Visualization of **Humanities Data**

Florian Windhager

VIS for DH – Lecture No. 1

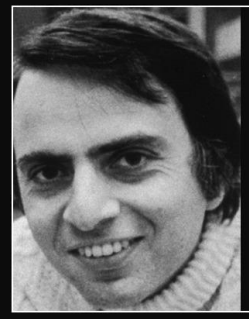
1. Human creations („culture“) & humanities data
2. Navigating the zoo of visualization methods
3. Maps & Charts
4. Graphs & Trees
5. Visualization networks - towards bigger pictures of humanities topics

Can we see „culture“ from outer space?

Thesis: The most important issues today are not visible to our naked eyes.

Culture is to know the best that has been said and thought in the world.

~ Matthew Arnold



The uniqueness of humans has been claimed on many grounds, but most often because of our tool-making, culture, language, reason and morality. We have them, the other animals don't, and -- so the argument goes -- that's that.

— Carl Sagan —

AZ QUOTES

Culture, as we're practicing it is causing a lot of pain.

Terence McKenna

artful things

useful things

awful things

human culture

Culture is one of the two or three most complicated words in the English language.

— Raymond Williams —

AZ QUOTES

? →

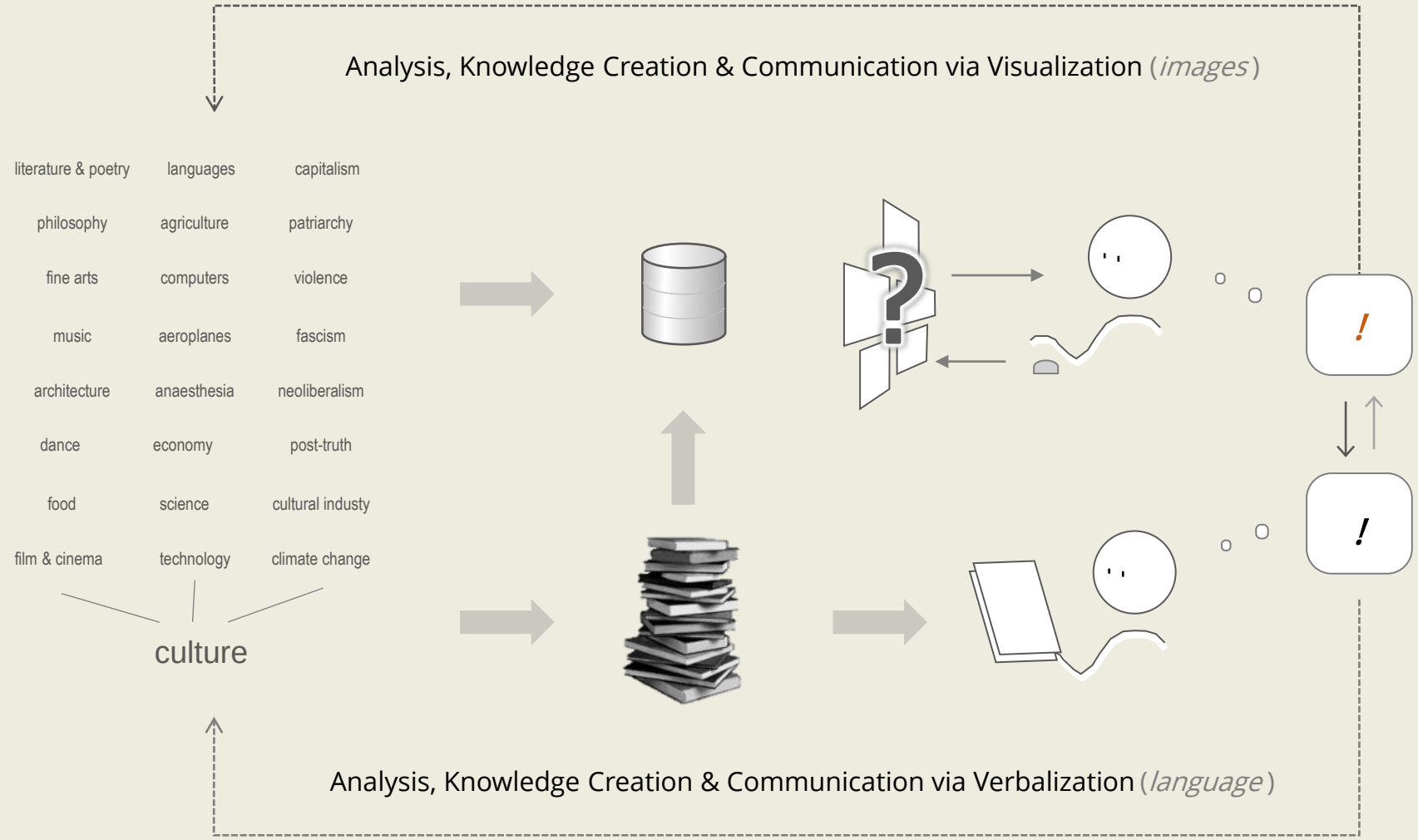
Guiding question: Can we actually **see** the things and subject matters of the arts & humanities?

>> please visit bit.ly/VIS4DH-L1

literature & poetry	languages	capitalism
philosophy	agriculture	patriarchy
fine arts	computers	violence
music	aeroplanes	fascism
architecture	anaesthesia	neoliberalism
dance	economy	post-truth
food	science	cultural industy
film & cinema	technology	climate change
artful things	**useful things**	**awful things**

human creations

Two major modes of information processing



Visualization as modern cultural practice –

between useful & artful

Distributed Cognition as a Theoretical Framework for Information Visualization

Zhicheng Liu, Nancy J. Nersessian, and John T. Stasko, *Member, IEEE*

Abstract—Even though information visualization (Info…matured in recent years, it is generally acknowledged that the field still lacks supporting, encompas…distributed cognition framework can be used to substantiate the theoretical foun… theoretical constructs of the distributed cognition approach, based … uss how the distributed cognition framework can have a… earchers. Our contributions are as follows. Firs… a property of the human mind. Second, w… s may not be useful in informing Info… lications on how we perform eva…

Index Ter…

1 INTRODU…

Card, Mackin… , which "attempt to (InfoVis) as "th… eria other than testa- *sentations* of abs… ories of the world, and Representation an… ocess in a way that per- portant dimensions …3]. To this definition, we tion in InfoVis resea… tial to lead to new hypothe- mentally central to these… perimental research. We argue ceived far less scrutiny and … make cognition a research agenda cognitive scientists. Although t… can have implications on the research tive science, there are no fully develo… thodologies we adopt. and predict the effectiveness (or lack thereo…) sign, which InfoVis research can simply adopt. Existing … lines such as Shneiderman's mantra "overview first, zoom and filter, then details-on-demand" [27] have been useful in offering guidance to practitioners, but they also lack the rigor and power that formal theories or methodologies should provide [9]. As InfoVis designers, we often have only our own intuition and experience to depend on. Even though InfoVis research has matured technically in recent years, an important problem for the field remains the lack of an underlying theory or even a systematic framework for guiding design and investi-

…BUTED COGNITION: AN ENVIRONMENTAL PERSPECTIVE

DCog is a theoretical framework that originated largely with the work of Edwin Hutchins and colleagues at UCSD in the mid 1980s. In this section we provide a short description of the origin and the philosophy of this framework. We begin with the conventional view that tools amplify cognition. Visual tools, ranging from multiplication on pencil and paper to diagrams, do seem to amplify cognition as we can accomplish more with than without them. From this traditional per-

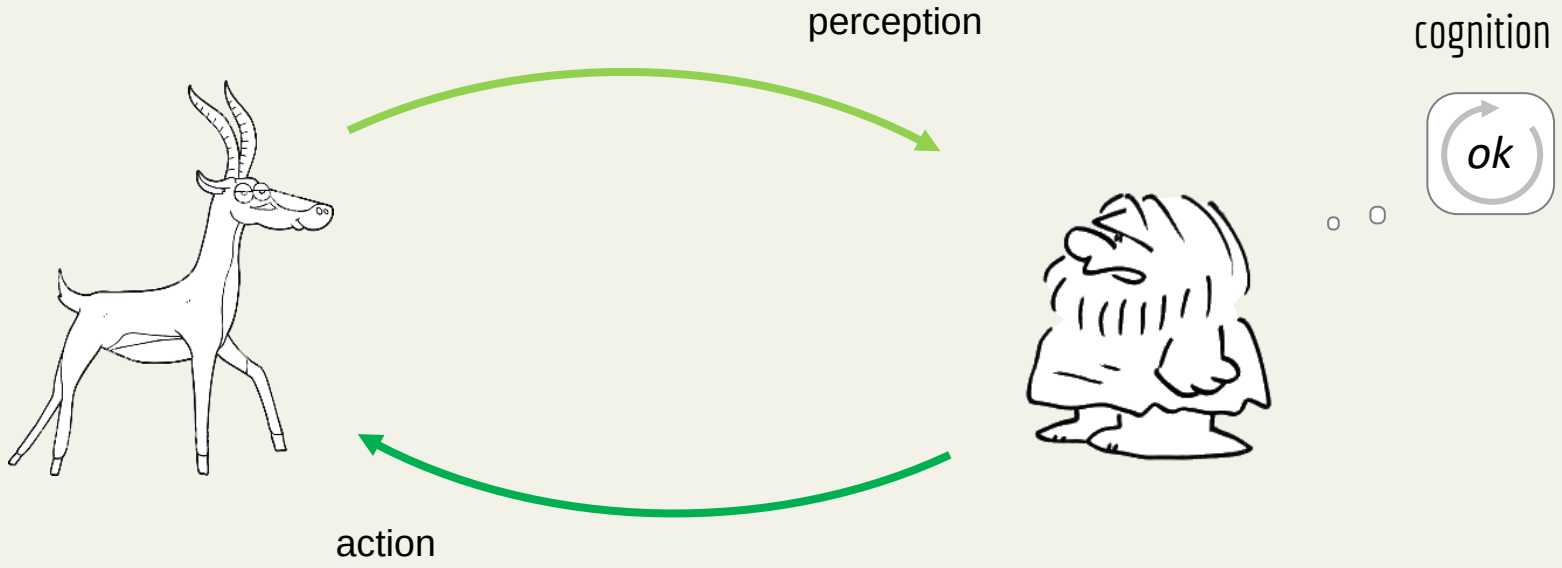
image [c] jantoo

Liu, Nersessian & Stasko (2008) link

Brains evolved to deal with

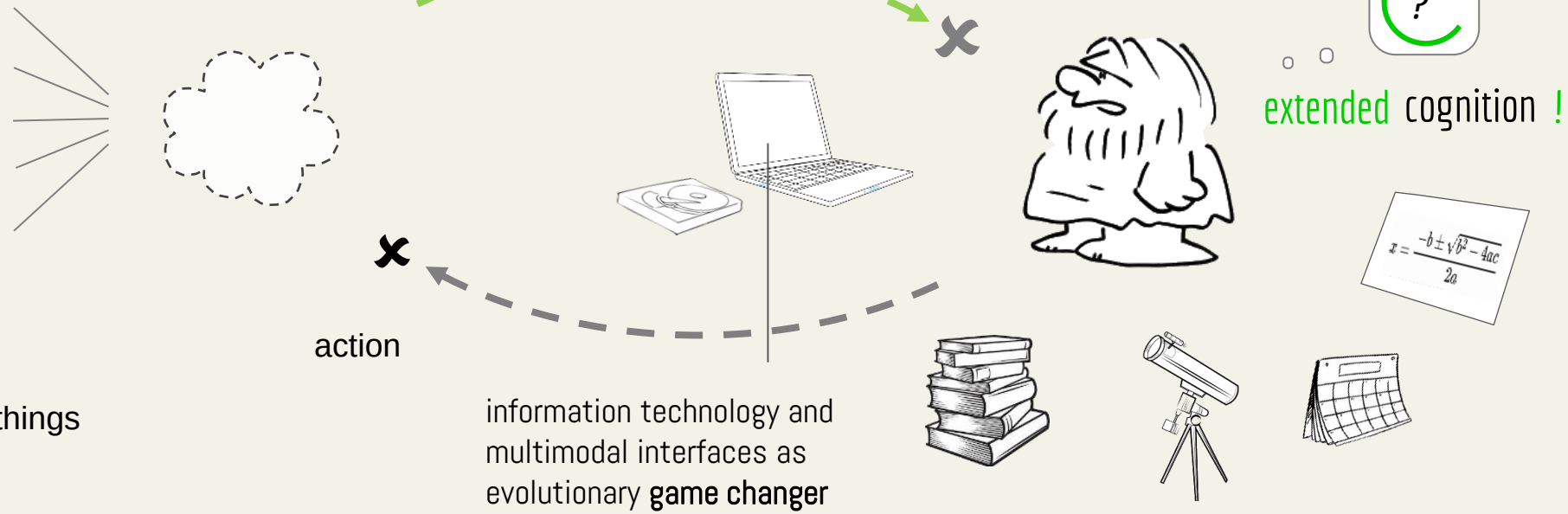
common cognitive challenges

- *things to run at*
- *things to eat*
- *things to run away from*
- *etc.*



Modernity brings a multitude of
severe cognitive challenges

- nation states
- political systems
- science
- microbes
- stock markets
- climate change
- inequality
- pandemics
- …
- history
- complex cultural things



▸▸ to which cultures respond with cultural & technical
augmentations & extensions of cognition

Visualization Definitions

"visual representations of datasets intended
to help people carry out some task more effectively ."

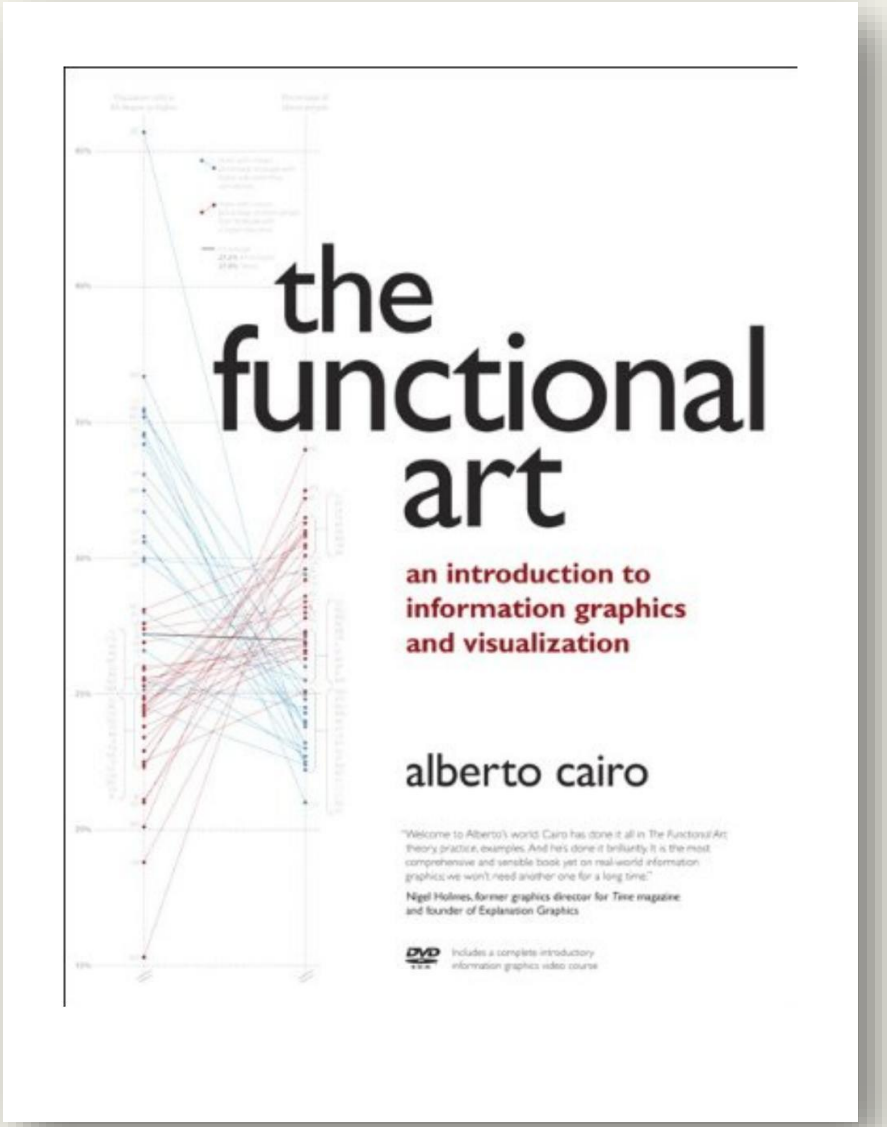
(Munzner, 2014)

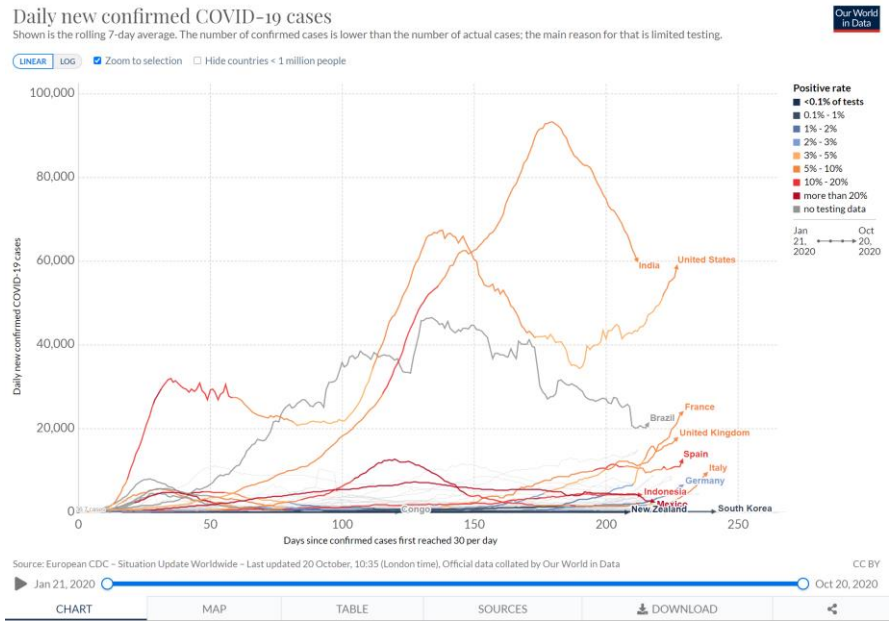
"The use of computer-suppported, interactive, visual representations of abstract data to amplify cognition."

(Card, 1999)

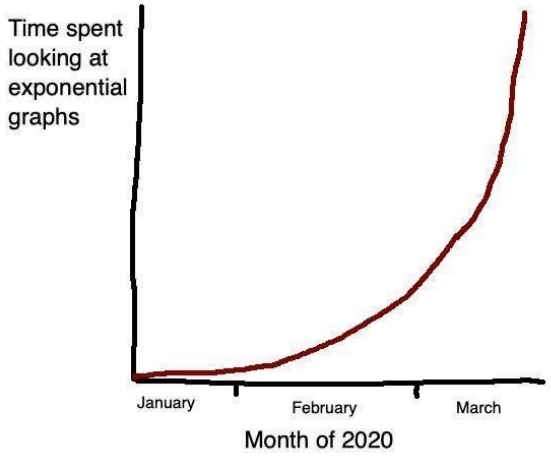
"Let's face it: The most important
issues today are not photographable."

(Stefaner, 2014, link)





https://ourworldindata.org/coronavirus



Source: Internet

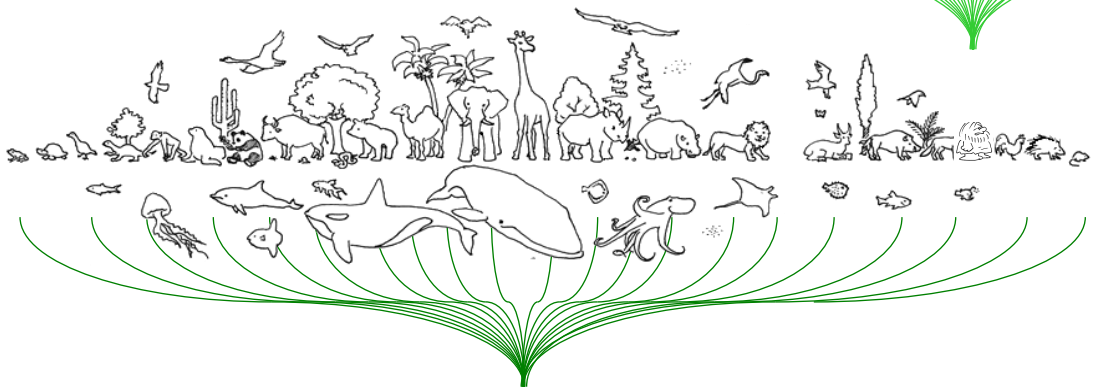
Human creations & humanities data

Verbalization of Human Knowledge & Culture

New Analysis & Communication Options by Visualization

See e.g., Heer et al., (2010). “A Tour through the Visualization Zoo” (link)

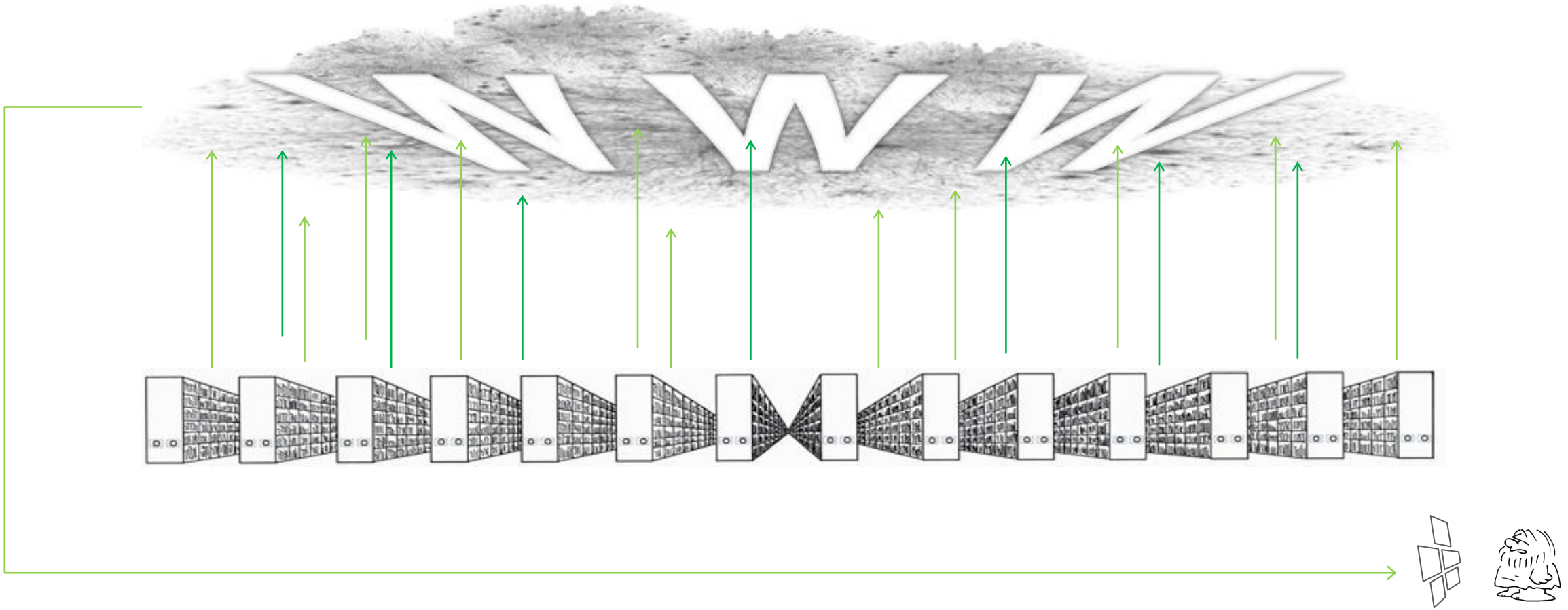
“Culture”



“Nature”

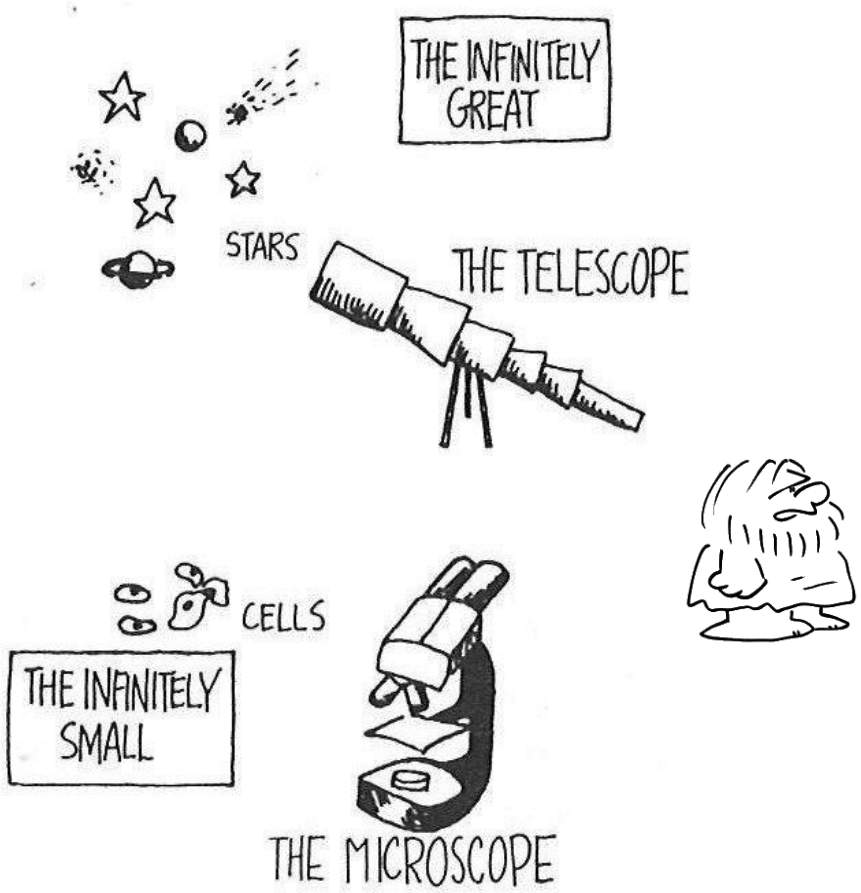
„World“

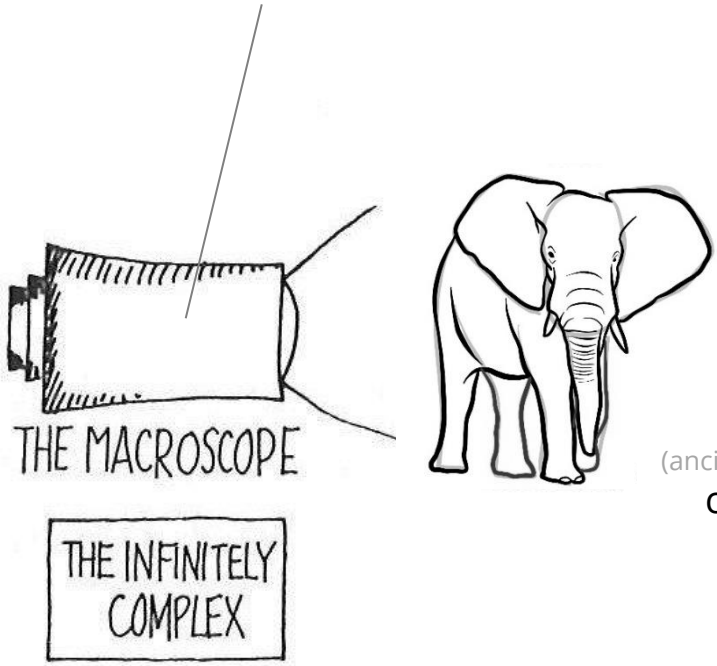
Large Scale Digitization of historical & cultural sources



Availability of large amounts of digital historical & cultural sources enables & requires new types of cognition support & sensemaking strategies

Visualizations as „Macroscopes“ (DeRosnay, 1979)





(ancient metaphor for)
complex things

Stefaner, 2014: http://well-formed-data.net/archives/1027/worlds-not-stories

Source: Pat Whitehead In Monets' Garden

„Close Reading & Viewing“

Google Arts & Culture Experiments: Freefall

„Distant Reading & Viewing“

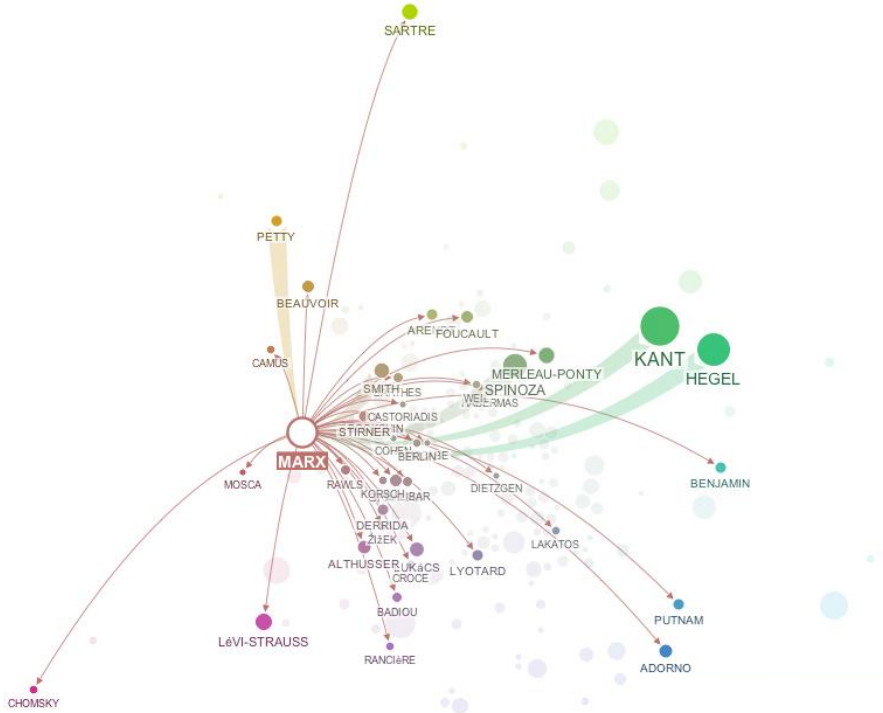
Visualization examples

EdgeMaps

by Dörk et al., 2011

https://mariandoerk.de/edgemaps/

visualization method:
network diagram



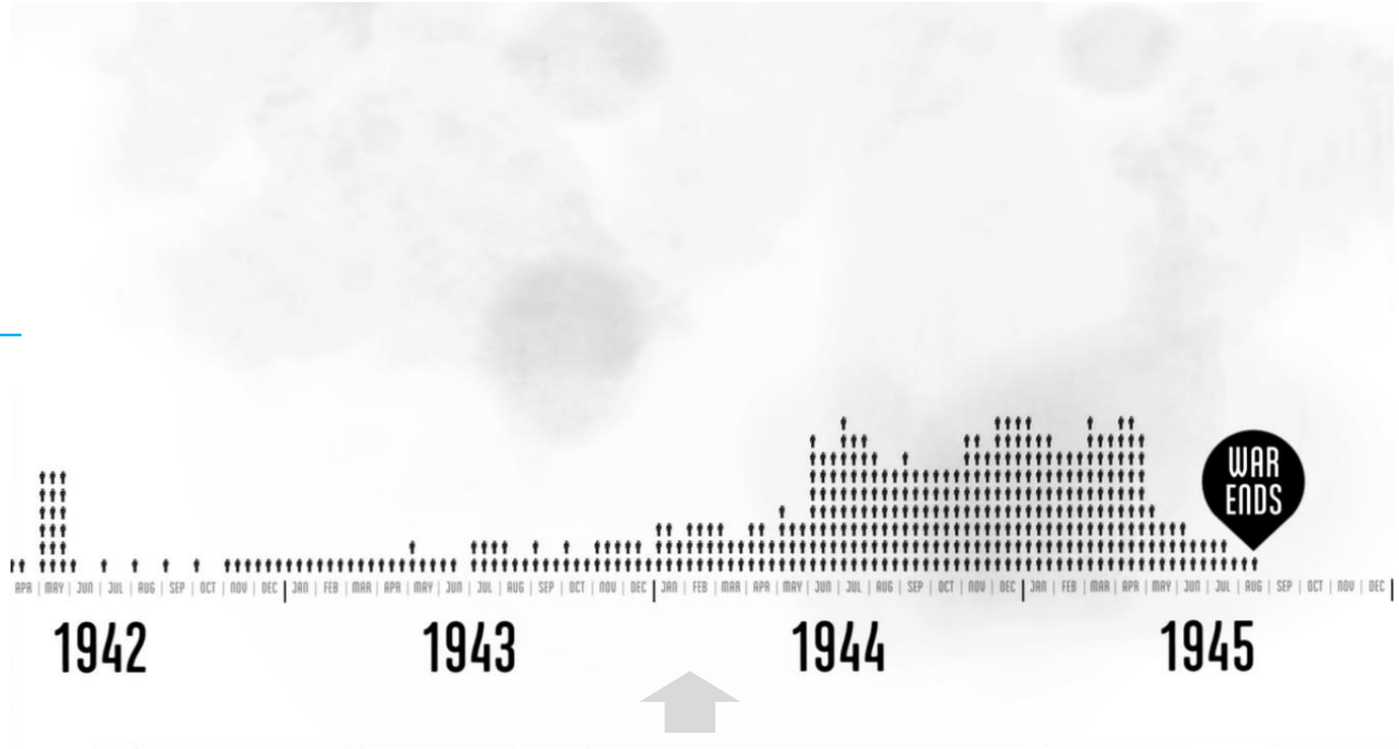
Philosophical Discourse

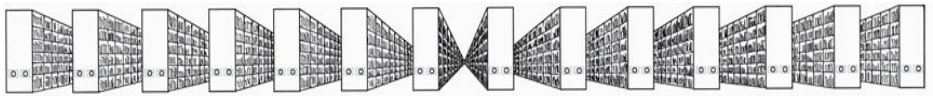
The Fallen of WW2

by Neil Halloran

http://www.fallen.io/ww2/

visualization method:
bar chart





World War II + Human Violence

Navigating the zoo of visualization methods

DOI:10.1145/1743546.1743567

Article development led by acmqueue
queue.acm.org

A survey of powerful visualization techniques, from the obvious to the obscure.

BY JEFFREY HEER, MICHAEL BOSTOCK, AND VADIM OGIEVETSKY

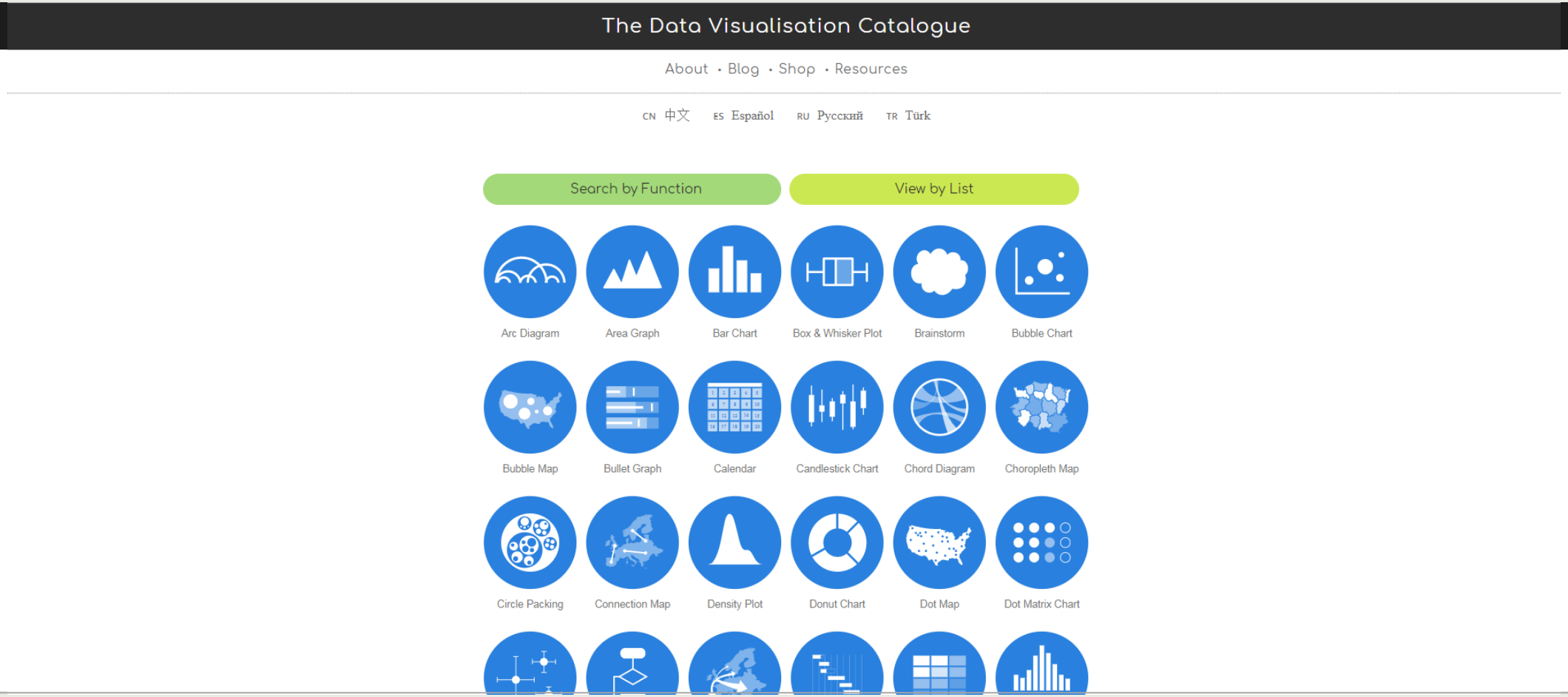
A Tour Through the Visualization Zoo

THANKS TO ADVANCES in sensing, networking, and data management, our society is producing digital information at an astonishing rate. According to one estimate, in 2010 alone we will generate 1,200 exabytes—60 million times the content of the Library of Congress. Within this deluge of data lies a wealth of valuable information on how we conduct our businesses, governments, and personal lives. To put the information to good use, we must find ways to explore, relate, and communicate the data meaningfully.

The goal of visualization is to aid our help engage more diverse audiences in exploration and analysis. The challenge is to create effective and engaging visualizations that are appropriate to the data.

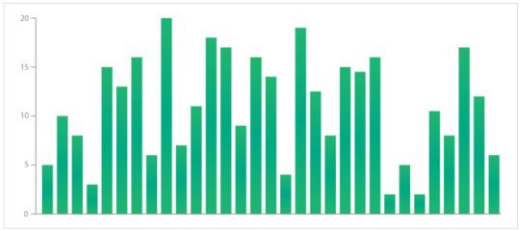
Creating a visualization requires a number of nuanced judgments. One

Interactive Text-/ Picture Books



Interactive Overviews: https://datavizcatalogue.com/ or https://datavizproject.com/

Bar Chart



Description

As known as *Bar Graph* or *Column Graph*.

The classic Bar Chart uses either horizontal or vertical bars (column chart) to show discrete, numerical comparisons across categories. One axis of the chart shows the specific categories being compared and the other axis represents a discrete value scale.

Bars Charts are distinguished from Histograms, as they do not display continuous developments over an interval. Bar Chart's discrete data is categorical data and therefore answers the question of "how many?" in each category.

One major flaw with Bar Charts is that labelling becomes problematic when there are a large number of bars.

Anatomy

Functions

Comparisons Patterns

Similar Charts

Histogram Multi-set Bar Chart Population Pyramid

Radial Bar Chart Radial Column Chart Stacked Bar Graph

More Bar Chart variations in this blog post.

Tools to Generate Visualisation

MS Excel & Apple Numbers
AnyChart (code)
amCharts (code)
D3 (code)
DataHero
Datamatic
Datawrapper
Google Charts (code) & Google Docs
Infogram
JSCharting (JS Library)
OnlineChartTool.com
Python Graph Gallery (code)
Slemma
Vega (code)
ZoomCharts

Examples

Countries With Most Oil Reserves [2017-18] - FusionCharts

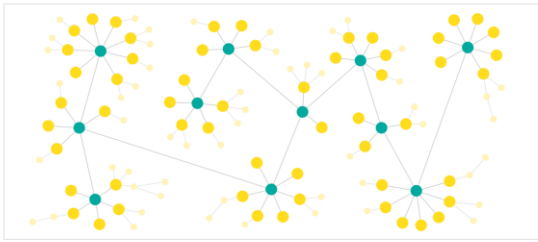
Bar Chart Gallery - AnyChart

Want your work linked on this list? Click Here

Need to access this page offline? Download the eBook from here.

Merchandise & other related dataviz products can be found at the store

Network Diagram



Description

Also known as *Network Graph, Network Map, Node-Link Diagram*.

This type of visualisation shows how things are interconnected through the use of nodes / vertices and link lines to represent their connections and help illuminate the type of relationships between a group of entities.

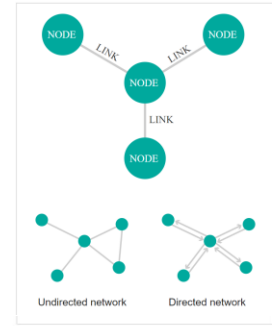
Typically, nodes are drawn as little dots or circles, but icons can also be used. Links are usually displayed as simple lines connected between the nodes. However, in some Network Diagrams, not all of the nodes and links are created equally: additional variables can be visualised, for example, by making the node size or link stroke weight proportion to an assigned value.

By mapping out connected systems, Network Diagrams can be used to interpret the structure of a network through looking for any clustering of the nodes, how densely nodes are connected or by how the diagram layout is arranged.

The two notable types of Network Diagram are "undirected" and "directed". Undirected Network Diagrams only display the connections between entities, while directed Network Diagrams show if the connections are one-way or two-way through small arrows.

Network Diagrams have a limited data capacity and start to become hard to read when there are too many nodes and resemble "hairballs".

Anatomy



Functions

Relationships

Similar Charts

Arc Diagram Brainstorm Chord Diagram (non-ribbon)

Connection Map Tree Diagram

Tools to Generate Visualisation

amCharts
Any Chart
Cytoscape
Datamatic
Gephi
Graph-tool
Mike Bostock's Block (D3 code)
Plot.ly (code)
Python Graph Gallery (code)
sigmajs
Vega (code)
yFiles (code) or yEd live
ZoomCharts

Maps

Charts

Graphs

Trees

Time

Maps for Humanities Data

Maps

Interactive Catalogue:
https://datavizcatalogue.com/

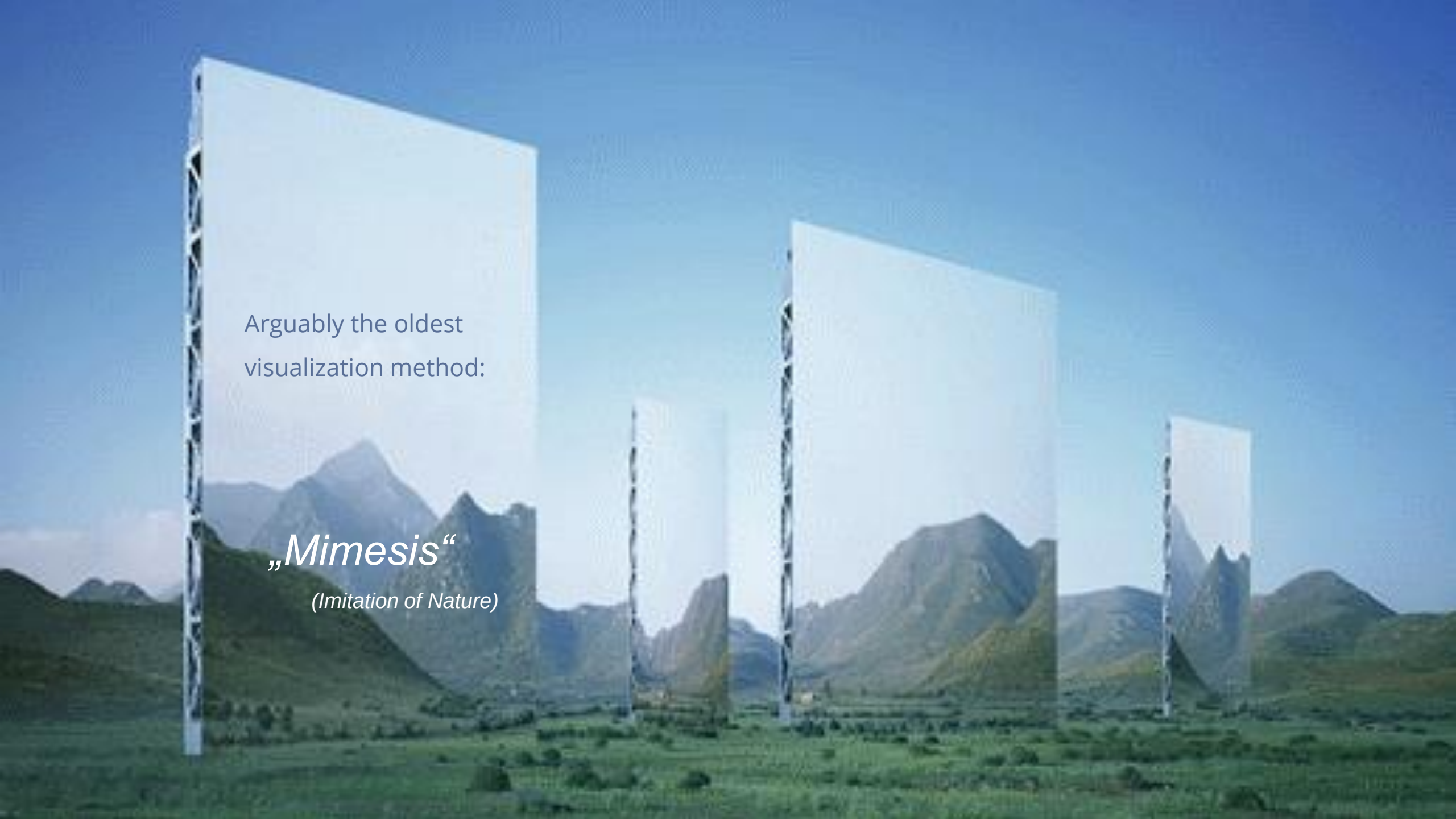
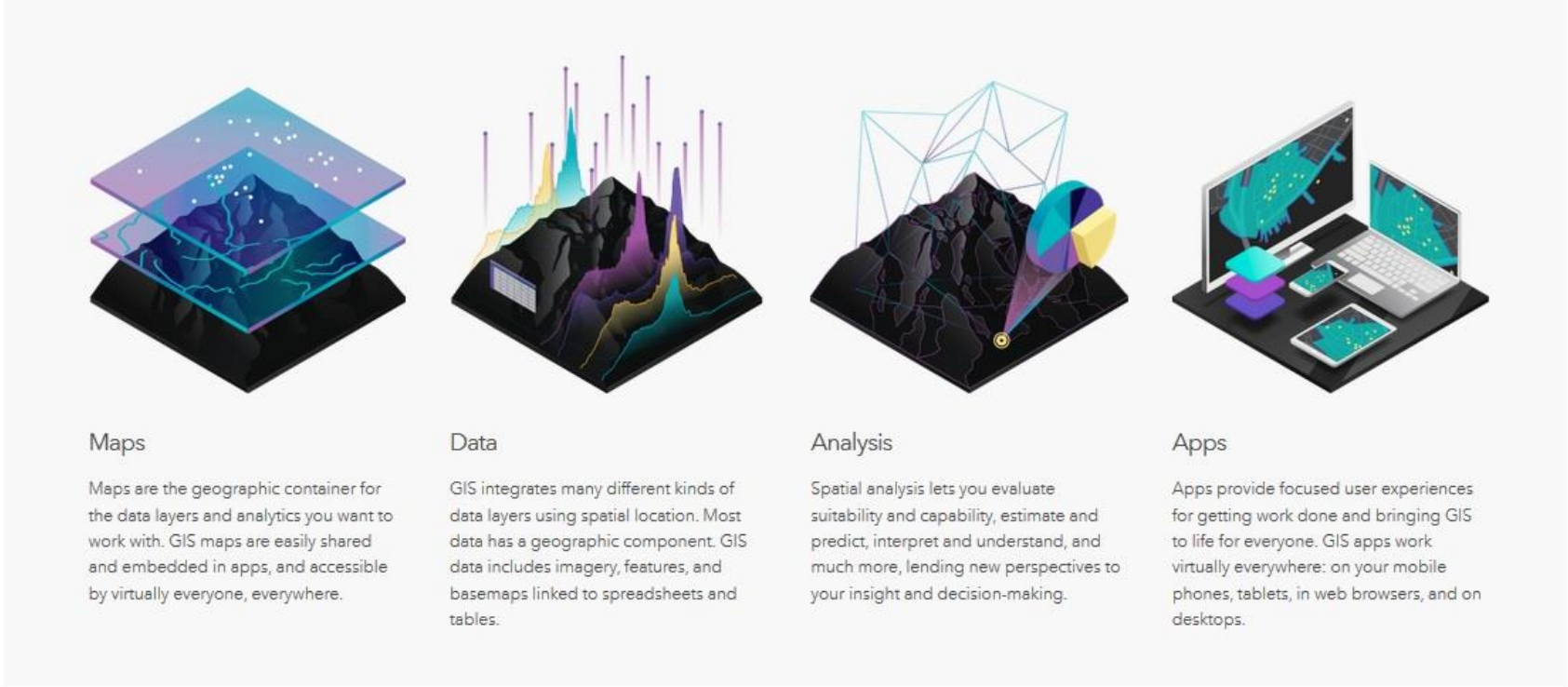

Arguably the oldest
visualization method:
„Mimesis“
(Imitation of Nature)

Image: https://www.intelligence-airbusds.com/imagery/constellation/#gallery-1

Geo Information Systems (GIS)



Source: www.esri.com

Concept: Dr. Maria Vargha

What GI?

„any information that refers to a location on the Earth's surface"

What is GIS?

„In simple terms, a GIS is a computer package that is designed to represent Geographical Information effectively. It is, therefore, a system that allows us to handle information about the location of features or phenomena on the Earth's surface. This is usually done by combining a database management system (DBMS) with a computer mapping system."

(Gregory-Ell 2007)

Hybrid / Multi Layer Maps

adding data layers on top of geographic grids

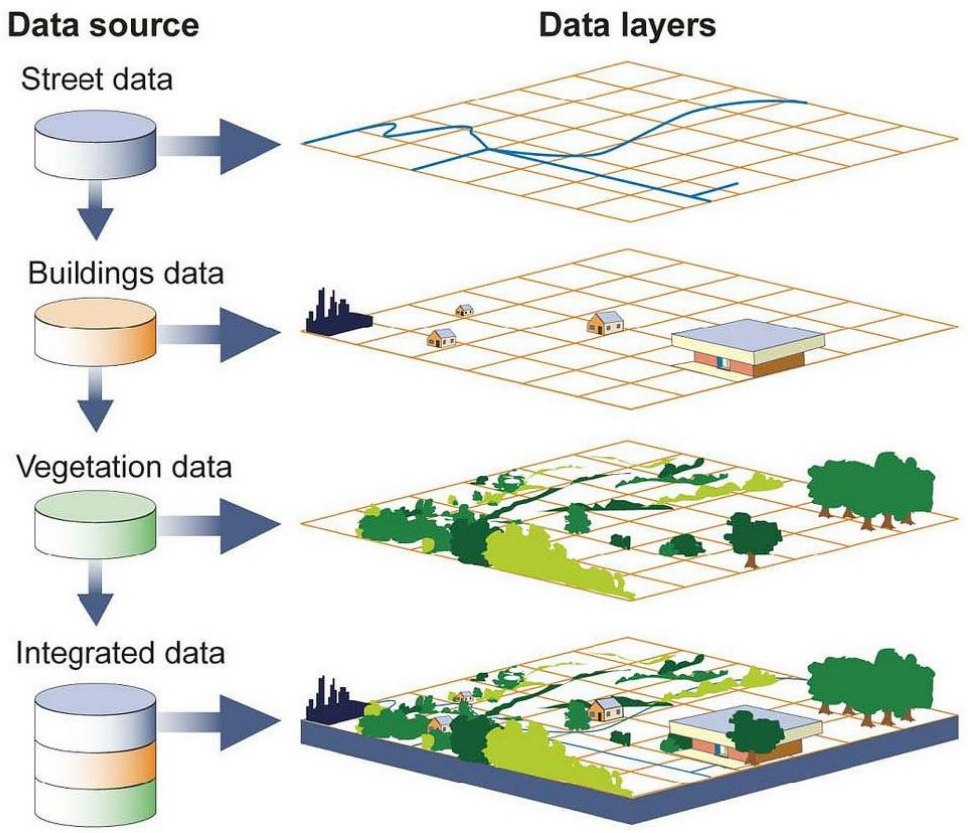
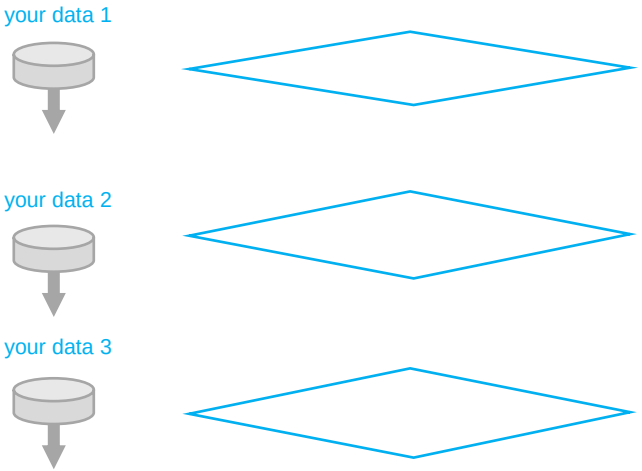
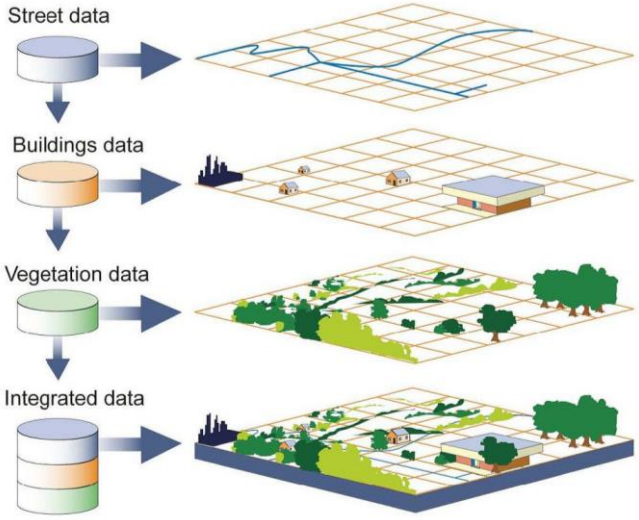
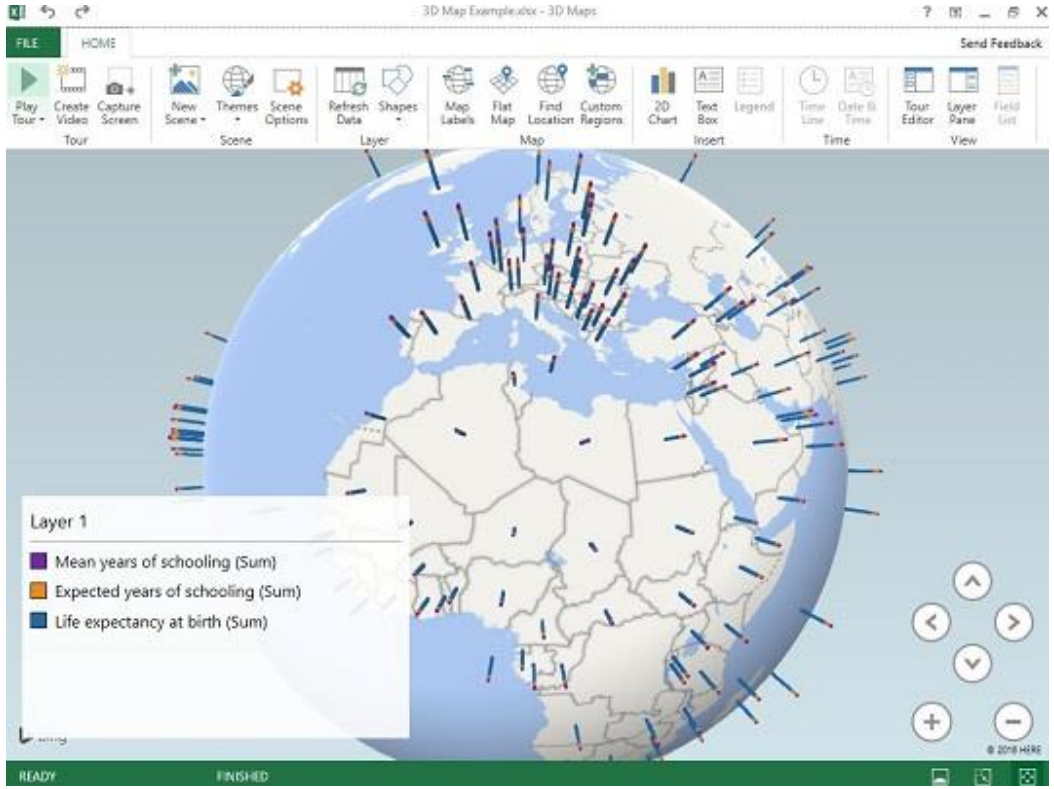


Image: https://www.nationalgeographic.org/encyclopedia/geographic-information-system-gis/



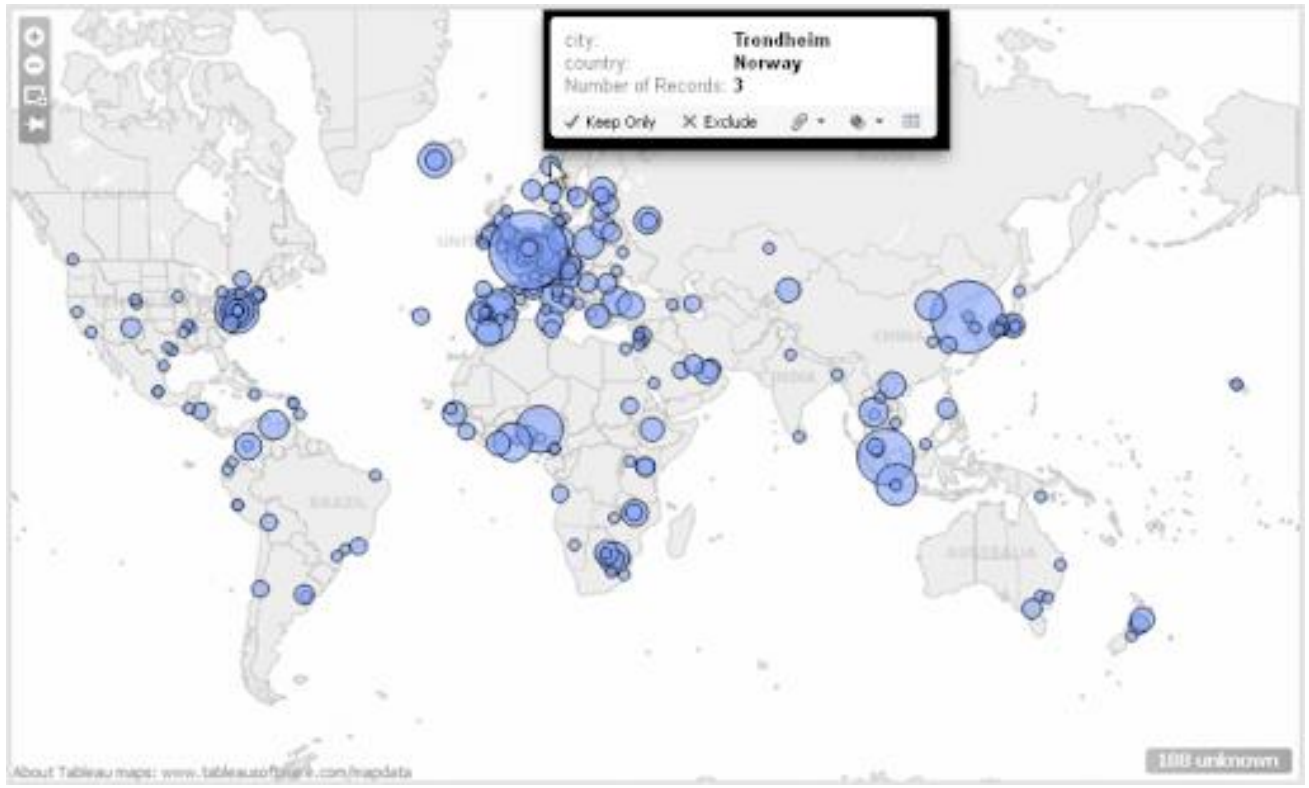


Mapping Tools I



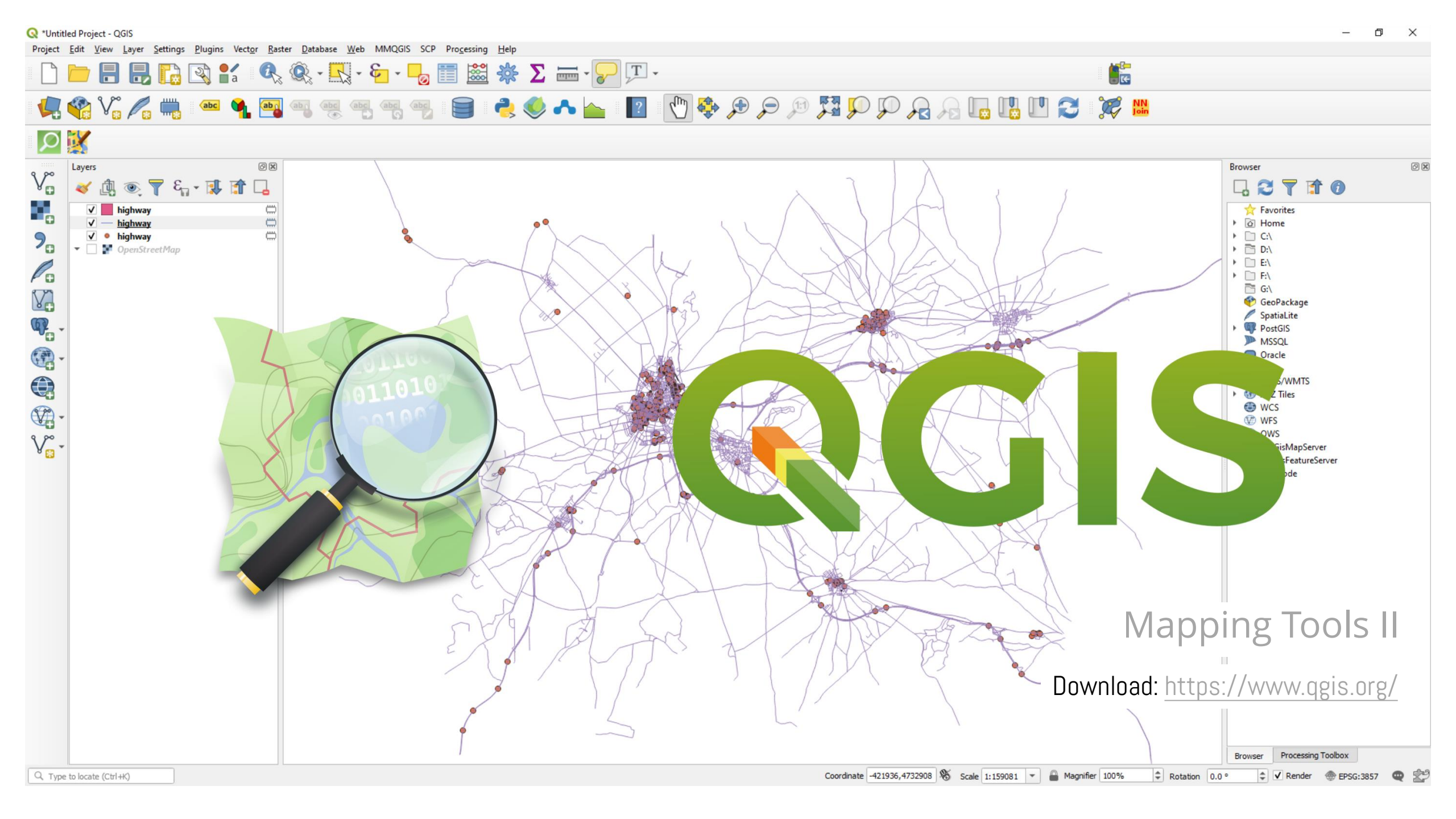
MS Excel

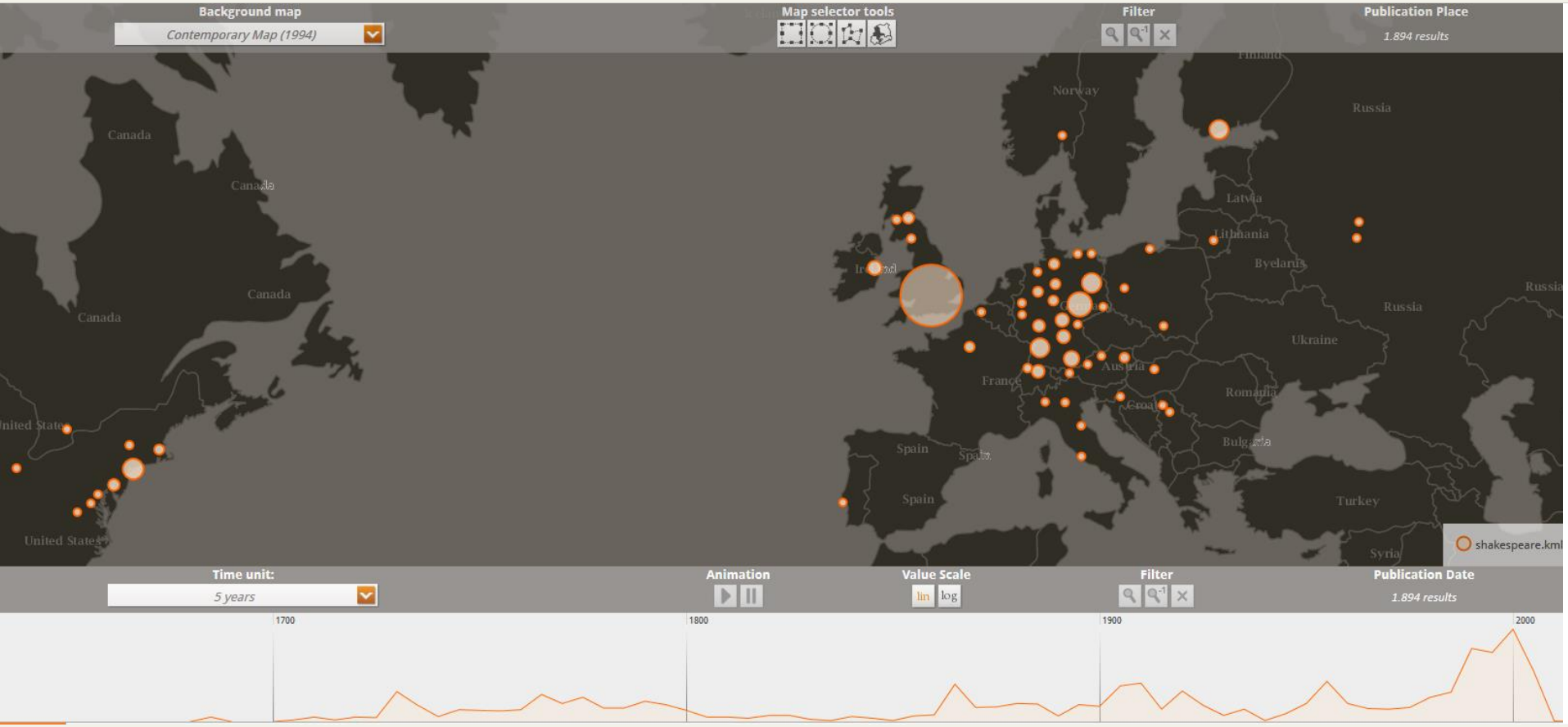
Download: https://www.microsoft.com/en-us/microsoft-365/excel



Tableau

Download: https://www.tableau.com/academic/students


*Untitled Project - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web MMQGIS SCP Processing Help
Layers
highway
highway
highway
OpenStreetMap
Browser
Favorites
Home
C:\
D:\
E:\
F:\
G:\
GeoPackage
SpatiaLite
PostGIS
MSSQL
Oracle
WCS
WFS
Browser
Processing Toolbox
Type to locate (Ctrl+K)
Coordinate -421936,4732908
Scale 1:159081
Magnifier 100%
Rotation 0.0 °
Render
EPSG:3857
QGIS
Mapping Tools II
Download: https://www.qgis.org/



Bubble Map

Jänicke et al. | Dariah Geobrowser | http://geobrowser.de.dariah.eu/

Schich et al., 2014 | Charting Culture https://www.youtube.com/watch?v=4gIhRkCcD4U

+ TIME

space-time cube

coordinated timeline

animation

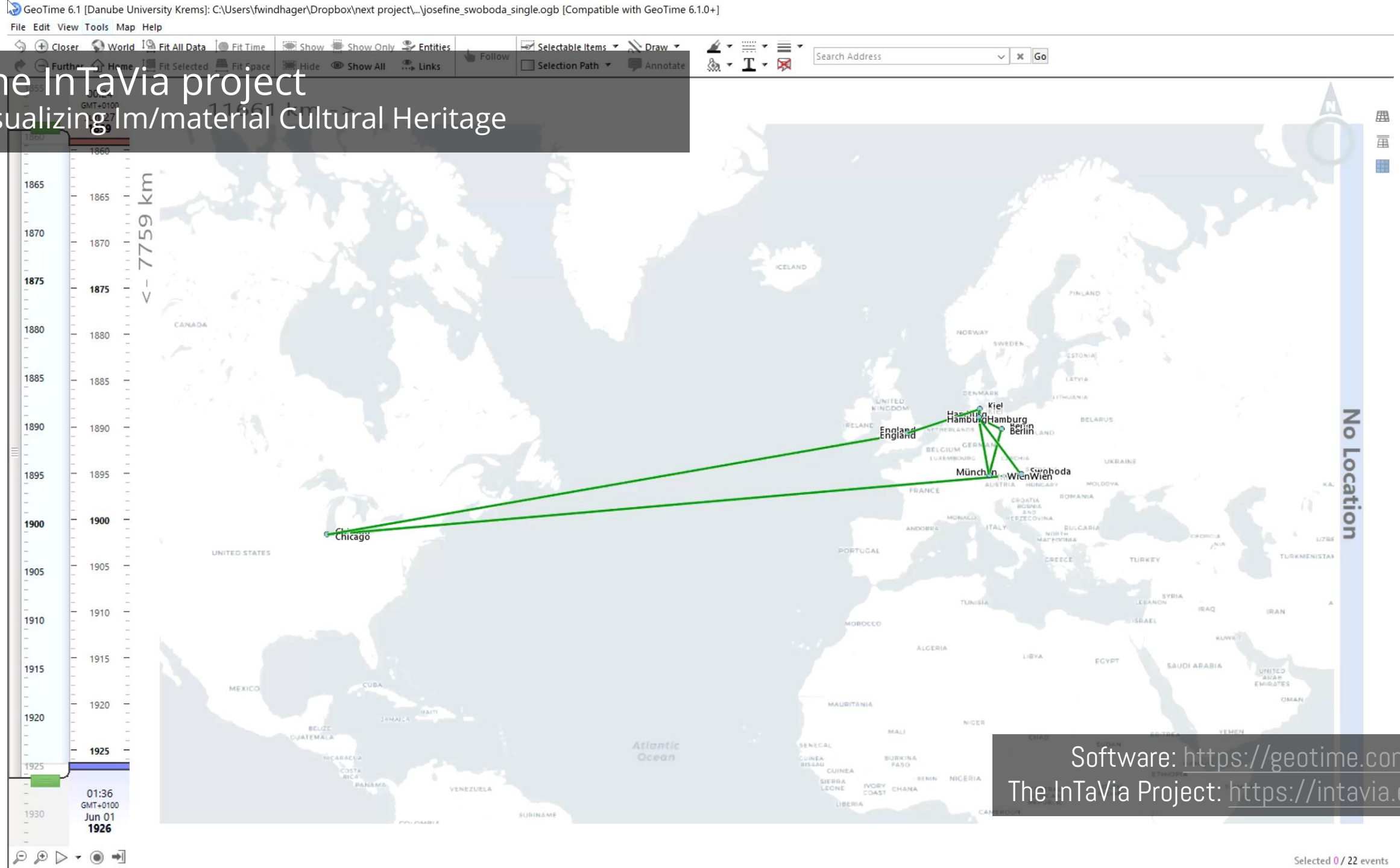
color-coding

data comics

many other options

The InTaVia project

Visualizing Im/material Cultural Heritage



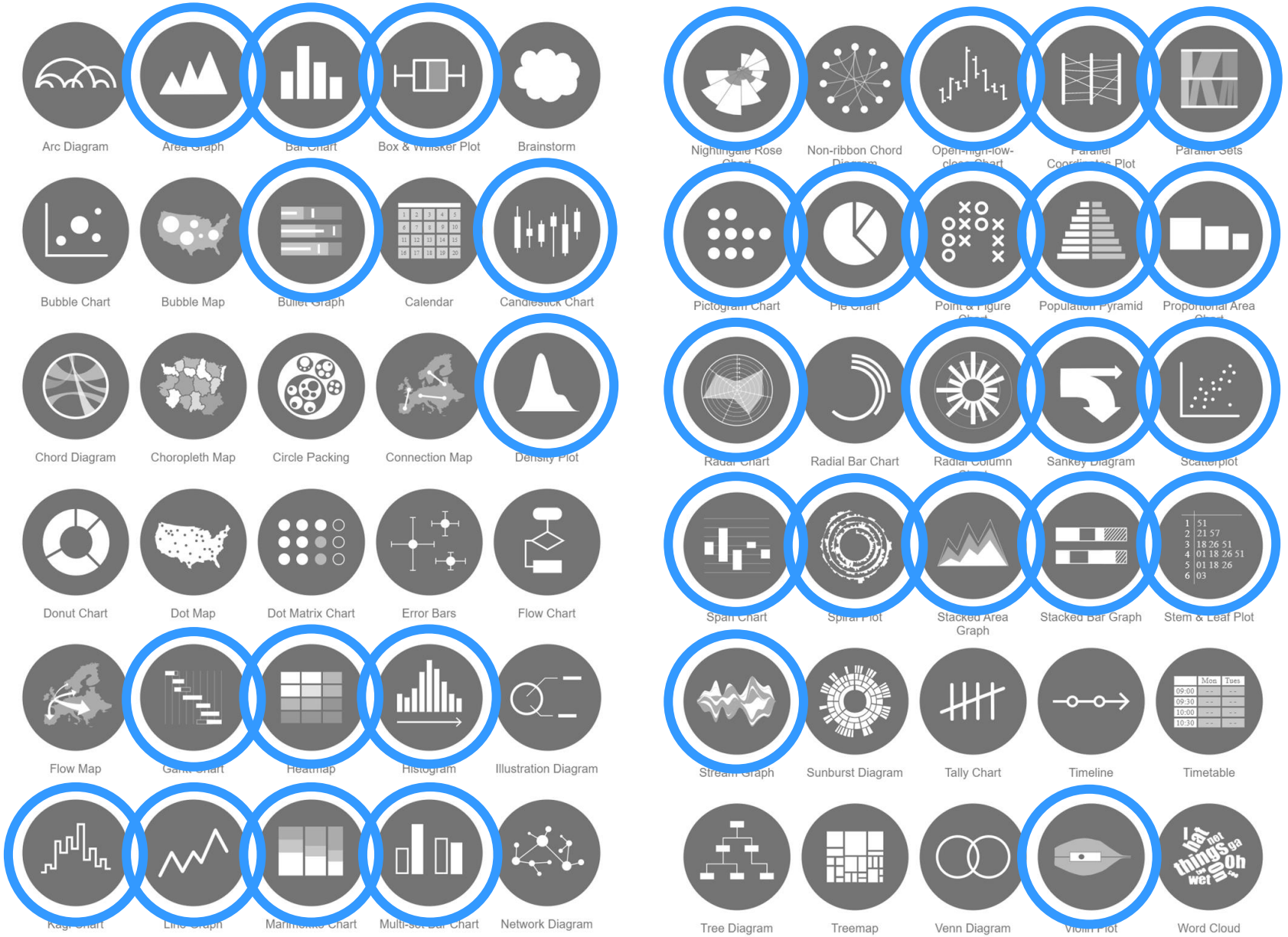
Software: https://geotime.com/

The InTaVia Project: https://intavia.eu

Tools to map arts & humanities data:

- Dariah Geobrowser (https://de.dariah.eu/geobrowser)
- Storymap (https://storymap.knightlab.com/)
- Palladio (http://palladio.designhumanities.org/)
- My Maps (https://www.google.com/maps/d/) (Google)
- MS Excel, Tableau, D3.js, etc

Charts for Humanities Data



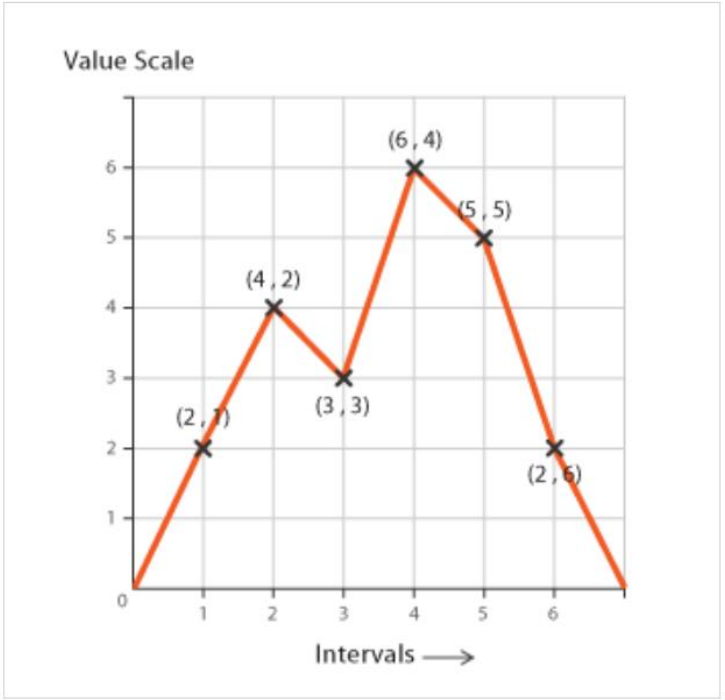
Charts

Interactive Catalogue:
https://datavizcatalogue.com/

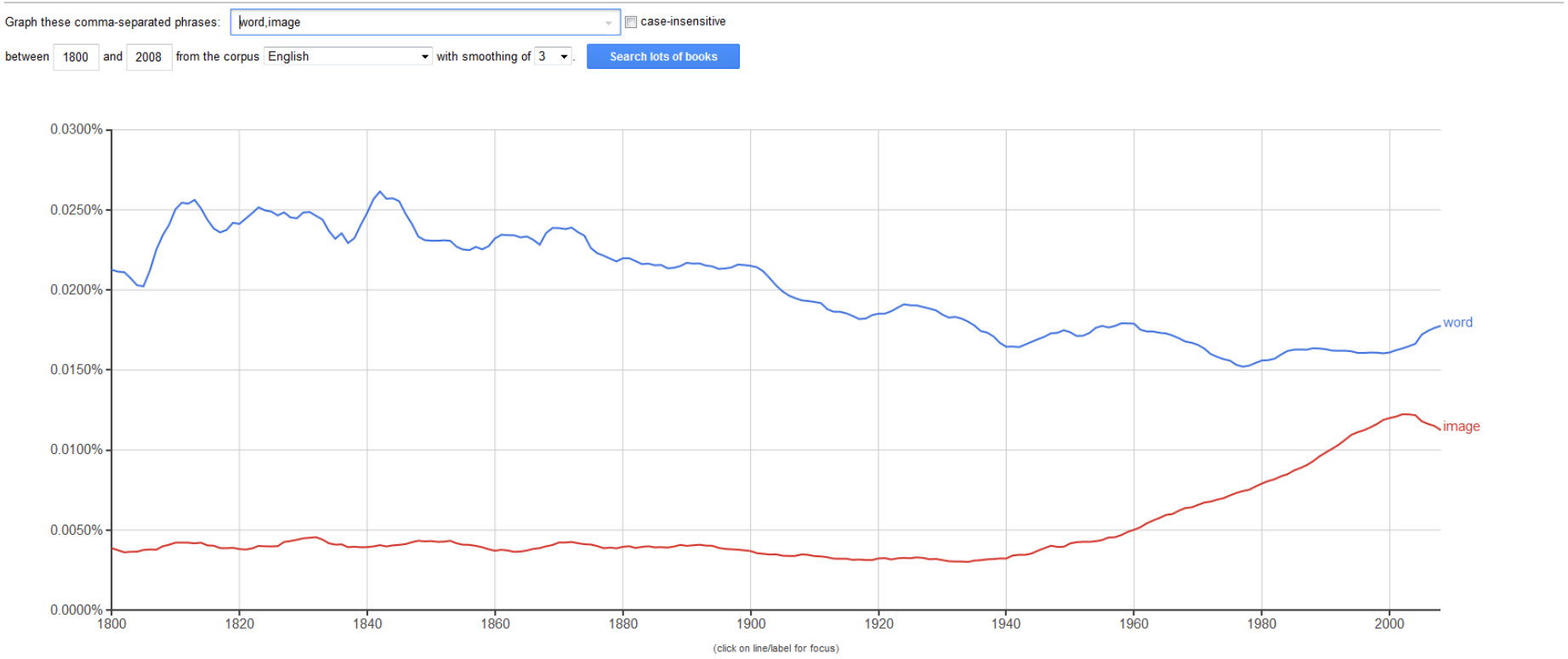
Line Graphs are used to display quantitative values over a continuous interval or time period. A Line Graph is most frequently used to show trends and analyse how the data has changed over time.

Line Graphs are drawn by first plotting data points on a Cartesian coordinate grid, then connecting a line between all of these points. Typically, the y-axis has a quantitative value, while the x-axis is a timescale or a sequence of intervals. Negative values can be displayed below the x-axis.

Anatomy



https://datavizcatalogue.com/methods/line_graph.html



Google Books | Ngram Viewer https://books.google.com/ngrams

Graphs & Trees for Humanities Data

Graphs

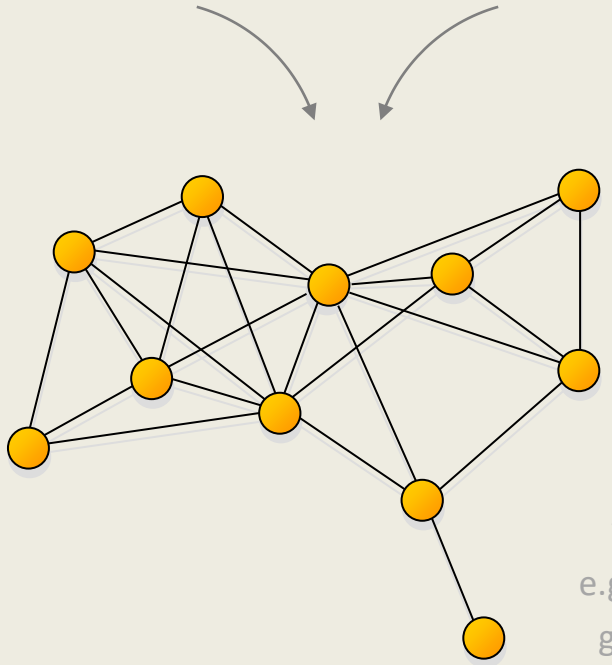
Trees

Interactive Catalogue:
https://datavizcatalogue.com/

nodes

e.g.,
persons
roles
teams
companies
nations
signs
texts
computer
websites
cities
networks
…

A graph (or network) is a sets of **nodes** (vertices) and a set of lines (links, edges) that connect these **nodes**.



visual representation
e.g. with a **force-directed** graph drawing method

links

e.g.,
communications
interactions
collaborations
trade relations
international relations
syntactic relations
citations
fibre optic cables
hyperlinks
streets
edge bundles
…

graphs can be modeled mathematically to calculate and compute various of their properties and structural metrics

(graph theory)

$$C_D(j) = \sum_{j=1}^{n} A_{ij}$$

(degree centrality)

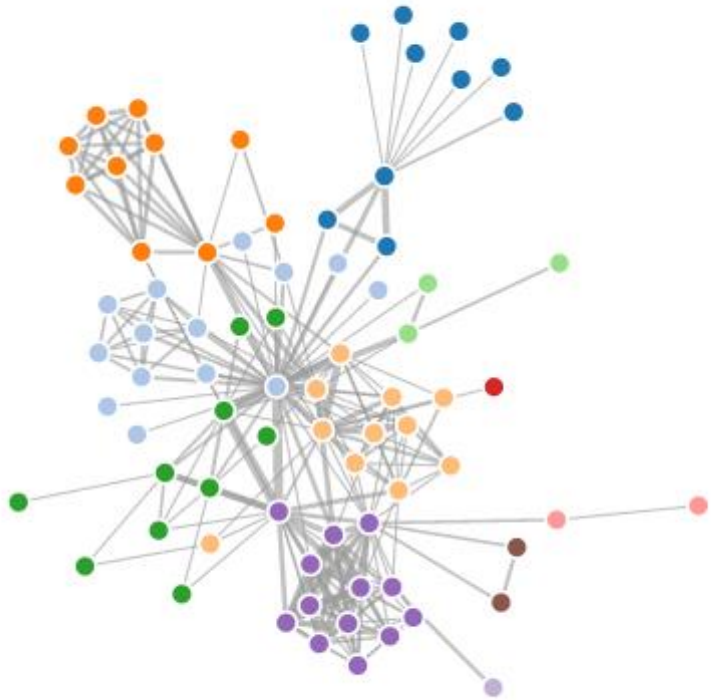
and graphs can be drawn to disclose some of their properties and structural metrics to a visual analyst

(graph drawing)

mesh bus ring star tree

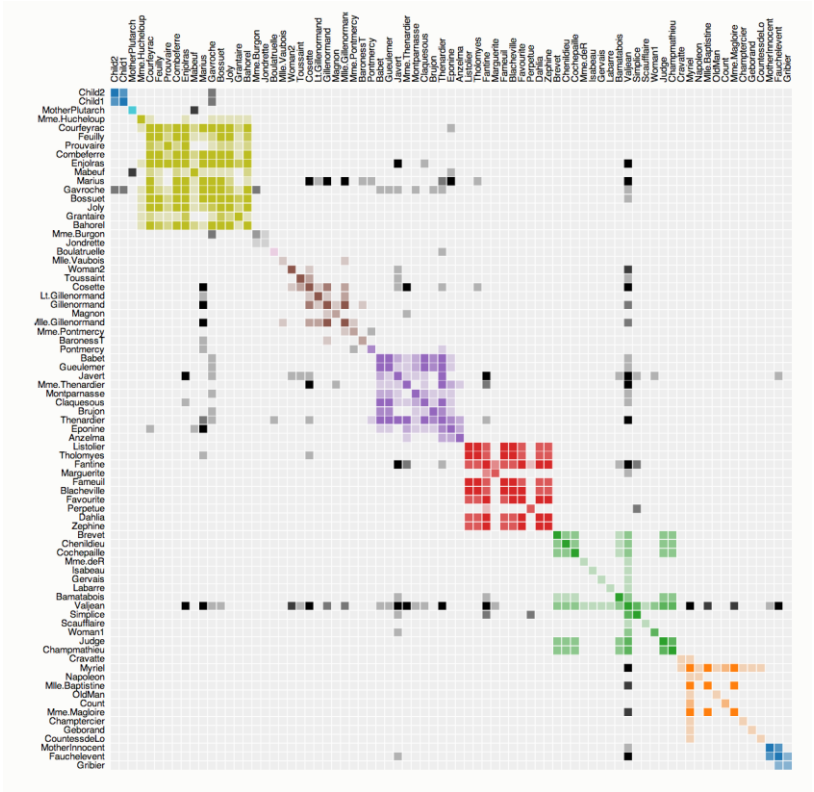
There are multiple ways to visually represent graphs, including

Force-directed Layout



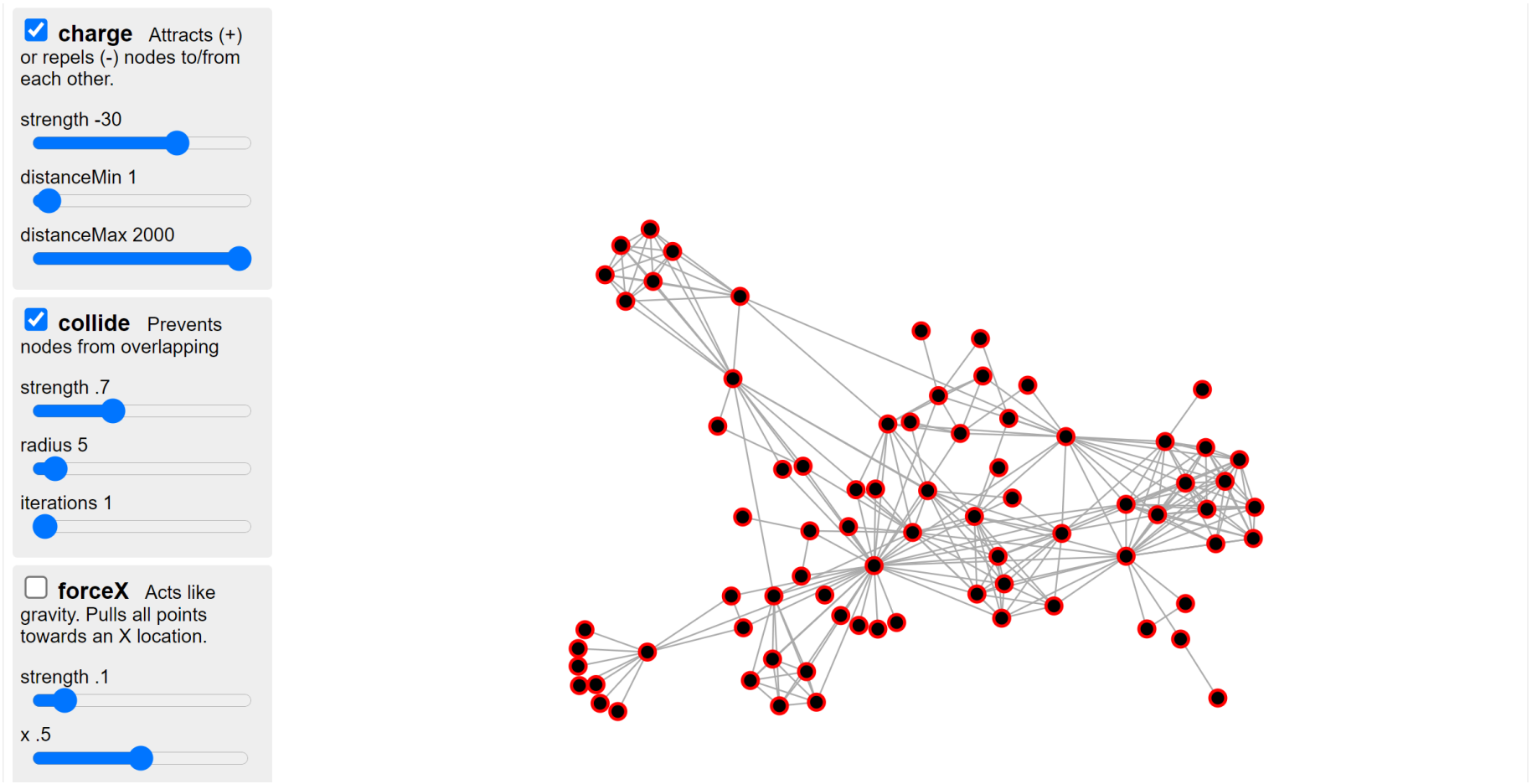
Mike Bostock: http://bl.ocks.org/mbostock/4062045

Adjacency Matrix



Mike Bostock https://bost.ocks.org/mike/miserables/

Interactive Introduction I: Force-directed Layout

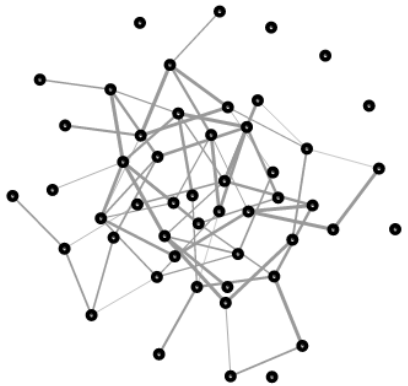


Steve Haroz | D3 Force Testing Ground
https://bl.ocks.org/steveharoz/8c3e2524079a8c440df60c1ab72b5d03

Interactive Introduction II

Gravity 0.03
Attraction (Links) 50 Weight
Charge (Nodes) -60 Size
Size by Nothing Degree Labels
Directed
Delete some edges | Delete some nodes | Add nodes

Canvas gravity draws all nodes toward the center of the canvas, preventing them from flying out of view.



Fork me on GitHub

Introduction to Network Analysis and Representation

Elijah Meeks and Maya Krishnan

Introduction
Models
Random | Dramatic | Transportation
Correspondence | Genealogy | Administration
Layouts
Force-Directed | Plot | Hierarchical
Metrics
Centrality | Clustering Coefficient
Community Detection
Exploration
Pathfinding | Ego Network | Spatial Problem

Random Graph

Random graphs are used as control sets to compare variation with the studied network. Random isn't the best word for these networks, as they can be generated by very specific and complicated rules, so as better to model the phenomena being studied.

If plotted, each node will be placed according to random xy values generated when the network is first created.

This graph is created with 50 nodes and a 2.5% for each node to be connected to each node (in both directions, so a 5% chance total if treated as an undirected graph). You can create a new random graph with different settings:

Nodes: 50 Link Chance .025 New Graph

Citations

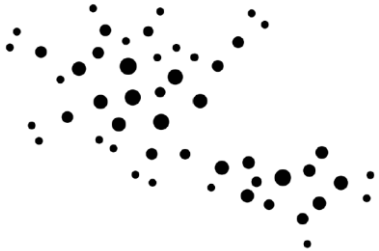
Albert, R., and Barabasi, A-L. (2002). Statistical mechanics of complex networks. Reviews of modern physics 74 (January): 47-97.
...

Code

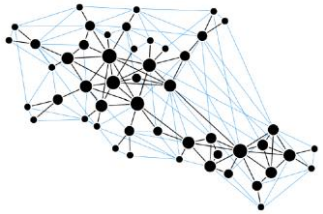
```
Some code that produces a random graph, with random links and random spatial
characteristics.
```

Q: What can we **see** when we look at force-directed graphs?

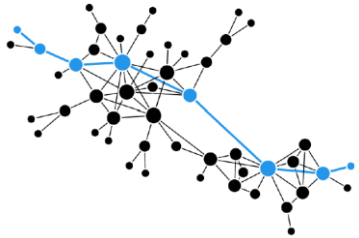
A: Various **visual patterns**, disclosing information about the structure of relational phenomena



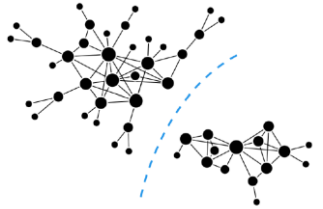
graph size
number of nodes and edges in total



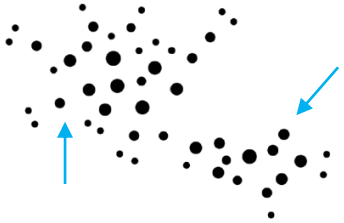
graph density
overall connectivity



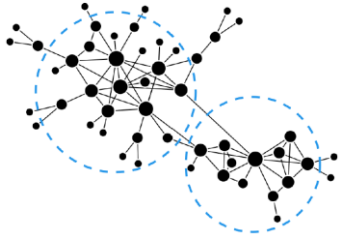
diameter
distance of most distant nodes



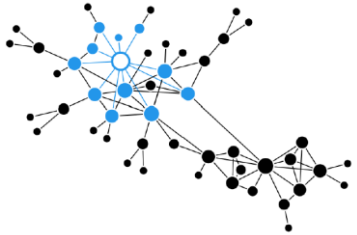
connectedness
onve vs multiple graph components



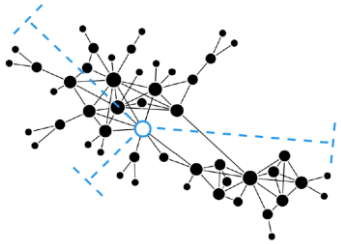
clusters & cliques
groups of more densely connected nodes



node positions
individual positions & collective positions



degree centrality
connections of a specific node



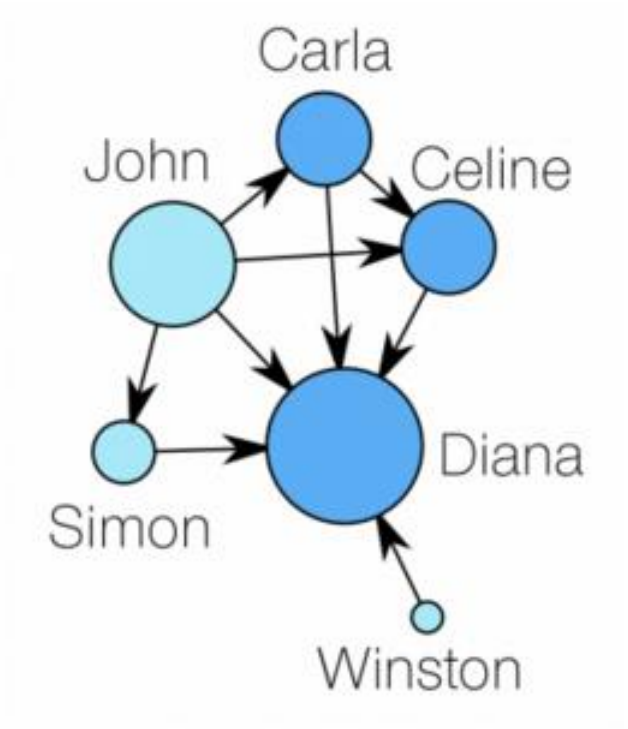
closeness centrality
being in the middle of the network

Grandjean & Jacomy, 2019 – Translating Networks
https://halshs.archives-ouvertes.fr/halshs-02179024/document (key reading)

Network Data

Nodes	Edges
Id,Label,Attribute	Source,Target
1,John,1	1,2
2,Carla,2	1,3
3,Simon,1	1,4
4,Celine,2	1,6
5,Winston,1	2,4
6,Diana,2	2,6
	3,6
	4,6
	5,6

Network Visualization



Martin Grandjean | Introduction to Network Visualization with GEPHI
http://www.martingrandjean.ch/introduction-to-network-visualization-gephi/

Movie Galaxies http://moviegalaxies.com/

BABEL

2006

TRAFFIC

2000

TWIN PEAKS

1990

THE BIG LEBOWSKI

1998

THE GODFATHER

1972

2001: A SPACE ODYSSEY

1968

ANNIE HALL

1977

THE MATRIX

1999

PRETTY WOMAN

1990

THE EVIL DEAD

1981

Spotify Artist Network

beta software!

This website allows for exploring networks of related artists on Spotify. It uses sigma.js for graph visualization and chroma.js for colors.

First search for an artist via the text field below. Click on an artist to load the network. Loading the network may take over a minute on the first run.

Double click to zoom into the network and click on a node for details. Node size and color indicate popularity.

I make other stuff as well.

download network as gdf file (for visualization in gephi)

Tosca

go

Wax Tailor

open with spotify

popularity: 57

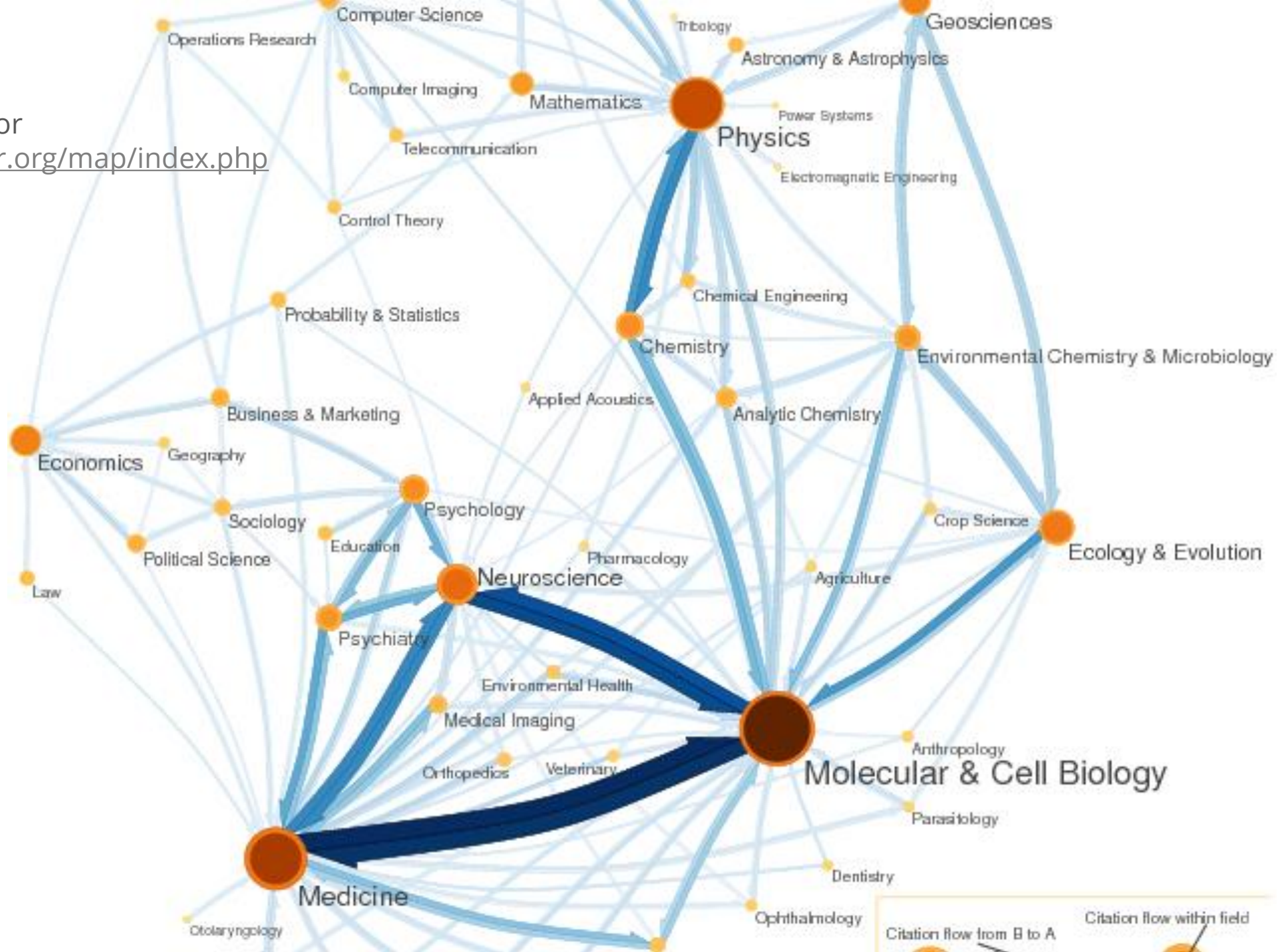
followers: 94046

Bernhard Rieder | Spotify Artist Network

http://labs.polsys.net/playground/spotify/

Smoke City
Sofa Surfers
Peace Orchestra
Esthero
Kruder & Dorfmeister
Bitter Sweet
Nitin Sawhney
Terranova
Kid Loco
Zero 7
Martina Topley-Bird
Morcheeba
Hooverphonic
Lamb
Sneaker Pimps
Télépopmusik
Archive
Portishead
I Monster
Lemon Jelly
Air
Massive Attack
Nightmares On Wax
Moloko
Tricky
The Cinematic
FC Kahuna
DJ Cam
Röyksopp
UNKLE
Groove Armada
The Future Sound Of London
Red Snapper
Death In Vegas
Aim
Bonobo
Underworld
Hybrid
Fluke
Orbital
Leftfield
DJ Krush
The Chemical Brothers
Mint Royale
Timo Maas
Propellerheads
The Avalanches
Quantic
Xploding Plastix
Gramatik
Funki Porcini
The Herbaliser

Jevin West | Eigenfactor
http://www.eigenfactor.org/map/index.php

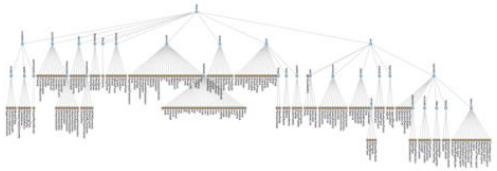


Trees (Directed, acyclic graphs)

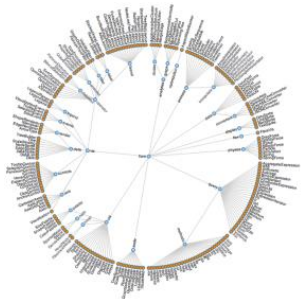
Multiple Options to visualize tree-like, hierarchical data

node link diagrams

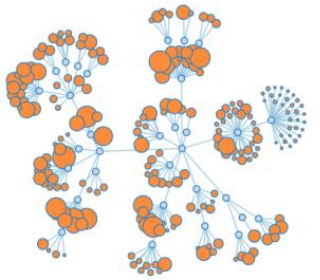
intended tree



cluster dendrogram



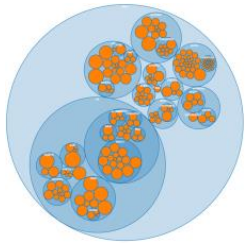
circular dendrogram



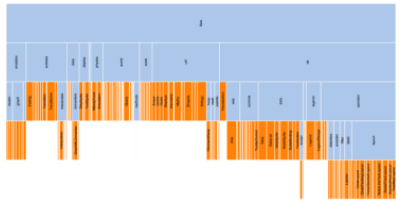
collapsible force

adjacency diagrams

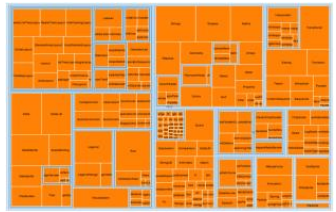
sunburst



circle packing



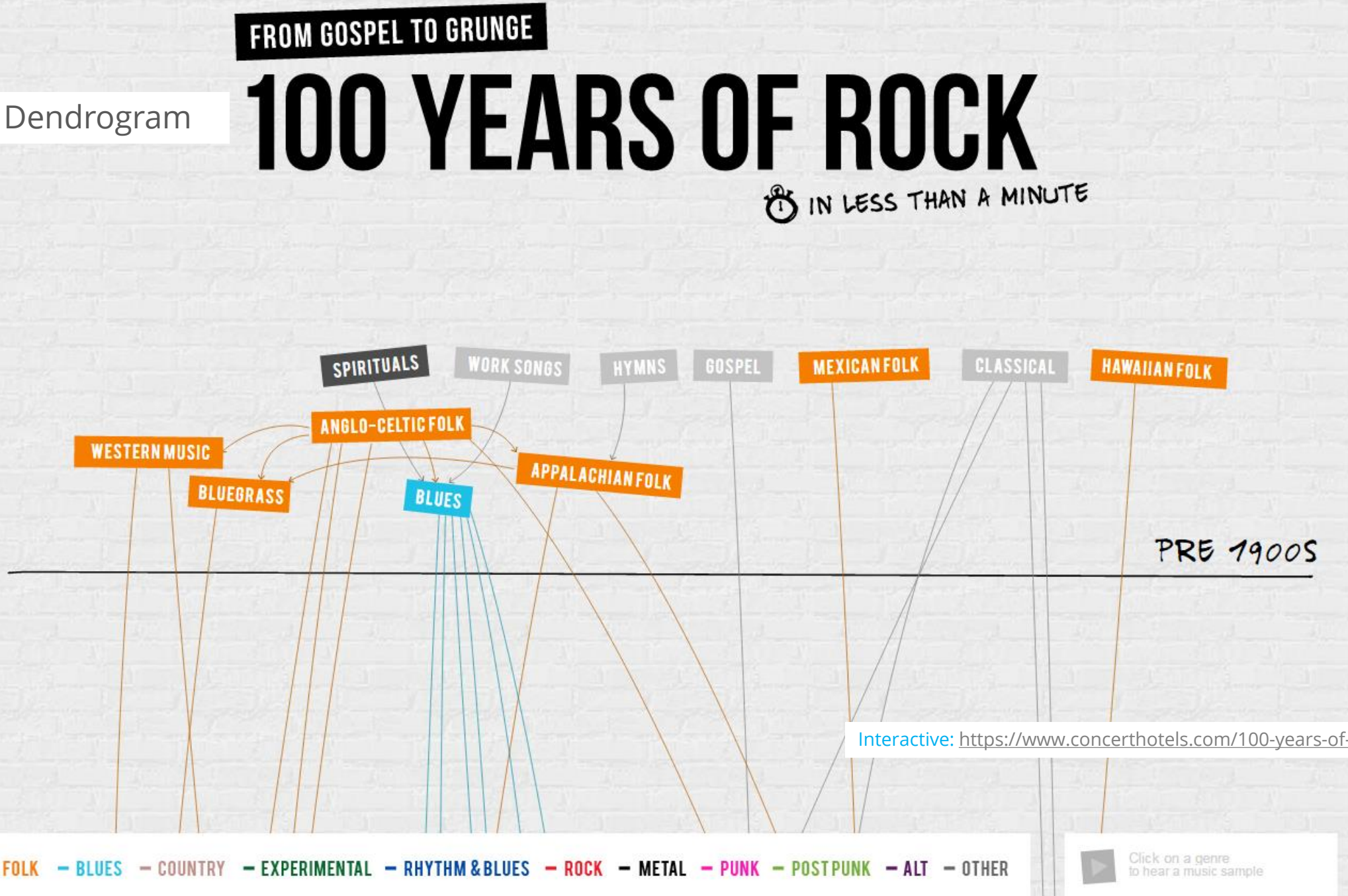
icicle layout



treemap

(D3.js, 2014)

Example Dendrogram



Interactive: https://www.concerthotels.com/100-years-of-rock/

Example: Treemap

PANTHEON
MAPPING HISTORICAL CULTURAL PRODUCTION

METHODS VISION TEAM

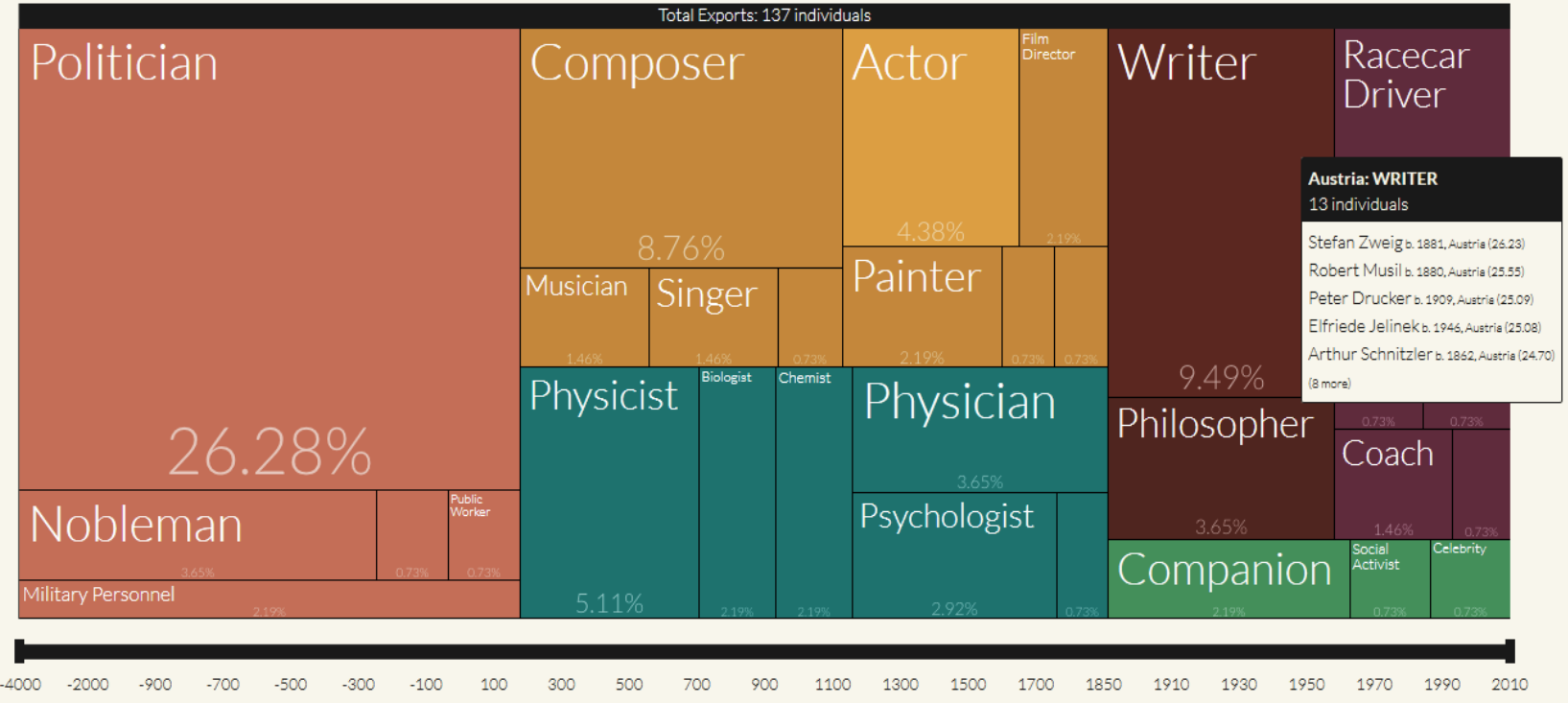
Who are the globally known people born within present day **Austria***?
[4000 B.C. – 2010]

VISUALIZATIONS

TREEMAPS
By Place of Birth
By Domain
MATRICES
SCATTERPLOTS
MAPS

PARAMETERS

PLACE OF BIRTH*
Austria
FROM 4000 B.C.
TO 2010
DATA
Pantheon
INDEX
HPI | L
15



Austria: WRITER
13 individuals

Stefan Zweig b. 1881, Austria (26.23)
Robert Musil b. 1880, Austria (25.55)
Peter Drucker b. 1909, Austria (25.09)
Elfriede Jelinek b. 1946, Austria (25.08)
Arthur Schnitzler b. 1862, Austria (24.70)
(8 more)

RANKINGS

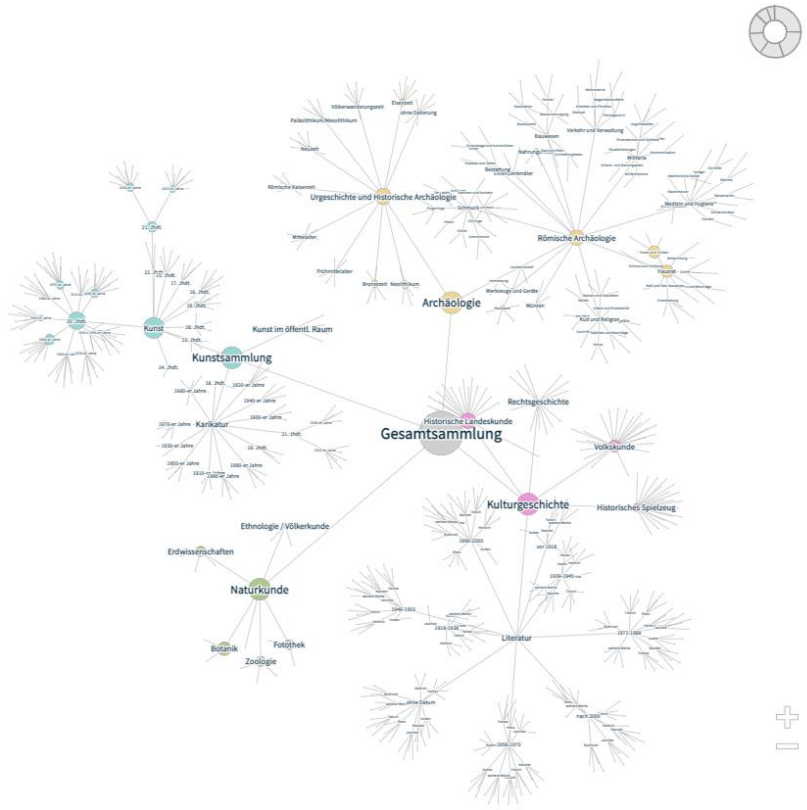
1. Adolf Hitler
POLITICIAN, b. 1889 (30.58)
2. Wolfgang Amadeus Mozart
COMPOSER, b. 1756 (30.51)
3. Marie Antoinette
NOBLEMAN, b. 1755 (28.45)
4. Joseph Haydn
COMPOSER, b. 1732 (28.08)
5. Ludwig Wittgenstein
PHILOSOPHER, b. 1889 (27.52)
6. Karl Popper
PHILOSOPHER, b. 1902 (27.20)
7. Franz Schubert
COMPOSER, b. 1797 (27.14)
8. Maria Theresa of Austria
POLITICIAN, b. 1717 (27.11)
9. Franz Liszt
MUSICIAN, b. 1811 (27.07)
10. Theodoric the Great
POLITICIAN, b. 454 (26.99)

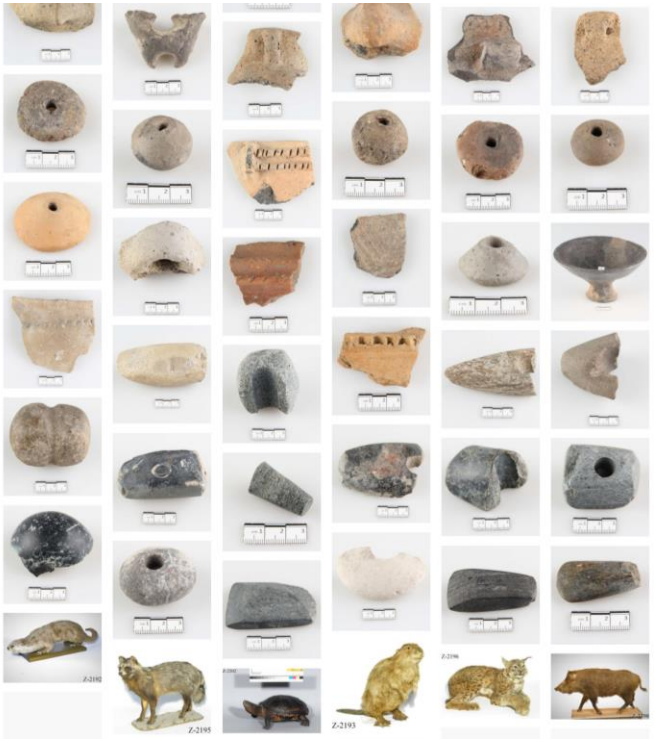
Go to Full Ranking List

MIT Media Lab | PANTHEON - Interactive: https://pantheon.world/explore/viz

Sunburst Diagram

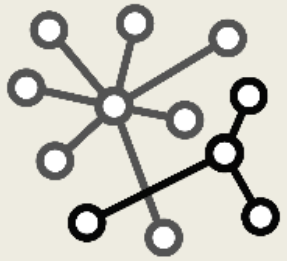
Dendrogram





Smuc et al. | NÖ Landessammlungen auf einen Blick
https://https://visualisierung.landessammlungen-noe.at/

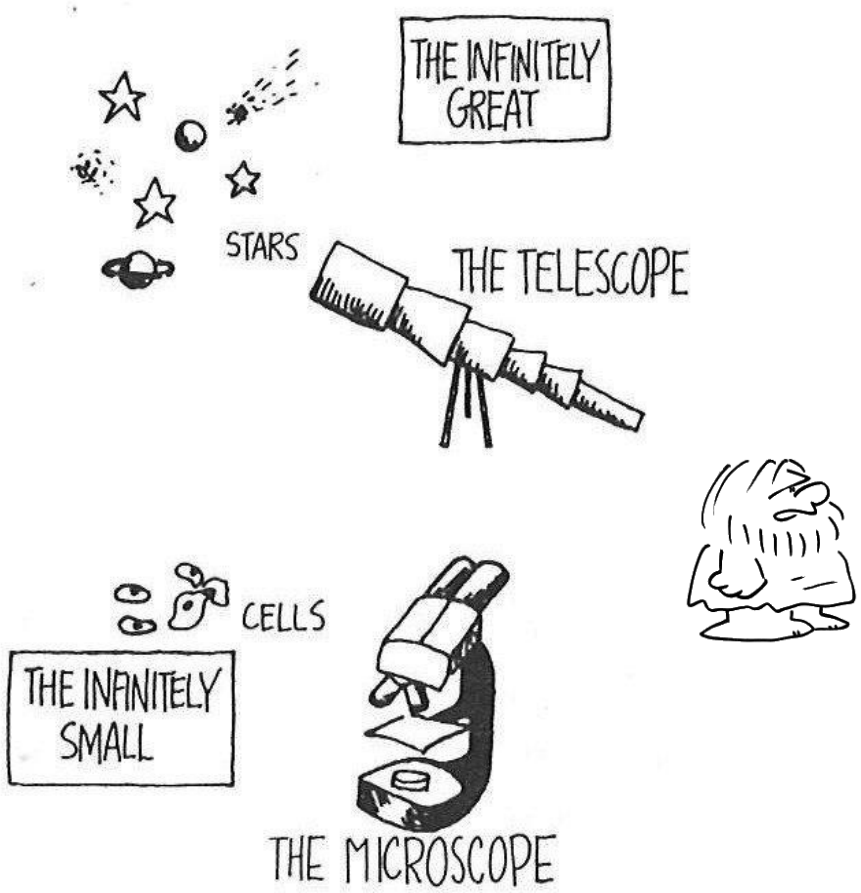
Tools for Visualizing Graphs & Trees:

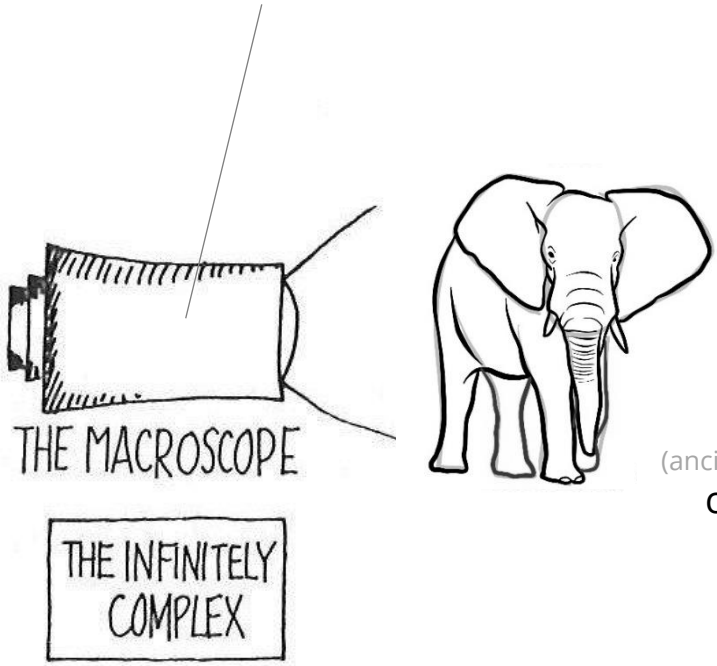


- Gephi (https://gephi.org/)
- Visone (https://visone.info/)
- Palladio (https://hdlab.stanford.edu/palladio/)
- Excel, Tableau, D3.js, etc.

How to create „bigger pictures“ ?

Visualizations as „Macroscopes“ (DeRosnay, 1979)





(ancient metaphor for)
complex things

Stefaner, 2014: http://well-formed-data.net/archives/1027/worlds-not-stories

Oftentimes, one visualization is not enough!

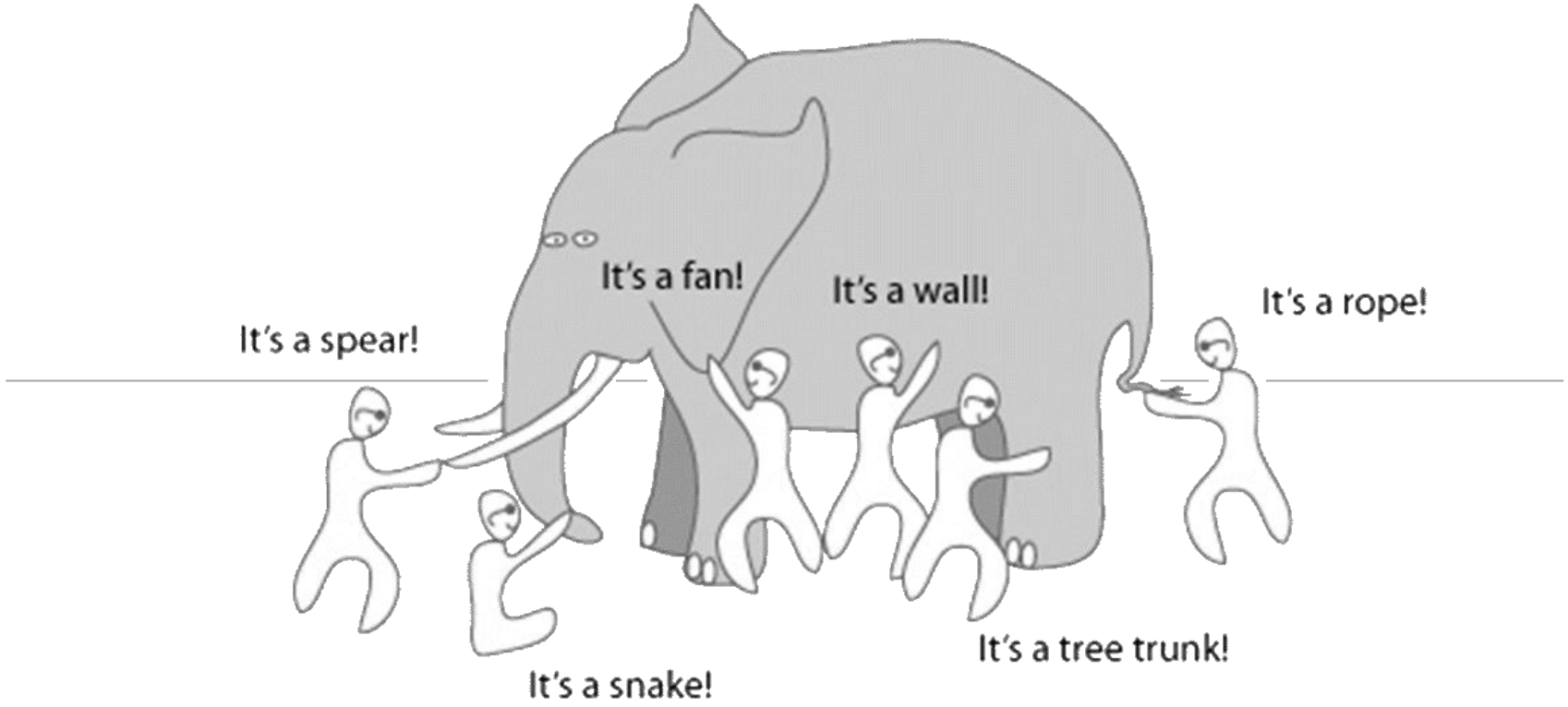
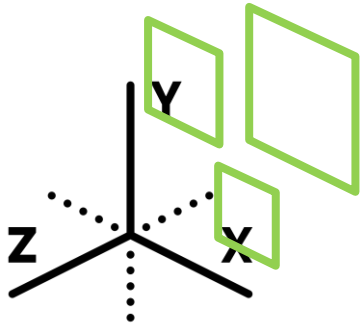


Image: http://www.sexandquantumphysics.com/2011/07/paradigms-shift/

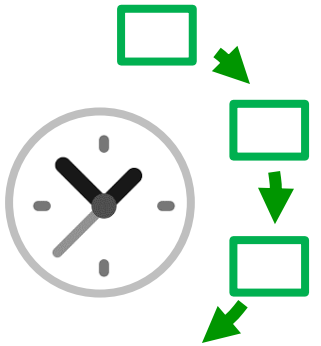
Techniques to create compound visualizations as bigger pictures of complex things

utilizing space



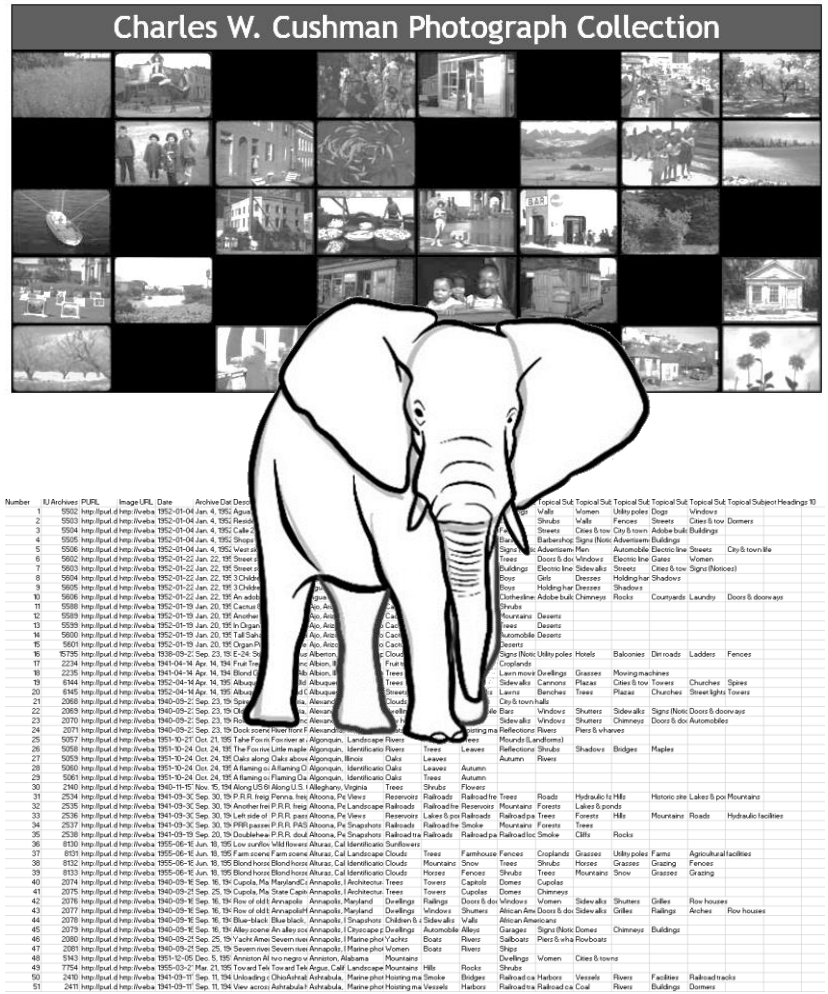
- Coordinated Multiple Views (Roberts, 2007, link)
- Synoptic Encoding (Mayr et al., 2018, link)
- Rich Pictures (Monk & Howard, 1998, link)
- Data Comics (Bach et al., 2017, link)
- Polycubism (Windhager et al., 2020, link)

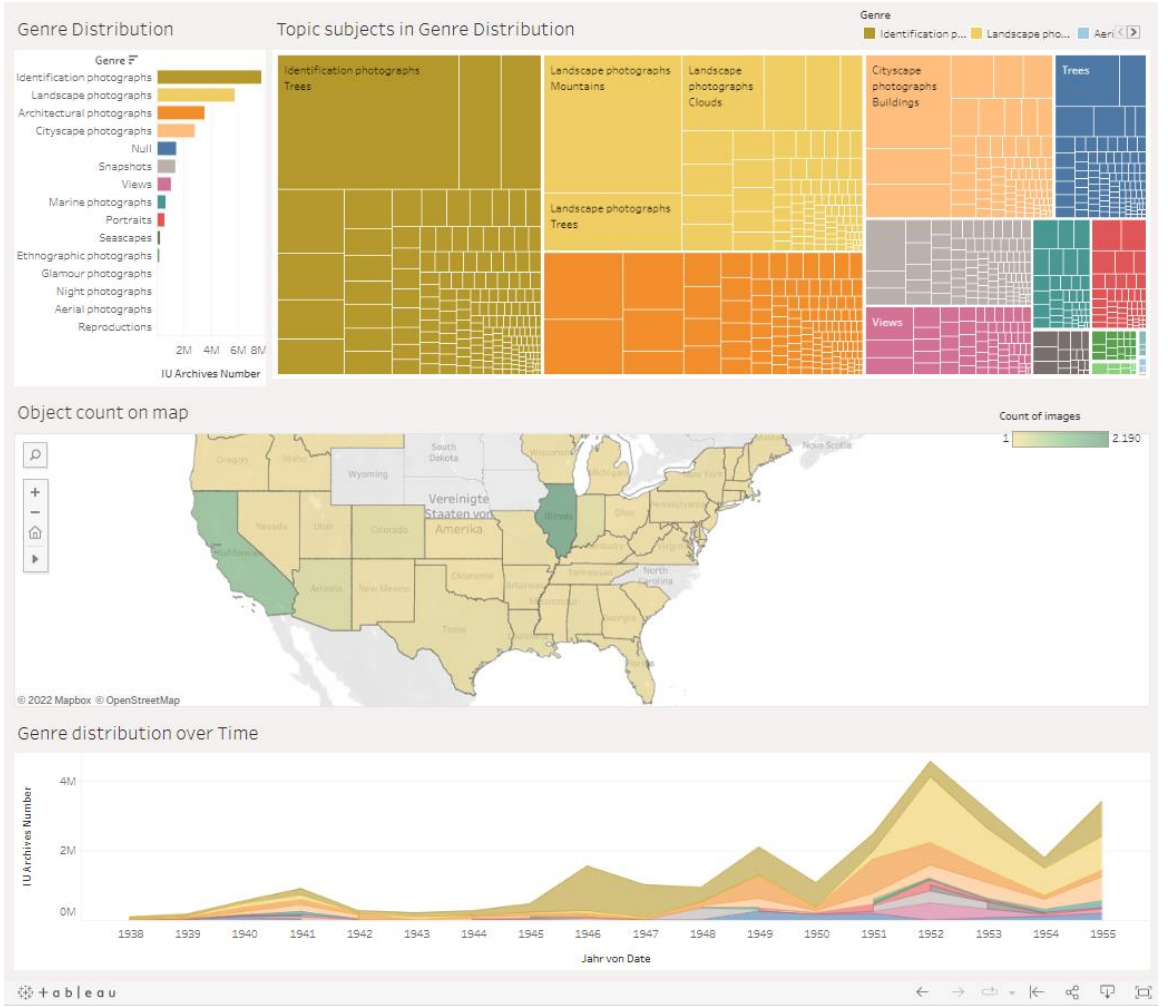
utilizing time



- Narrative Visualization (Segel & Heer, 2010, link)
- Sequential Presentation (standard technique)
- Film / Animation (standard technique)
- Interaction (standard technique)
- Animated transitions (Heer & Robertson, 2007, link)

The case of cultural collections

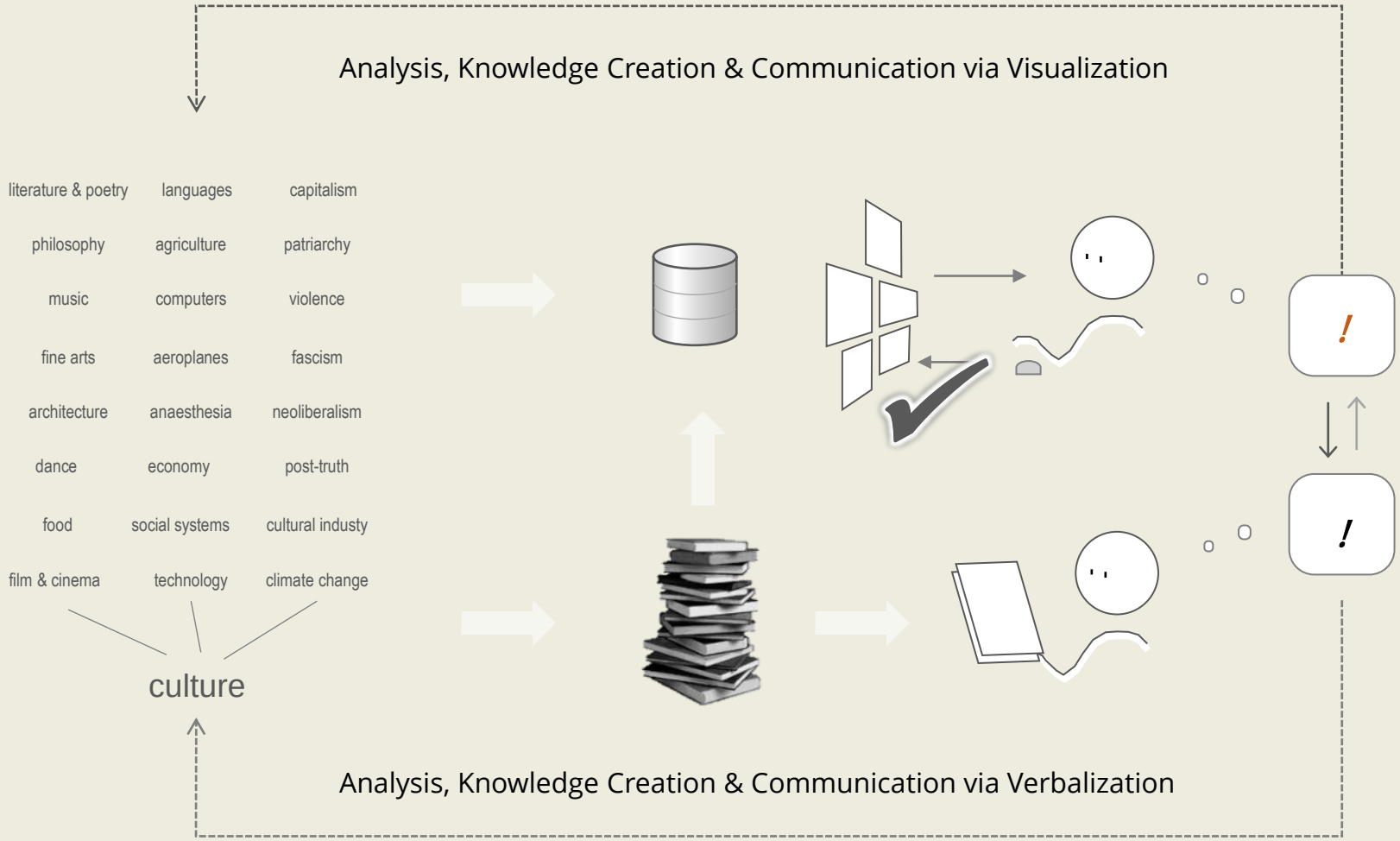




Closing the circle:

By the means of visualization techniques, we can utilize computers and our highly evolved perception to see complex cultural things anew.

As a new cultural and scientific practice, multiple / faceted views provide us with a plurality of visual perspectives to augment our language-based knowledge and access to our natural and cultural environments.



Thank you!