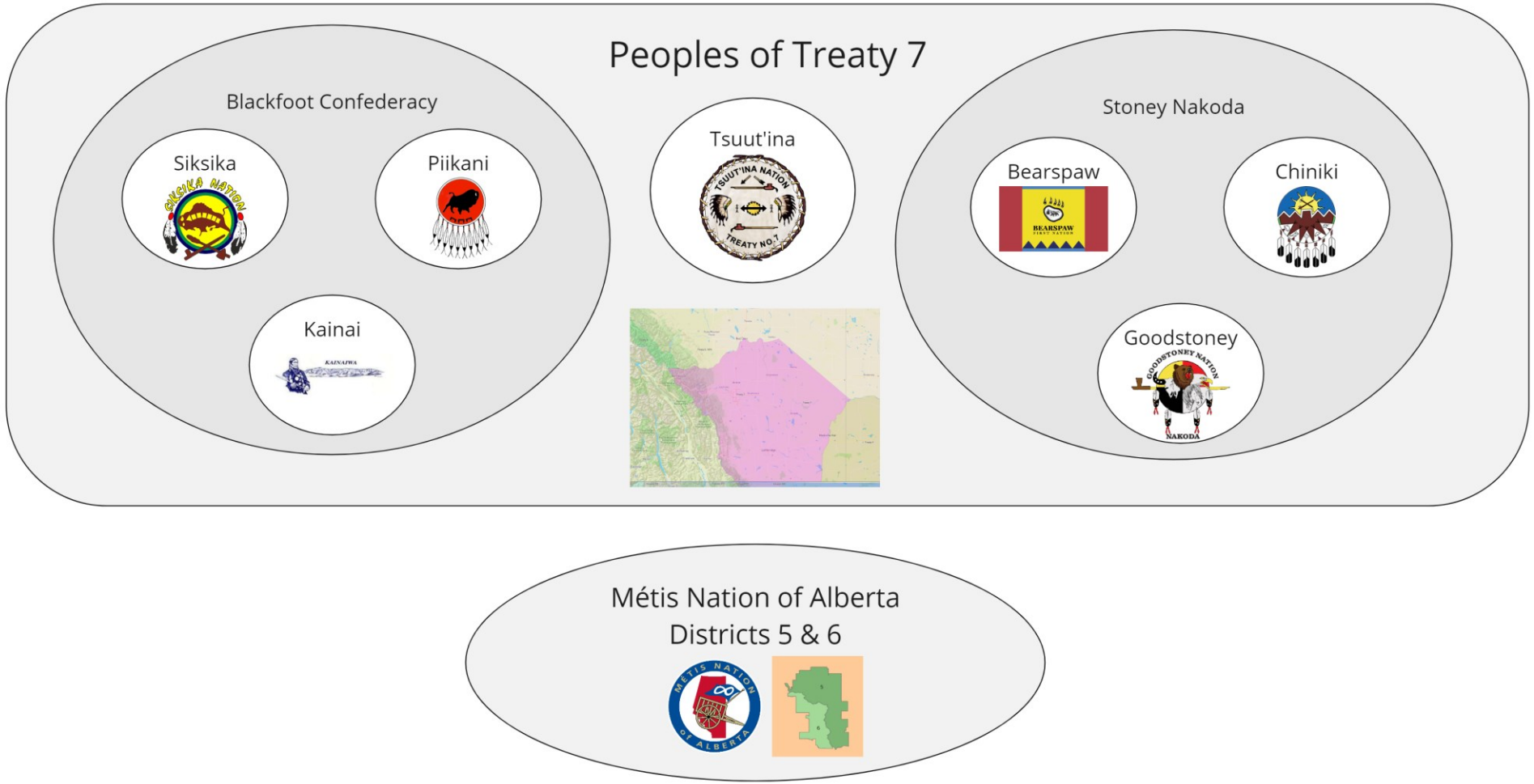
January 10, 2025

Visualization for Qualitative Data

John Brosz, PhD
Data and Visualization Curator
Libraries & Cultural Resources

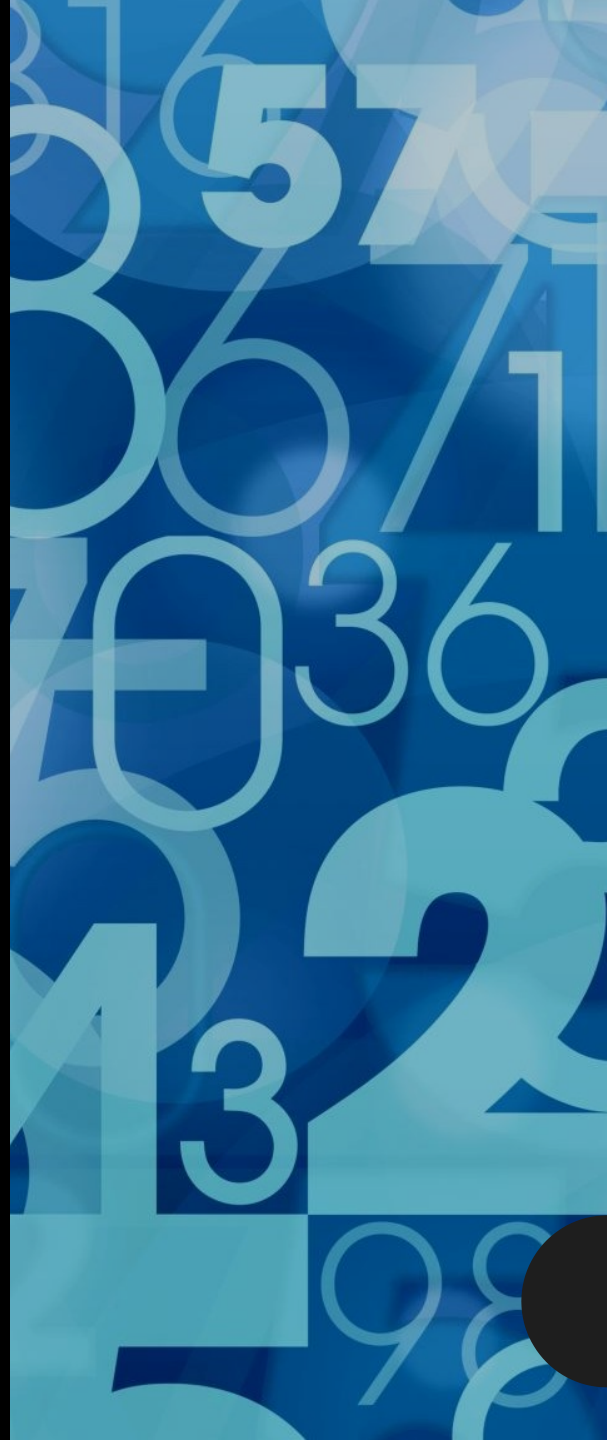
SLIDES: BROSZ.CA/SLIDES



The University of Calgary, located in the heart of Southern Alberta, both acknowledges and pays tribute to the traditional territories of the peoples of Treaty 7, which includes the Blackfoot Confederacy (comprised of the Siksika, the Piikani, and the Kainai First Nations), the Tsuut’ina First Nation, and the Stoney Nakoda (including the Chiniki, Bearspaw, and Goodstoney First Nations). The City of Calgary is also home to the Métis Nation of Alberta (Districts 5 and 6).

I want to visualize my data but it is not made up of numbers

Qualitative Visualization



SLIDES: BROSZ.CA/SLIDES

Topics

Qualitative Data

Visualization

Techniques

- Coloured Text/Heatmaps
- Hybrid Qual/Quant
- Word Clouds
- Sets – Venn & Upset
- Timelines
- Graphs/Networks, Trees

Qualitative Data

Data that is non-numerical in nature. Collected through observations, interviews, focus groups, written records & notes, and similar methods.

Interviews

Observations

Narratives & stories

Documents

Diaries

Photographs or other Audio & Video recordings

Processes / Relationships

Qualitative Data

What's Different?

Not as simple as dumping data into software

Framing is critical

Impact of stories versus numbers

Finding Qualitative Datasets

UK Data Service

https://beta.ukdataservice.ac.uk/datacatalogue

Offers a Data Type filter for Qualitative and Mixed Methods data

The Syracuse Qualitative Data Repository

https://data.qdr.syr.edu/

Kaggle - https://www.kaggle.com/

Data science focus, but has some qualitative data as well
E.g., Kentucky animal bites dataset
https://www.kaggle.com/datasets/rtatman/animal-bites

Sage Research Methods

https://methods.sagepub.com/datasets

Interview transcripts, drawings, images, letters, log files, diary entries, case notes, audio, reflexive notes visual art pieces

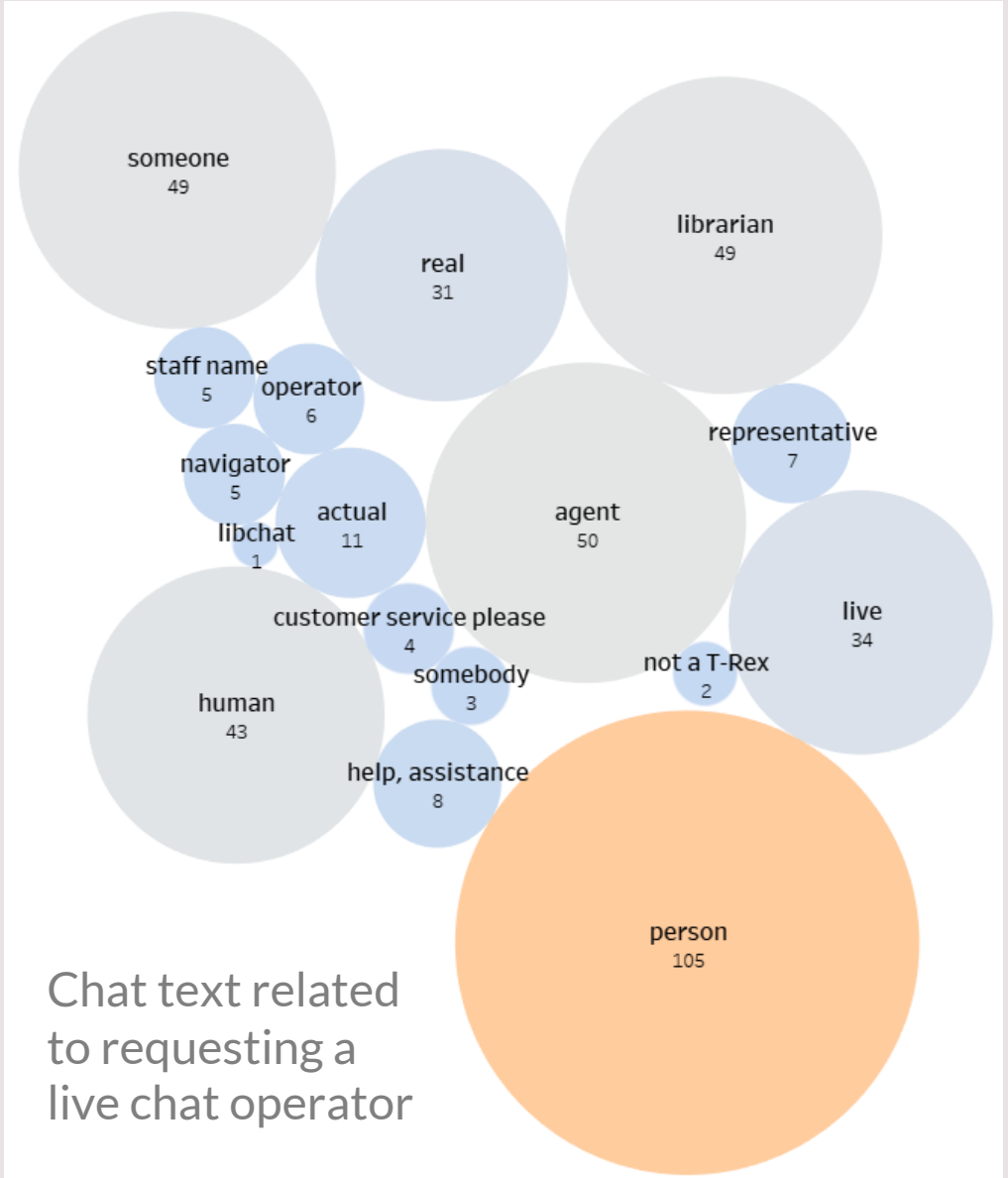
Data Visualization & Qualitative Data

Two approaches

Techniques designed for qualitative data	Count things (codes, groups, categories) and use quantitative visualization techniques.

Techniques designed for qualitative data

Count & use quantitative techniques



Chat text related to requesting a live chat operator

Data Visualization

Definitions

Visual representation of data

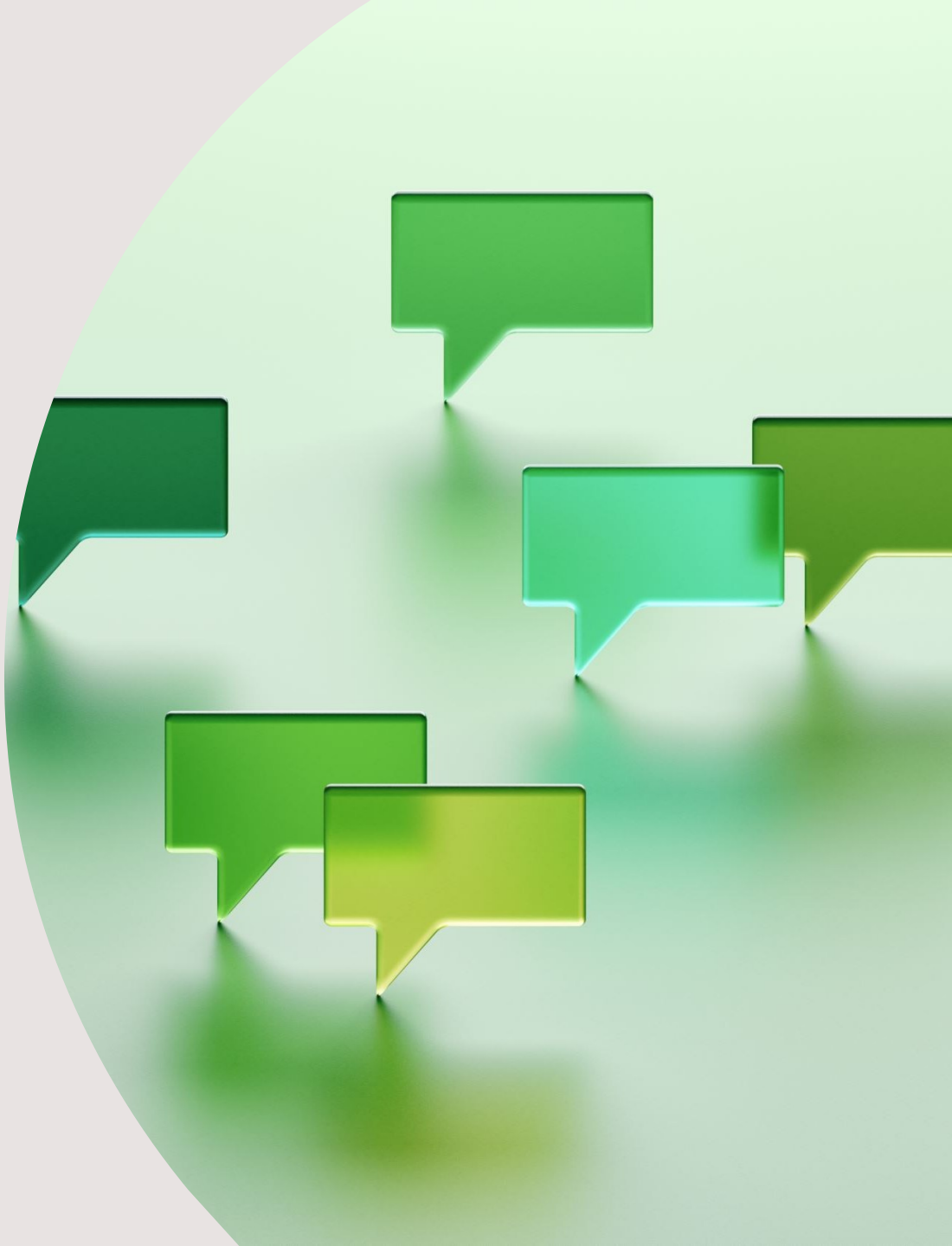
"Transformation of the symbolic into the geometric" [McCormick et al, 1987]

"... artificial memory that best supports our natural means of perception" [Bertin, 1967]

"Use of computer-generated, interactive, visual representations of data to amplify cognition" [Card, Mackinlay, & Shneidermann, 1999]

Why do we visualize data?

- Amplify cognition



Amplify Cognition

Paradoxes – Lewis Carroll

Paradoxes – Epimenides

Paradoxes – Self-ref

Paradoxes – Infinity

Epimenides – Self-ref

Epimenides – Tarski

Halting Problem – Decision Procedure

Halting Problem – Turing

Infinity – Halting problem

Infinity – Recursion

Infinity – Zeno

Infinity – Lewis Carroll

Zeno – Lewis Carroll

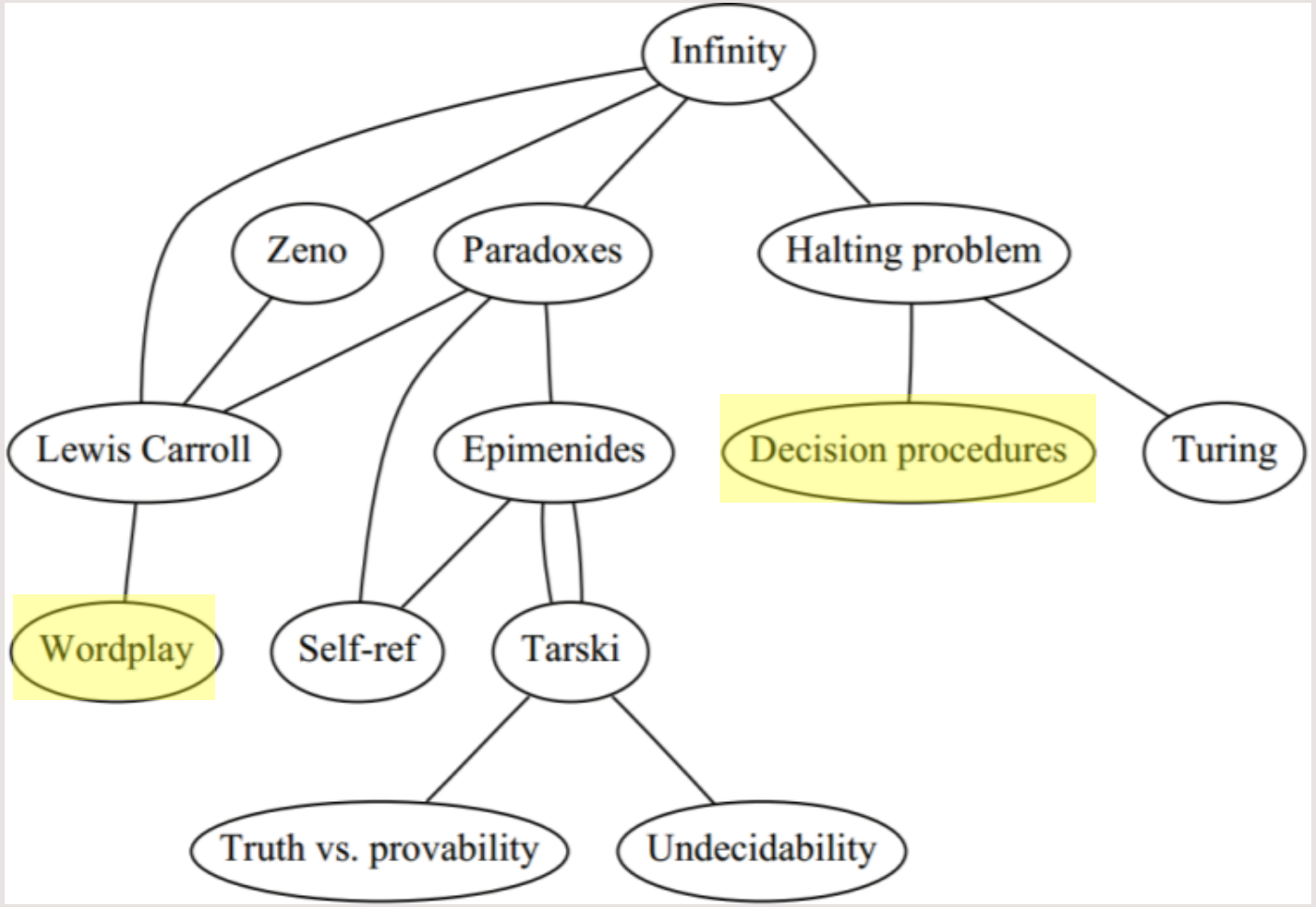
Lewis Carroll – Wordplay

Tarski – Epimenides

Tarski – Truth vs Provability

Tarski - Undecidability

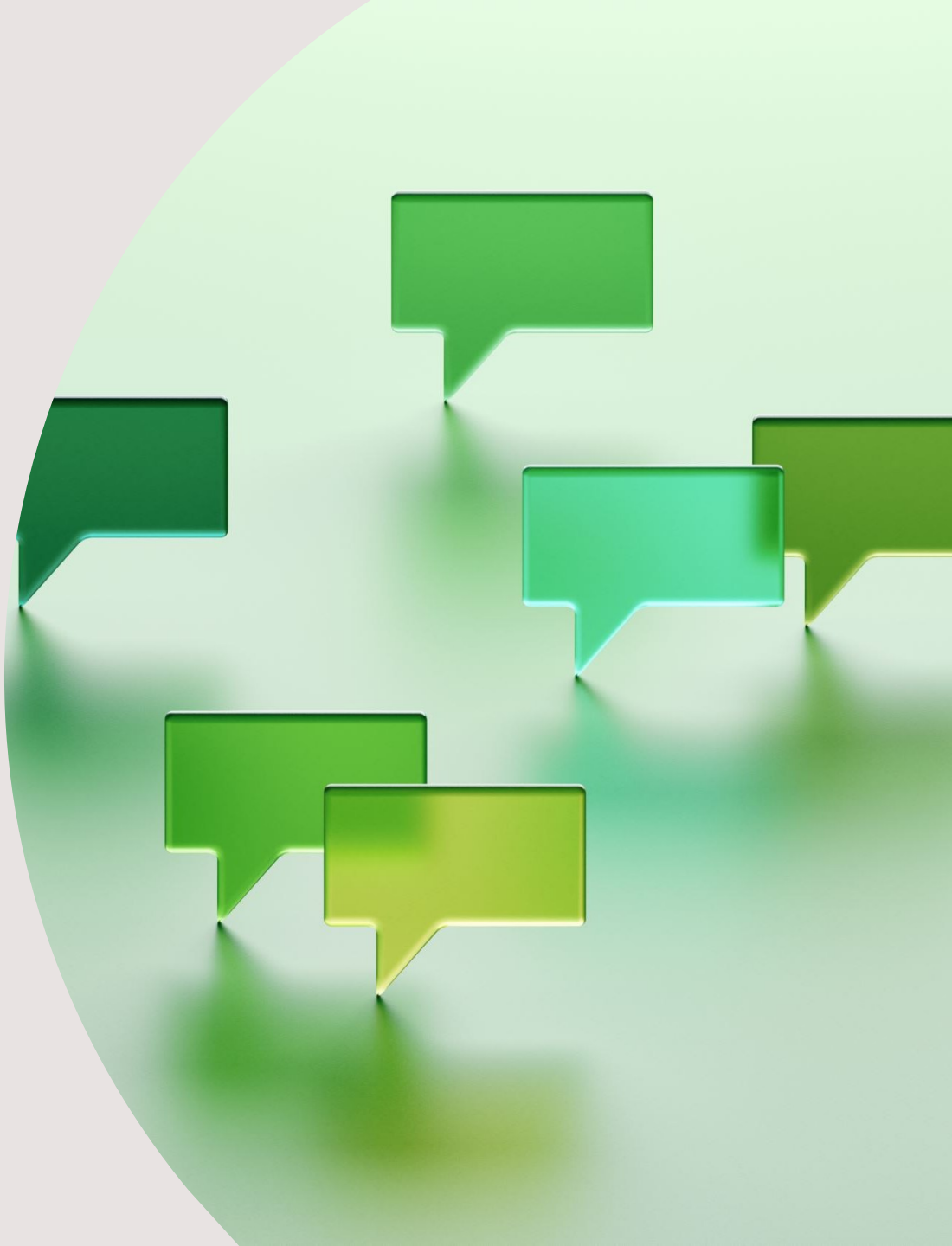
Amplify Cognition



Why do we visualize data?

- Amplify cognition
- Sometimes just break up text
- Different outlook on your data

No better way to find data problems



Visualization

Building Blocks

Encoding

Marks: Colour, Lightness, Size, Position, Orientation, Shape, Motion, Texture

Relationships between marks: Connections & Containment

Purpose, Audience, Medium, Goal/Story, Data Properties

Visual Cortex vs Cerebral Cortex

Purpose

Exploratory

Increase your understanding

Find patterns

Reflect your way of thinking

Explanatory

Capture attention

Increase speed of understanding

Maximize memorability

Stimulate action

Charts & Techniques

Text Extracts & Colours

Mixing Quantitative & Qualitative

Word Clouds

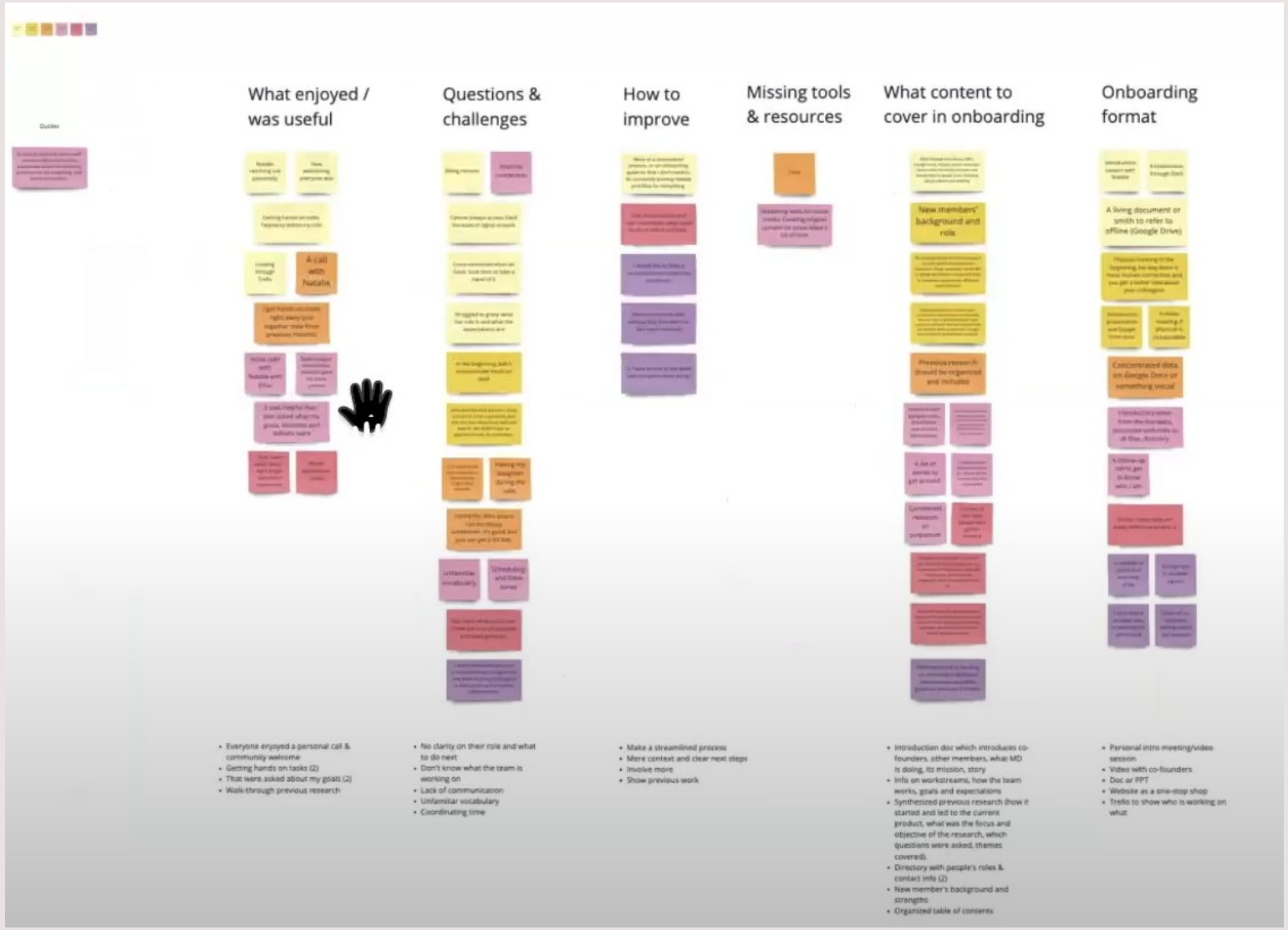
Venn Diagrams

Timelines

Networks & Graphs

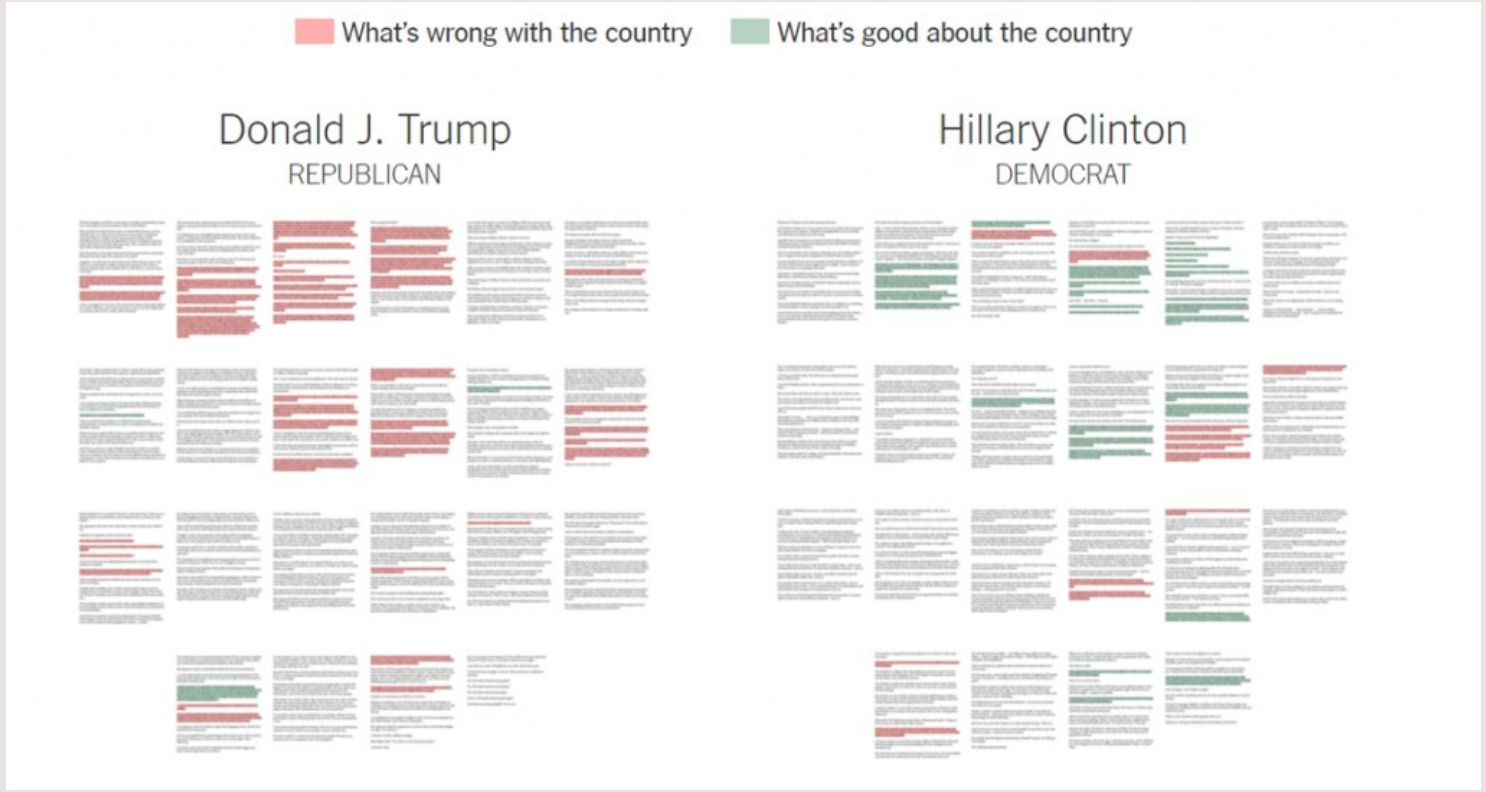
Trees & Hierarchies

Extracted, Grouped, and Coloured Text



https://miro.com/index/

Mural is another useful, similar app https://www.mural.co/



"Our convention occurs at a moment of crisis for our nation. The attacks on our police, and the terrorism in our cities, threaten our very way of life."

"How great are our police? And how great is Cleveland?"

"The irresponsible rhetoric of our president, who has used the pulpit of the presidency to divide us by race and color, has made America a more dangerous environment than frankly I have ever seen and anybody in this room has ever watched or seen."

"There's a lot of work to do. Too many people haven't had a pay raise since the crash. There's too much inequality. Too little social mobility. Too much paralysis in Washington. Too many threats at home and abroad."

"We have the most dynamic and diverse people in the world. We have the most tolerant and generous young people we've ever had."

"So don't let anyone tell you that our country is weak. We're not."

https://www.nytimes.com/interactive/2016/07/29/us/elections/trump-clinton-pence-kaine-speeches.html

Every time Ford and Kavanaugh answered the question — and didn't answer the question

Click on any part of the transcript to expand

Answers question Does not answer question

Ford

Kavanaugh

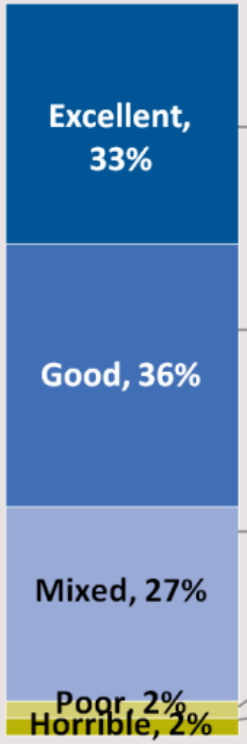
https://www.vox.com/policy-and-politics/2018/9/28/17914308/kavanaugh-ford-question-dodge-hearing-chart

How Willing Are Researchers to Adopt AI in Publishing?					
	Rarely	Unlikely	Fairly	Most Likely	Certain
Life Sciences	Low	Low	Low	Medium	Medium
Chemical Sciences	Low	Low	Low	Medium	Medium
Humanities	Low	Low	Medium	Medium	Medium
Computer Sciences	Low	Low	Medium	Medium	Extreme
Engineering	Low	Medium	Medium	Extreme	Extreme
Artificial Intelligence	Low	Medium	Extreme	Extreme	Extreme

From https://www.enago.com/academy/data-visualization-techniques-qualitative-research/

Mix Quantitative & Qualitative

17. If your organization worked with an external evaluator in 2011, what was your experience?



18/19. Why do you think the experience was/wasn't positive?

"[The external evaluator] was very involved, great listener, knowledgeable about evaluations, very willing to change and edit evaluation tool so that it worked for our programs. She met with us whenever necessary."

"Our external evaluators increased our capacity to do in-depth case studies on strategies and provided a valuable independent perspective."

"While we knew we needed a 'makeover' and major staff changes, it was stressful to open up to outsiders and admit the need."

"Backed out because of the amount of work involved as compared to the payment."

"Need more $ to pay them."

BEFORE

When responding to the question "What could have been improved?", customers' top 3 responses were complaints with A/V (poor video quality), the presenter's visuals and an unclear agenda.

Complaints with A/V quality:

- "The main challenge was logging into the webinar."
- "The AV quality is was difficult."
- "Sometimes the presenter's voice was very garbled for me. Very distracting to the flow and when asking a question, I didn't know if I was heard/understood."

The presenter's visuals were lacking:

- "Technical issues impacted my ability to see the presenter's visuals."
- "Many times the PPT froze or didn't work."

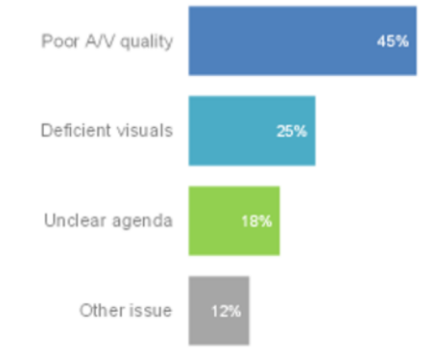
The agenda was unclear:

- "The webinar started late with a lot of time wasted."
- "Bad pre-sharing of information, no time keeper, no clear goal."
- "The agenda seemed very fluid and I was confused around the main point."

AFTER

Customers experienced issues with webinar

Top issues with webinar
% OF COMPLAINTS (n=900)



Sample verbatim comments

The main challenge was logging into the webinar | AV quality was difficult | Sometimes the presenter's voice was very garbled for me. Very distracting to the flow and when asking a question, I didn't know if I was heard/understood.

Technical issues impacted my ability to see the presenter's visuals | Many times the PPT froze or didn't work.

The webinar started late with a lot of time wasted | Bad pre-sharing of information, **no time keeper,** no clear goal | The agenda seemed very fluid and I was **confused around the main point.**

Lack of interest | Topic was not relevant | Unspecified

Let's discuss an immediate action plan for next quarter.

From https://www.storytellingwithdata.com/blog/2018/10/10/three-tips-for-storytelling-with-qualitative-data

Word Cloud

aka Tag Cloud, Wordle

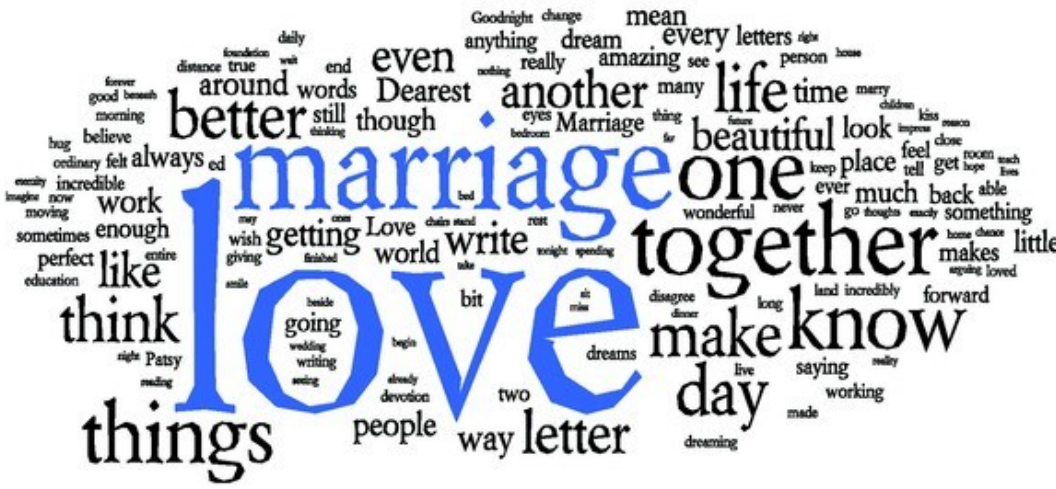
Collection of words, often sized by importance or frequency.

May also incorporate colour to highlight or reflect another value.

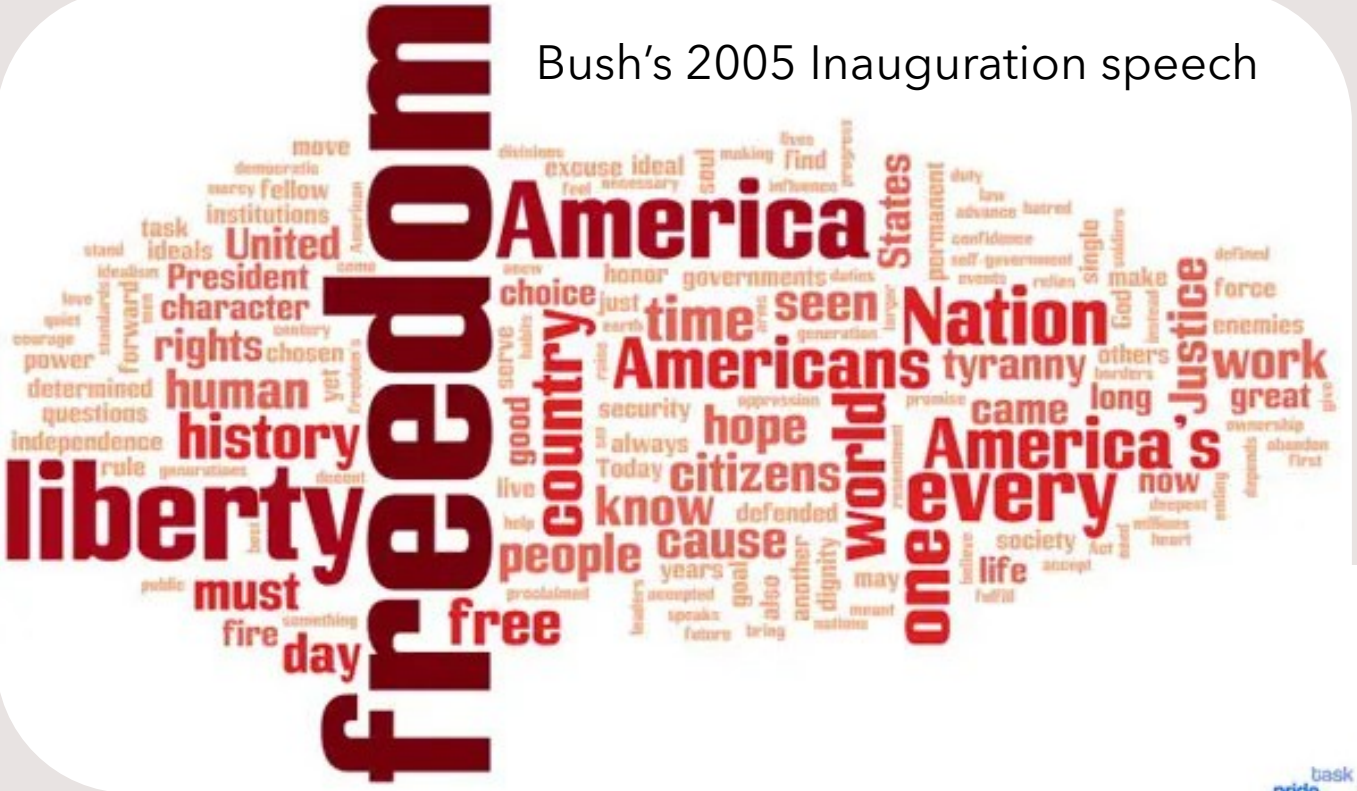
Memorable, engaging, and easy to read.

Conveys a general impression, not exact values.

Useful for quick group reflection on crowdsourcing, surveys, quizzes, summaries, or feedback.

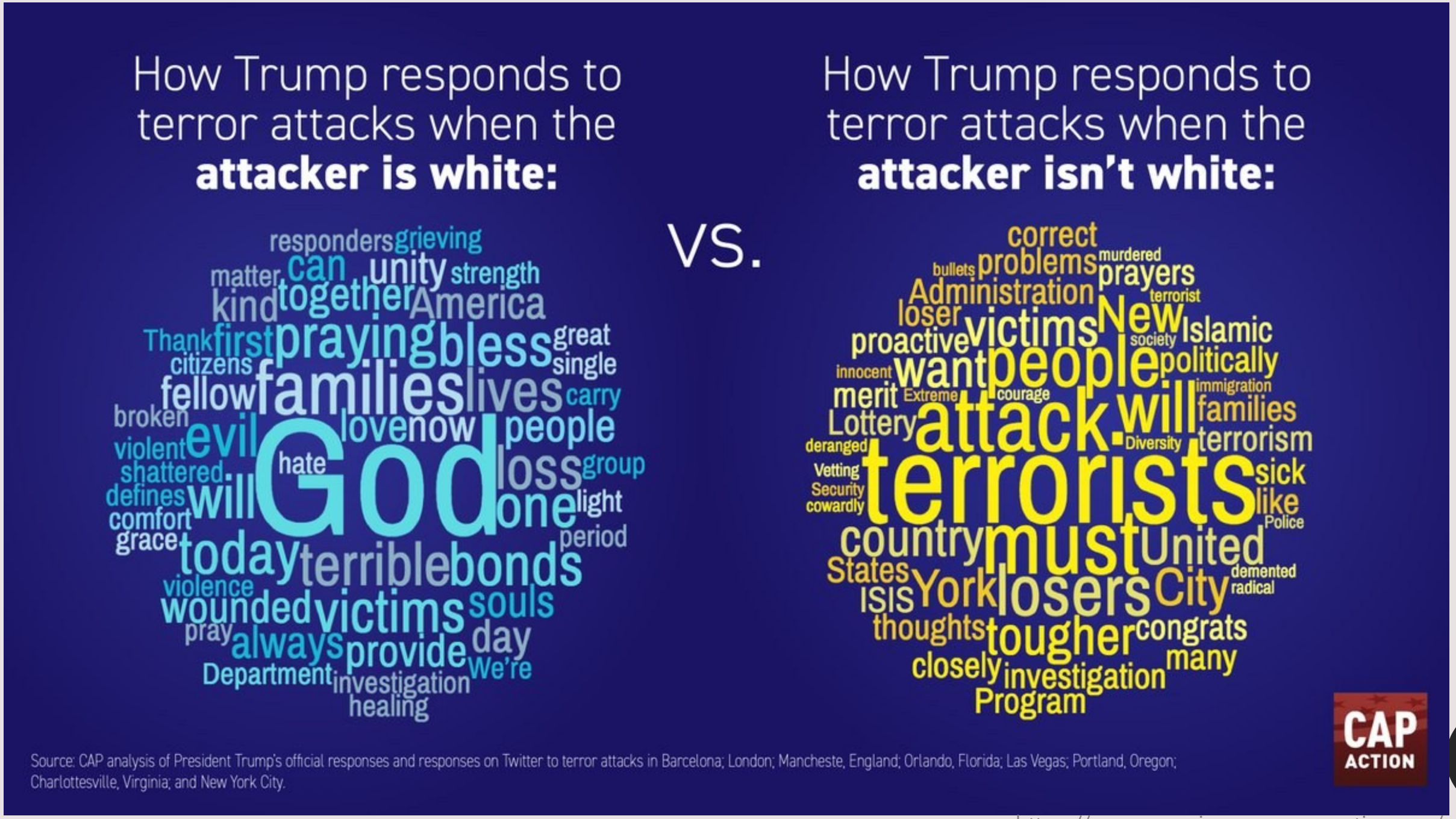


Bush’s 2005 Inauguration speech

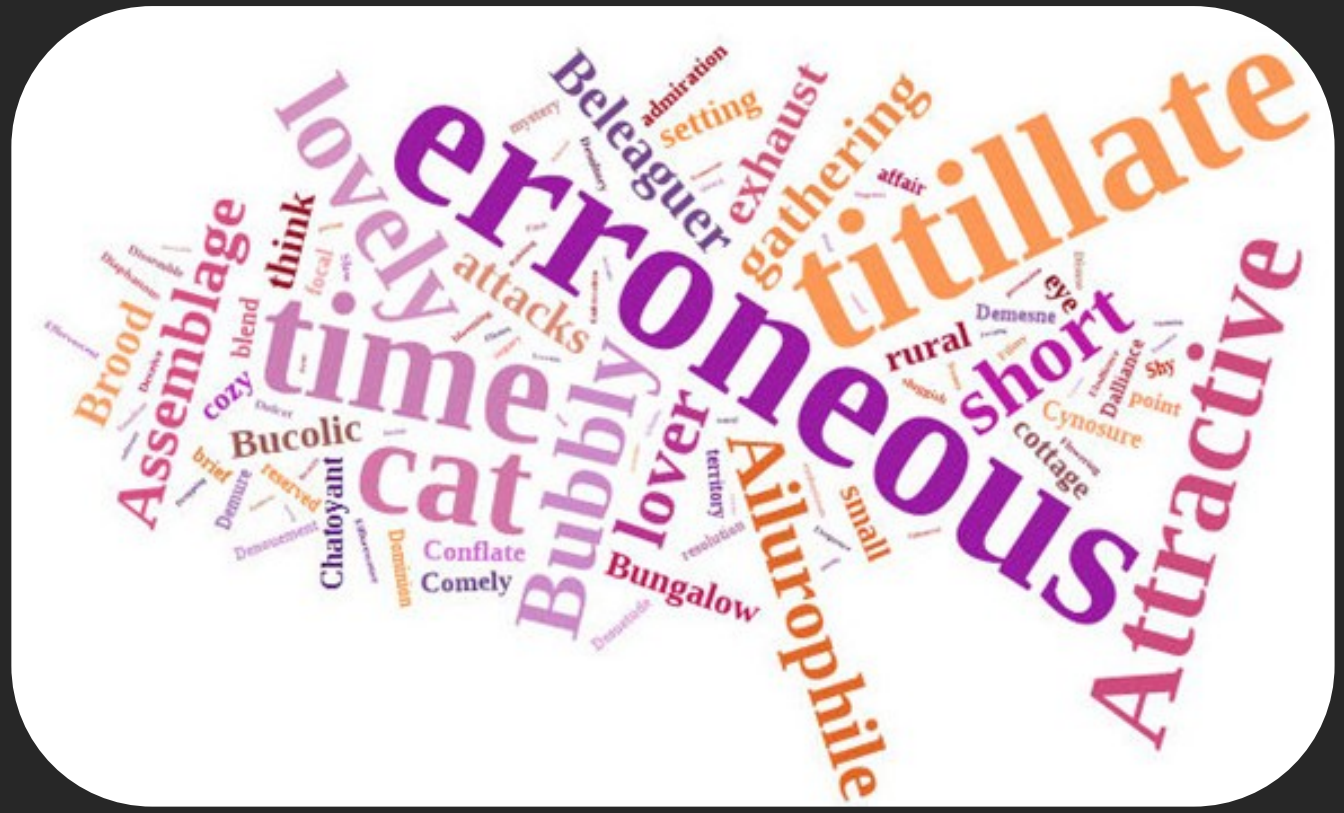


Obama’s 2013 Inauguration speech

From https://www.theguardian.com/world/gallery/2009/jan/21/obama-inauguration-word-clouds-gallery



https://www.americanprogressaction.org/



Word Cloud

Considerations

Beware of random colours

Pick easy to read fonts

Stop words & synonyms

Use important/unique words

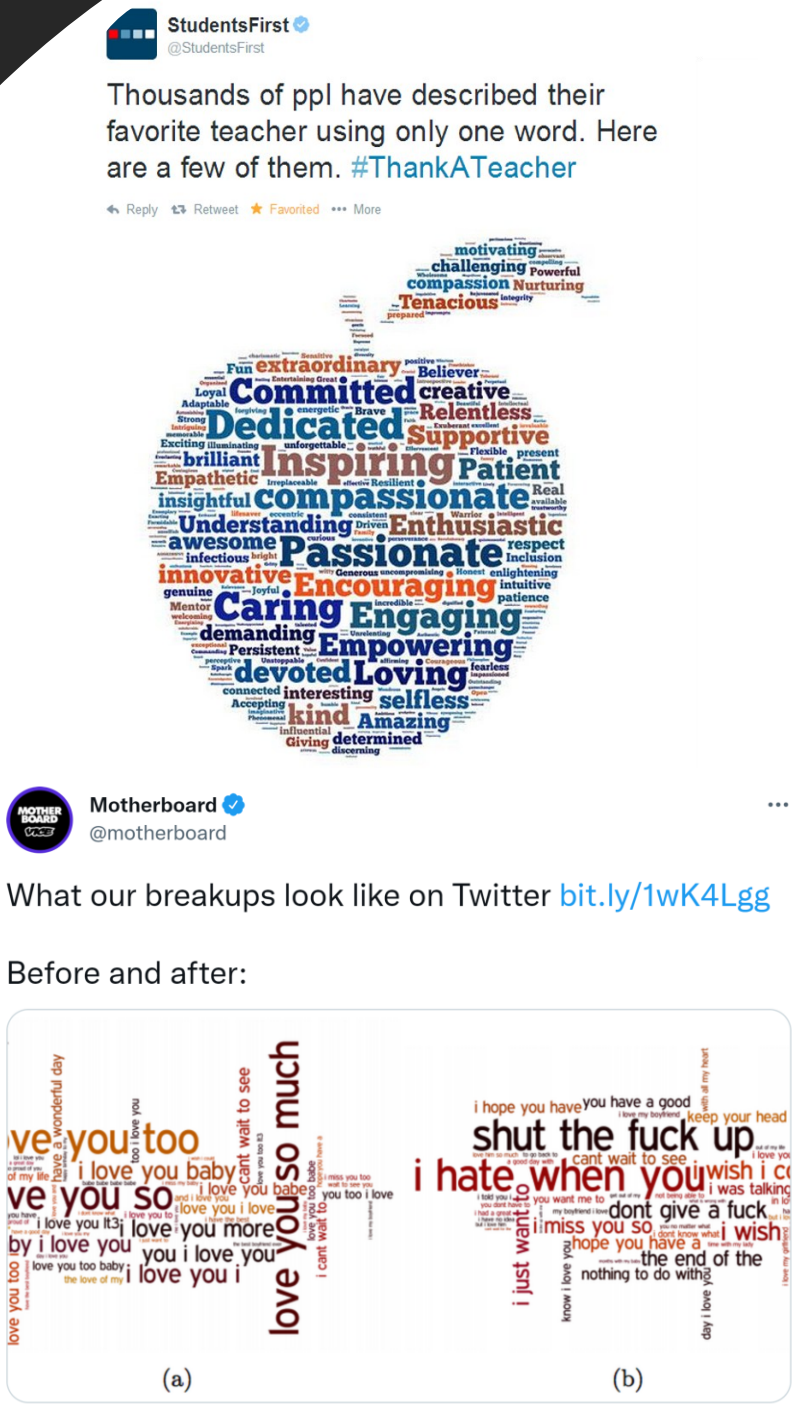
Context is removed

Can be overused

Word Cloud

Simply looking at how often a word or phrase appears in your dataset is *not* a sufficient way to analyze your data!

Be intentional. Example: add a question to your interview like what one word would you use to describe _____?



Word Cloud Generators

EdWordle http://www.edwordle.net/index.html
Basic word cloud.

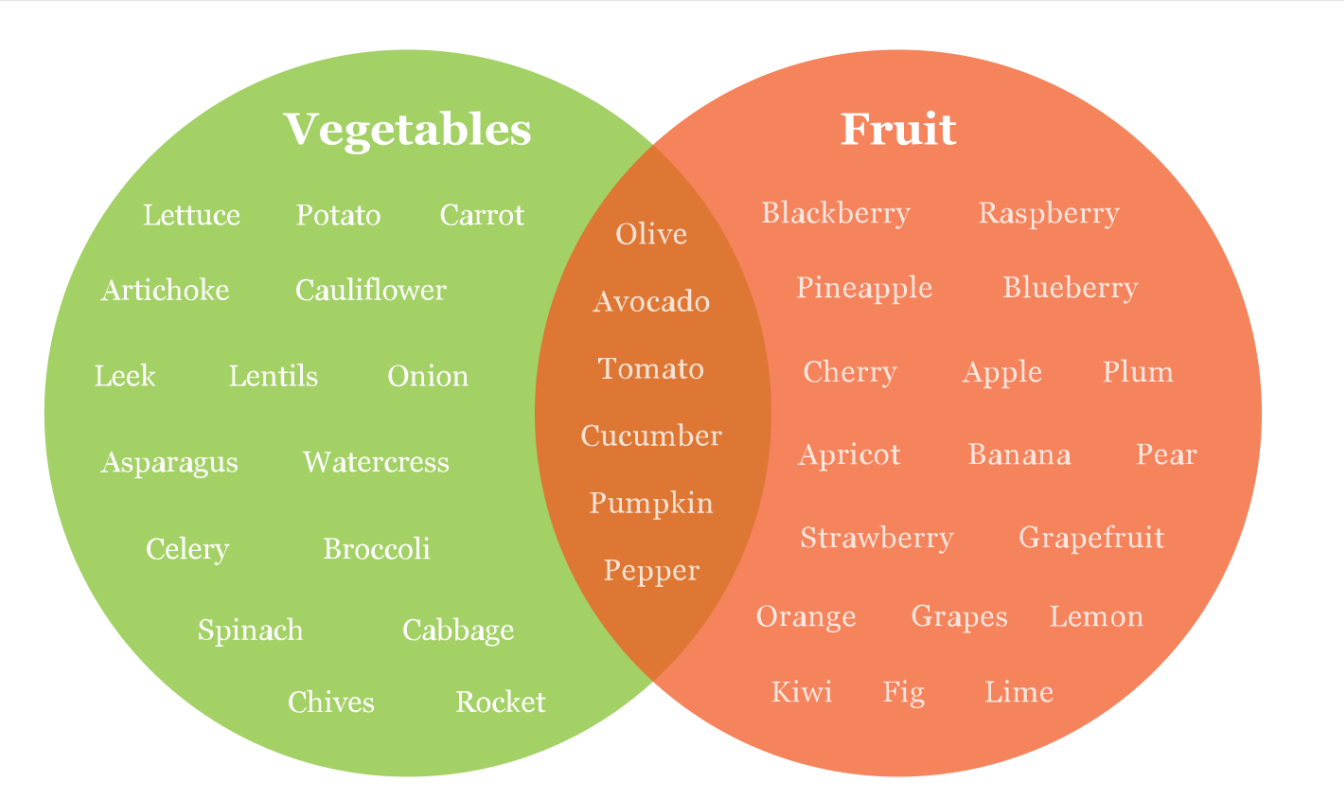
Jason Davies Word Cloud Generator
https://www.jasondavies.com/wordcloud/
Simple & web-based; has orientation controls.

Word Art https://wordart.com/create
Manually specify word size, fit cloud to shapes.

word_cloud
https://github.com/amueller/word_cloud
Python library for coding your own word cloud.

Venn Diagram

aka Set Diagram



Created in 1880 by John Venn

Intersection area can vary.

Some diagrams make the circles different sizes to reflect number of elements in each set.

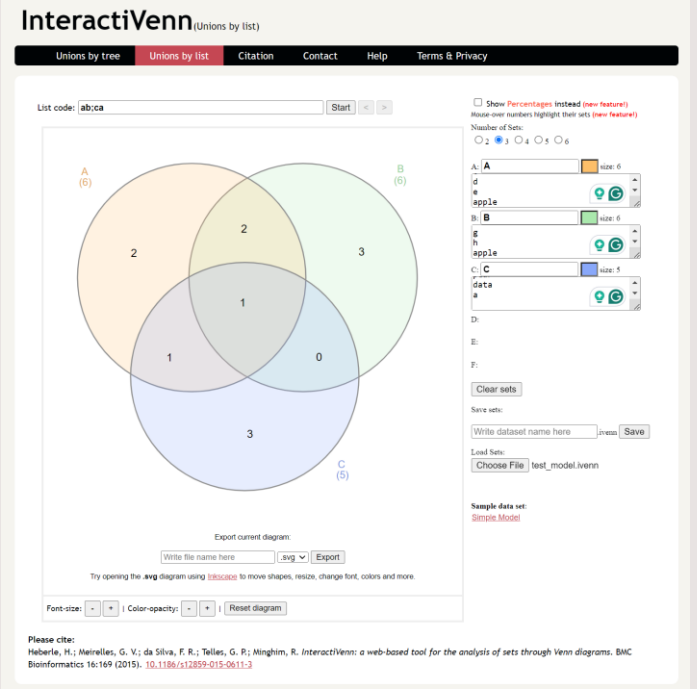
Venn Diagram Software

- Canva (canva.com) & Lucid Chart (lucidchart.com)

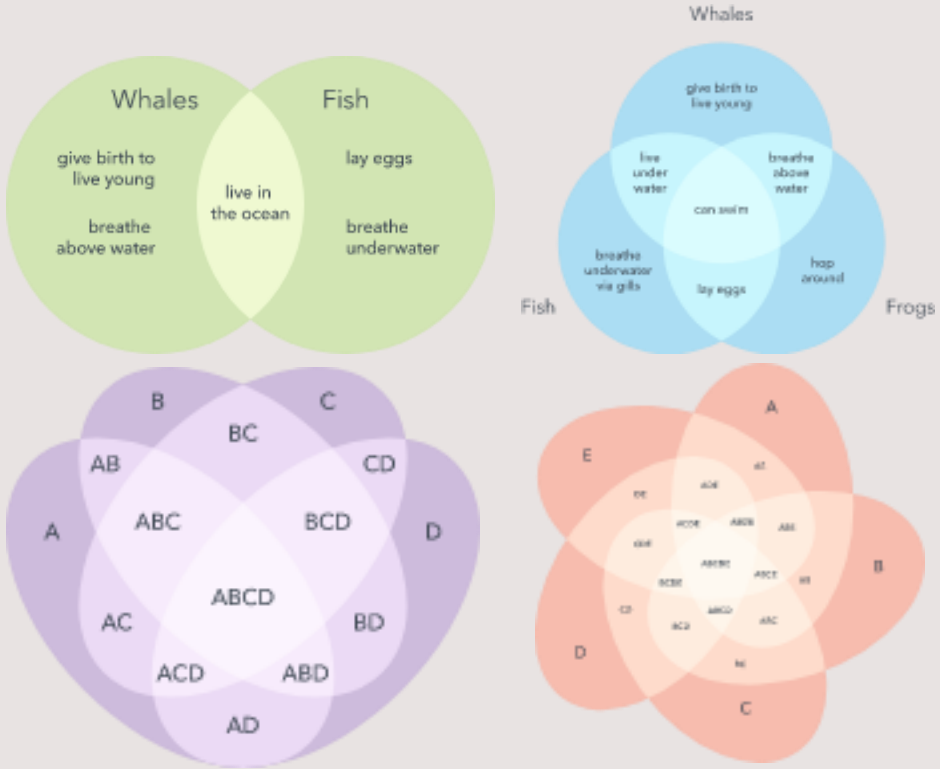
  Provide design templates with a variety of good-looking Venn diagrams. You'll have to add your own labels, sample items, change nodes sizes, etc.

- InteractiVenn (interactivenn.net)

  Web app where you can add lists of set members to produce a Venn Diagram. Limited features.

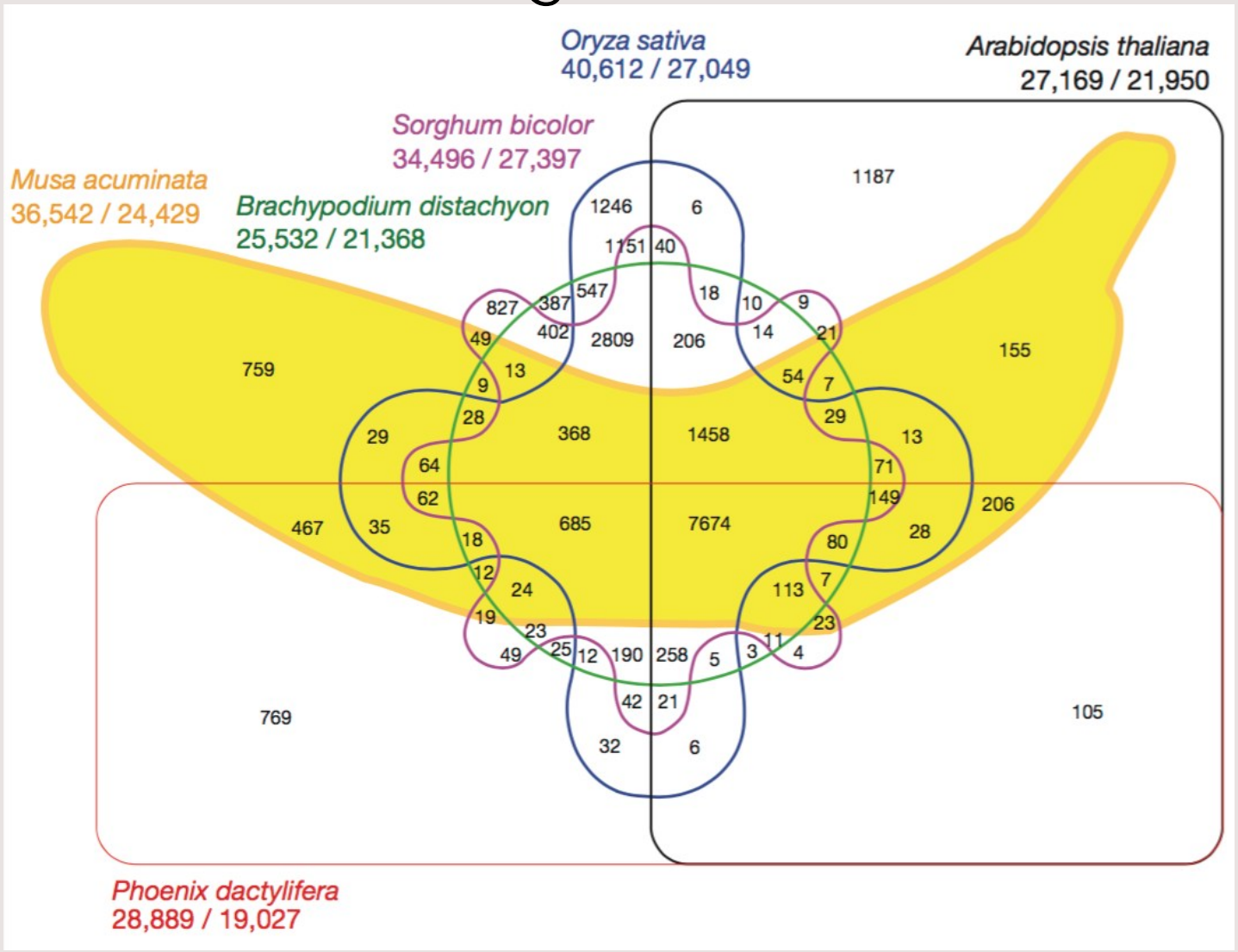


What about more than two sets?



From https://www.lucidchart.com/pages/tutorial/venn-diagram

A six-set venn diagram



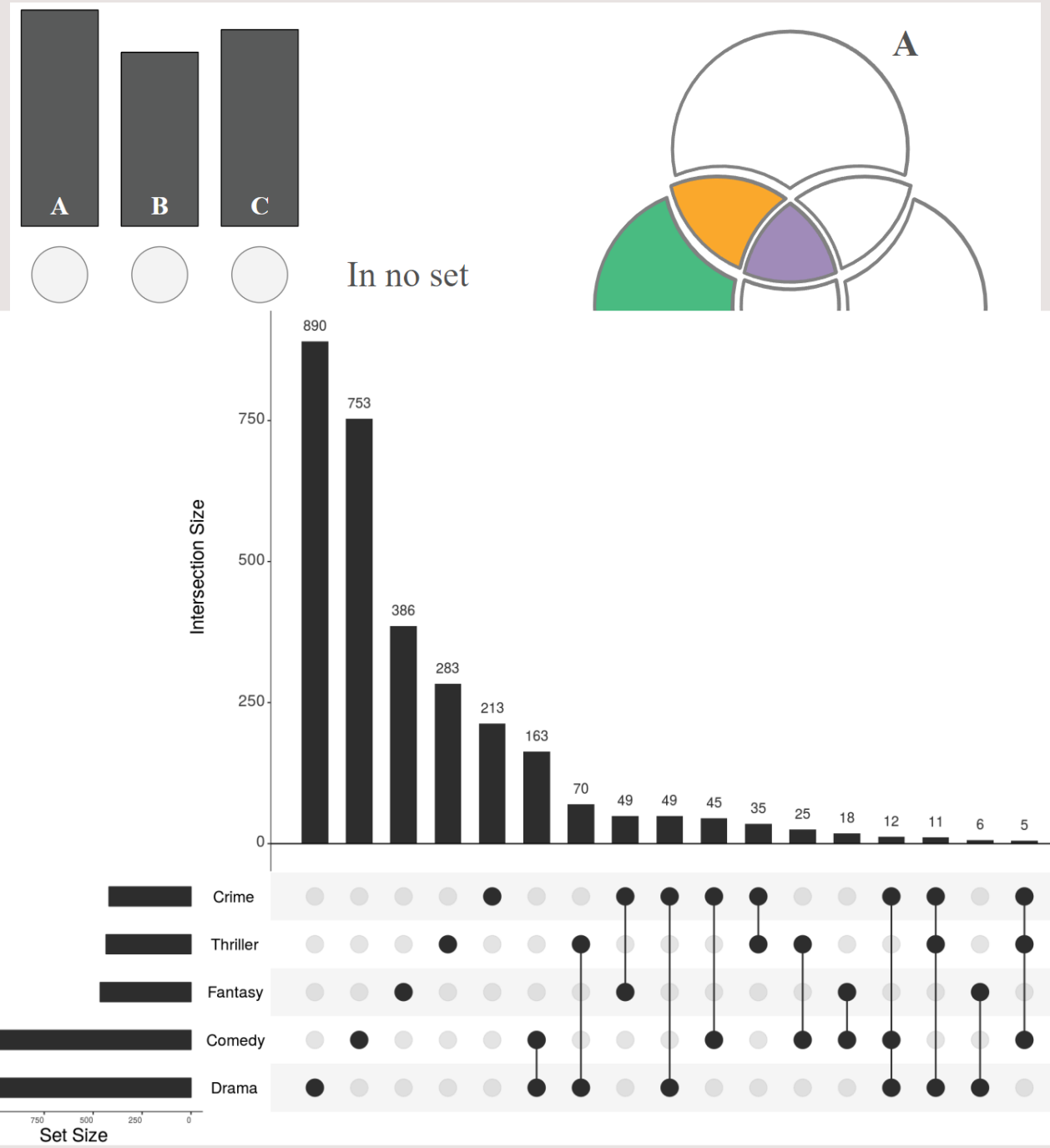
Overlap in genome between a banana and five other species.

D'Hont, A., Denoeud, F., Aury, JM. et al. The banana (Musa acuminata) genome and the evolution of monocotyledonous plants. Nature 488, 213–217 (2012). https://doi.org/10.1038/nature11241

UpSet

Web tool, R library, javascript library, etc.

https://upset.app/implementations/



Images from https://upset.app/

Timelines

In Gaza, a Pattern of Conflict

By JEREMY ASHKENAS, ARCHIE TSE and KAREN YOURISH JULY 31, 2014

Since Israel withdrew from the Gaza Strip in 2005, there have been periodic eruptions of violence between Israel and Hamas. Hamas launches volleys of rockets into Israel; Israeli warplanes strike targets in Gaza. This has escalated into an Israeli ground invasion twice. Below is a look at similarities and differences in the last three major conflicts.

After persistent rocket fire from Gaza, Israel **launches air campaign** and threatens ground invasion.

2008-9 **Day 1** of 23

Airstrikes kill more than 300 Palestinians — the highest one-day toll in decades. Israeli officials say the strikes are the start of an operation to force Hamas to end its rocket barrages and could include ground forces.

2012 **Day 1** of 8

Israel unleashes a ferocious air assault on Gaza, killing a top military commander of Hamas, and warns that a ground invasion could be coming.

2014 **Day 1**

Israel hits more than 200 targets — including senior Hamas officials, smuggling tunnels and rocket launch sites — and calls up army reserves in preparation for a possible ground invasion.

http://www.nytimes.com/interactive/2014/07/31/world/middleeast/in-gaza-a-pattern-of-conflict.html

https://www.behance.net/gallery/30395075/U10-Atlas-Regio-Utrecht

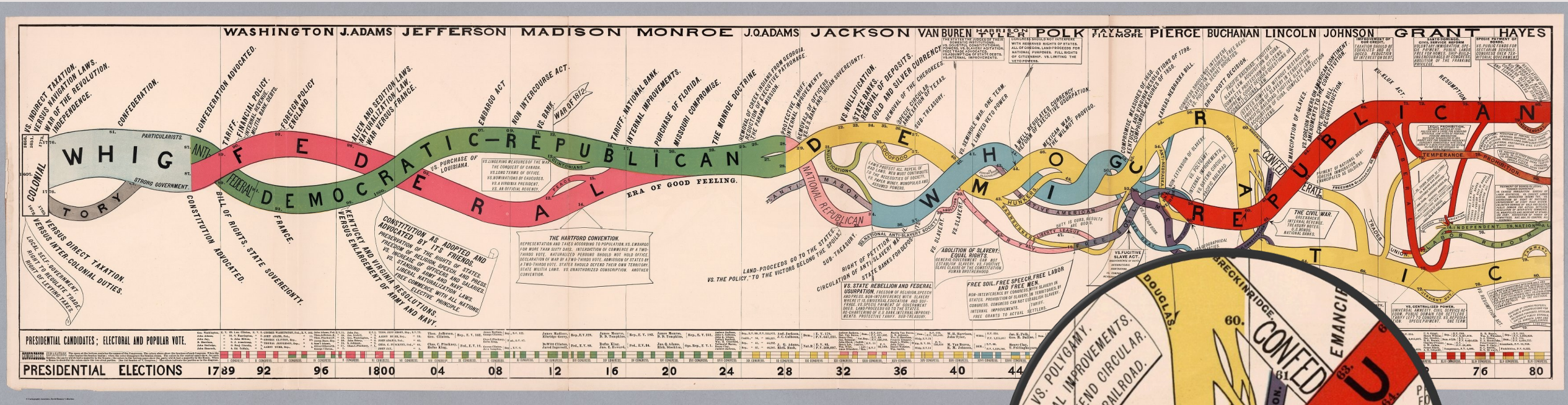
Diagram of the Rise and Fall of American Political Parties

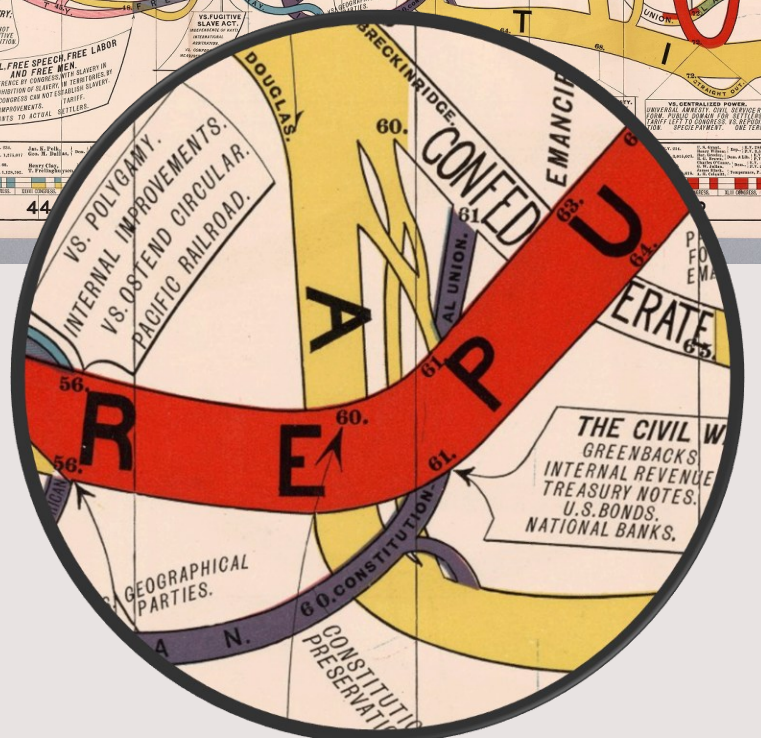
from 1789 to 1880, inclusive

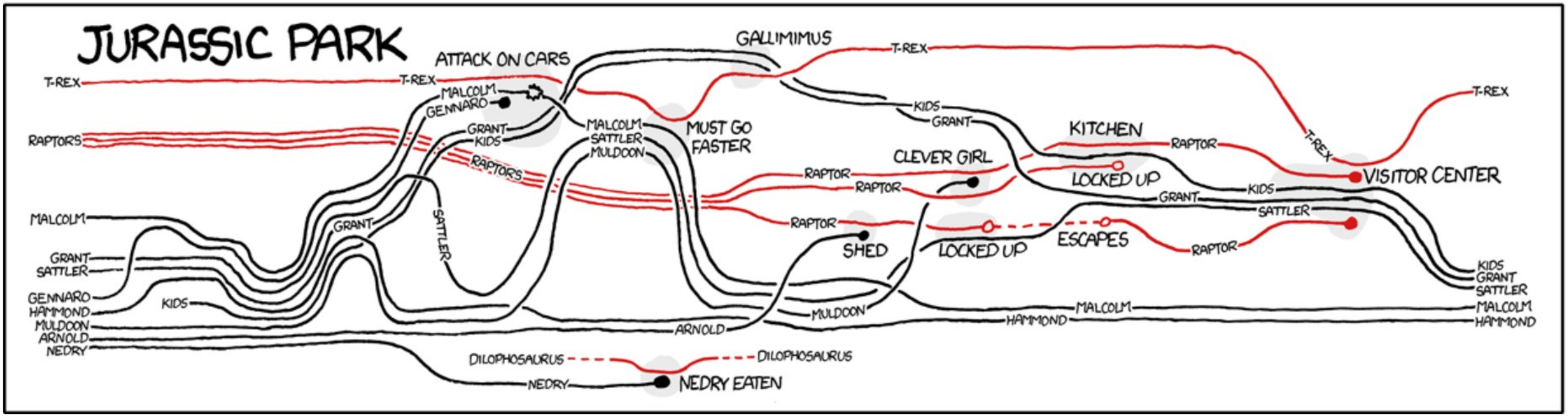
Walter R. Houghton

1880, Bloomington/Indianapolis

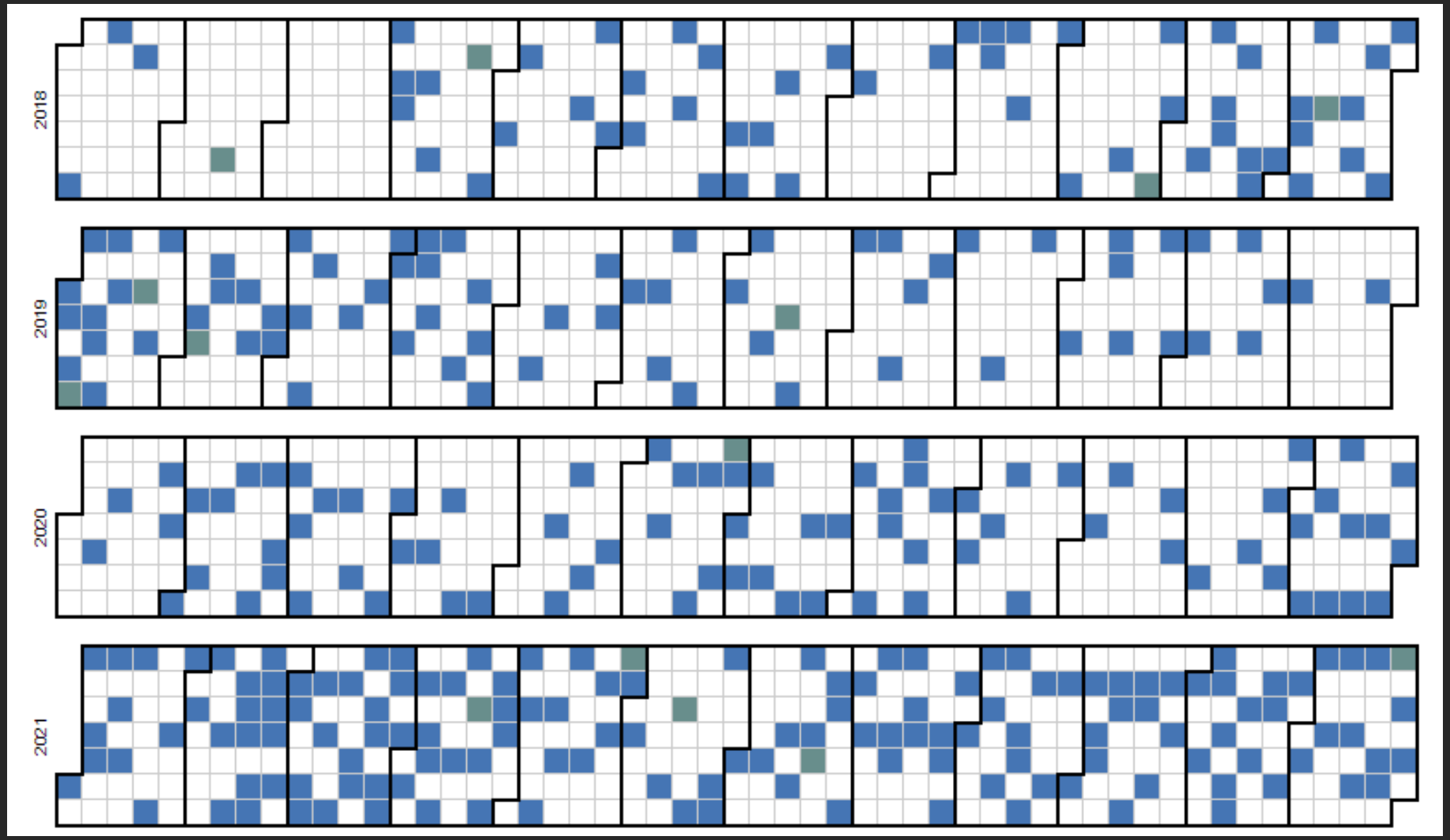
https://exhibits.stanford.edu/dataviz/catalog/824a049e3319f42ba83eccfdbd6adb41





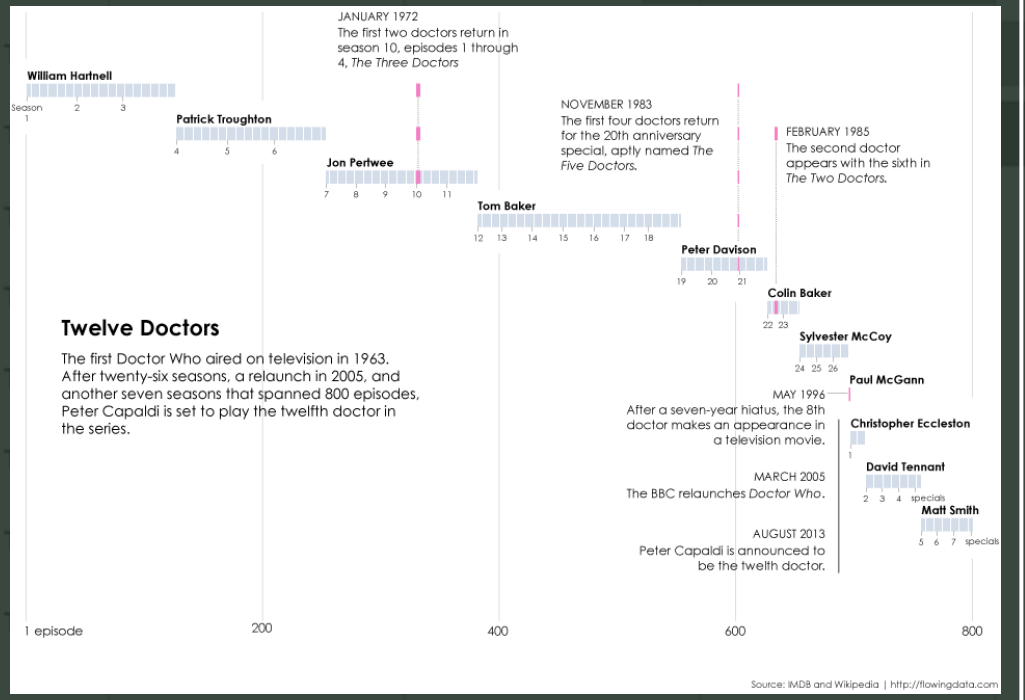


https://xkcd.com/657/large/

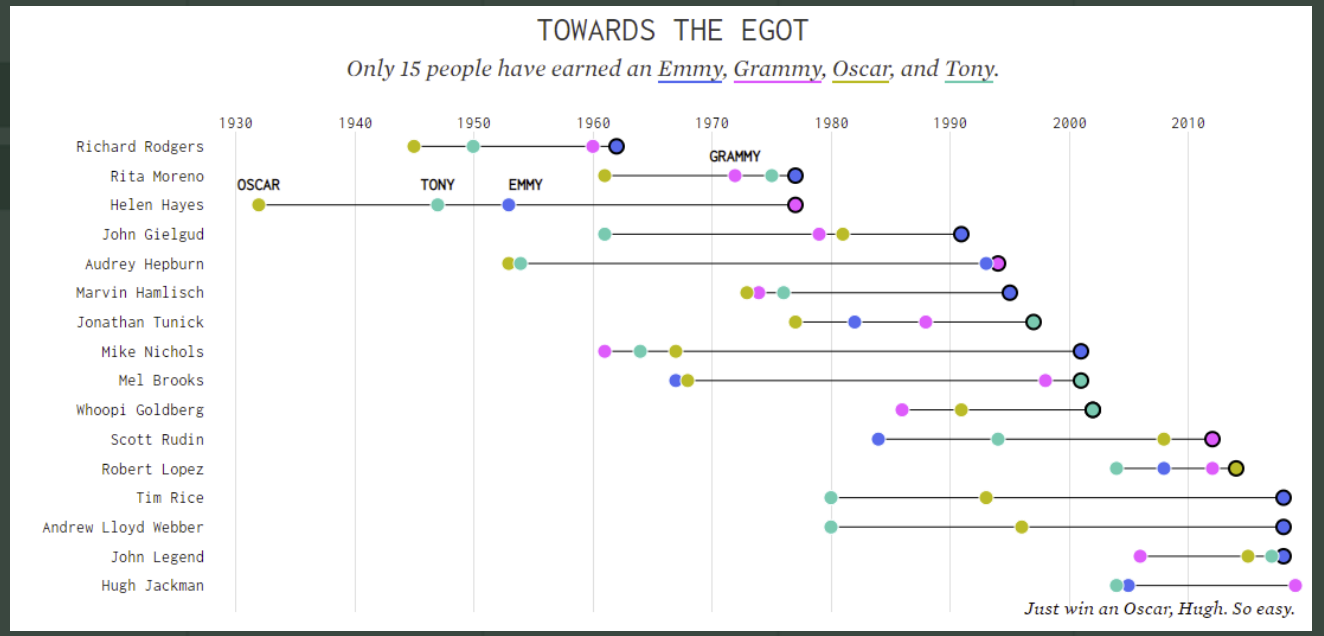

2018
2019
2020
2021

Gantt Chart

- Bar chart + timeline
- Often used in project management
- Each row is an activity, length of bar is duration
- Can be created in Excel, Tableau, and many other tools.



https://flowingdata.com/2013/08/05/a-look-back-at-past-doctor-whos/



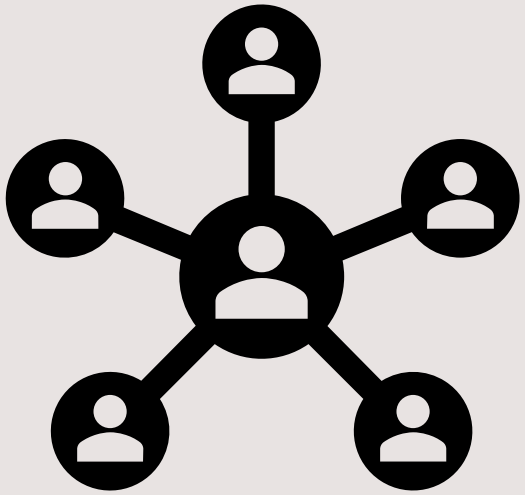
https://flowingdata.com/2019/02/12/egot-winner-timelines/

Timeline Tools

- TimelineJS
  http://timeline.knightlab.com/
- Fourish
  https://app.flourish.studio/templates#template-timeline
- Google Charts
  https://developers.google.com/chart/interactive/docs/gallery/timeline
- Tiki-Toki Online Timeline Maker
  https://www.tiki-toki.com/
- Time.graphics
  https://time.graphics/
- Vizzlo
  https://vizzlo.com/create/timeline (watermark)
- PowerBI (calendars)
- Powerpoint or MS Visio

Graphs

aka Networks
aka Node-Link Diagrams

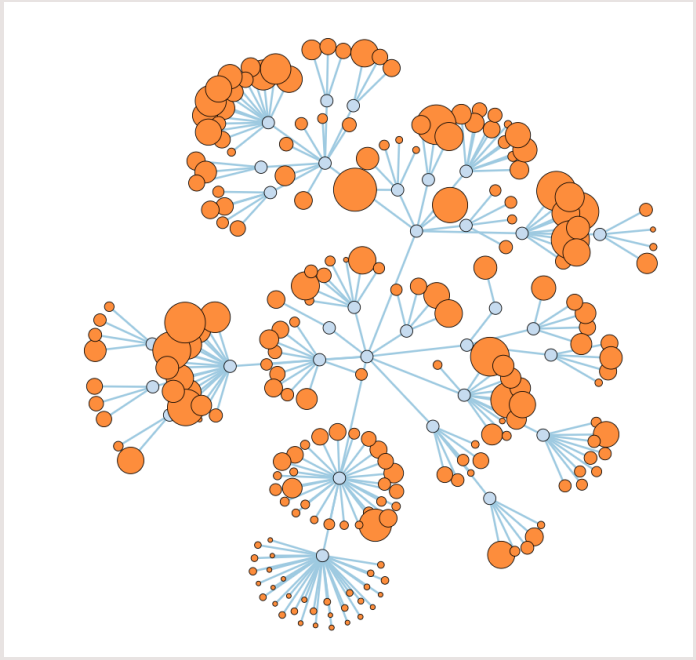


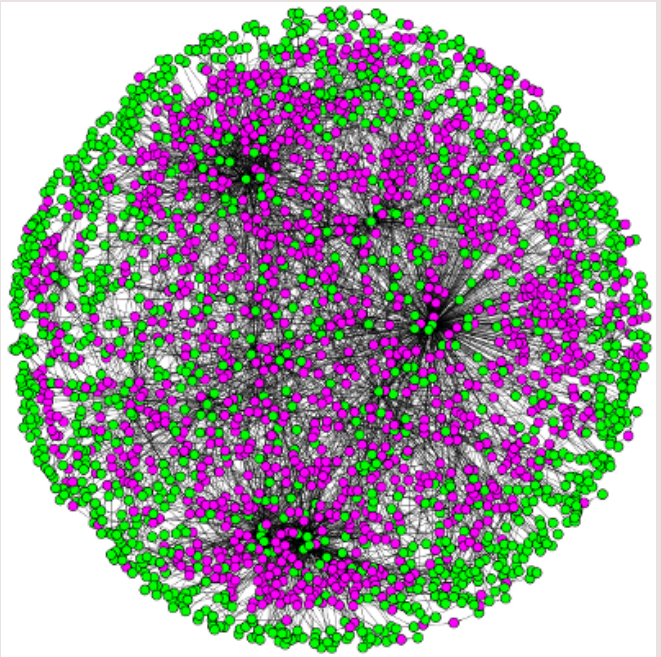
Used for visualizing relationships.

Nodes can be different shapes as well as coloured or sized based on data attributes.

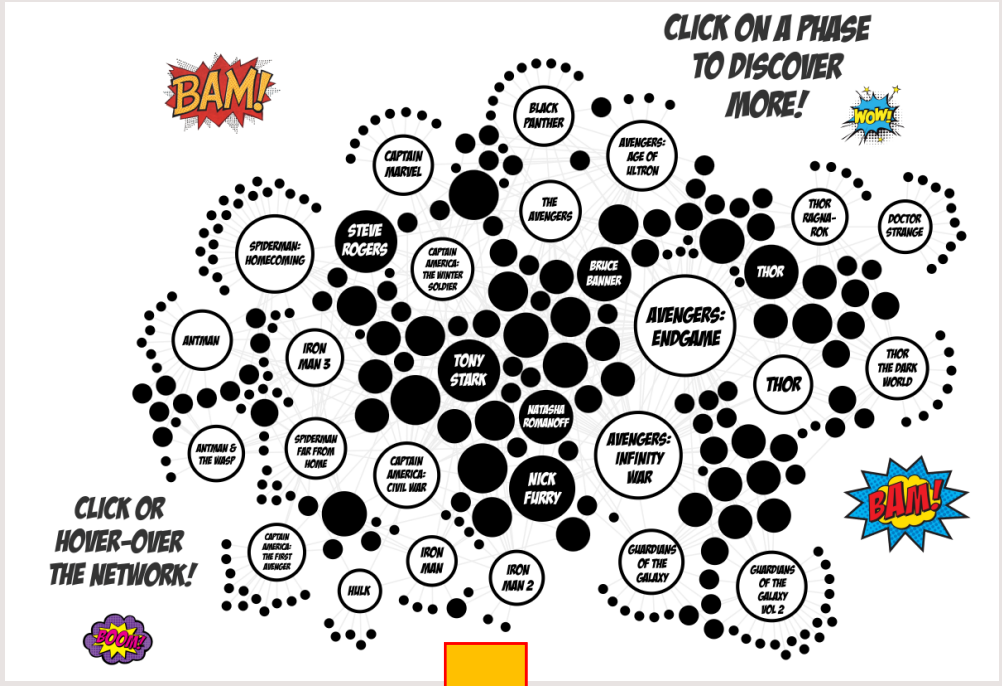
Edges can be directional and have different widths to reveal properties of the relationship.

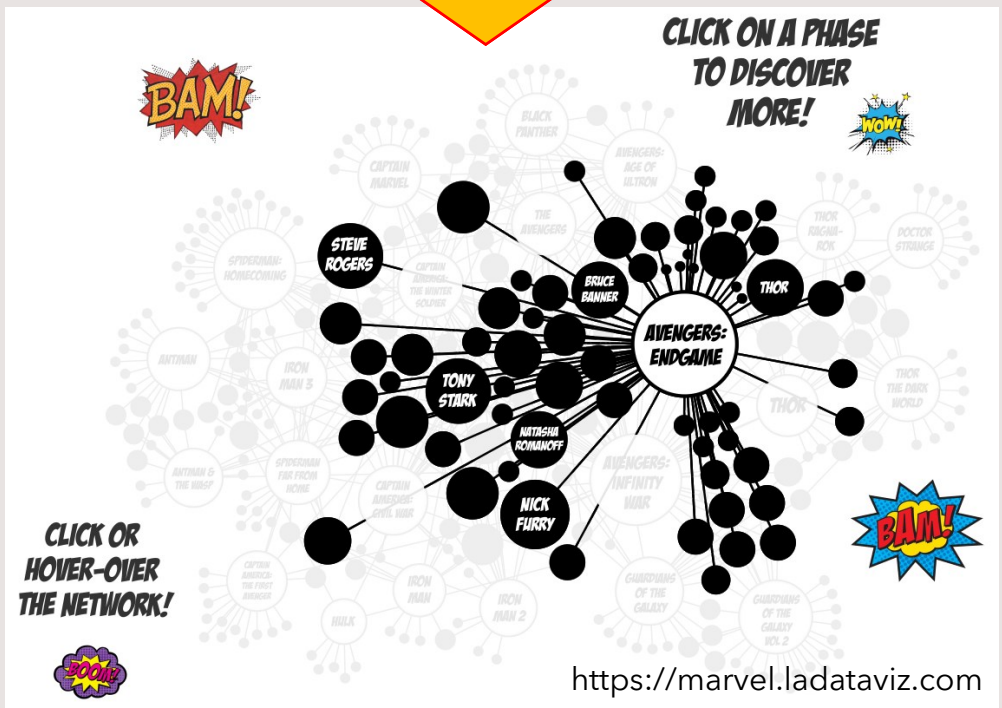
Positioning is critical.



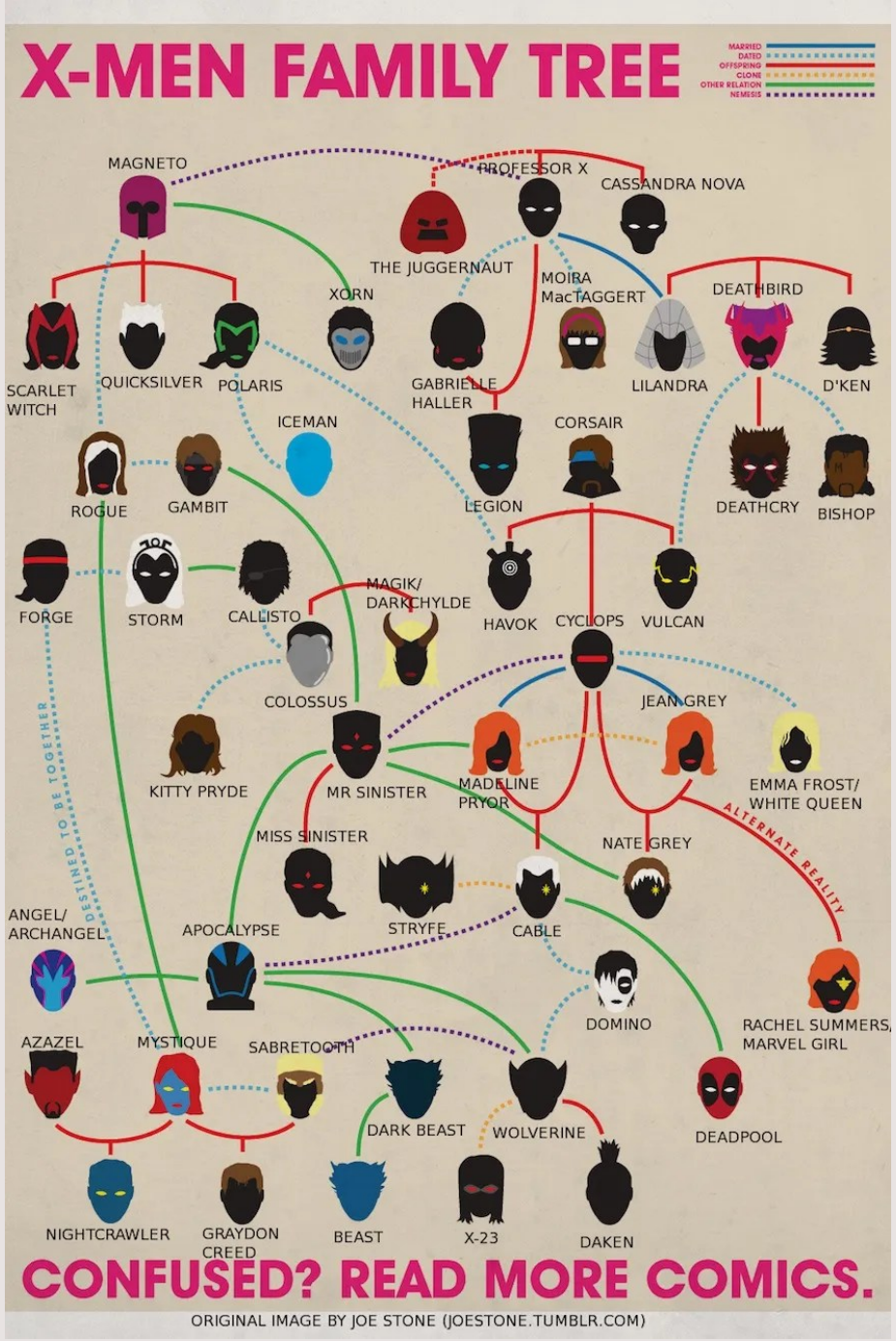


https://mbostock.github.io/d3/talk/20111116/force-collapsible.html

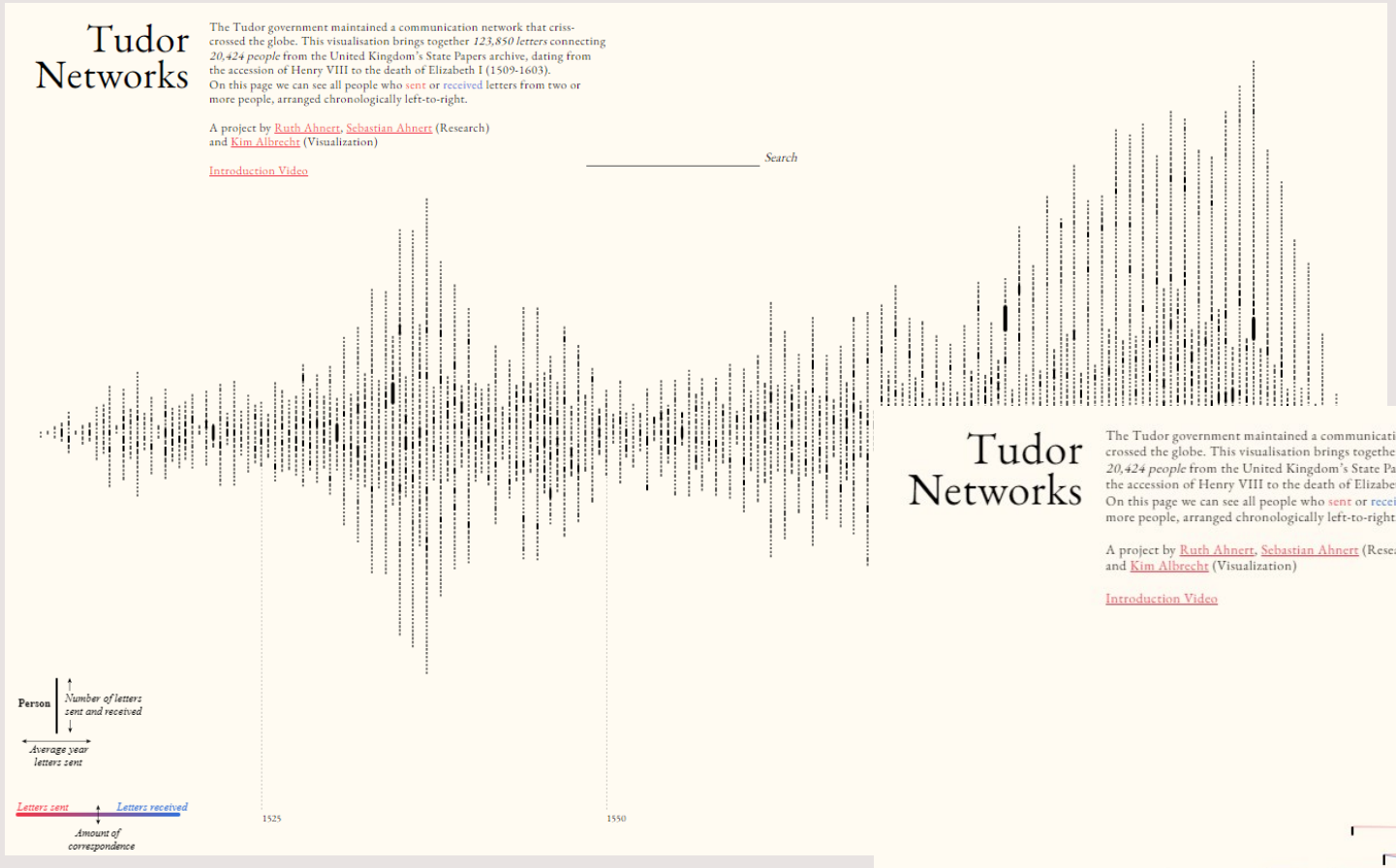


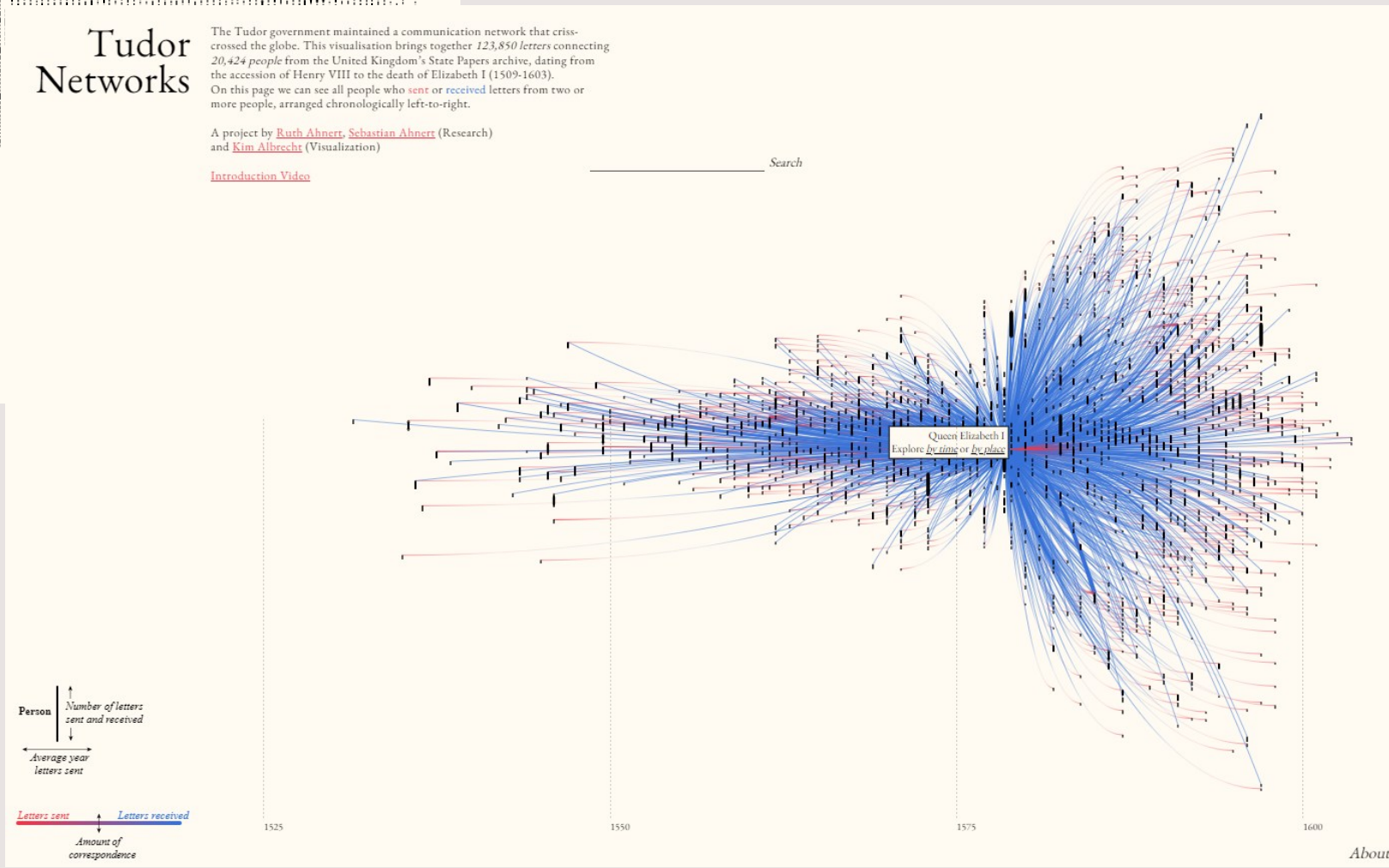


https://marvel.ladataviz.com



https://www.themarysue.com/an-annotated-x-men-family-tree/





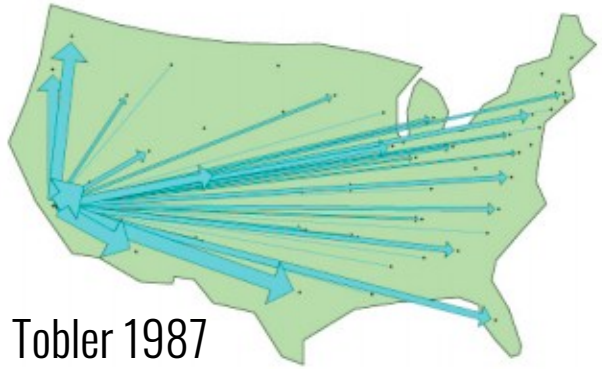
Ruth Ahnert, Sebastian Ahnert, and Kim Albrecht
https://tudornetworks.net/

Networks with Map Locations

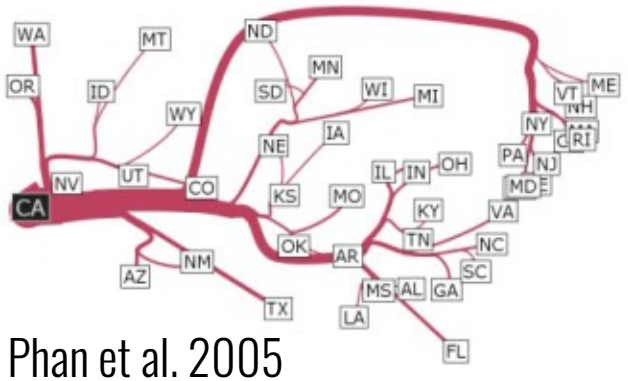
- Choice: Ignore exact locations or not
- If not, less freedom

  Focus on adjusting edge properties

Migration from California

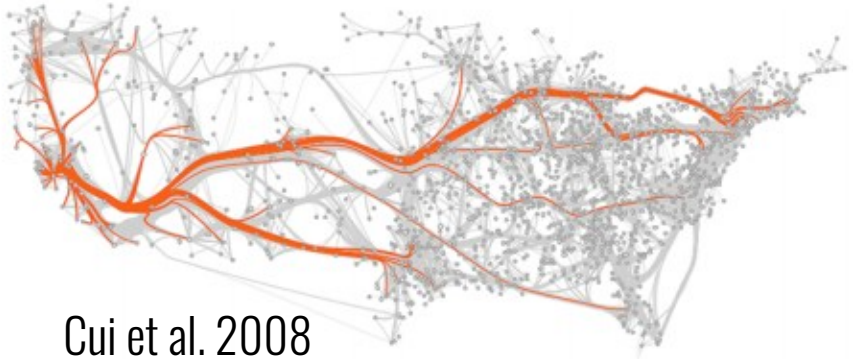


Tobler 1987

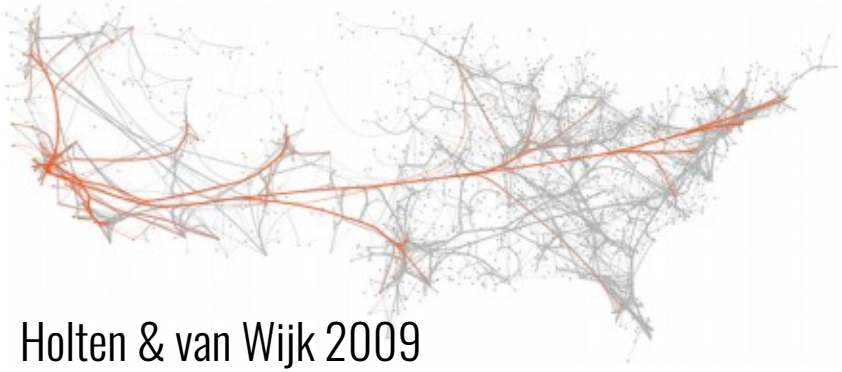


Phan et al. 2005

Verbeek et al. 2011

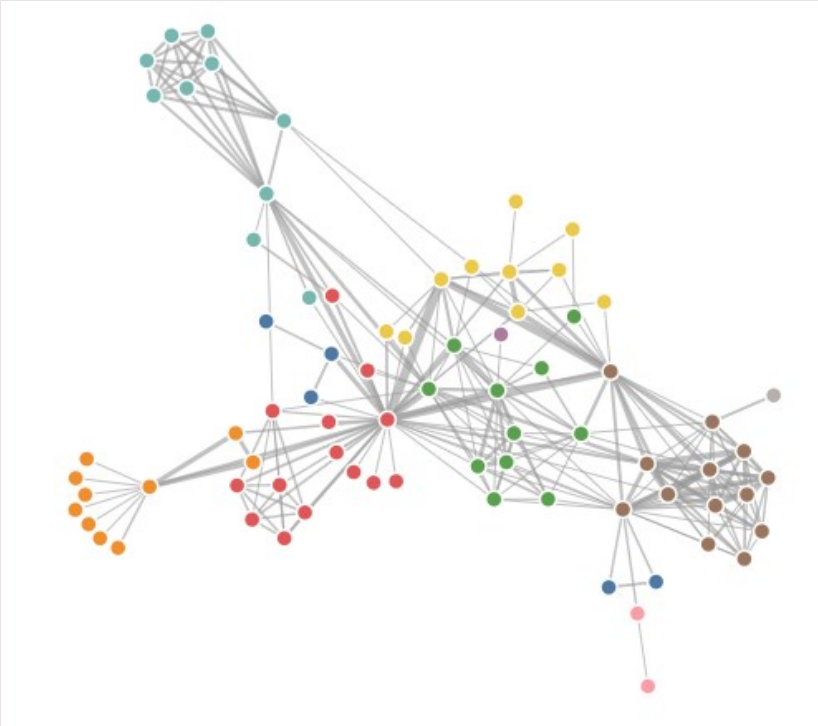


Cui et al. 2008



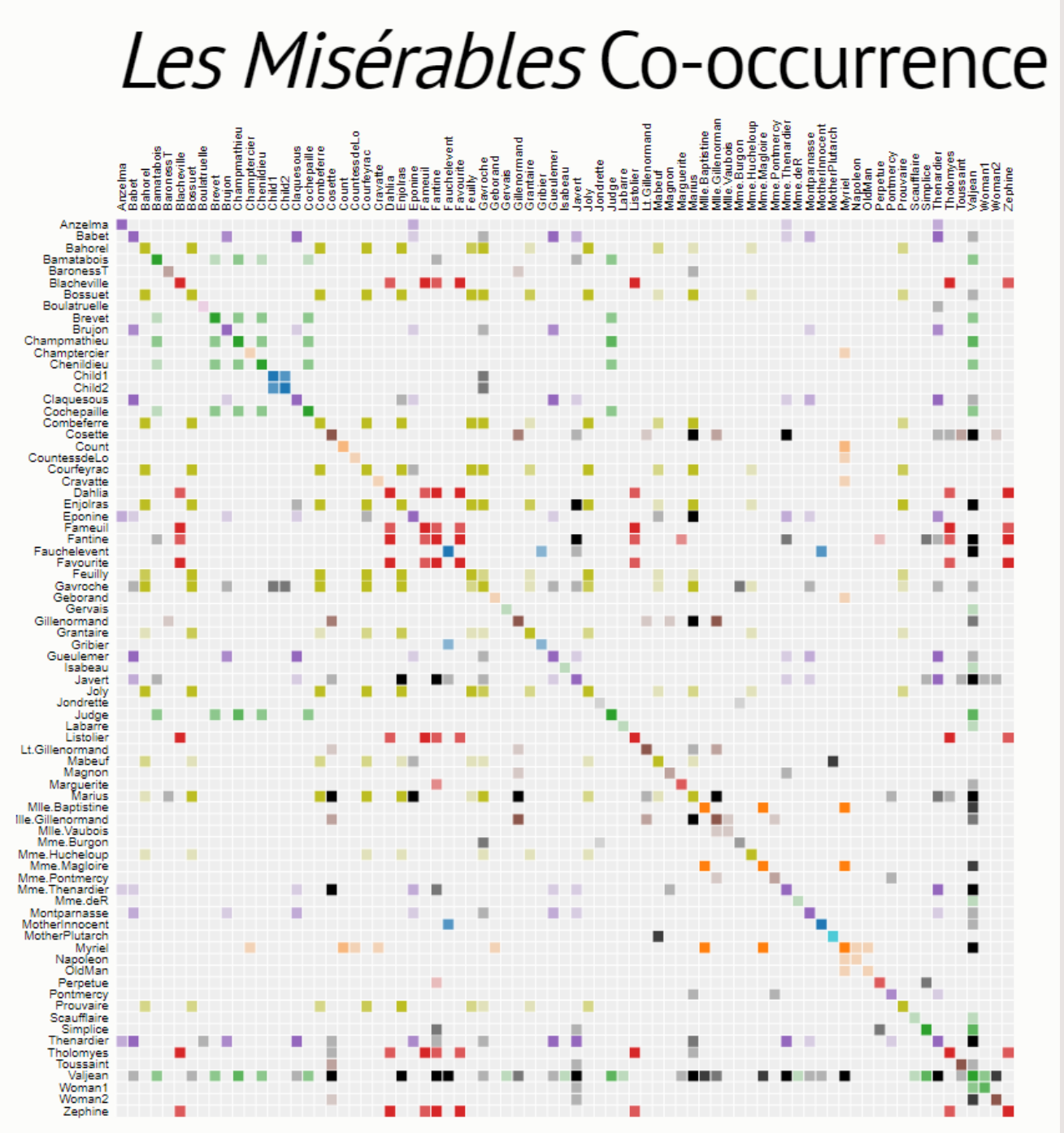
Holten & van Wijk 2009

Matrix Networks



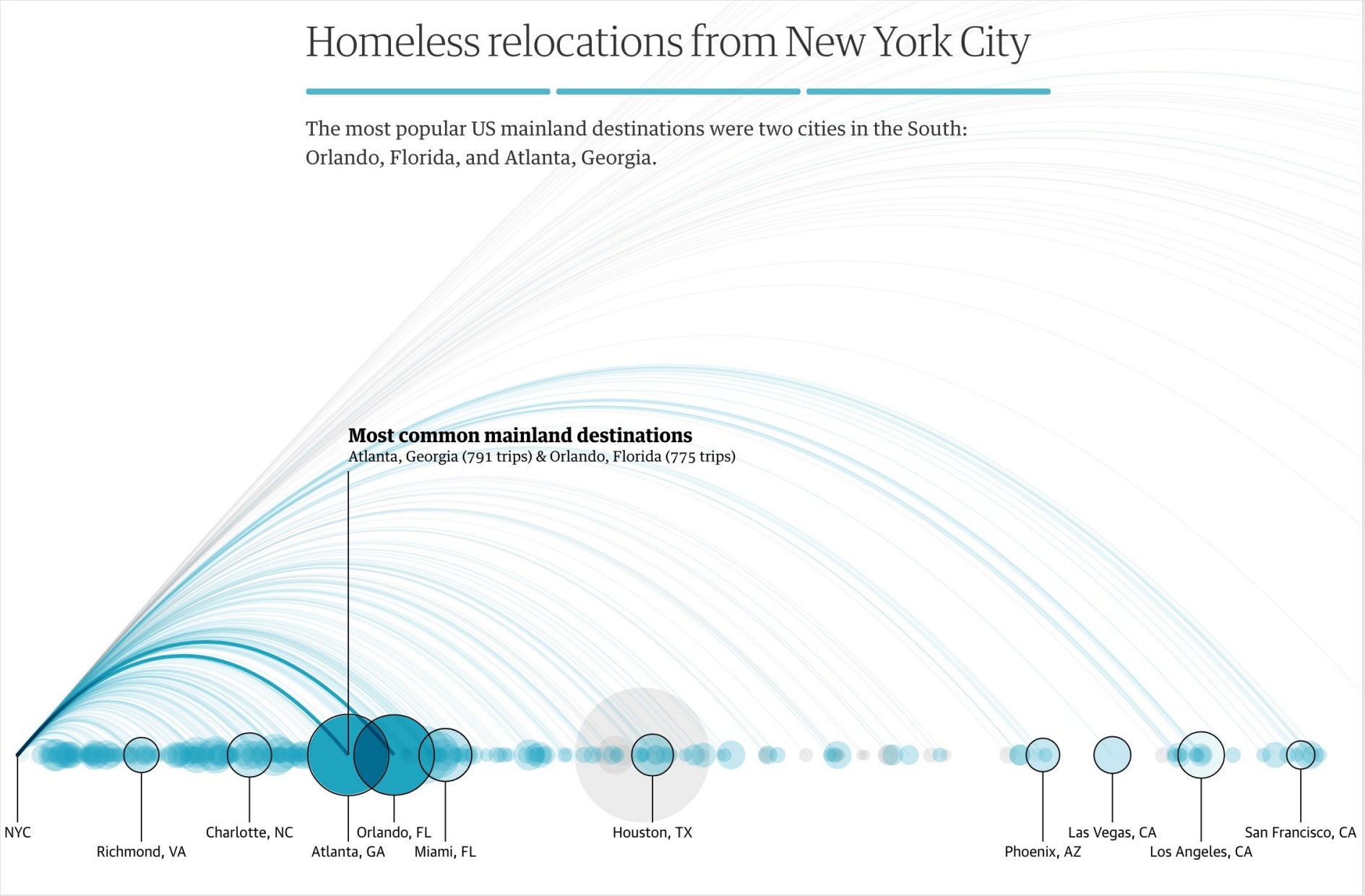
https://observablehq.com/@d3/force-directed-graph

Les Misérables Co-occurrence



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

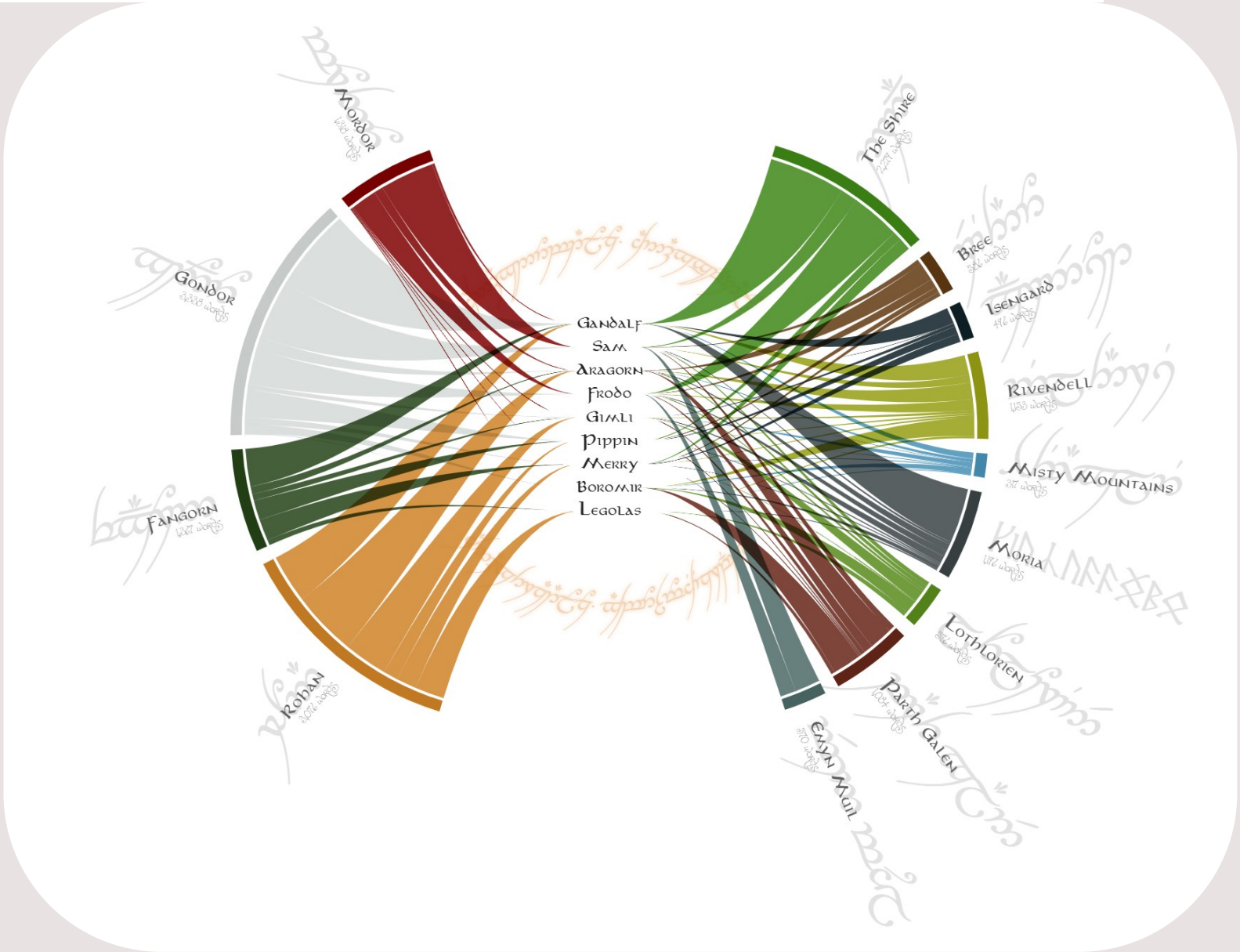
Arc Diagram



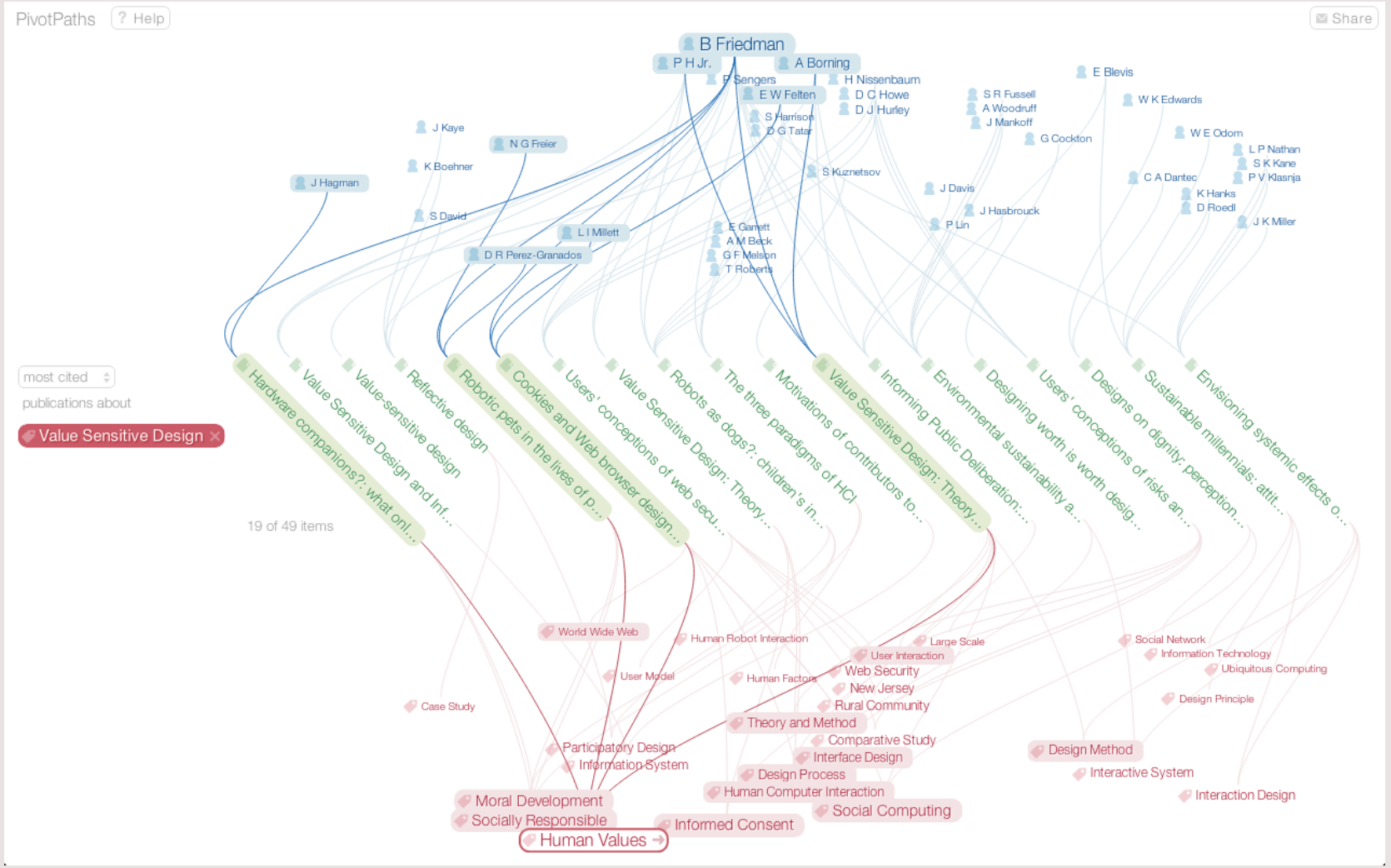
https://www.theguardian.com/us-news/ng-interactive/2017/dec/20/bussed-out-america-moves-homeless-people-country-study

Who's speaking in Middle Earth

How many words have the members of the Fellowship spoken across Middle Earth during all 3 extended editions of the Lord of the Rings



Do you need to see everything?



https://mariandoerk.de/pivotpaths/demo/#/1:0_497686

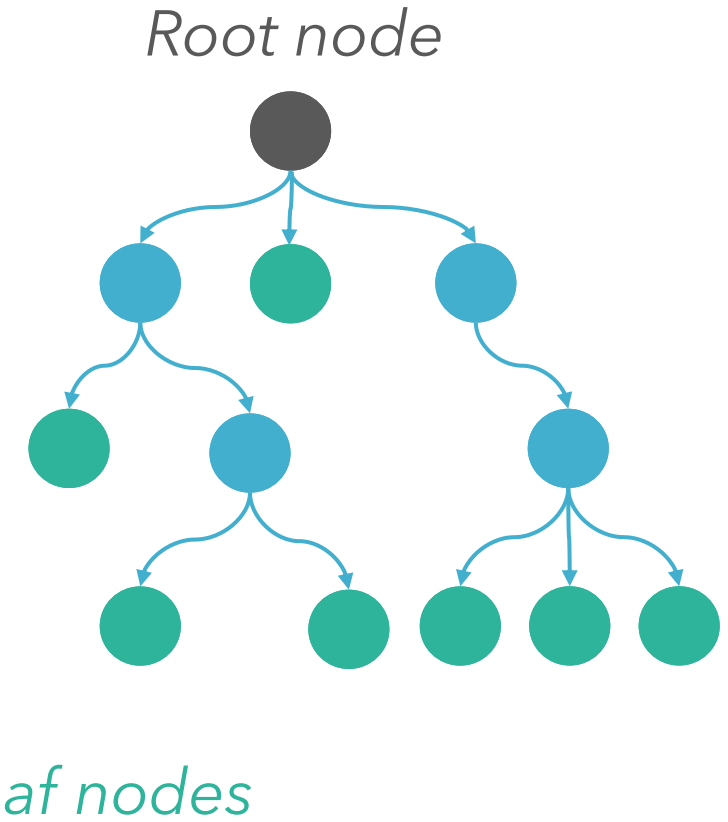
Trees

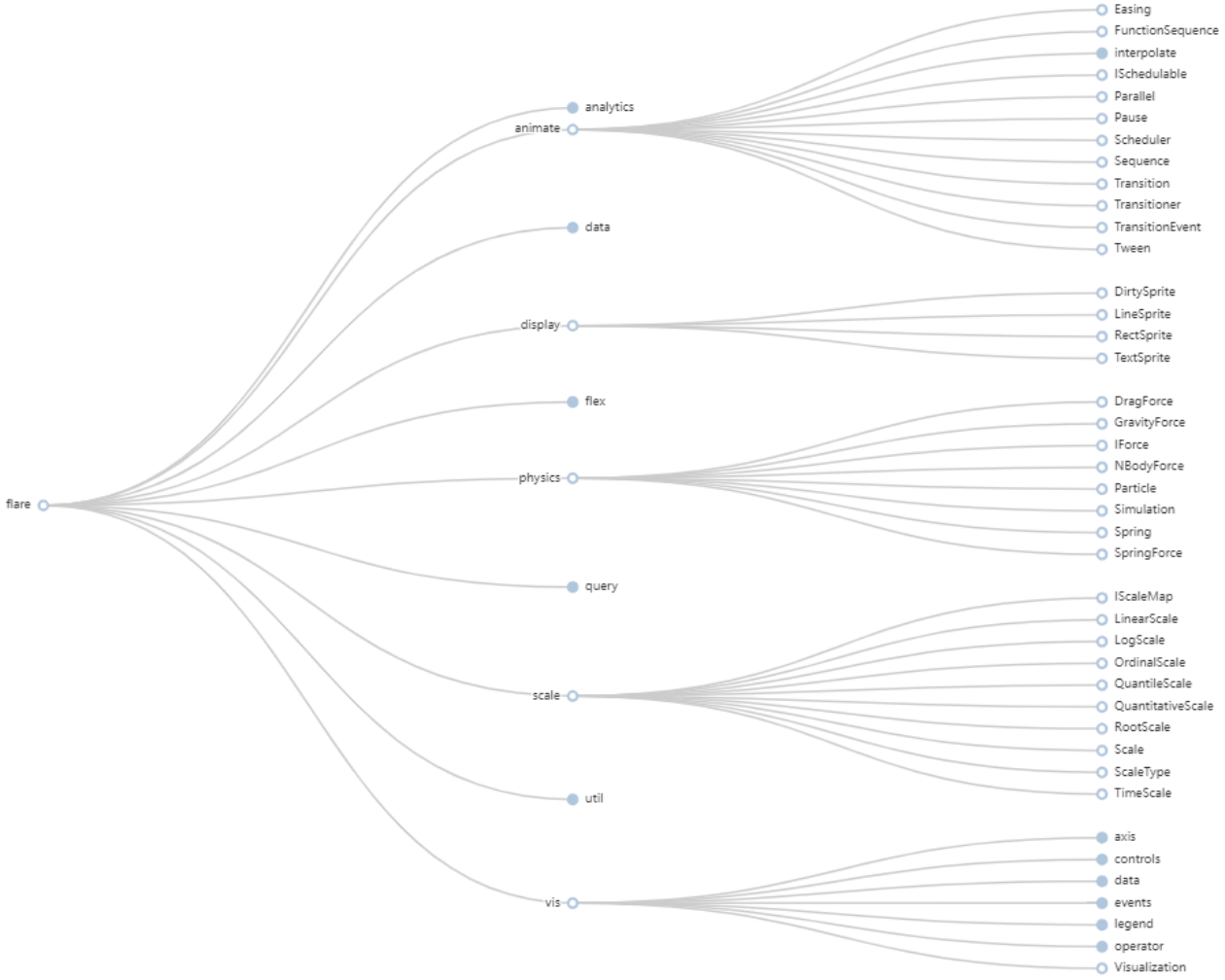
Simple(r) type of graph.
No cycles (loops).

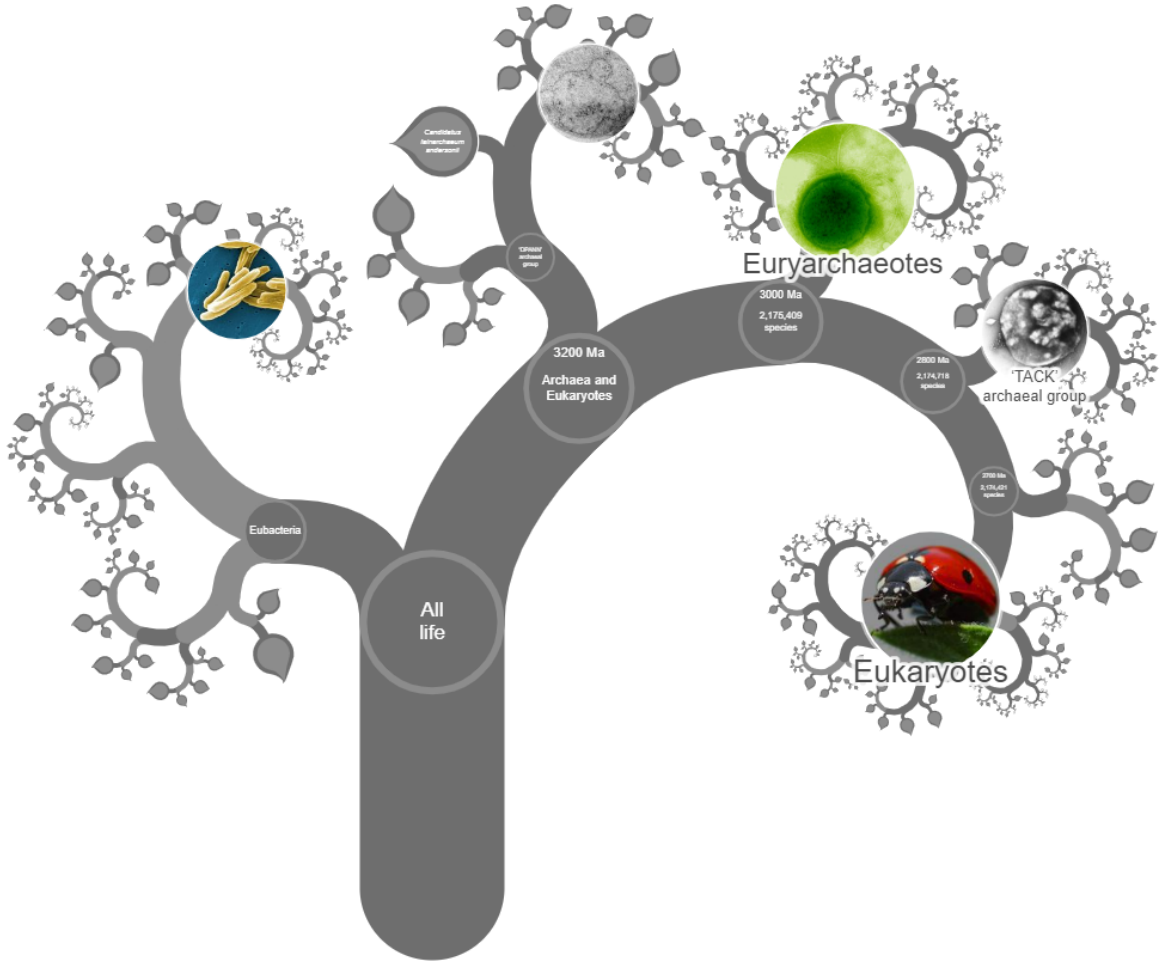
Often used to show hierarchies, family relations/descent, and classification structures.

Types of nodes:

- Root
- Parent/child
- Leaf

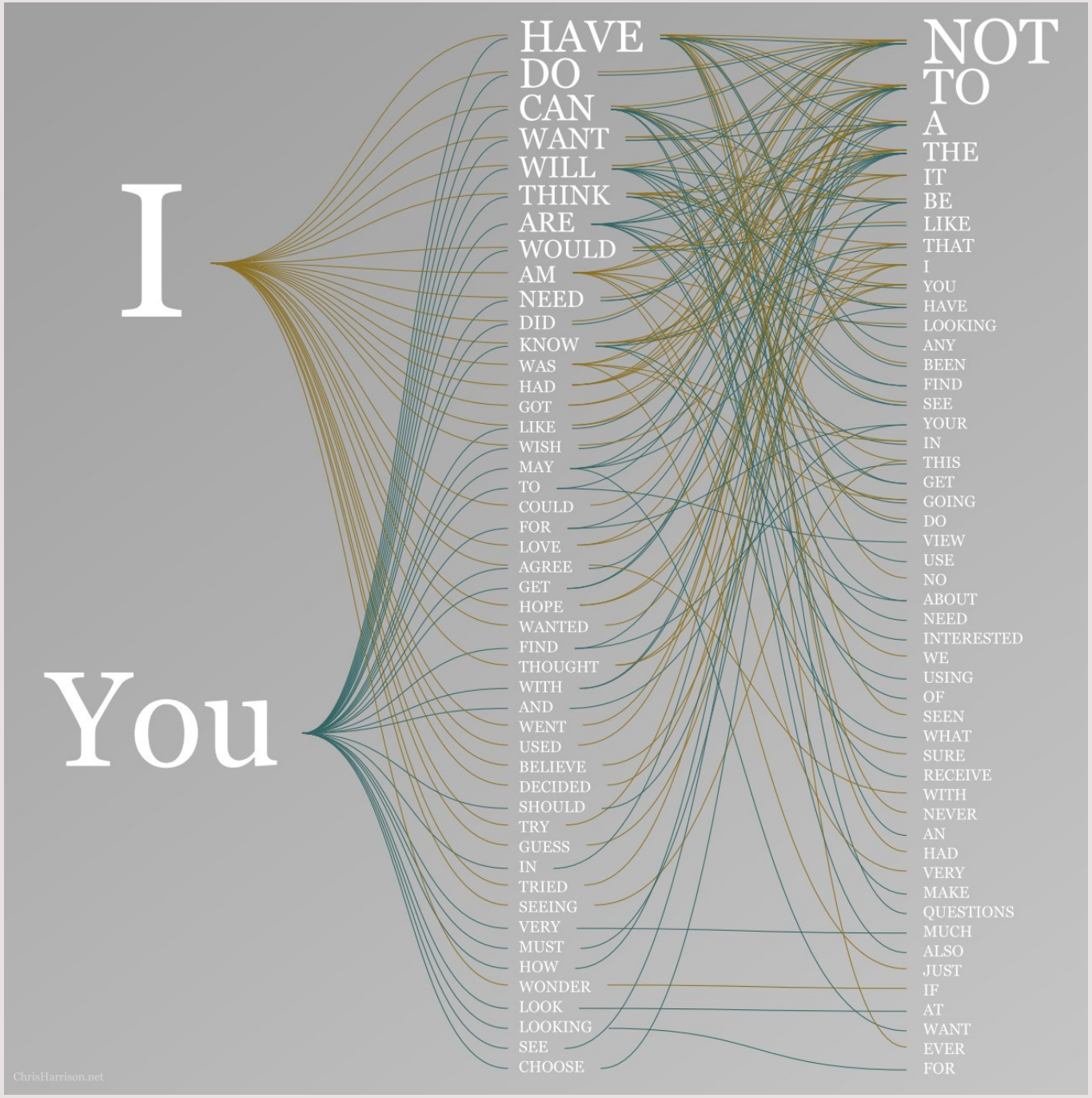






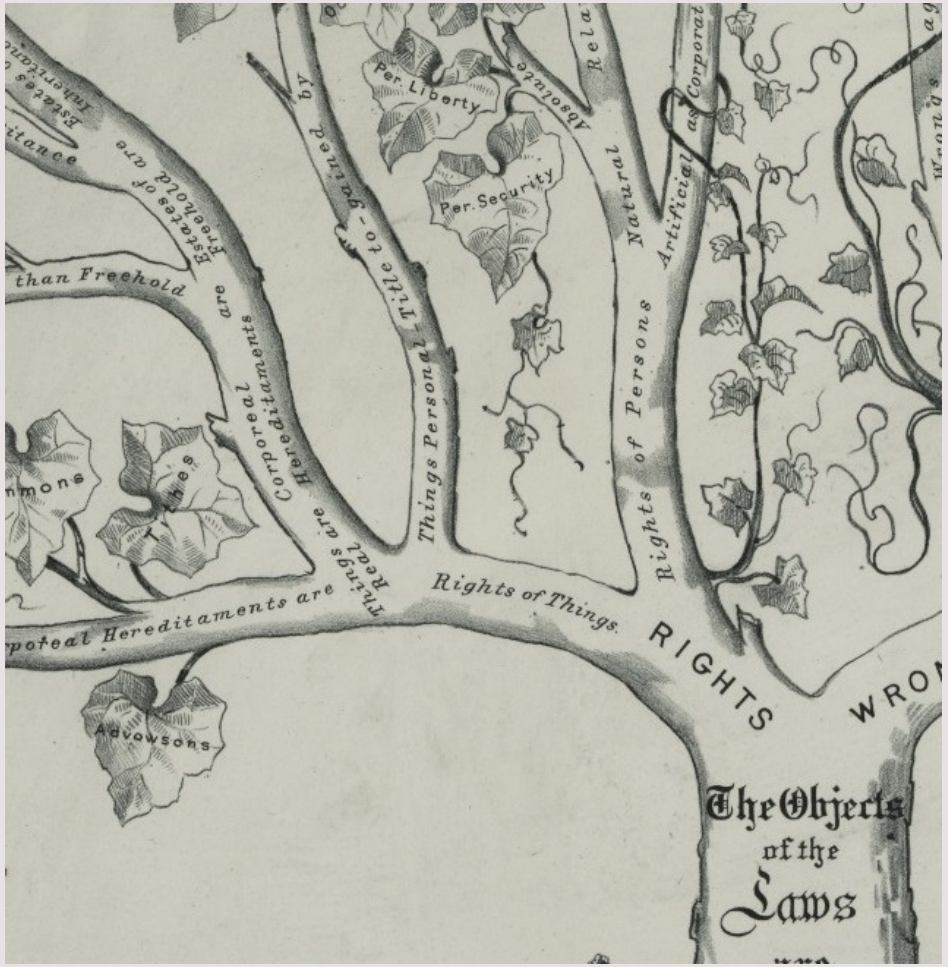
https://www.onezoom.org/

Linear Layouts



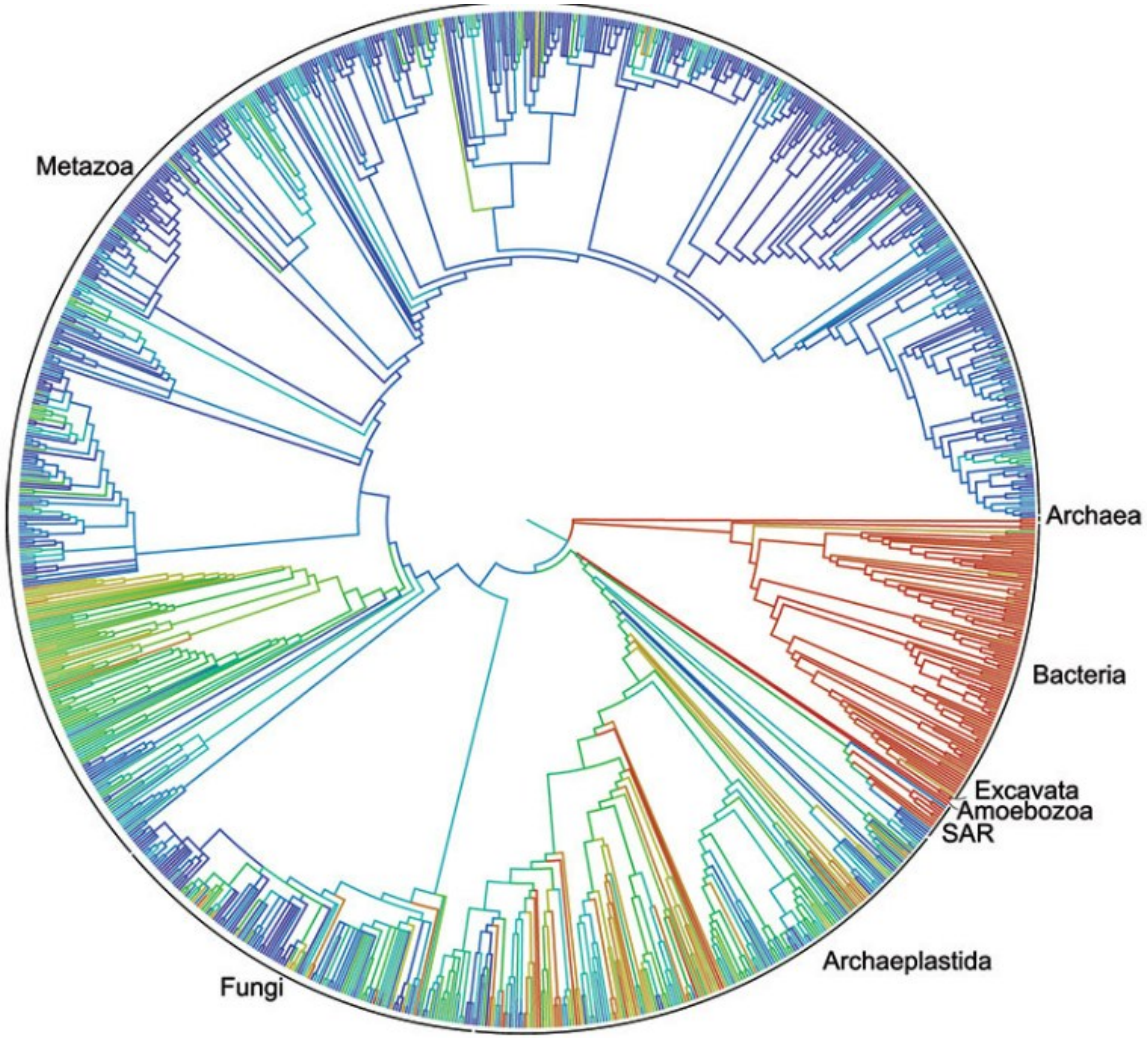
Tree created from Google's n-gram dataset (three-word phrases)

From https://www.chrisharrison.net/index.php/Visualizations/WebTrigrams

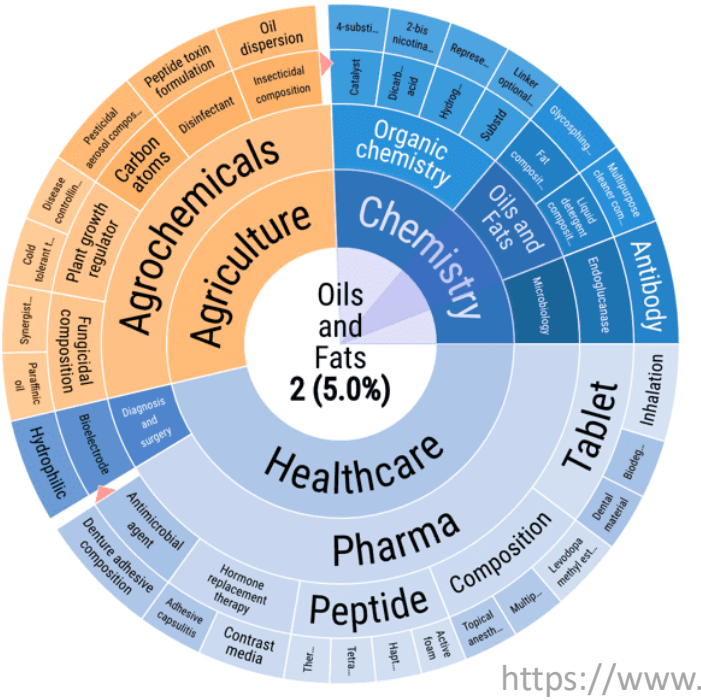


The common law in the similitude of a tree by R.C. Bierce, 1878. https://www.loc.gov/resource/pga.00530/

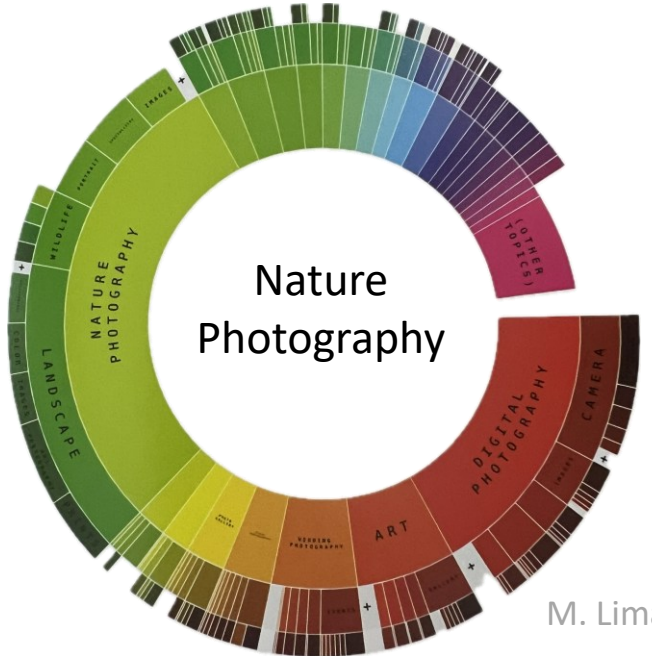
RADIAL LAYOUTS



https://www.science.org/content/article/first-comprehensive-tree-life-shows-how-related-you-are-millions-species



https://www.lexisnexisip.com/resources/must-have-patent-visualization-tools/

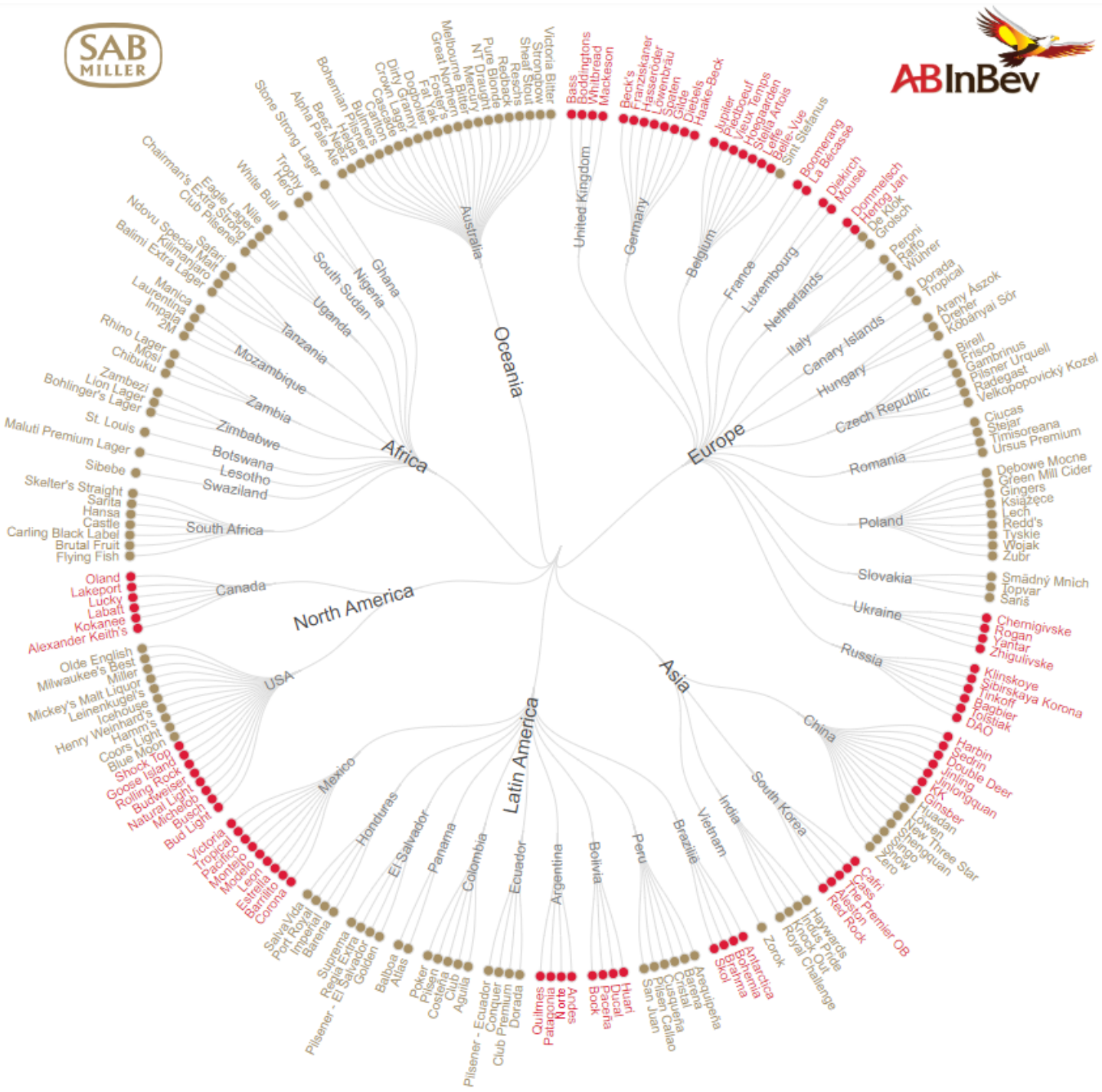


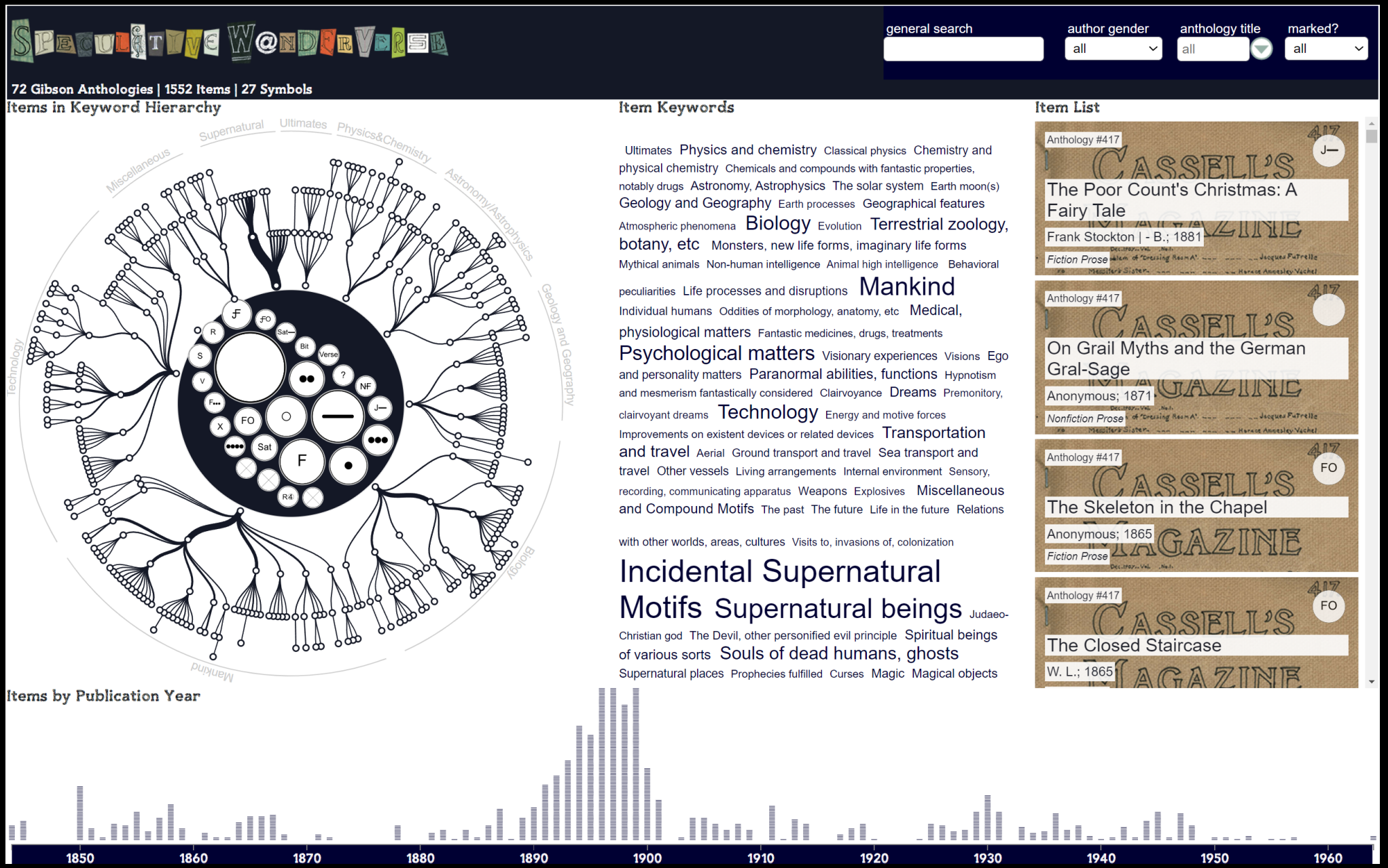
M. Lima, *Book of Trees,* p183.

RADIAL LAYOUT EXAMPLE

The 200+ Beers Brands of SAB InBev

https://www.maartenlambrechts.com/2015/10/16/interactive-the-200-beer-brands-of-sab-inbev.html

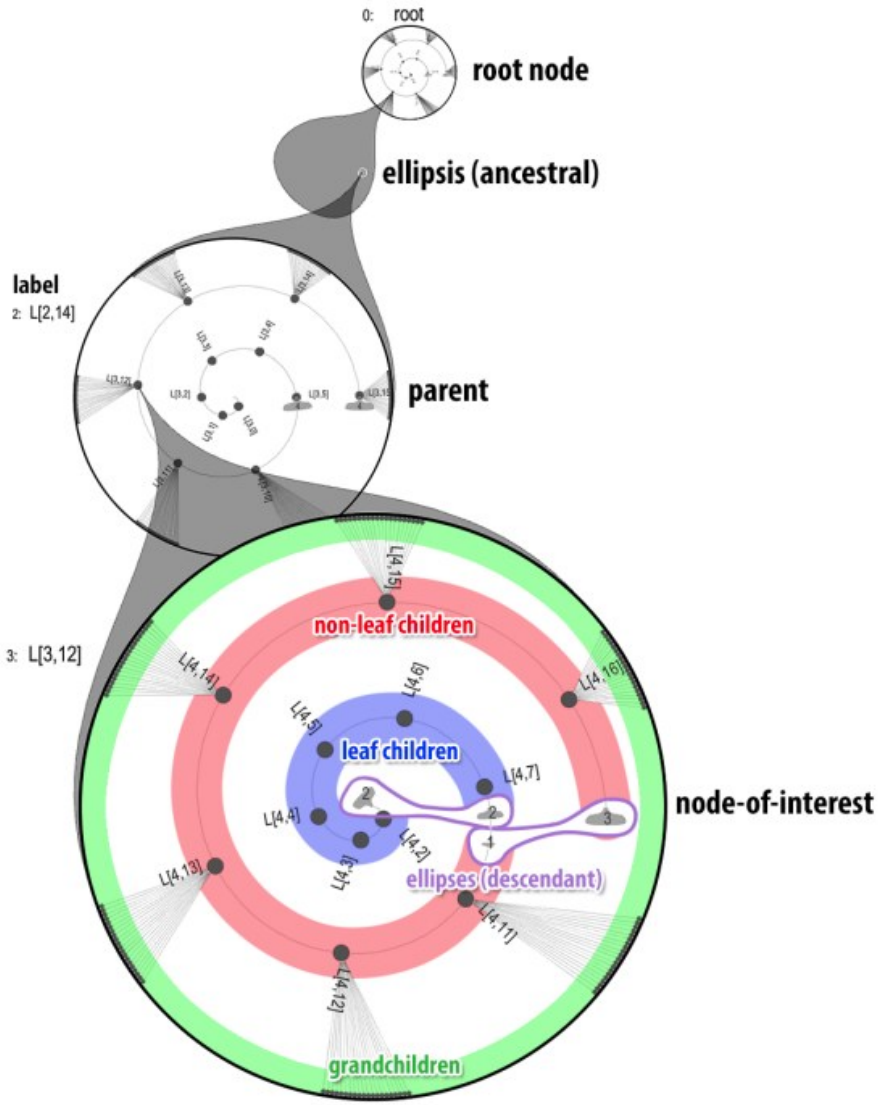


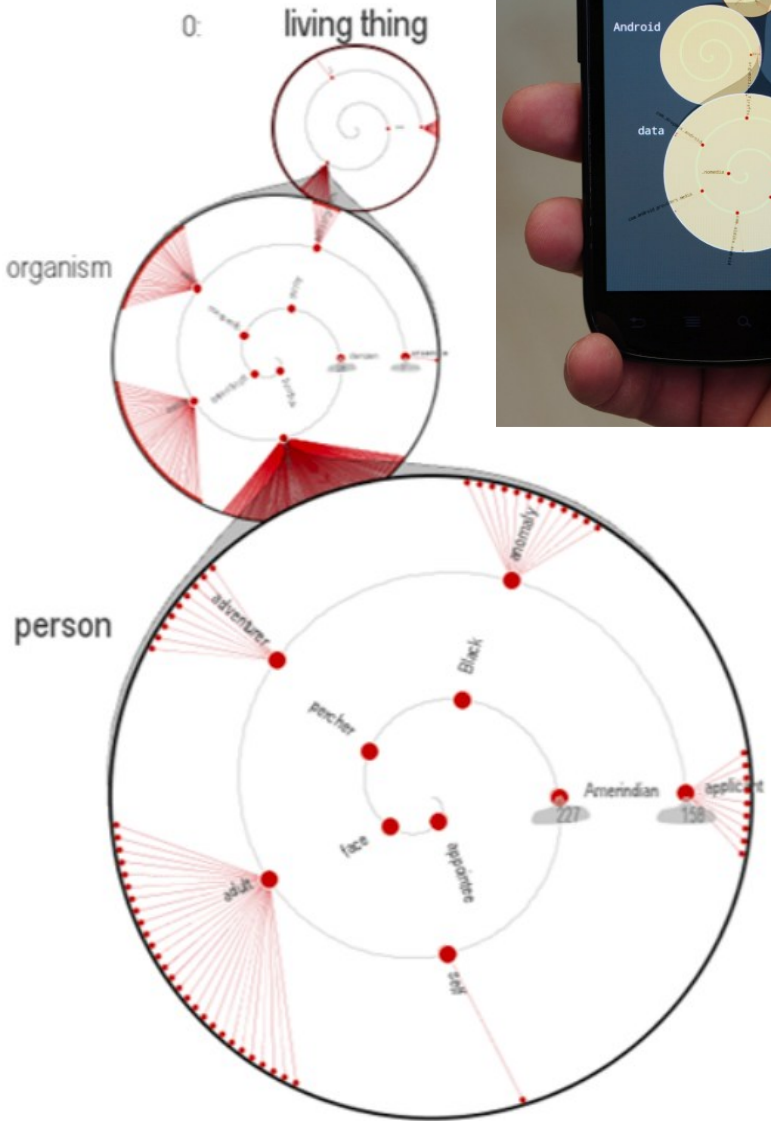


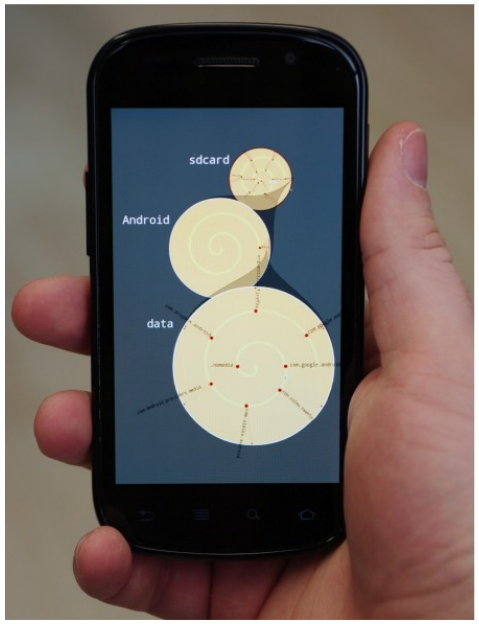
http://stuffofsciencefiction.ca/vis/Wonderverse/

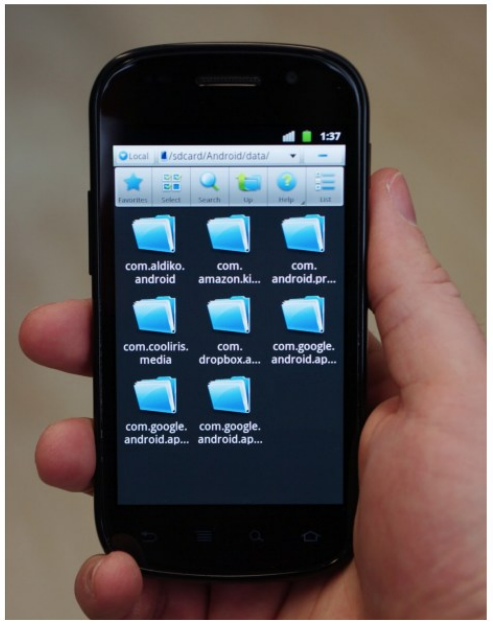
Uta Hinrichs, Stefania Forlini and Bridget Moynihan. Speculative Practices: Utilizing InfoVis to Explore Untapped Literary Collections. IEEE TVCG, 2016.

PAISLEY TREE





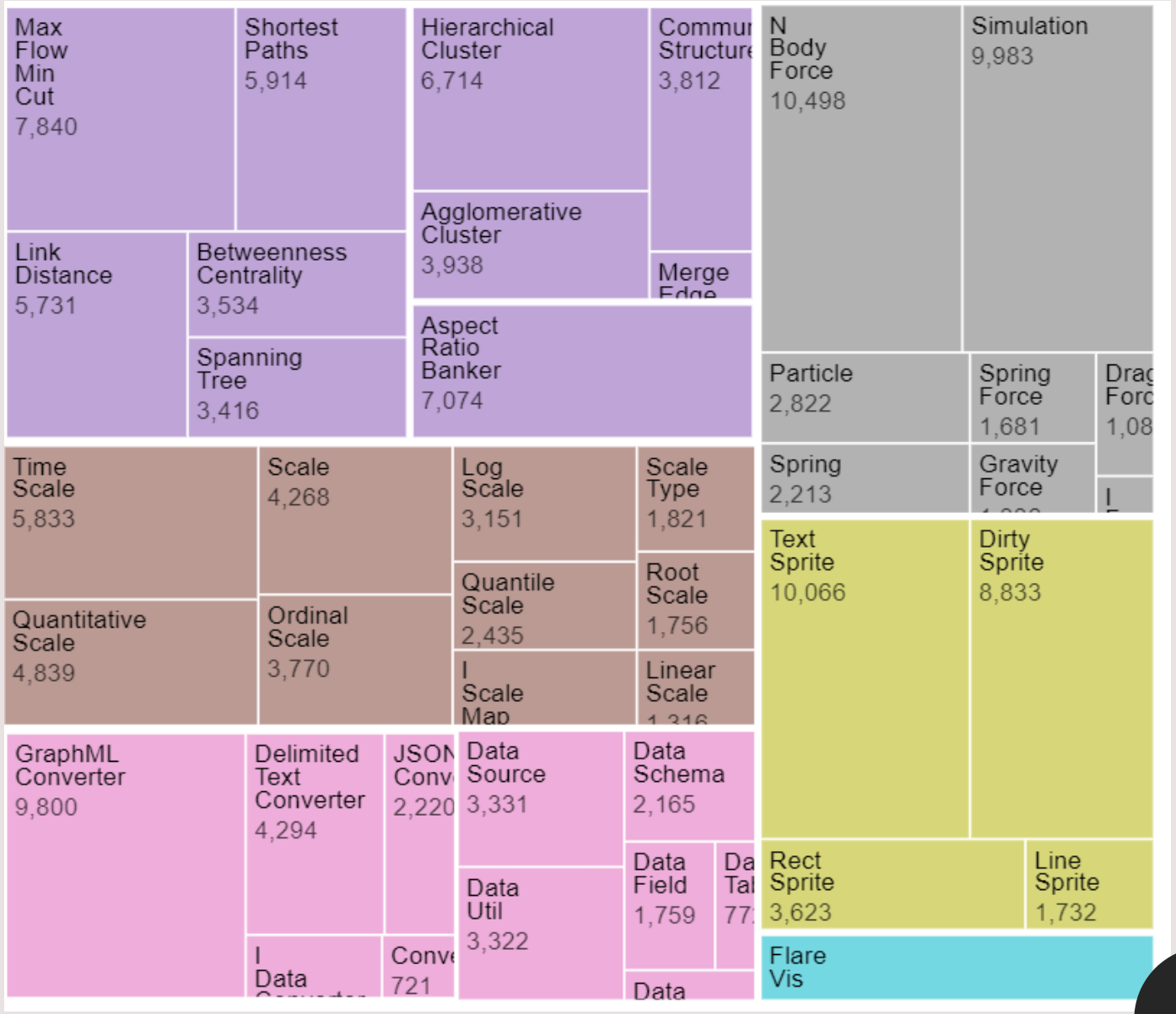




Etemad K, Baur D, Brosz J, Carpendale S, Samavati FF. PaisleyTrees: A Size-Invariant Tree Visualization. EAI Endorsed Trans Creat Tech. 2014.

Make trees by containing areas:

TREEMAP

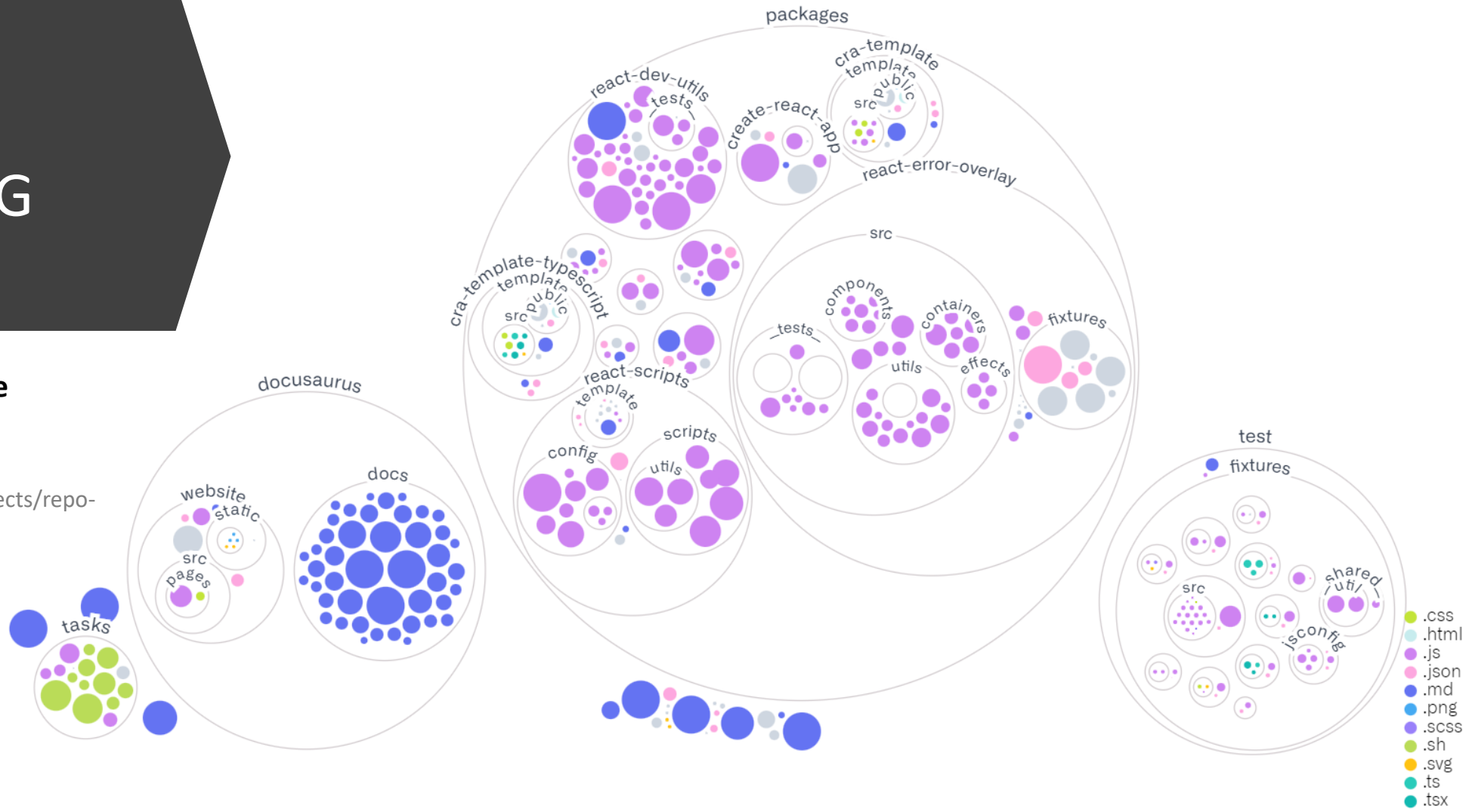


https://observablehq.com/@d3/treemap

CIRCLE PACKING

Visualizing a Software Repository

https://githubnext.com/projects/repo-visualization/



https://treevis.net/

Tools for Graphs & Trees

Applications / Webapps	*Gephi*, *Cytoscape*, *Circos*, *NanoHistory*, *Onodo*
Drawing Tools	*Powerpoint*, *MS Visio*, *Miro*, *Adobe Illustrator*
Code Libraries	*D3*, *ggraph*, *p5.js*, *Cytoscape.js*, *sigma.js*, *tldraw*, *Raphael*, *grano*
Treemaps	*Tableau*, *PowerBI*, *Flourish*, *Datawrapper*

* This is a very small selection of the many tools out there, keep searching if these don't meet your needs.

Topics

Qualitative Data

Visualization

Techniques

- Coloured Text/Heatmaps
- Hybrid Qual/Quant
- Word Clouds
- Sets – Venn & Upset
- Timelines
- Graphs/Networks, Trees

THANKS FOR COMING

FEEDBACK APPRECIATED!

John Brosz
JDLBROSZ@UCALGARY.CA

SLIDES: BROSZ.CA/SLIDES