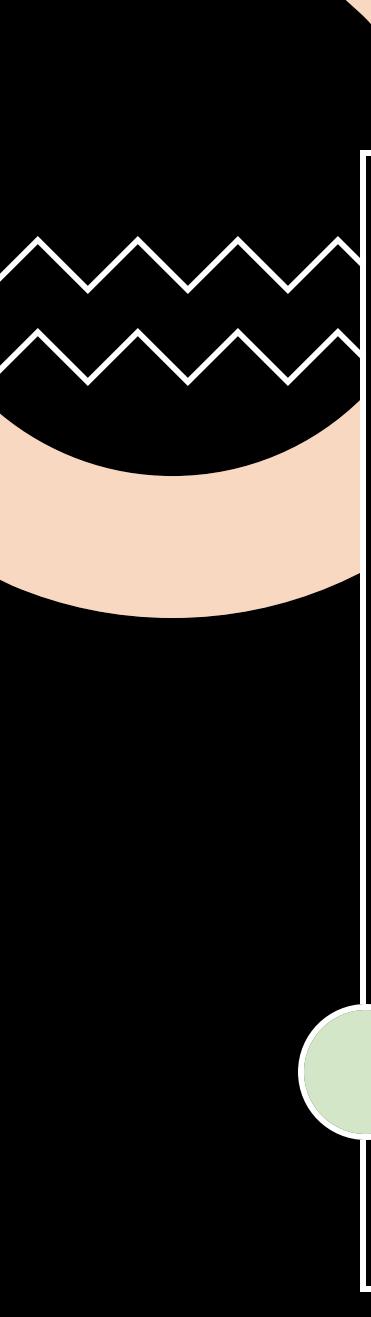
WORDCLOUD

DOUBLETREE

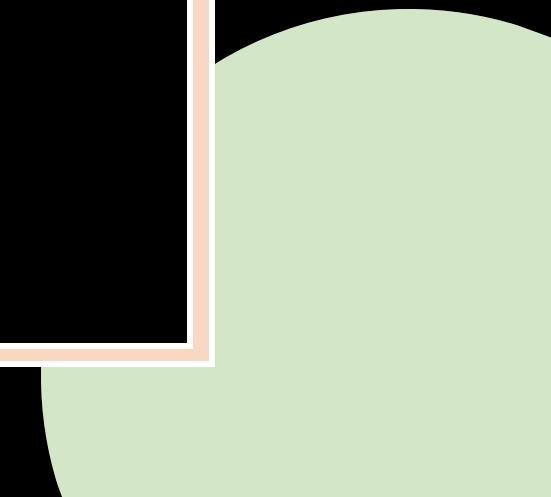
+ GitMA

(<https://catma.de/documentation/access-your-project-data/git-access/gitma/>)

= Python library to process, analyse and manipulate CATMA annotations



Data Collection and Organisation



BeautifulSoup

<https://beautiful-soup-4.readthedocs.io/en/latest/>

– “Python library for pulling data out of HTML and XML files”

```
from bs4 import BeautifulSoup
soup = BeautifulSoup(html_doc, 'html.parser')

print(soup.prettify())
# <html>
# <head>
#   <title>
#     The Dormouse's story
#   </title>
# </head>
# <body>
#   <p class="title">
#     <b>
#       The Dormouse's story
#     </b>
#   </p>
#   <p class="story">
#     Once upon a time there were three little sisters; and their names were
#     <a class="sister" href="http://example.com/elsie" id="link1">
#       Elsie
#     </a>
#     , Lacie
#     , and Tacie
#   </p>
```





Explore your research photos

[Download Tropy for Windows \(x64\)](#)[What's new in version 1.16.2](#)

A screenshot of the Tropy software interface. On the left, a sidebar shows a tree view of 'American Revolution' lists: Costs, Loyalists, American Strategy: South, British Strategy: South, Tactics, Last Import (selected), Deleted Items, and Tags. Under Tags, 'desertion' is selected. The main area displays a grid of 187 historical documents, each showing a thumbnail of a handwritten page from a historical source. A search bar at the top right says 'Search'.



Organisation, metadata enrichment and annotation of images

<https://tropy.org/>



OpenRefine

OpenRefine is a powerful free, open source tool for working with messy data: cleaning it; transforming it from one format into another; and extending it with web services and external data.

[Download](#)

Main features



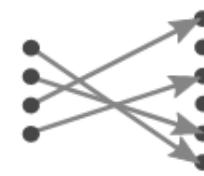
Faceting

Drill through large datasets using facets and apply operations on filtered views of your dataset.



Clustering

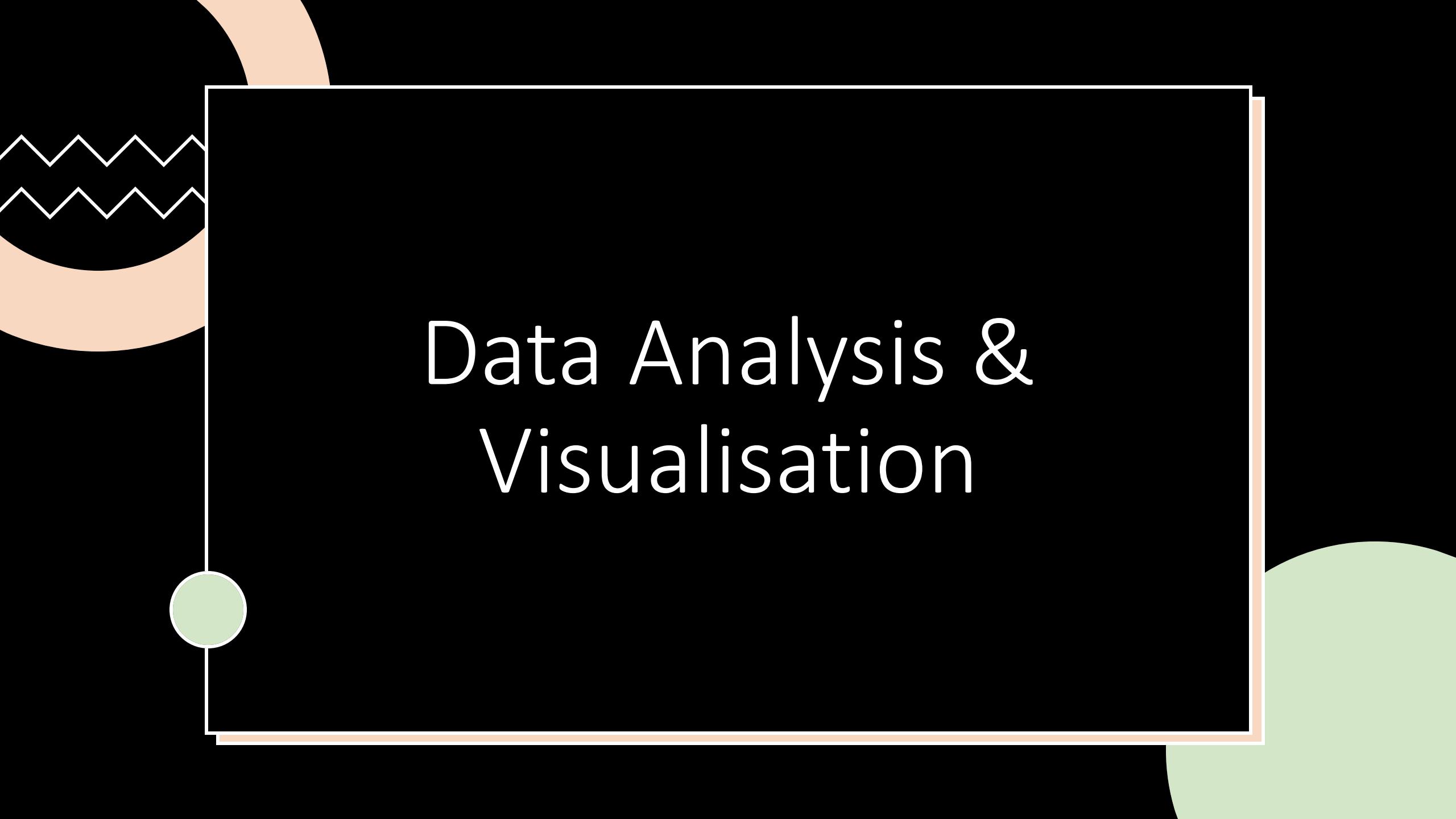
Fix inconsistencies by merging similar values thanks to powerful heuristics.



Reconciliation

Match your dataset to external databases via reconciliation services.

<https://openrefine.org/>



Data Analysis & Visualisation

text analysis tool



VOYANT TOOLS

<https://voyant-tools.org/>

Voyant Tools interface showing analysis of the text "I have just returned from a visit to my landlord..."

Summary: This corpus has 1 document with 1,949 total words and 933 unique word forms. Created 2 seconds ago.

Vocabulary Density: 0.479

Readability Index: 10.400

Average Words Per Sentence: 24.1

Most frequent words in the corpus:

- **mr** (10); **heathcliff** (6); **sir** (4); **little** (4); **joseph** (4)

Contexts:

Document	Left	Term	Right
I have j...	A perfect misanthropist's heaven: and	mr	. Heathcliff and I are such
I have j...	as I announced my name. '	mr	. Heathcliff? I said. A nod
I have j...	A nod was the answer. '	mr	. Lockwood, your new tenant, sir
I have j...	entered the court 'Joseph, take	mr	. Lockwood's horse; and bring up
I have j...	Heights is the name of	mr	. Heathcliff's dwelling. 'Wuthering' being a
I have j...	right time after dinner. But	mr	. Heathcliff forms a singular contrast

Bubblelines:

The chart displays the relative frequency of five words over 10 document segments. The y-axis represents 'Relative Frequency' from 0.0000 to 0.0016. The x-axis represents 'Document Segments' from 1 to 10. The words tracked are sir (cyan), mr (blue), little (orange), joseph (purple), and heathcliff (green). The 'mr' bubbleline starts at a high frequency (approx. 0.0015) in segment 1 and decreases rapidly. The 'heathcliff' bubbleline starts at a moderate frequency (approx. 0.0010) in segment 1 and remains relatively stable. Other words like 'sir', 'little', 'joseph', and 'heathcliff' show lower frequencies across the segments.

Trends:

The trend chart shows the document frequency of the same five words over 10 document segments. The y-axis is 'Relative Frequency' (0.0000 to 0.0016). The x-axis is 'Document Segments' (1 to 10). The trends are generally similar to the bubbleline chart but reflect the cumulative nature of document frequency. 'Mr' has the highest peak frequency, followed by 'heathcliff' and 'sir'.

Correlations:

Document	Left	Term	Right
I have j...	A perfect misanthropist's heaven: and	mr	. Heathcliff and I are such
I have j...	as I announced my name. '	mr	. Heathcliff? I said. A nod
I have j...	A nod was the answer. '	mr	. Lockwood, your new tenant, sir
I have j...	entered the court 'Joseph, take	mr	. Lockwood's horse; and bring up
I have j...	Heights is the name of	mr	. Heathcliff's dwelling. 'Wuthering' being a
I have j...	right time after dinner. But	mr	. Heathcliff forms a singular contrast

Voyant Tools - Stefan Sinclair & Geoffrey Rockwell (© 2023) Privacy v. 2.6.7

VOYANT

see through your text

Text hinzufügen

Geben Sie eine oder mehrere URLs in separate Zeilen ein oder fügen Sie einen Volltext ein.

Öffnen Hochladen Anzeigen

Voyant Tools ist eine webbasierte Lese- und Analyseumgebung für digitale Texte..
Deutsche Übersetzung von Jan Theurl

VOYANT

see through your text

Text hinzufügen

Geben Sie eine oder mehrere URLs

Öffnen

Ein Korpus wählen:

- Shakespeare's Plays
- Austen's Novels
- Mary Shelley's Frankenstein

✓ Öffnen

✓ Anzeigen

Öffnen Hochladen

Voyant Tools ist eine webbasierte Lese- und Analyseumgebung für digitale Texte..
Deutsche Übersetzung von Jan Theurl

Voyant Tools

Cirrus Corpus Terms Links ? Reader Wortrauben Trends Dokumentbegriffe ?



Introduction

The Publishers of the Standard Novels, in selecting Frankenstein for one of their series, expressed a wish that I should furnish them with some account of the origin of the story. I am the more willing to comply, because I shall thus give a general answer to the question, so very frequently asked me — How I, then a young girl, came to think of, and to dilate upon, so very hideous an idea? It is true that I am very averse to bringing myself forward in print; but as my account will only appear as an appendage to a former production, and as it will be confined to such topics as have connection with my authorship alone, I can scarcely accuse myself of a personal intrusion.

It is not singular that, as the daughter of two persons of distinguished literary celebrity. I should very early in life have thought of writing. As a child I scribbled: and

relative Frequenzen

0.00040
0.00035
0.00030
0.00025
0.00020
0.00015
0.00010
0.00005
0.00000

1 2 3 4 5 6 7 8 9 10

Dokument Segmente (Frankenstein or, the Modern Prometheus.)

monster man Zurücksetzen Darstellung

Zusammenfassung Dokumente Phrasen ? Kontexte Bubblelines Kollokationen

Dieses Korpus enthält 1 Dokument mit insgesamt 78,030 verschiedenen/ einzigartigen Wortformen. Erstellt vor ca. 2 Jahren.

Wortschatzdichte: 0.094

Readability Index: 11.653

Durchschnittliche Wortanzahl pro Satz: 22.9

Die häufigsten Begriffe im Korpus

- man (133); life (124); father (113); shall (109); eyes (107)

Elemente

Dokument	links	Begriff	rechts
Franken...	think that I was the	monster	that he said I was
Franken...	in daily fear lest the	monster	whom I had created should
Franken...	of your remaining friends. Abhorred	monster	I Fiend that thou art! The
Franken...	I was in reality the	monster	that I am, I was
Franken...	me. Was I, then, a	monster	, a blot upon the earth

monster ? 31 Kontext erweitern

Voyant Tools , Stéfan Sinclair & Geoffrey Rockwell (© 2024) Privacy v. 2.6.11



corpus analysis tool

AntConc

<https://www.laurenceanthony.net/software/antconc/>

File Edit Settings Help

Target Corpus KWIC Plot File View Cluster N-Gram Collocate Word Keyword Wordcloud

Name: AmE06_Learned Total Hits: 87 Page Size 100 hits 1 to 87 of 87 hits

Files: 80 Tokens: 161469

File	Left Context	Hit	Right Context
1 AmE...	It is, however, prompted by the need to place the	process	of taking moments in context. Moments of the Distribution
2 AmE...	f online distance education. Successful online teaching is a	process	of taking our very best practices in the classroom
3 AmE...	in their parents' homes. The findings demonstrate that the	process	of assimilation was not uniform for all groups. Some
4 AmE...	with the Communist Party of Indonesia, which was in the	process	of being eliminated by Soeharto's New Order government.
5 AmE...	acred texts," the canon of modern children's literature. The	process	of creating or augmenting professional identity relied parti
6 AmE...	eyes, you lack that protein. Now scientists are in the	process	of figuring out which proteins are coded for by
7 AmE...	: (2004), Donlan and Martin (2004), and Pysek et al. (2004).	Process	of invasion At one level, the issue of invasive
8 AmE...	or and the other participants is formed, through which the	process	of knowledge acquisition is collaboratively created. (See Ch
9 AmE...	express an inference. An inference , in turn, is a mental	process	of linking propositions by offering support to one proposition
10 AmE...	een theoretically appropriate for explaining the adaptation	process	of newcomers who arrived in America in the early 20
11 AmE...	motoric instructions, either unmodified or modified by the	process	of overlap. We postulate a parallel language-specific proces
12 AmE...	measured confounding is accounting for the findings, as the	process	of randomization makes the mathematical probability of su
13 AmE...	nt residue of mantle differentiation including the on-going	process	of seafloor spreading and building of island arcs. It
14 AmE...	kszentmihalyi's (1990) concept of "flow" is a more general	process	of self-actualization. In flow, a person's tasks
15 AmE...	a voice of "several strengths." Her voice thereby enacts a	process	of the black community speaking to itself and explores
16 AmF	is not. it is that the critical thinker takes the	process	of thinking seriously, consciously attends to that process an

Search Query Words Case Regex Results Set All hits Context Size 10 token(s)

process Adv Search

Sort Options Sort to right Sort 1 1R Sort 2 2R Sort 3 3R Order by freq

Time taken (creating kwic results): 0.3314 sec



<https://www.sketchengine.eu/nosketch-engine/>

NoSketch Engine

DASHBOARD

Susanne

CONCORDANCE

PARALLEL CONCORDANCE

WORDLIST

KEYWORDS

TRENDS

TEXT TYPE ANALYSIS

CORPUS INFO

You are using NoSketch Engine. These tools are only available in Sketch Engine

The dashboard interface for NoSketch Engine is shown. It features a sidebar with various icons and a main dashboard area. The dashboard includes sections for Concordance, Parallel Concordance, Wordlist, Keywords, Trends, and Text type analysis. A 'CORPUS INFO' button is located in the top right corner. The bottom of the page displays a message indicating the tools are only available in Sketch Engine.



SELECT CORPUS

FREE CORPORA

English Susanne 128,998 words

OPEN

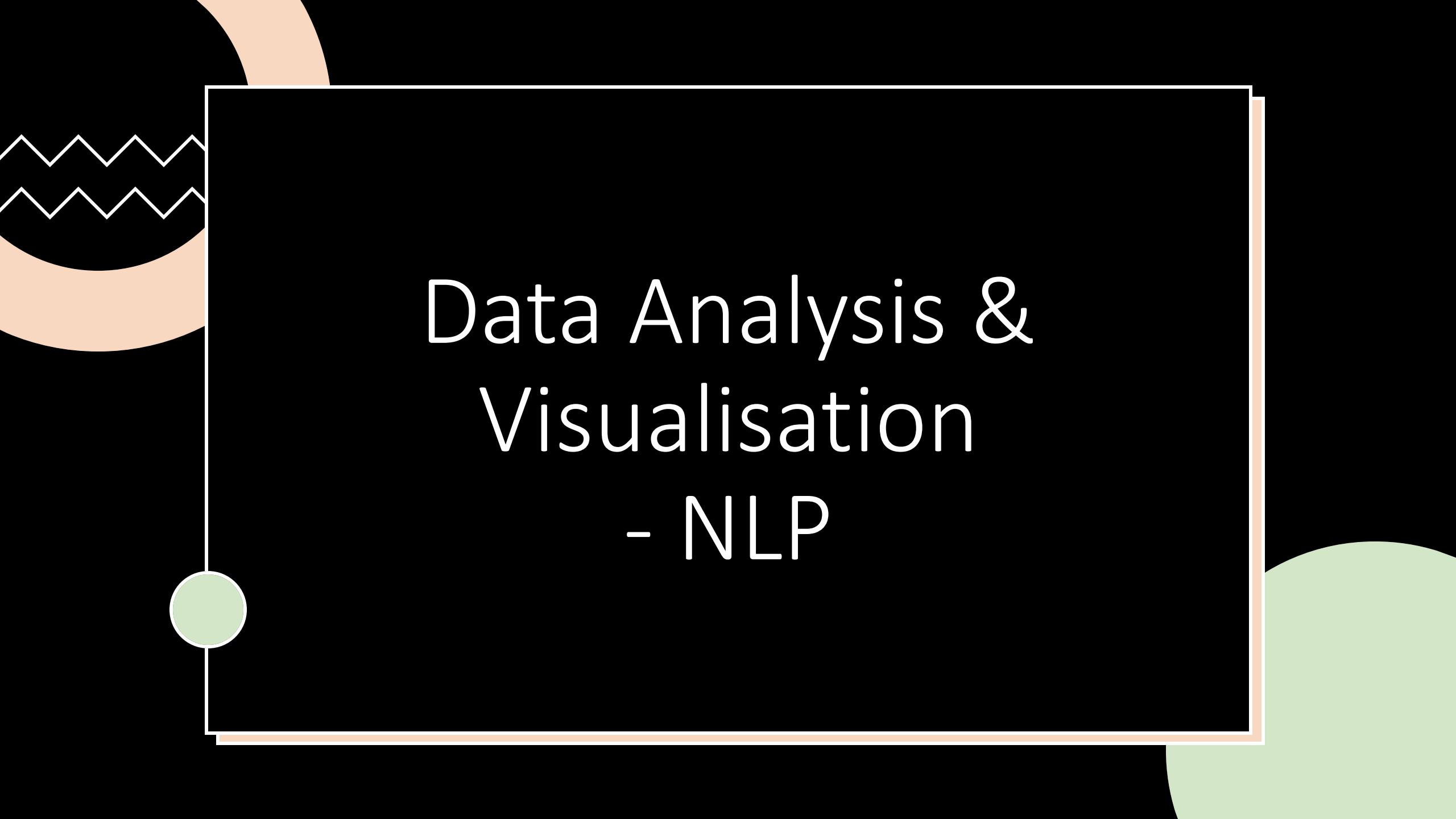
<https://no.sketchengine.eu/crystal/>

Tutorials

- Corpus Analysis with AntConc:

<https://programminghistorian.org/en/lessons/corpus-analysis-with-antconc>

- Getting started with Voyant Tools: <https://voyant-tools.org/docs/#!/guide/start>



Data Analysis & Visualisation - NLP

NLP = Natural Language Processing

Exemplary NLP areas and tasks:

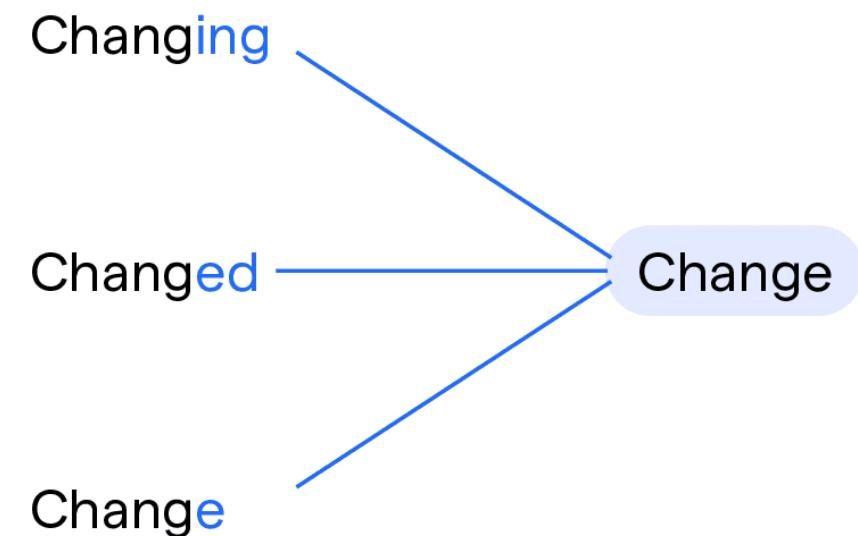
- Tokenization



NLP = Natural Language Processing

Exemplary NLP areas and tasks:

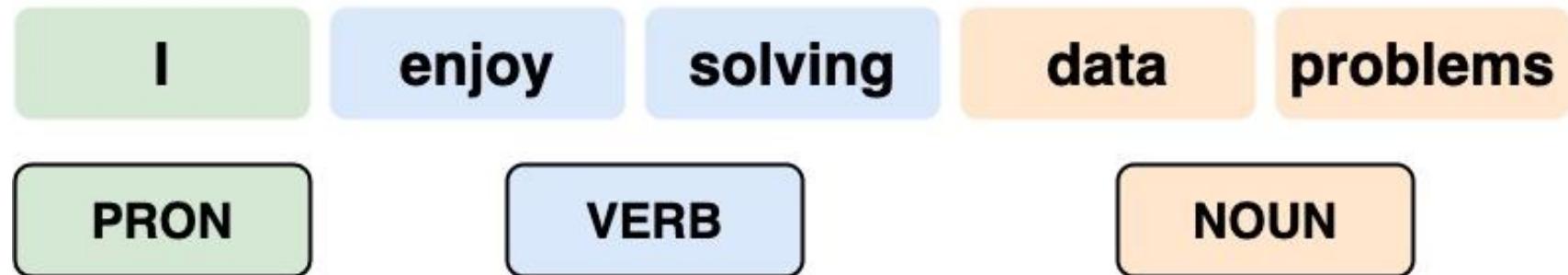
- Tokenization
- Lemmatization



NLP = Natural Language Processing

Exemplary NLP areas and tasks:

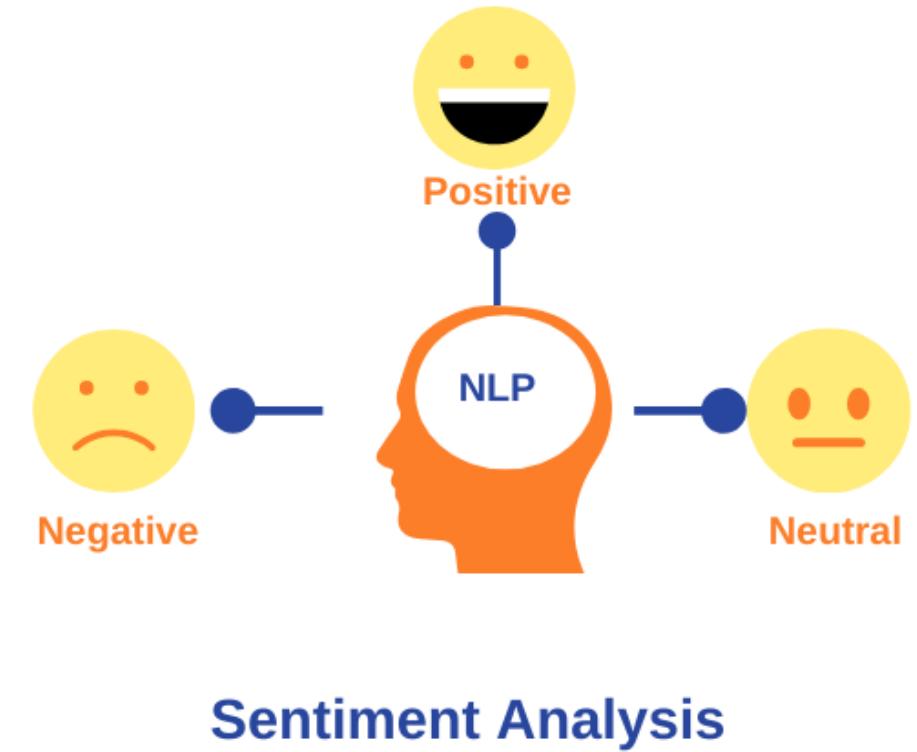
- Tokenization
- Lemmatization
- Part-of-Speech-Tagging (PoS-Tagging)



NLP = Natural Language Processing

Exemplary NLP areas and tasks:

- Tokenization
- Lemmatization
- Part-of-Speech-Tagging (PoS-Tagging)
- Sentiment Analysis



NLP = Natural Language Processing

Exemplary NLP areas and tasks:

- Tokenization
- Lemmatization
- Part-of-Speech-Tagging (PoS-Tagging)
- Sentiment Analysis
- Named Entity Recognition (NER)

On the 15th of September DATE , Tim Cook PERSON announced that
Apple ORG wants to acquire ABC Group ORG from New York GPE
for 1 billion dollars MONEY

Natural Language Toolkit (NLTK)

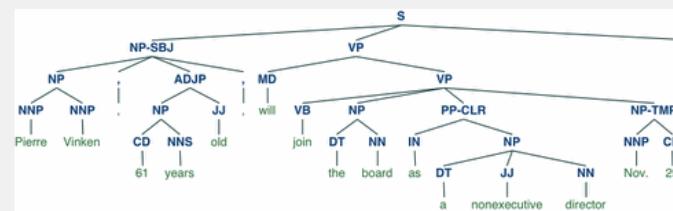
- Python library for NLP
- <https://www.nltk.org/>

Tokenize and tag some text:

```
>>> import nltk  
>>> sentence = """At eight o'clock on Thursday morning  
... Arthur didn't feel very good."""  
>>> tokens = nltk.word_tokenize(sentence)  
>>> tokens  
['At', 'eight', 'o\'clock', 'on', 'Thursday', 'morning',  
'Arthur', 'did', "n't", 'feel', 'very', 'good', '.']  
>>> tagged = nltk.pos_tag(tokens)  
>>> tagged[0:6]  
[('At', 'IN'), ('eight', 'CD'), ("o'clock", 'JJ'), ('on', 'IN'),  
('Thursday', 'NNP'), ('morning', 'NN')]
```

Display a parse tree:

```
>>> from nltk.corpus import treebank  
>>> t = treebank.parsed_sents('wsj_0001.mrg')[0]  
>>> t.draw()
```



Identify named entities:

```
>>> entities = nltk.chunk.ne_chunk(tagged)  
>>> entities  
Tree('S', [(['At', 'IN'], ('eight', 'CD'), ("o'clock", 'JJ'),  
          ('on', 'IN'), ('Thursday', 'NNP'), ('morning', 'NN')),  
      Tree('PERSON', [(['Arthur', 'NNP'])]),  
      ('did', 'VBD'), ("n't", 'RB'), ('feel', 'VB'),  
      ('very', 'RB'), ('good', 'JJ'), ('.', '.')])
```

spaCy

<https://spacy.io/>

Python library for Natural Language Processing
(Tokenization, Pos-Tagging, NER, Sentiment Analysis, ...)



```
# pip install -U spacy
# python -m spacy download en_core_web_sm
import spacy

# Load English tokenizer, tagger, parser and NER
nlp = spacy.load("en_core_web_sm")

# Process whole documents
text = ("When Sebastian Thrun started working on self-driving cars at "
        "Google in 2007, few people outside of the company took him "
        "seriously. "I can tell you very senior CEOs of major American "
        "car companies would shake my hand and turn away because I wasn't "
        "worth talking to," said Thrun, in an interview with Recode earlier "
        "this week.")

doc = nlp(text)

# Analyze syntax
print("Noun phrases:", [chunk.text for chunk in doc.noun_chunks])
print("Verbs:", [token.lemma_ for token in doc if token.pos_ == "VERB"])

# Find named entities, phrases and concepts
for entity in doc.ents:
    print(entity.text, entity.label_)
```

RUN

Industrial-Strength Natural Language Processing

IN PYTHON



Hugging Face



Models



Datasets



Spaces

<https://huggingface.co/>

The screenshot shows the Hugging Face website. On the left, there's a dark sidebar with a "Create Assistants in HuggingChat" button and a yellow emoji. The main content area features the text "The AI community building the future." Below it is a paragraph about the platform being where the machine learning community collaborates. On the right, there's a navigation bar with "Tasks", "Libraries", "Datasets", "Languages", "Licenses", and "Other". The "Models" section is currently selected. It shows a search bar and a list of categories: Multimodal (Text-to-Image, Image-to-Text, Text-to-Video, Visual Question Answering, Document Question Answering, Graph Machine Learning), Computer Vision (Depth Estimation, Image Classification, Object Detection, Image Segmentation, Image-to-Image, Unconditional Image Generation, Video Classification, Zero-Shot Image Classification), Natural Language Processing (Text Classification, Token Classification, Table Question Answering, Question Answering, Zero-Shot Classification, Translation, Summarization, Conversational, Text Generation, Text2Text Generation, Sentence Similarity), and Audio (Text-to-Speech, Automatic Speech Recognition, Audio-to-Audio, Audio Classification, Voice Activity Detection). To the right of the categories, there's a list of models with their names, descriptions, and statistics like updates, downloads, and likes.

NEW Create Assistants in HuggingChat

The AI community building the future.

The platform where the machine learning community collaborates on models, datasets, and applications.

Tasks Libraries Datasets Languages Licenses Other

Models 469,541 Filter by name

Multimodal

- Text-to-Image
- Image-to-Text
- Text-to-Video
- Visual Question Answering
- Document Question Answering
- Graph Machine Learning

Computer Vision

- Depth Estimation
- Image Classification
- Object Detection
- Image Segmentation
- Image-to-Image
- Unconditional Image Generation
- Video Classification
- Zero-Shot Image Classification

Natural Language Processing

- Text Classification
- Token Classification
- Table Question Answering
- Question Answering
- Zero-Shot Classification
- Translation
- Summarization
- Conversational
- Text Generation
- Text2Text Generation
- Sentence Similarity

Audio

- Text-to-Speech
- Automatic Speech Recognition
- Audio-to-Audio
- Audio Classification
- Voice Activity Detection

Models 469,541

meta-llama/Llama-2-70b
Text Generation • Updated 4 days ago • 25.2k • 64

stabilityai/stable-diffusion-xl-base-0.9
Updated 6 days ago • 2.01k • 393

openchat/openchat
Text Generation • Updated 2 days ago • 1.3k • 136

llyasviel/ControlNet-v1-1
Updated Apr 26 • 1.87k

cerspense/zeroscope_v2_XL
Updated 3 days ago • 2.66k • 334

meta-llama/Llama-2-13b
Text Generation • Updated 4 days ago • 328 • 64

tiiuae/falcon-40b-instruct
Text Generation • Updated 27 days ago • 288k • 899

WizardLM/WizardCoder-15B-V1.0
Text Generation • Updated 3 days ago • 12.5k • 332

CompVis/stable-diffusion-v1-4
Text-to-Image • Updated about 17 hours ago • 448k • 5.72k

stabilityai/stable-diffusion-2-1
Text-to-Image • Updated 2 days ago • 1.3k • 136

[easyh/de_fnhd_nerdh](#)

like 0

[Token Classification](#) [spaCy](#) [German](#) [Eval Results](#)[Model card](#) [Files and versions](#) [Community](#)

Deutsche NER-Pipeline für frühneuhochdeutsche Texte (2.Version)

Feature	Description
Name	de_fnhd_nerdh
Version	0.0.2
spaCy	>=3.4.1,<3.5.0
Default Pipeline	tok2vec, ner
Components	tok2vec, ner
Vectors	500000 keys, 500000 unique vectors (300 dimensions)
Sources	n/a
License	n/a

[Edit model card](#)

Downloads last month

0

Hosted inference API ⓘ[Token Classification](#)

Your sentence here...

[Compute](#)

This model can be loaded on the Inference API on-demand.

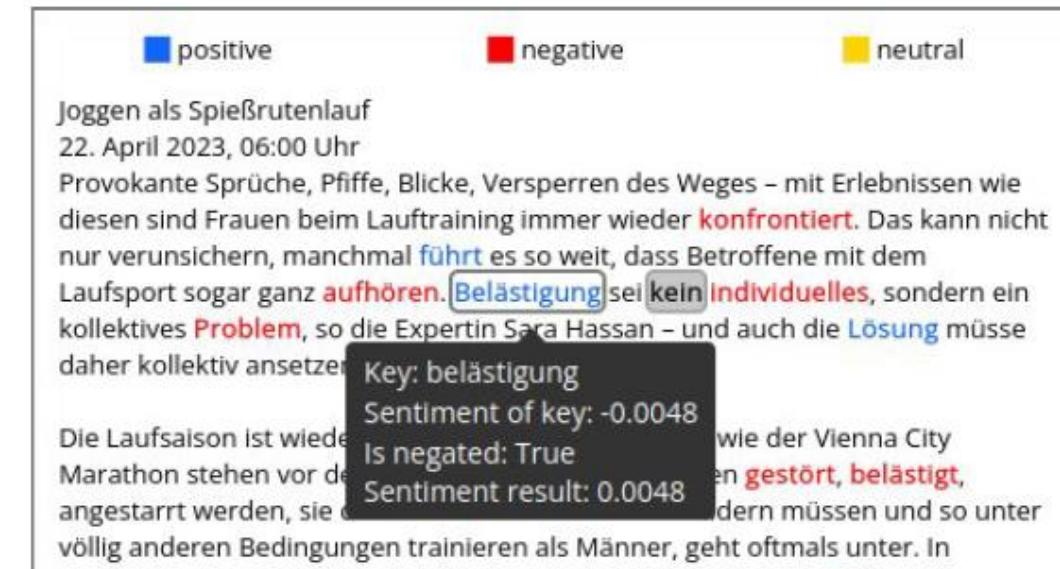
[JSON Output](#)[Maximize](#)[Space using easyh/de_fnhd_nerdh](#) 1[easyh/NerDH_Visualisierer](#)[Evaluation results](#) ⓘ

SentText

<https://thomasschmidtur.pythonanywhere.com/>

File: article.txt

Visualization of submitted data ⓘ		Evaluation and correction of submitted data ⓘ		
Document-level Document Score: -0.003142 (Normalized by number of words) (Absolute Score/Number of words) → document is negative Absolute: -3.5537 (decimal number) (The sum of all sentiment values in this document/folder) Number of words: 1131		<p>positive negative neutral</p> <p>Joggen als Spießrutenlauf 22. April 2023, 06:00 Uhr Provokante Sprüche, Pfiffe, Blicke, Versperren des Weges – mit Erlebnissen wie diesen sind Frauen beim Lauftraining immer wieder konfrontiert. Das kann nicht nur verunsichern, manchmal führt es so weit, dass Betroffene mit dem Laufsport sogar ganz aufhören. Belästigung sei kein individuelles, sondern ein kollektives Problem, so die Expertin Sara Hassan – und auch die Lösung müsse daher kollektiv ansetzen.</p> <p>Die Laufsaison ist wieder voll im Gange. Laufbewerbe wie der Vienna City Marathon stehen vor der Tür. Dass Frauen beim Laufen gestört, belästigt, angestarrt werden, sie deshalb ihre Gewohnheiten ändern müssen und so unter völlig anderen Bedingungen trainieren als Männer, geht oftmals unter. In sozialen Netzwerken finden sich reihenweise Berichte über negative Erlebnisse. Auch Patrizia, die seit Jahren regelmäßig in Wien laufen geht, erlebt immer wieder Pfiffe oder Rufe, wenn sie im Sommer mit kurzen Hosen trainiert. Im Gespräch mit ORF Topos (siehe Video) schildert sie bedrohliche Situationen, etwa wenn Autos langsam neben ihr herfahren, das Fenster runtergeht und sie einfach nur angestarrt wird. Anspruch auf öffentlichen Raum Belästigung im öffentlichen Raum hängt mit der Frage zusammen, wem der öffentliche Raum gehört, erläutert die Journalistin und Publizistin Sara Hassan, die zu Belästigung und Machtmisbrauch forscht. Hassan sieht darin einen Ausdruck von Anspruchsdenken und verweist in diesem Kontext auf eine Beobachtung aus der Soziologie: Männer neigen eher dazu, geradeaus zu gehen und entgegenkommenden Passantinnen und Passanten kaum oder gar nicht Platz zu machen. Frauen hingegen sind darauf konditioniert, anderen Personen im öffentlichen Raum auszuweichen. Dieses Phänomen zeigt gut, wie das Spektrum gestaltet ist, das diese Anspruchshaltung auf den öffentlichen Raum abbildet, so Hassan: „Hier wird klar, wer automatisch den Raum für sich beansprucht und wem die anderen Platz zu machen haben. Der öffentliche Raum gehört aber allen, und es sollten sich alle frei bewegen können. Das sehen allerdings nicht alle so. Belästigung wird dann zum Instrument, um Macht zu demonstrieren und diesen Anspruch geltend zu machen. „Der öffentliche Raum werde damit zum Aushandlungsort von Machtdynamik und der Frage von Zugänglichkeit, Fokus auf sich richten.“</p>		



Tutorials and further resources

- Tutorial for sentiment analysis:

<https://programminghistorian.org/en/lessons/sentiment-analysis>

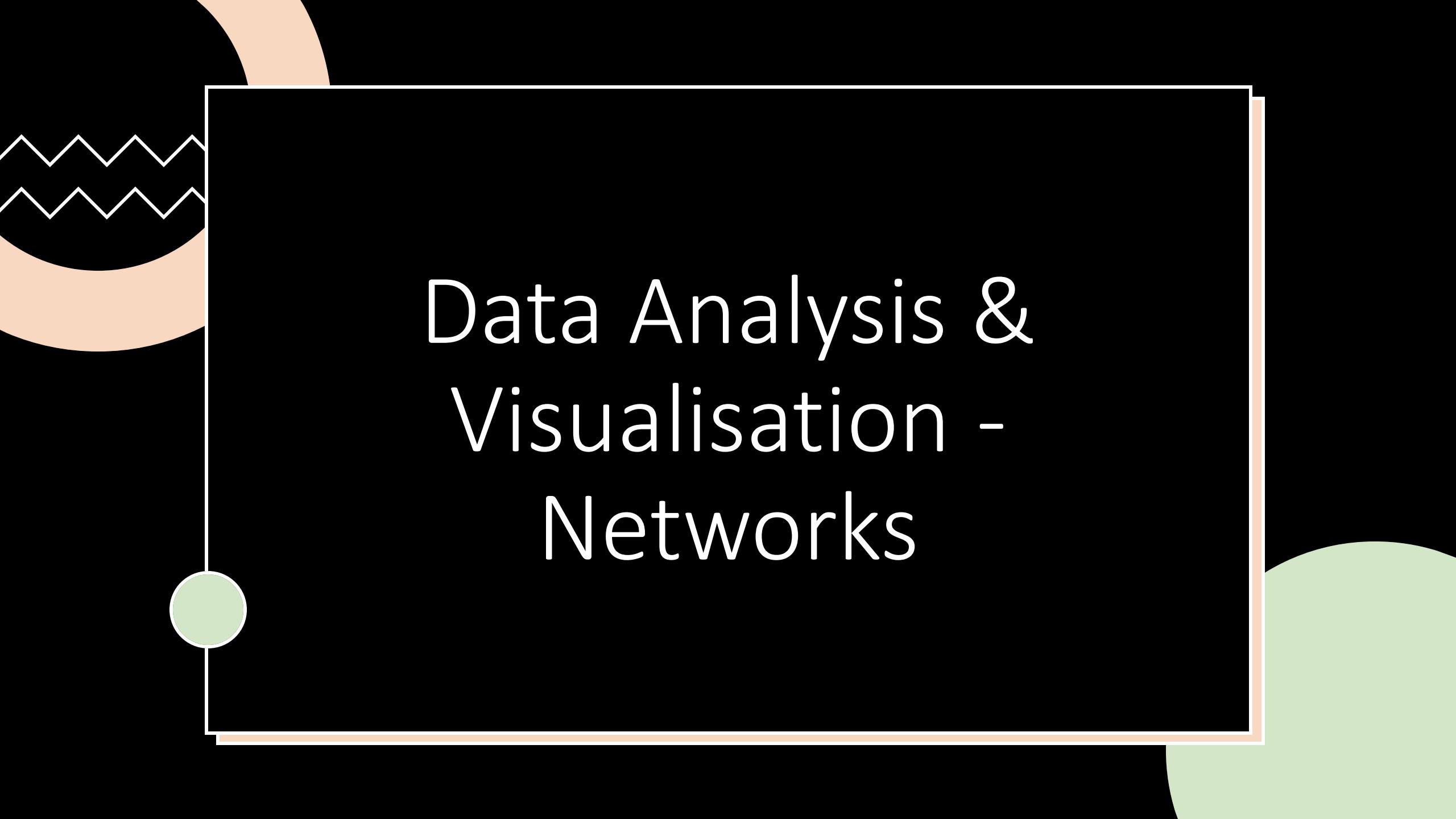
- Tutorial for spacy:

<https://programminghistorian.org/en/lessons/corpus-analysis-with-spacy>

- Tutorial for SentText: <https://fortext.net/tools/tools/senttext>

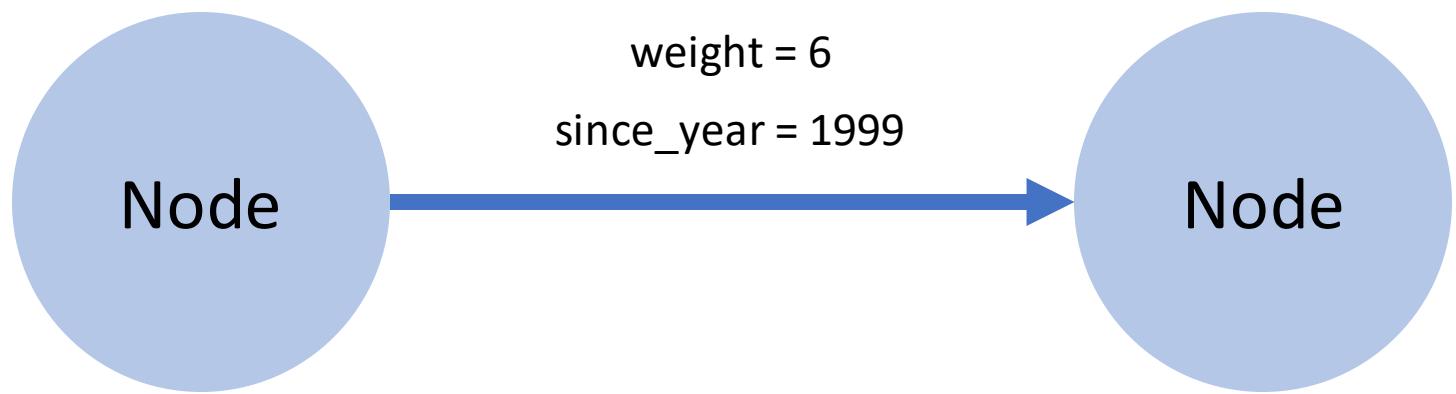
- Tutorial for NLTK: <https://www.nltk.org/book/ch01.html>

- Collection of tools for corpus analysis: <https://corpus-analysis.com/>



Data Analysis & Visualisation - Networks

Nodes and edges



Graph

Ezlinavis

<https://ezlinavis.dracor.org>

Easy Linavis

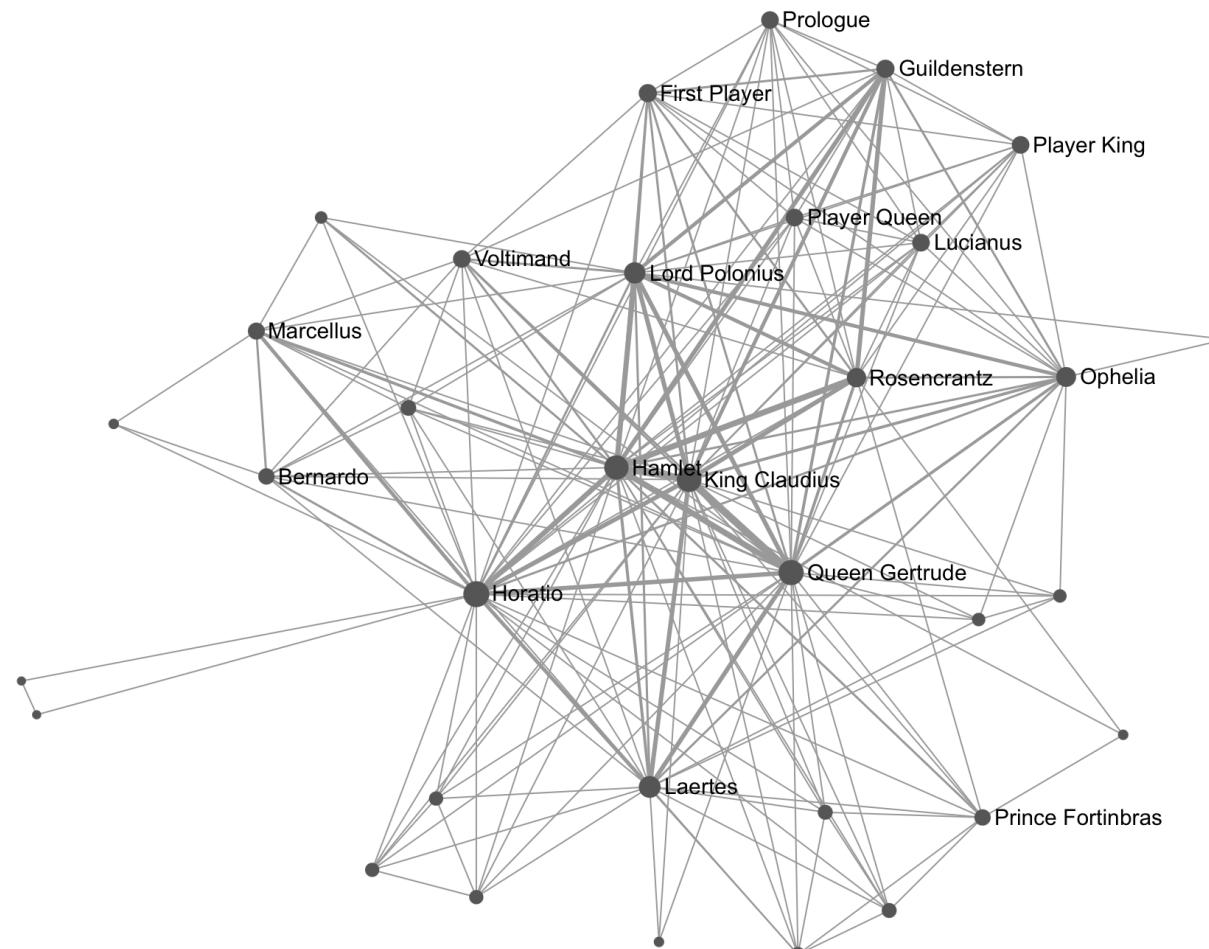
Examples ▾ Graph ▾ About

Shakespeare: The Tragedy of Hamlet,
Prince of Denmark (c. 1599/1602)
Characters/speakers per scene, in
order of appearance
Extracted from this version:
<http://shakespeare.mit.edu/hamlet/full.html>
By Frank Fischer, October 2017
Formalisation starts with first hashtag:

ACT I
SCENE I. Elsinore. A platform before the castle.
Bernardo
Francisco
Horatio
Marcellus
SCENE II. A room of state in the castle.
King Claudius
Cornelius
Voltimand
Laertes
Lord Polonius
Hamlet
Queen Gertrude
Horatio
Marcellus
Bernardo
SCENE III. A room in Polonius' house.
Laertes
Ophelia
Lord Polonius
SCENE IV. The platform.
Hamlet
Horatio
Marcellus
SCENE V. Another part of the platform.
Hamlet
Ghost
Marcellus
Horatio

[download CSV](#)

Source,Type,Target,Weight
Bernardo,Undirected,Cornelius,1
Bernardo,Undirected,Francisco,1
Bernardo,Undirected,Hamlet,1
Bernardo,Undirected,Horatio,2
Bernardo,Undirected,King Claudius,1
Bernardo,Undirected,Laertes,1
Bernardo,Undirected,Lord Polonio,1
Bernardo,Undirected,Marcellus,2
Bernardo,Undirected,Queen Gertrude,1
Bernardo,Undirected,Voltimand,1
Captain,Undirected,Hamlet,1
Captain,Undirected,Prince Fortinbras,1
Captain,Undirected,Rosencrantz,1
Cornelius,Undirected,Hamlet,1
Cornelius,Undirected,Horatio,1
Cornelius,Undirected,King Claudius,1
Cornelius,Undirected,Laertes,1
Cornelius,Undirected,Lord Polonio,1
Cornelius,Undirected,Marcellus,1
Cornelius,Undirected,Queen Gertrude,1
Cornelius,Undirected,Voltimand,1
Danes,Undirected,Gentleman,1
Danes,Undirected,Horatio,1
Danes,Undirected,King Claudius,1
Danes,Undirected,Laertes,1
Danes,Undirected,Ophelia,1
Danes,Undirected,Queen Gertrude,1
First Ambassador,Undirected,Hamlet,1
First Ambassador,Undirected,Horatio,1
First Ambassador,Undirected,King Claudius,1
First Ambassador,Undirected,Laertes,1
First Ambassador,Undirected,Lord Polonio,1

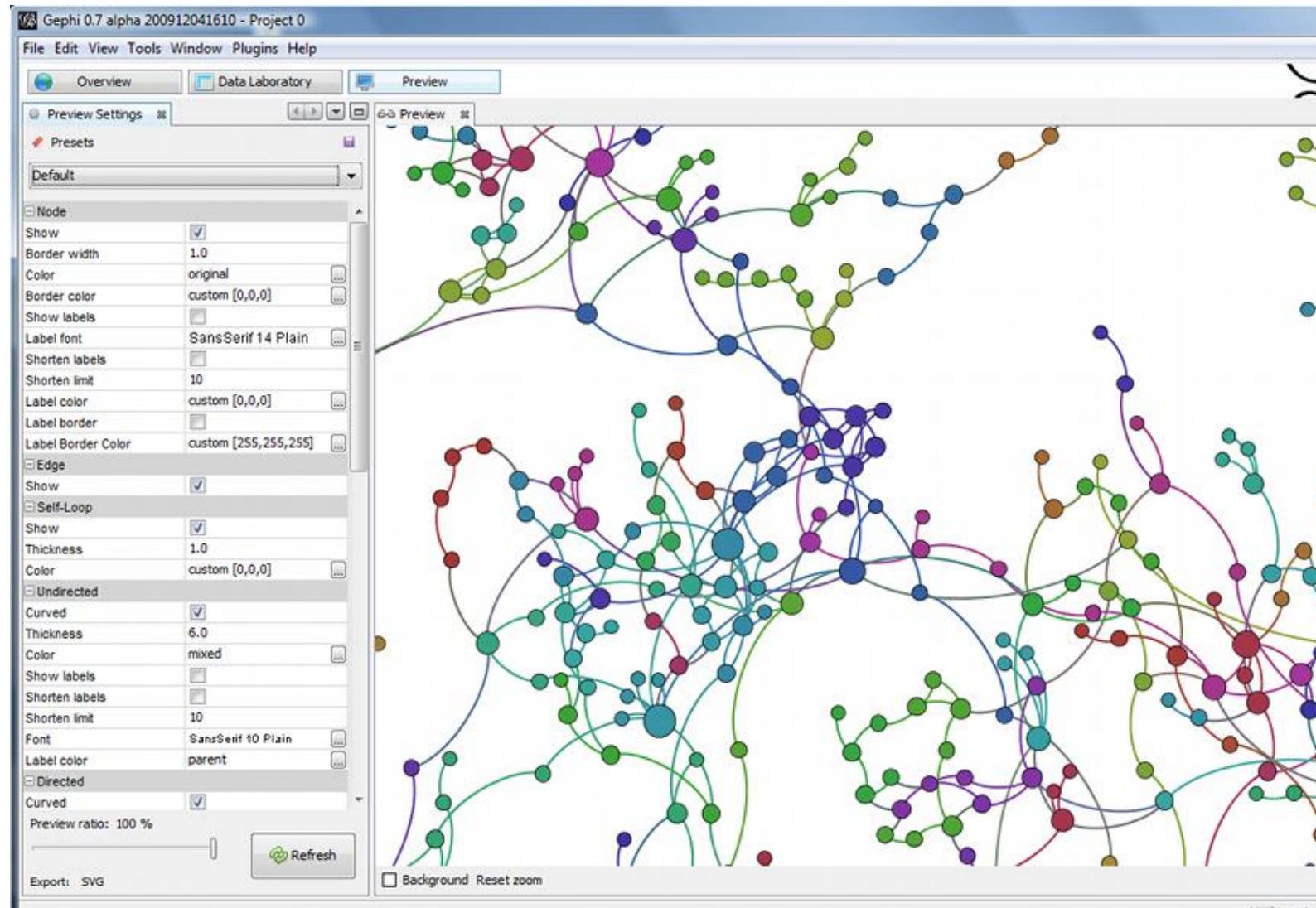


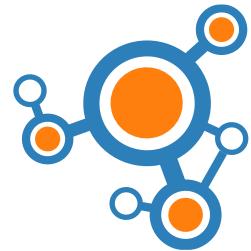
Gephi

Textual and spatial encoding

- ❖ LOD
- ❖ Relational data/graphs
- ❖ Various downloadable formats

Online version: <https://gephi.org/gephi-lite/>





NetworkX

Network Analysis in Python

```
[8] nx.degree(G)
```

```
... DegreeView({'Aemon': 5, 'Grenn': 4, 'Samwell': 15, 'Aerys': 4, 'Jaime': 24, 'Robert'
```

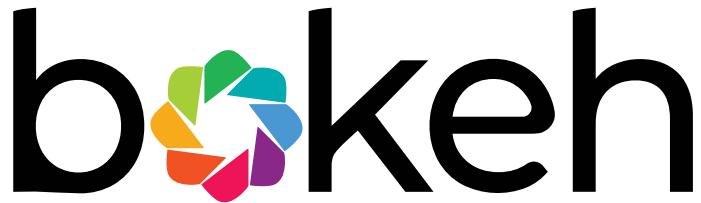
```
[9] degrees = dict(nx.degree(G))
nx.set_node_attributes(G, name='degree', values=degrees)
```

```
[13] degree_df = pd.DataFrame(G.nodes(data='degree'), columns=['node', 'degree'])
degree_df = degree_df.sort_values(by='degree', ascending=False)
degree_df.head()
```

	node	degree
6	Tyrion	36
21	Jon	26
50	Sansa	26
48	Robb	25
4	Jaime	24

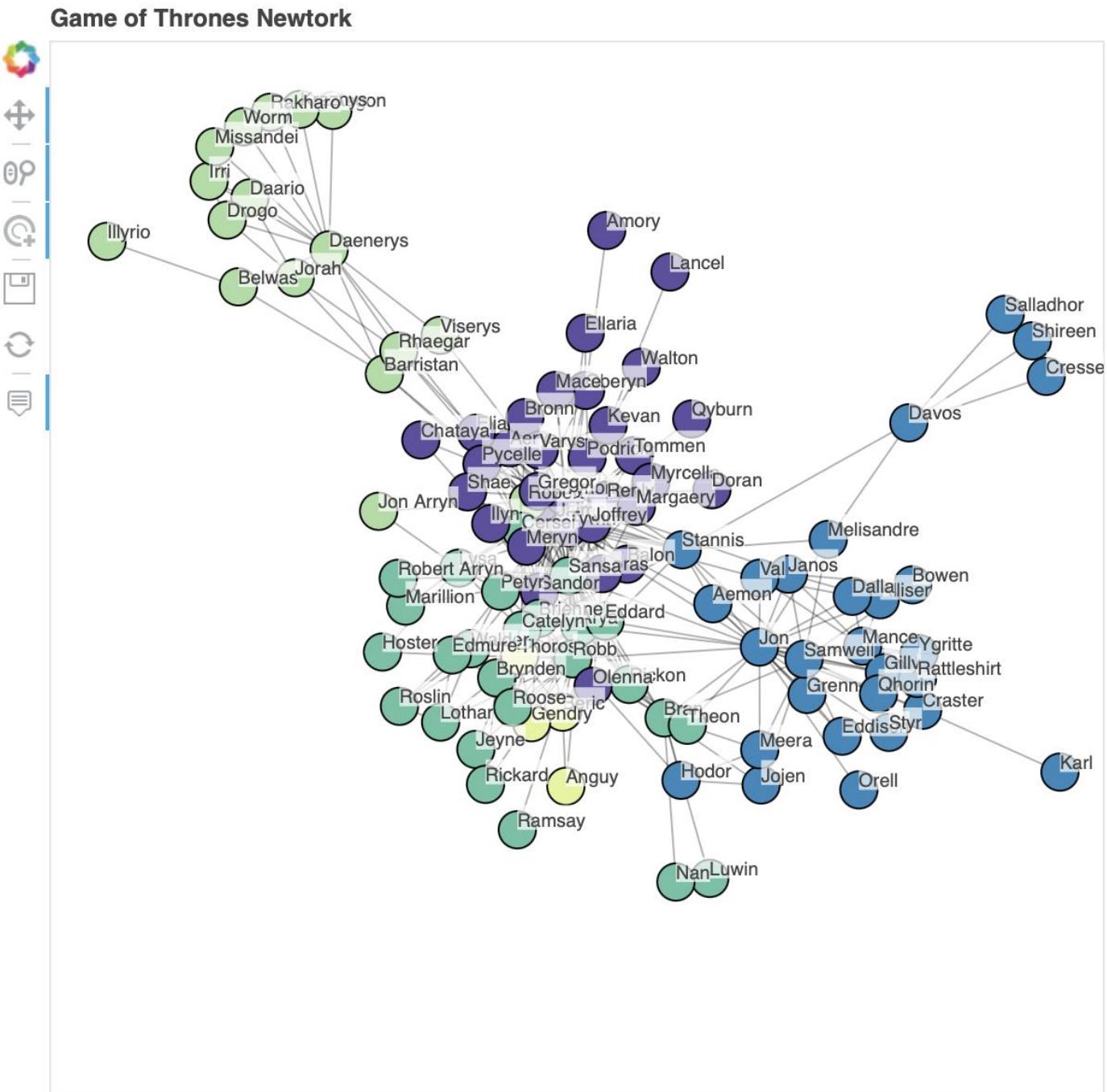
Documentation

- Create **graphs** from scratch or from pre-existing data
- Run **analyses** on them, e.g.:
 - How many connections?
 - How strong are the connections?
- Basic **visualisation** methods



<http://bokeh.org>

- Python library
 - Create **interactive** visualisations
 - Use these visualisations in **web pages**
 - See an example [here](#) (based on *Game of Thrones*)

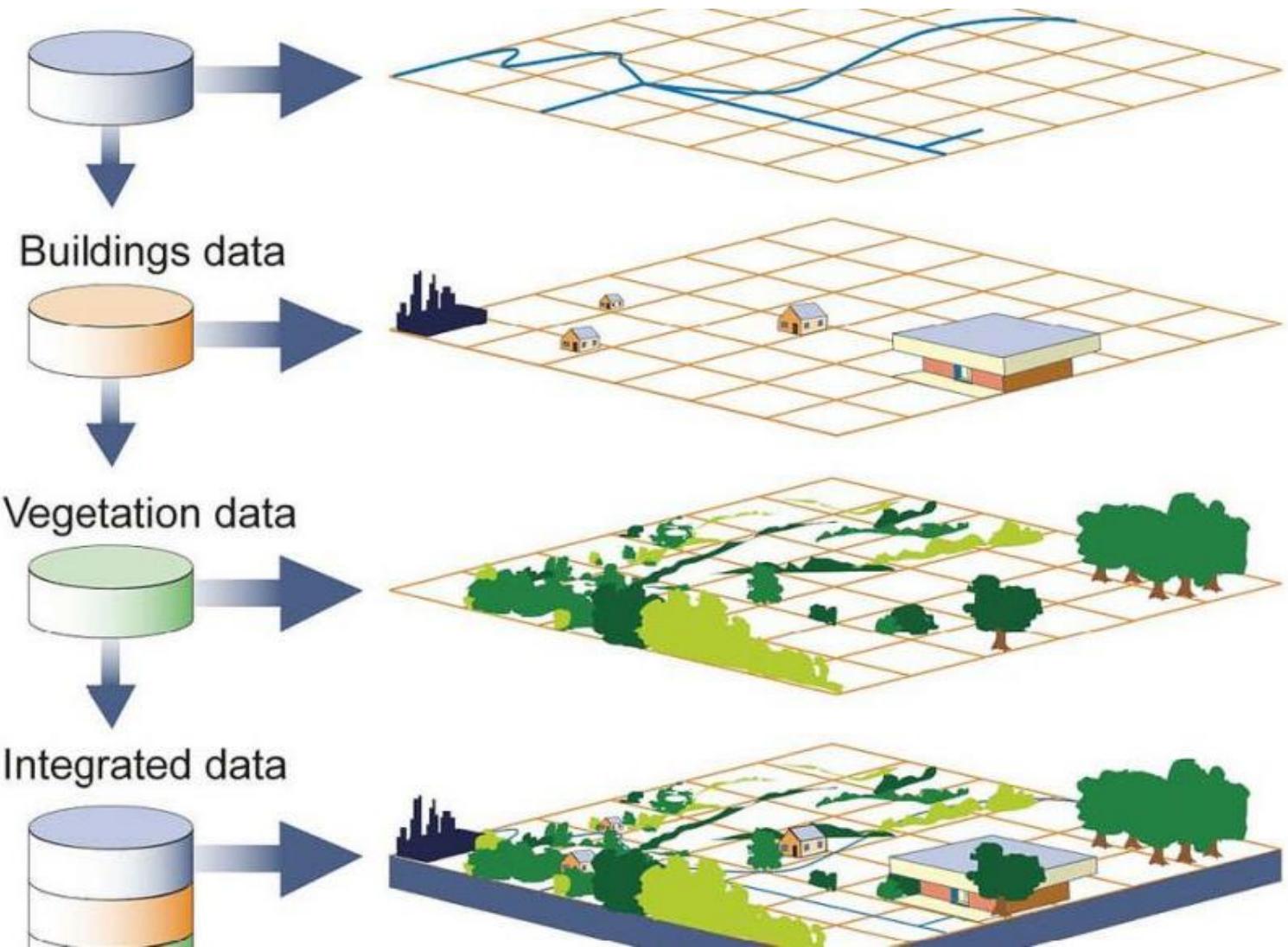


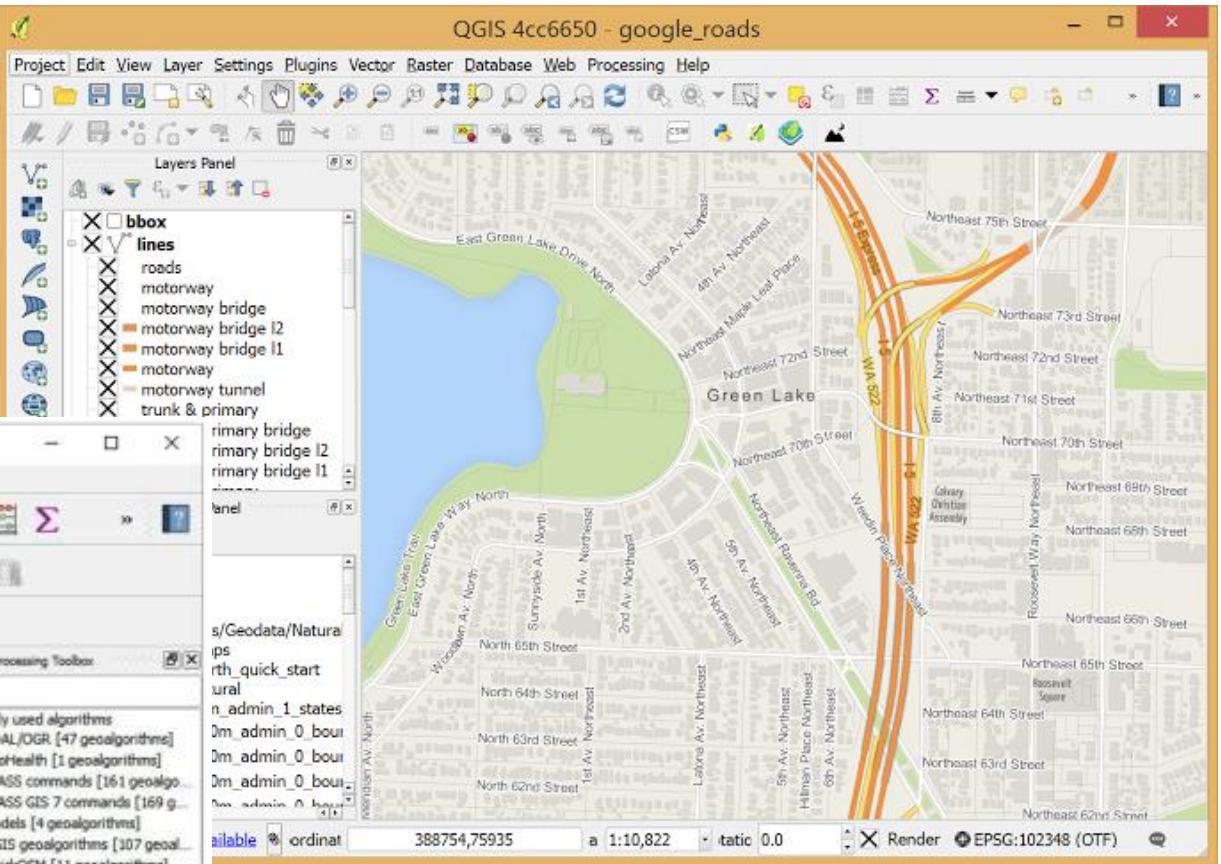
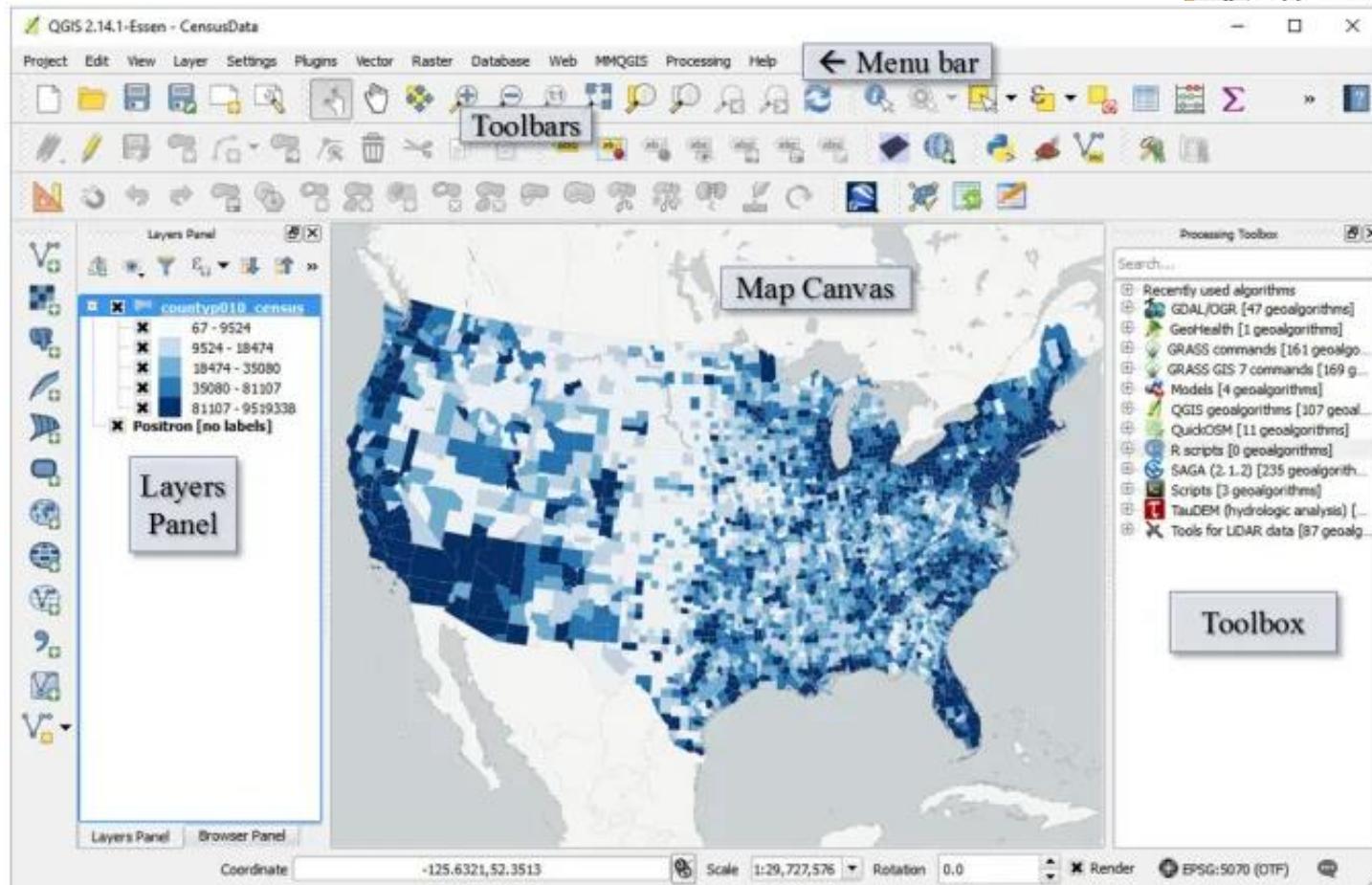


GIS

GIS = Geographic Information System(s)

- computer-based tools used to store, visualize, analyze, and interpret geographic data





<https://qgis.org/de/site/>

Open-Source GIS software
(spatial analysis, map
visualisations, ...)

Exemplary Use Case

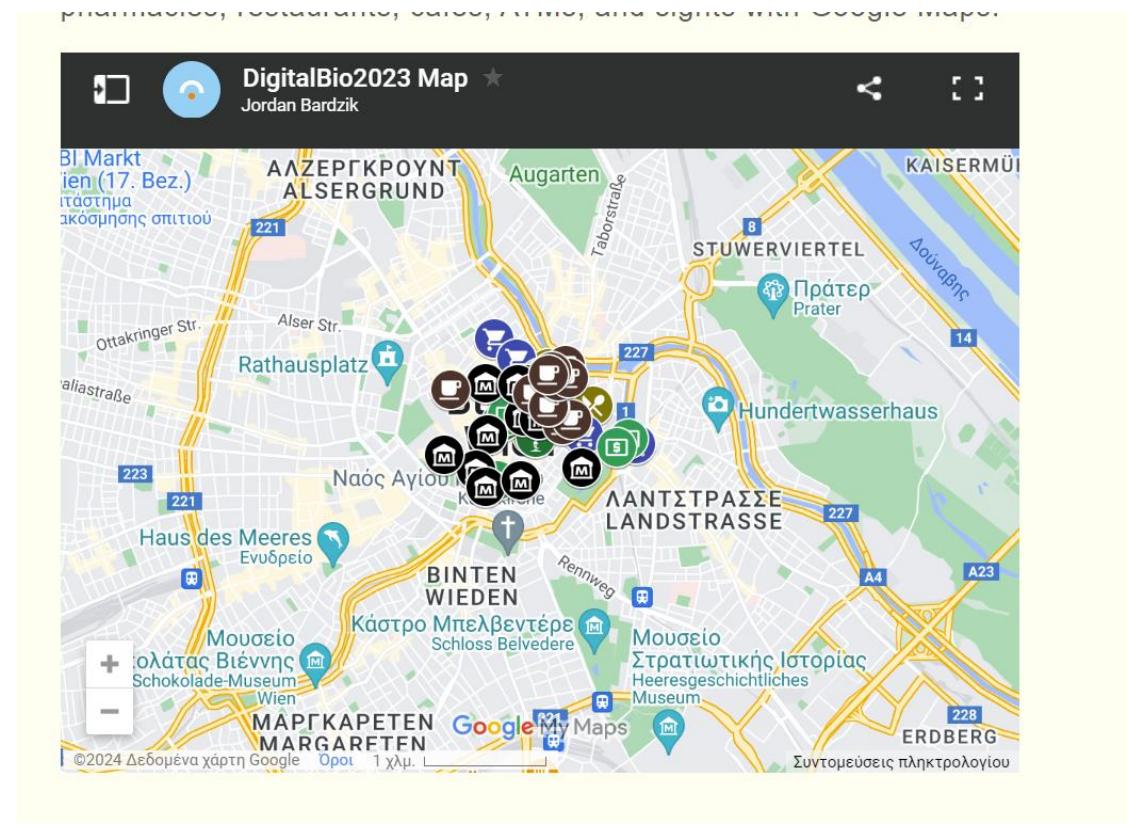


- ACDH-CH Project "Visiting Vienna"
- Mapping of arriving persons in early 18th century Vienna based on arrival list data
- Use of georeferenced historic plans of Vienna



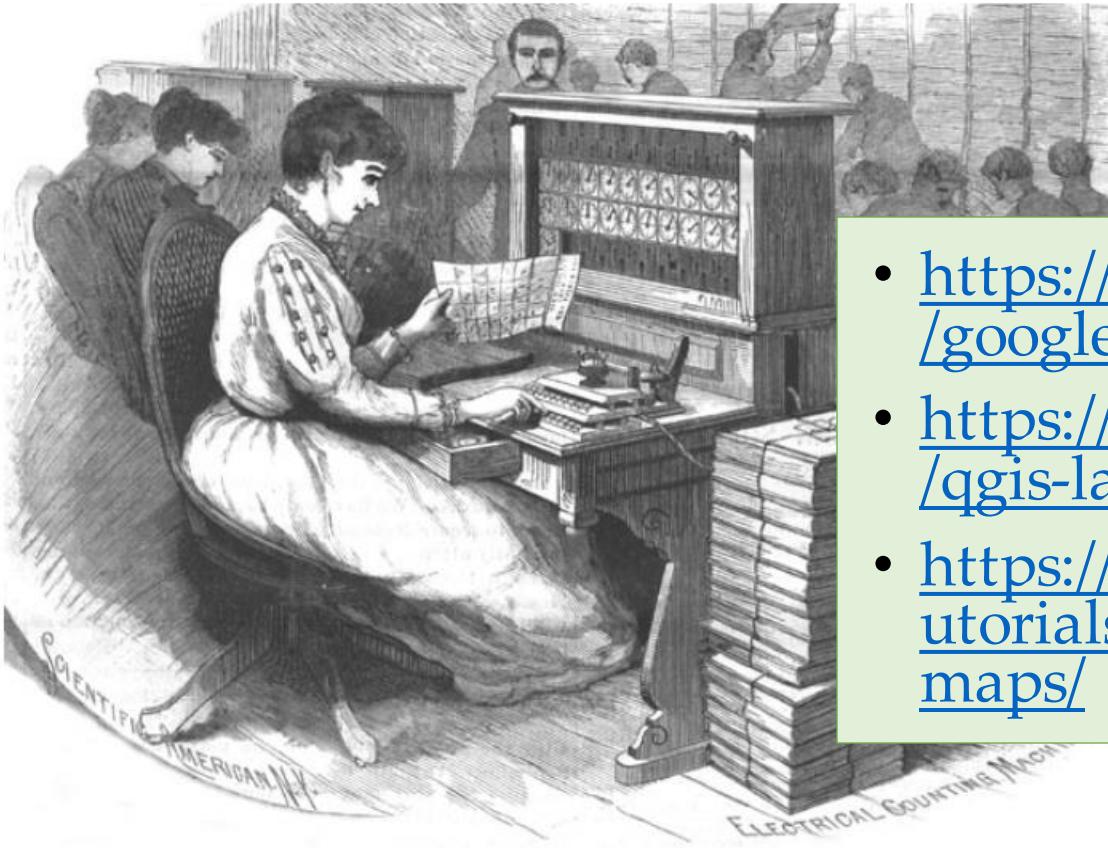


Google Maps



https://www.google.com/maps/d/edit?mid=1o1p_2DPFs62j6OXx00ShFAvqg1JCIQw&usp=sharing

Tutorials



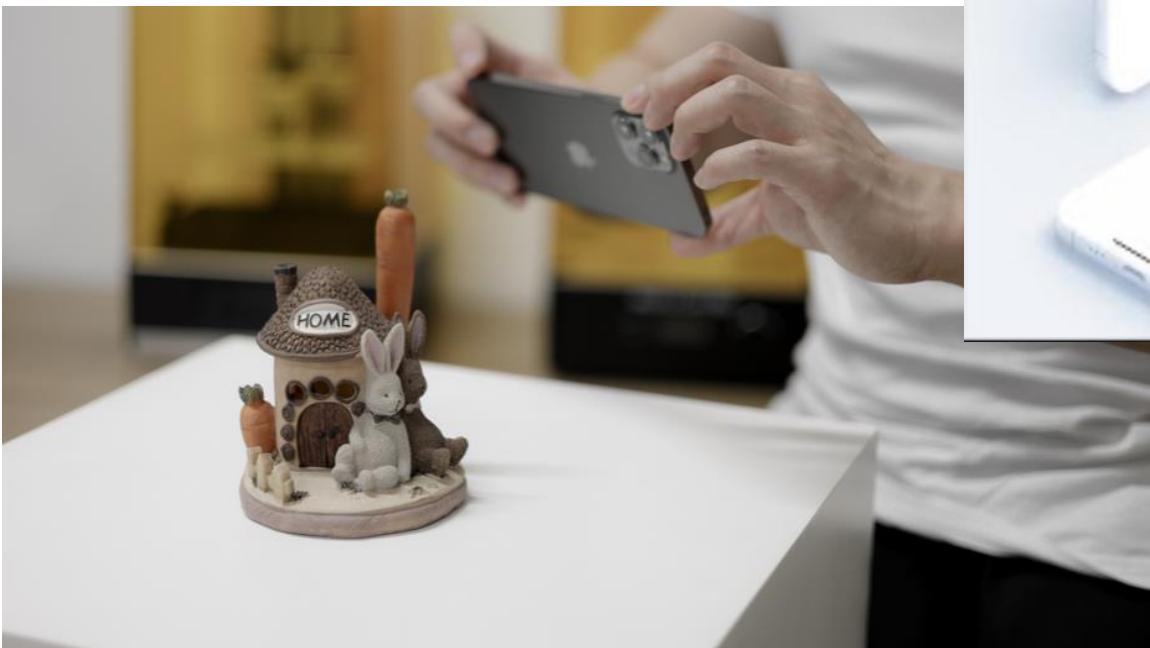
- <https://programminghistorian.org/en/lessons/googlemaps-googleearth>
- <https://programminghistorian.org/en/lessons/qgis-layers>
- <https://multimedia.journalism.berkeley.edu/tutorials/create-custom-maps-in-google-maps/>

3D Visualisation



KIRI Engine

<https://www.kiriengine.app/>



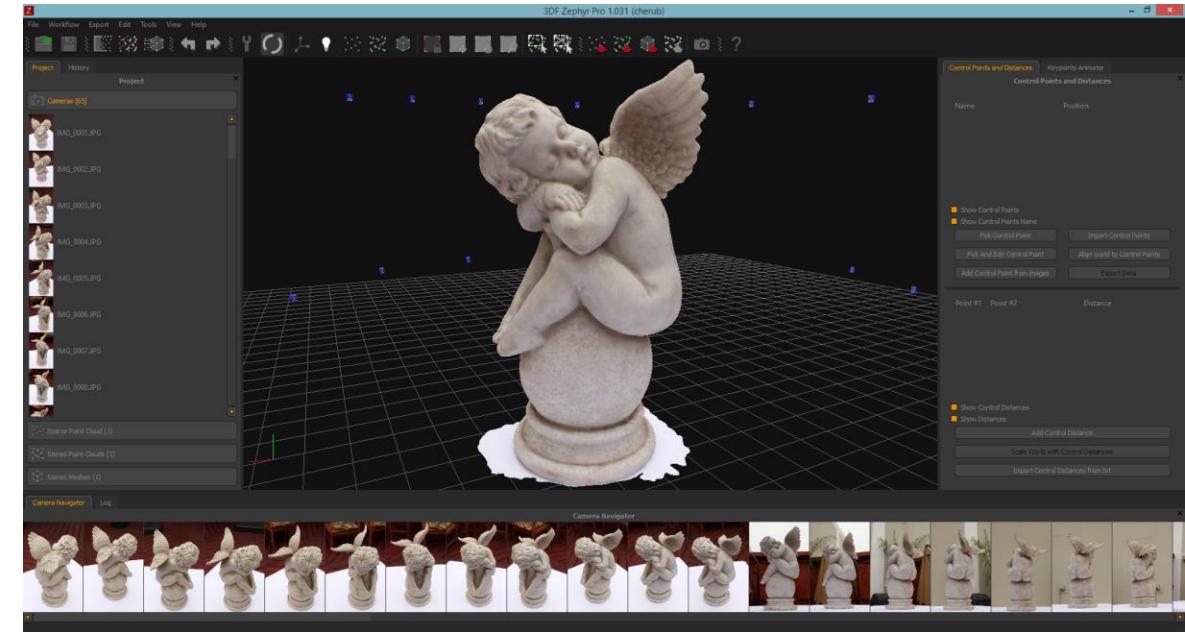
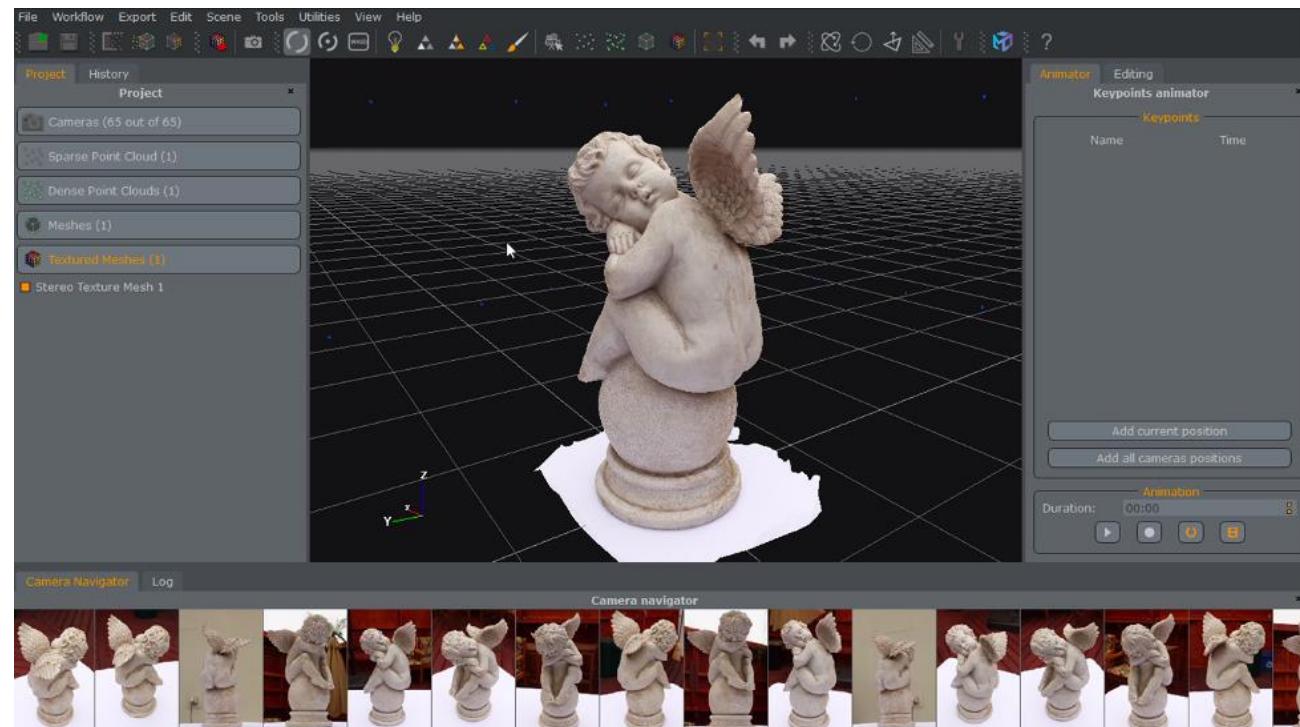
An advertisement for the KIRI Engine app. It features three white smartphones arranged in a triangular formation. The top phone shows a login screen with a teal button. The bottom-left phone shows a 3D wireframe model of a red sports car from a top-down perspective. The bottom-right phone also shows the same 3D model. To the right of the phones, the text "TURN STILL IMAGES INTO 3D MODELS" is written in large, bold, blue capital letters. Below this text are two download links: "Download on the App Store" with the Apple logo, and "GET IT ON Google Play" with the Google Play logo. At the very bottom right, the website address "www.kiriengine.com" is displayed.



3DF ZEPHYR

The Complete Photogrammetry Solution

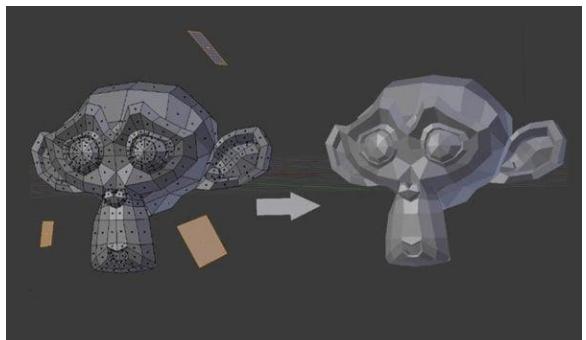
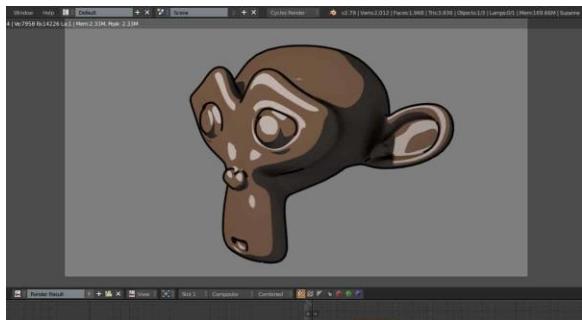
<https://www.3dflow.net/3df-zephyr-photogrammetry-software/>



Free Photogrammetry tool (3D representations)



<https://www.blender.org/>



Free and Open source 3D creation suite.
It supports the entirety of the 3D pipeline (e.g. modeling, rigging, animation, simulation, rendering, compositing and motion tracking, even video editing and game creation.)

Semantic Web



WIKIPEDIA
The Free Encyclopedia

<https://en.wikipedia.org/wiki/Vienna>



Vienna

Article

Talk

Read

[View source](#)

[View history](#)

Tools

From Wikipedia, the free encyclopedia

Coordinates: 48°12'30"N 16°22'21"E

This article is about the capital of Austria. For other uses, see Vienna (disambiguation).

"Wien" redirects here. For other uses, see Wien (disambiguation).

Not to be confused with Vienne (disambiguation).

This article contains **weasel words**: vague phrasing that often accompanies **biased or unverifiable information**. Such statements should be clarified or removed. (September 2023)

Vienna (viːnə/ ⓘ) ⓘ vee-EN-ə-[9][10] German: **Wien** [viːn]

ⓘ ⓘ; Austro-Bavarian: **Wean** [veːn] is the capital, largest city, and one of nine federal states of Austria. Vienna is Austria's most populous city and its primate city, with about two million inhabitants^{[3][11]} (2.9 million within the metropolitan area,^[12] nearly one-third of the country's population), and its cultural, economic, and political center. It is the fifth-largest city proper by population in the European Union and the largest of all cities on the Danube river by population.

The city lies on the eastern edge of the Vienna Woods (Wienerwald)—the northeasternmost foothills of the Alps, that separate Vienna from the more western parts of Austria—at



<https://www.wikidata.org/wiki/Q1741>



Vienna

capital of and state in Austria

Wien | Vienna, Austria | W

▼ In more languages

Configure

Language	Label	Description	Also known as
English	Vienna	capital of and state in Austria	Wien Vienna, Austria W
German	Wien	Bundeshauptstadt der Republik Österreich	AT-9 AT VIE W
Bavarian	Wean	Hauptstod vo Österreich	
French	Vienne	capitale de l'Autriche	Wien

All entered languages

Statements

instance of

federal capital

▼ 0 references

statutory city of Austria

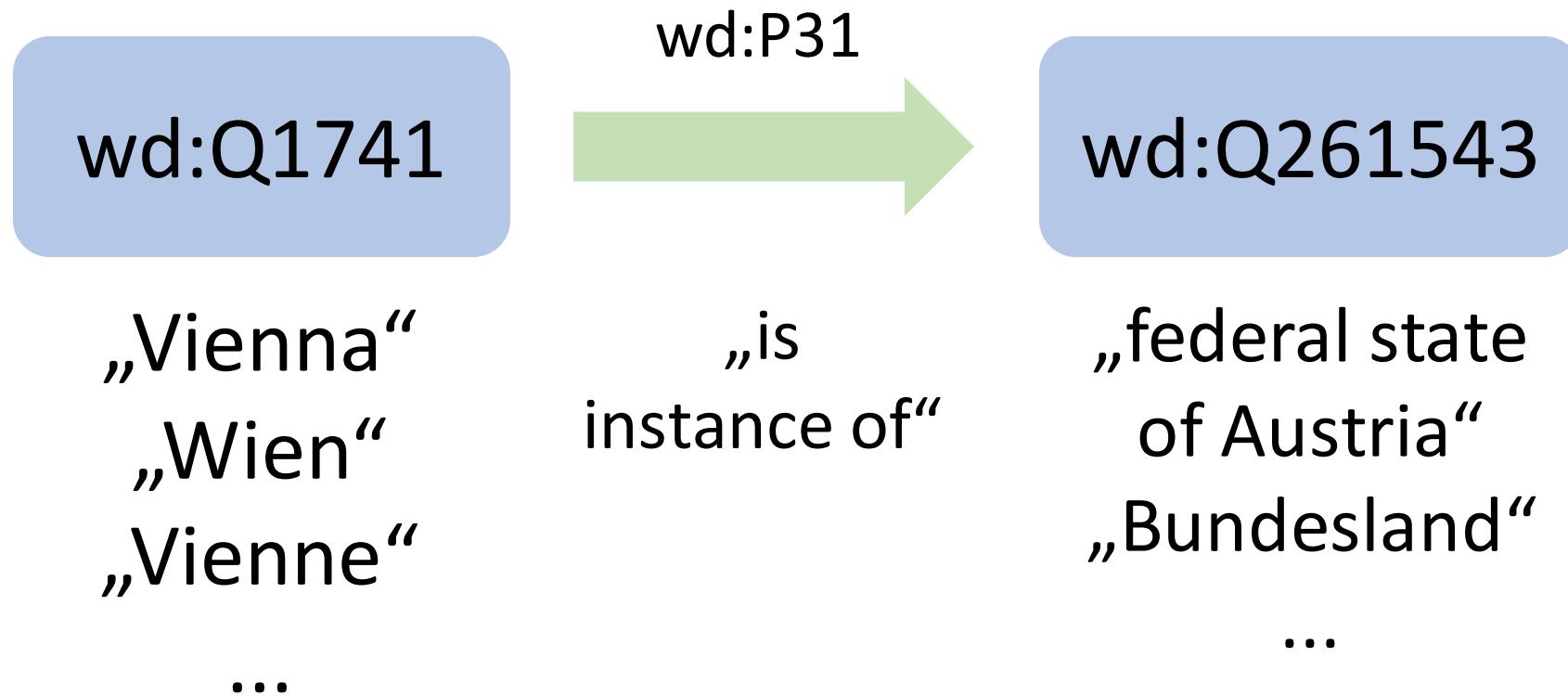
▼ 0 references

federal state of Austria

start time 18 November 1920 Gregorian
an

How does the Semantic Web work?

Vienna ([/viˈenə/](#) ⓘ vee-EN-ə;^{[9][10]} German: *Wien* [viːn] ⓘ; Austro-Bavarian: *Wean* [veːn]) is the capital, largest city, and one of nine federal states of Austria. Vienna is Austria's most populous city and its primate city,



Google Knowledge Panel

Google harrison ford X 🔍

All Images News Videos Books More Tools SafeSearch ▾

About 81.500.000 results (0,40 seconds)

Harrison Ford

American actor

Overview Movies Relationships Videos

Ethan Miller/FilmMagic

Age 81 years 13 Jul 1942

Spouse Calista Flockhart (m. 2010), Melissa Mathison (m....)

YouTube CO INDIANA JONES

7:27 10 Jul 2023

SR Screen Rant

Jake Gyllenhaal's Upcoming Remake Of \$221 Million ...

... Harrison Ford film looks like it could be even better. Jake Gyllenhaal in Road...

3 days ago

Global and Unified Access to Knowledge Graphs

Announcement, Google Summer of Code

GSoC 2024 – Call for Contributors

March 1, 2024 by Julia Holze

Are you a student looking for a summer experience that combines coding skills with open source development? Then look no further than the Google Summer of Code program 2024, where you can join



[Faceted browser](#)

[Query interface](#)
(advanced)

<https://dbpedia.org/page/Vienna>

Controlled vocabularies



Vocabularies About Feedback Help | Interface language: English ▾

UNESCO Thesaurus

Content language English ▾ Search

Alphabetical	Hierarchy	Groups
A Å	B C	D E F G H I J K
L M N	O P Q R S T U V W	X Y Z

Abadi → Awadhi
Abandoned children
Abbreviations
Ability
Ability grouping
Abohi → Awadhi
Aborigines → Indigenous peoples
Abortion
Absenteeism → Leave
Abstract journals → Abstracts
Abstract reasoning → Reasoning
Abstracting
Abstracting and indexing services → Bibliographic services
Abstracts
Abuse of human rights → Human rights violations
Abuse of power → Oppression
Academic achievement
Academic admission → Admission requirements
Academic buildings
Academic degrees → Degrees
Academic equipment → Educational equipment
Academic facilities → Educational facilities
Academic failure
Academic fraud
Academic freedom
Academic grouping → Educational grouping
Academic laboratories → University laboratories
Academic libraries

Vocabulary information

TITLE	UNESCO Thesaurus
DESCRIPTION	The UNESCO Thesaurus is a controlled and structured list of terms used in subject analysis and retrieval of documents and publications in the fields of education, culture, natural sciences, social and human sciences, communication and information. Continuously enriched and updated, its multidisciplinary terminology reflects the evolution of UNESCO's programmes and activities.
IDENTIFIER	http://vocabularies.unesco.org/thesaurus
PUBLISHER	UNESCO
RIGHTS HOLDER	UNESCO
RIGHTS	CC-BY-SA
LICENSE	http://creativecommons.org/licenses/by-sa/3.0/igo/
CREATED	Saturday, January 1, 1977 00:00:00
LAST MODIFIED	Tuesday, January 16, 2024 16:39:28
SUBJECT	http://id.loc.gov/authorities/subjects/sh85029027 http://id.loc.gov/authorities/subjects/sh85062913 Social sciences http://id.loc.gov/authorities/subjects/sh85034755 http://id.loc.gov/authorities/subjects/sh85040989 http://id.loc.gov/authorities/subjects/sh85118553

- Define **terms** that should be used in a domain / for a task
- Allow all applications (or users) to understand what we are talking about **unequivocally!**
- Usually based on the **SKOS** model
- Some examples include the **UNESCO Thesaurus** (on the left)

Some ways to go...

- Search for any Wikipedia page that might interest you, and click on „Tools > Wikidata item“... you can compare the structured data to the original page or browse through the entities by clicking on the links
- (More advanced) You can load a sample Neo4j graph database by using the [Neo4j Sandbox](#), for example a database on „Movies“
- Have a look at the ACDH-CH services, [ARCHE](#) and [Vocabs](#), which provide data according to Semantic Web recommendations



Further resources to educate yourself



Programming Historian

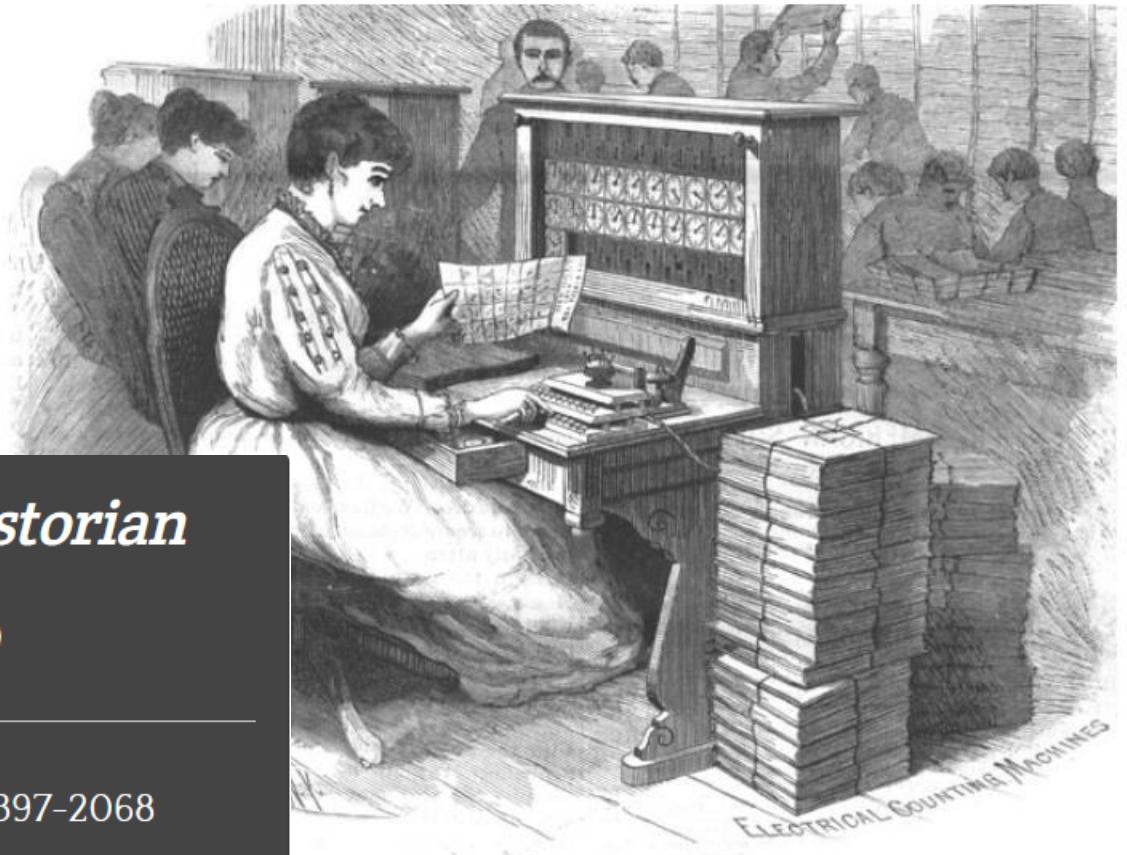
<https://programminghistorian.org/>

[PH]

Enter *Programming Historian*
(The initial English version)

101 lessons ISSN: 2397-2068

We publish novice-friendly, peer-reviewed tutorials that help humanists learn a wide range of digital tools, techniques, and workflows to facilitate research and teaching.



<https://campus.dariah.eu/>



Looking for learning resources?

DARIAH-CAMPUS is a discovery framework and hosting platform for learning resources.

ACDH-CH-HowTo

<https://howto.acdh.oeaw.ac.at/>

**Share and expand
your knowledge
in Digital
Humanities**

Welcome to ACDH-CH Learning Resources. This website gathers interactive learning material, practical HowTo articles and best practices on a wide range of Digital Humanities topics, methodologies and infrastructures. These resources can be used for self-guided learning as well as for teaching in higher education. It is part of our mission to actively transfer knowledge from ongoing research into the wider Digital Humanities Community as well as to educate a new generation of humanities researchers with digital methods.

Posts

[Corpus Query Language im Austrian Media Corpus](#)

Diese Resource stellt den Austria Media Corpus (amc) und seine Nutzungsmöglichkeiten vor. Im Besonderen wird die Corpus Query Language (CQL) erklärt.

[Read more →](#)

[Lesson 2: Meet the data. Examples of data perspectivity, with some thoughts on shopping for groceries.](#)

This is part of a blog post series documenting the SemanticKraus research project at the ACDH-CH.

[Read more →](#)

[Lesson 1: Make a plan, then don't stick to it](#)

Fine-tuning a workflow and a question millions of years old. This is part of a blog post series documenting the SemanticKraus research project at the ACDH-CH.

[Read more →](#)

[How to do a HowTo - A short manual for the ACDH-CH Learning Resources Platform](#)

This article is a short introduction to the HowTo-Platform of the ACDH-CH and covers how to plan and structure content, write an article and publish it via the Content Management System.

[Read more →](#)



ACDH-CH TOOL GALLERIES

**ACDH-CH TOOL
GALLERY 9.2**
TRANSKRIBUS FÜR ANFÄNGER:INNEN
14. JUNI 2023 (09:00-12:30)

**ACDH-CH
TOOL GALLERY 9.1**
XSLT FÜR DIGITAL HUMANISTS
12 APRIL 2023, (10.30-15.30)

[https://www.oeaw.ac.at/acdh/events/
event-series/acdh-ch-tool-galleries](https://www.oeaw.ac.at/acdh/events/event-series/acdh-ch-tool-galleries)

The ACDH-CH Tool Galleries, which take place **three times a year**, are workshops dedicated to providing information and **hands-on training in the use of specific digital humanities tools**. Participants are given the opportunity to engage with the tools and to learn from experts with first hand experience in the application of digital methods in the field of humanities research. They are a service **open to everyone** but directed, in particular, at humanities scholars who would like to learn about and engage with computer-based methods in order to implement them in their own research.

Summer Schools

- European Summer University in Digital Humanities „Culture & Technology“
 - <https://digihubb.centre.ubbcluj.ro/>
 - Edirom Summer School
 - <https://ess.uni-paderborn.de/>
 - Helsinki Digital Humanities Hackathon (DHH)
 - <https://www.helsinki.fi/en/digital-humanities/helsinki-digital-humanities-hackathon-2023-dhh23>
 - Digital Humanities @Oxford Summer School (DHOX)
 - <https://digital.humanities.ox.ac.uk/digital-humanities-oxford-summer-school>
- (application for) financial support possible in most cases

ACDH-CH INTERNSHIPS

SUMMER 2024 (JULY & AUGUST) AT THE ÖAW CAMPUS ACADEMY IN VIENNA



© Klaus Pichler / ÖAW

Are you a student looking for hands-on work experiences at an academic research institute?

Are you thinking of an academic career in the humanities?

Would you like to know how to apply digital methods and tools to make existing research more insightful and ask new research questions?

Then the internship at the **Austrian Centre for Digital Humanities and Cultural Heritage** is for you!

WHO WE ARE

The ACDH-CH is a leading research institution in digital humanities (DH) and part of the Austrian Academy of Sciences (ÖAW). We work closely with international research networks and support local research interests.

WHAT WE OFFER

During your time with us we want you to:

- gain hands-on work experience
- get insights into the workings of an academic research institute
- find out more about the focus on DH scholarship
- understand what a career as a DH researcher may entail

We offer internships of 80 hours (20 hrs a week for 4 weeks), 120 hours (30 hrs a week for 4 weeks), and 160 hours (20hrs for 8 weeks).

In addition to getting to know a project, we offer our **Meet the Researchers** sessions: researchers from different areas of expertise offer additional **workshop sessions** to expand your DH knowledge.

Also, depending on your university, you can be awarded **ECTS points** that count towards your degree.

WHO ARE WE LOOKING FOR

In general: students from DH and the fields of research that we work on!

In particular: applicants with backgrounds in **philology** (particularly **German philology**), **linguistics**, **literary studies**, **musicology**, or **other paradigms of philological, cultural or sociological paradigms**. Some familiarity with

ACDH-CH Internships in July and August 2024

Apply now!

Deadline:

15 May 2024

<https://www.oeaw.ac.at/acdh/education/acdh-ch-internships>

Wap-pap, pa-da-di-da-da

Gleich treffen sich Bänke im Einkino, besser bekannt als Bibi H., was ich normalerweise nicht mehr merke. Ich kann mir nicht vorstellen, dass sie nicht den Fehler gemacht, meinen Töchtern davon zu erzählen. Ich gebe zu, ich wollte etwas angeben, als ich beim Abendessen sagte: „Übermorgen gehen wir auf die Fahrt.“ Mit dem Kindern beginnen mich die Angewandtschulmeile und investigativen Erwähnungen von zwei halben Berliner Schulklassen. Leistungsdruck, verdammt.

Aberseits der Aufmerksamkeit der Erwachsenen prägt Bibi H. seit Jahren die Jugendkultur. Jetzt sorgt sie mit einem Song für Aufsehen. Ein Besuch

KUNSTWERK
Aufblasbare Ballerina
von Svenja Schulte

Der Meister des Kitsches, Jeff Koons, hat in New York sein neues Kunstwerk vorgestellt: eine aufblasbare Hölleläuferin. Bis zum 2. Juni dauerst du 14 Meter hohe Tänzerin in Platz vor dem Rockefeller Center auf dem der 62-jährige Koons das zwei Mal zuvor Werke präsentiert hat. Die Installation soll auch der vermissten Kinderwagen, der im Mai

Mittelägypten

200

Absurdes
Aufmerksamkeit der
Erwachsenen prägt
Bibi H. seit Jahren
die Jugendkultur.
Jetzt sorgt sie mit
einem Song für
Aufsehen. Ein Besuc

Arten, wie man sich an den niedlichen Nachbarn rammischen kann?

Aber dann fällt mir gerade noch rechtzeitig ein, dass ich mich früher meinen Nachbarn sehr viele Fragen in mein Notizbuch geschrieben habe, und die werden ich Ihnen jetzt stellen, obwohl es eigentlich keine einzige Frage auf dieser Welt geben dürfte, die Ihnen nicht gewandt beantwortet hat. Sie bekommen sofort jeden Worten, das Ihnen in den Kopf kommt, gerund um Ihr Leben, und zwar ohne dass jemand fragt, was Ihnen Erfolg versprochen hat.

Ein paar Daten: Bibi ist 24 Jahre alt, ihr Kanal gibt es seit 2012, er hat über 1,5 Millionen Abonnenten und über 1,85 Millionen Aufrufe. Sie ist die erfolgreichste deutsche Frau auf YouTube. Und sie ist jenseitig Feministin: Sie ist ein Lehrer, der sich für die Rechte von Frauen einsetzt und bei gesungen hat, hinter ihr steht eine sehr junge Frau aus Köln im Paradies, die universität studiert und Karriere macht. Ihr Song wurde in der Woche des 1. Mai auf der Bühne der Deutschen Meisterschaft der Beatboxer auf der Universität Regensburg aufgeführt. Einem Maßstab nach ist nicht leicht zu verstehen, was das alles bedeutet, aber mit kleinen Teenagern wie Bibi ist es so auch das Gemeinsame.

Das Gemeinkörper geht ungefähr so: Bei

ist peinlich, Bibi kann nicht singen, I hat den Song geklaut. Alles Vorwürfe, vielleicht stimmen, vielleicht nicht. Wahrheit geht es natürlich um etwas

Wahrheit geht es darum, dass die Leute darüber reden. Es geht um Neid. Auf YouTube wird über das Video gesprochen, und das ist wahrscheinlich noch mehr als im sonstigen Leben.

Das Gemecker hat, wie so oft, eine Aufmerksamkeit auf sich gezogen als der Erfolg Das Video schaffte es mit 1,9 Millionen Dislikes in wenigen Tagen in eine entsprechende Top-Ten-Liste, und zwar weltweit – noch niemand in Deutschland hat so etwas geschafft, vor allen

hat etwas Ähnliches gemacht, nicht die deutlich weniger erfolgreiche deutsche YouTube-Konkurrenz, die nach Tagen verzweifelt versucht, auch ein bisschen von Bihls Erfolg zu profitieren. Sogar die Parodien des Liedes ziehen Millionen von Klicks. „Einige haben mir gut gefallen“, sagt Bihl. Und es ist ja wohl klar, dass sie das auch gewollt haben. „Ich weiß so was nur sagt, wer wirklich drüberstellt.“

Also auch an Miley Cyrus vorbei auf Platz 6. Bisher war Miley Cyrus das Vorbild, und zwar „sehr immer“. Das verrät Biri mir jetzt, hier auf dem Sofa. Na ja, verraten klingt zwar gut, wenn ich an die Vermarktung des Geprächs gegenüber meinen Tochterkindern denke, aber vielleicht ein bisschen übertrieben, denn sicherlich hat Biri das mit Miley Cyrus auch schon x-mal erzählt. Wie sich sagen will, ist eigentlich nur das: Es gibt keinen besonderen Grund, sich für Platz sechs auf dieser Liste zu schämen.

Auch darum: Es ist geldechnisch völlig Banane, ob jemand ein Video gern anschaut, der von ihm kein Ansehen hat.

schaut. Sobald Bibbi auf den Bildschirm schaut, wird. Der Klick ist eine wertvolle Re-
chengabe. Es gibt für Bibbi also nicht
nur einen Grund, sich zu ärgern, son-
dern Millionen Gründe, sich zu freuen.
Und Bibbi sagt, dass sie das auch tut.
Aber gleichzeitig kriegen YouTuber pro 1000
Abonnenten 80 Euro von YouTube, was Bibbi
weder bestätigt noch abstreitet, sondern
ganz nicht kommentiert. Sie weiß nicht, ob
die anderen Freunde berichten, ob sie
Emojis schicken, ob sie überhaupt nicht
ihre YouTuber-Seite besuchen. Sie weiß
nicht, warum sie nicht mehr auf dem Bild-
schirm ist. Sie weiß nicht, ob sie sich
auf dem Bildschirm wiederfindet. Sie weiß
nicht, ob sie sich auf dem Bildschirm wiederfindet.
Wie kann sie das wissen? Sie hat
einen Mund, um darüber zu reden, und
sie hat eine Seele, die darüber zu reden
möchte, aber sie kann es nicht.

Auf dem Weg nach Hause: Auruf merkt minderjährige Töchter. „Mama, Mama hat dich in ihre Snapchatgruppe aufgenommen – du bist Interview Nummer sieben.“ Ich bin Interview Nummer sieben! In Babis Windschatten habe auch ich es unter die Top Ten geschafft.

eum im Leben unerlässiges Sorgen. Gesc-
tern zum Beispiel kam ich mir ein biss-
chen schäelig vor und setzte sich auf den Besucher-
sessel, wie die Zeitschriften er und die
Pausenunterhaltungen. Er saß so
herausgeworfen vor mir und frag-
te mich, was es mit der Sache
Theorie auf sich habe. Sein Freund
Punkt für Punkt aufgelistet, hörte davon ge-
fragt und er fragte sich, ob das
wirklich stimmte, denn dann habe er
ein Problem. Die 1000-Schüssel-Theo-
rie besagt, dass ein Mann im Leben
genau ein tausendstes Organ entdecken
kann. Danach ist Fehlerfrei. Mit
diesem Kreislauf werden die Jungen von On-
kologen behandelt werden können. Es kann
aber auch sein, dass diese Formel
nur von Verbrauch Deutschlands
mathematischer erfunden wurde, damit die
Leute in dieser Zeit mit Zahlen vor
sich her schreien. Ich kann mir vorstellen,
dass es eine Menge von Leuten gibt, die Vier-
undzwanzigstundenarbeit erstatthalb wünschen,
die es nicht gewohnt sind, ihre Familienpla-
nungen bereits um Pflichten herum zu
drehen, immer bedient zu sein und ihrem
Kinderwunsch zu entsagen.

Ebenso gut gefällt mir der Gedanke, dass es solche Grenzen auch bei anderen Errichtungen geben könnte. Jeder Mensch kann nur 500 Fußballspiele gucken, dann wird er blind. Nach 300 Tafeln Schokolade fallen die Zahne aus. Man muss sich vorstellen, was Menschen unterhalb erhöhen. Der Organhandel blühte wie Raps auf dem Kartoffelfeld. Und die globale Wirtschaft brummte wie verkateter Liederkranich.

rückt, wenn wir bloß 10.000 hörten könnten, bevor die Ohren abfielen. Ich erging mich in on, und Nick hörte auf.

fragte er: „Also...“

„Natürlich ist es des Wettschießens.“ und hörte, wie der Abpfiff. Er stand auf, sah aus dem Fenster, schaute der Tag zu, schaute er fröhlich in die Tageszeitung in sein Zimmer,

Staying up to date

Websites

- DHd blog: <https://dhd-blog.org/>
- dha website: <https://digital-humanities.at>
- H-Net (Humanities and Social Sciences Online): <https://networks.h-net.org/> (identify specific networks interesting to you)

Mailing lists

DHd-Mailingliste: <https://lists.uni-hamburg.de/mailman/listinfo/dhd>



Explore!