

*CSE512 :: 27 Feb 2014*

# Final Project

# Progress Presentation

*@EEBo37*

CSE512 Students University of Washington

D5

# D5

## *Data Driven Documents Driving Development*

Fahad Pervaiz & Trevor Perrier

# Data Driven Development




# Data Driven Development

BILL & MELINDA  
GATES *foundation*



# DHIS2: District Health Information System 2

 Kenya HIS

MaintenanceServicesHelpLog out

Report

Standard Report  
Dataset Report  
Data Completeness Report  
Static Report  
Distribution Report  
Tally Sheet Generator  
Report Table

Chart

Chart

Pivot Table

Pivot Table

Download as Excel

Download as CSV

Download as PDF

Download as Jasper

Download as JRXML

Back

Immunization Numbers

Period	Organisation unit	Fully immunised	DPT3 doses given	Measles doses given	DPT1 doses given
October 2010	Bamba Medical Clinic	0.0	0.0	0.0	0.0
October 2010	Uzima MC (Kilifi)	0.0	0.0	0.0	0.0
October 2010	Mwanzo Medical Clinic	0.0	0.0	0.0	0.0
October 2010	Menengai Health Care Serv	0.0	0.0	0.0	0.0
October 2010	Mtwapa Dispensary	153.0	127.0	153.0	117.0
October 2010	Vipingo Health Centre	76.0	102.0	76.0	107.0
October 2010	Union Medical	0.0	0.0	0.0	0.0
October 2010	Amani Family Medical Clin	0.0	0.0	0.0	0.0
October 2010	Michaela Denis MC	0.0	0.0	0.0	0.0
October 2010	Ganze Dispensary	29.0	0.0	29.0	76.0
October 2010	Vutakaka Medical Clinic	2.0	0.0	2.0	0.0
October 2010	Chumani Medical Clinic	2.0	8.0	2.0	9.0

# DHIS2: District Health Information System 2



Current report visualizations are great for summary data analysis but not data exploration.

- No way to drill down into provincial level from national level.
- No interactive maps
- Complicated menu system for building reports.

# DHIS2: District Health Information System 2

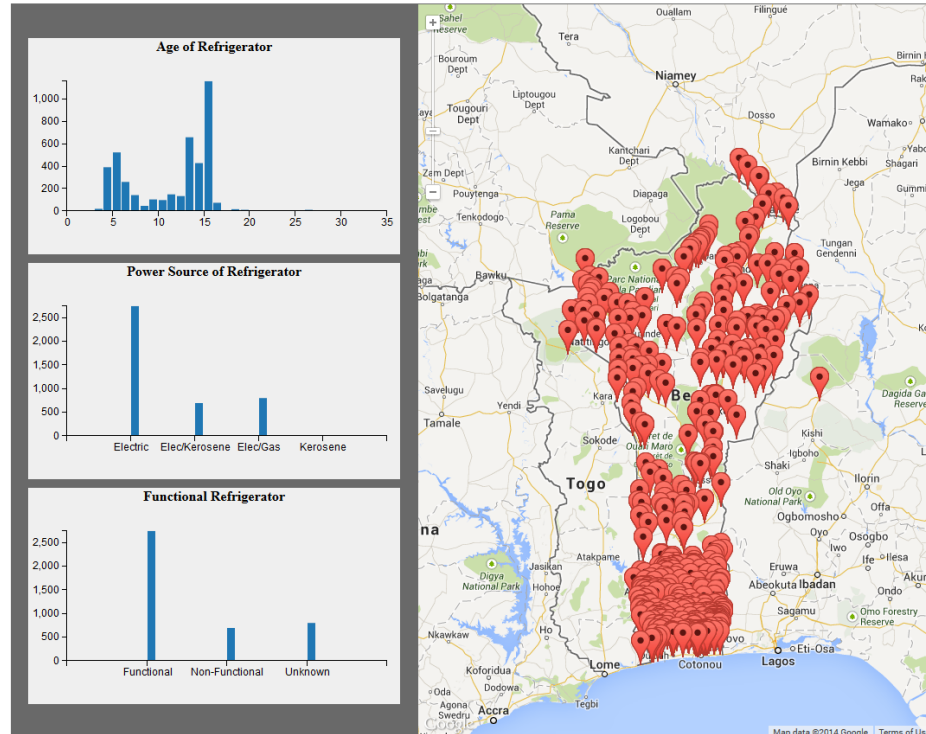


Current report visualizations are great for summary data analysis but not data exploration.

- No way to drill down into provincial level from national level.
- No interactive maps
- Complicated menu system for building reports.

Interactivity is key to data exploration, but this must be tied to the proper search queries to the underlying database.

# Example: A3



# What we plan to do.....

- Build a data exploration application to run on top of DHIS2
  - Focus on interactivity
    - Drill down from country  $\Rightarrow$  province  $\Rightarrow$  district  $\Rightarrow$  clinic
    - Easily add and remove data from the graph
    - Use brushing to filter graphed data.
  - Easily switch between views of data
  - Work with the DHIS2 app store so that any DHIS implementation can use our visualization tool.

# Feedback?

1. What kind of data exploration are generally done by managers?
2. What data types are common for this exploration?
3. What are the most common ways to drill-down information on a visualization?
4. What drill-down operations works best on multidimensional dataset and are not complicated for our target population?
5. How to handle data density when user drills down on one dimension or branch and expands the visualization?

(Presented by Trevor Perrier and Fahad Pervaiz)

# Dynamic News Timelines



NOV  
2008

Michael Beswetherick & Joe Kohlmann



FEB  
2014



Currently

The New York Times Provides  
Aggregate News Pages for Certain  
Subjects (such as the POTUS)



## Chronology of Coverage

FEB. 26, 2014

Pres Obama tells Afghan Pres Hamid Karzai that he has instructed the Pentagon to begin planning for a complete withdrawal of American troops from Afghanistan by the end of 2014; Obama also says the United States is still open to leaving a limited military force in the country, message aimed less at Karzai than at whoever will replace him. [MORE »](#)

FEB. 25, 2014

News analysis; Pres Obama has approached 2014 revolution in Ukraine with clinical detachment aimed at avoiding instability; approach is in stark contrast to Pres George W Bush's democracy-promoting response to Ukraine's Orange Revolution of 2004; while sympathetic to Ukraine's pro-Western protestors, Obama has not made global aspirations of democracy the animating force of his presidency. [MORE »](#)

**FEB. 25, 2014**

Louisiana Gov Bobby Jindal's critical comments about Pres Obama in a news conference just after annual meeting of nation's governors at the White House prompts immediate partisan arguing among dozen or so governors assembled just steps from the Oval Office. [MORE »](#)

FEB. 24, 2014

More than 500 leaders of a national network of young immigrants, frustrated that House Republicans say they will not move on immigration in 2014, have decided to turn their protests on Pres Obama in an effort to pressure him to act unilaterally to stop deportations. [MORE »](#)

FEB. 24, 2014

Bill Keller Op-Ed column examines how Pres Obama has lived up to the high expectations that he would reform the broken criminal justice system; contends that so far, the Obama administration has been unimpressive on the issue; expresses hope that the administration will do more in the remaining years of Obama's presidency. [MORE »](#)

**Show More**

[illegible]

## NEWS ANALYSIS

## By PETER BAKER

November 8, 2012 | US | NEWS ANALYSIS

## 4 Years Later, Scarred but Still Confident

By PETER BAKER

September 6, 2012 | US | NEWS

## By ADAM NAGOURNEY

November 5, 2008 | US | NEWS

## Newest First | Oldest First

**Page:** [1](#) | [2](#) | [3](#) | [4](#) | [5](#) | [6](#) | [7](#) | [8](#) | [9](#) | [10](#) | [Next >>](#)

## By MARK LANDLER

President Obama will press Prime Minister Benjamin Netanyahu of Israel to agree to a “framework” for a conclusive round of peace talks, officials



© 2014 The New York Times Company | [Site Map](#) | [Privacy](#) | [Your Ad Choices](#) | [Advertise](#) | [Terms of Sale](#) | [Terms of Service](#) | [Work With Us](#) | [RSS](#) | [Help](#) | [Contact Us](#)  
[Site Feedback](#)

**Newest First** | [Oldest First](#)

**Page:** [1](#) | [2](#) | [3](#) | [4](#) | [5](#) | [6](#) | [7](#) | [8](#) | [9](#) | [10](#) | [Next >>](#)


February 27, 2014, Thursday

**Obama's Pipeline**

By RAÚL M. GRIJALVA

Keystone XL shows how business still drives environmental policy.

February 27, 2014, Thursday

An illustration of an oil pumpjack (nodding donkey) and a pipeline. The pumpjack is on the right, with a long pipe extending from it towards the left. The background is a light blue sky with a few clouds. The entire illustration is enclosed in a thin black rectangular border.

By RAÚL M. GRIJALVA

**Keystone XL shows how business still drives environmental policy.**

February 27, 2014, Thursday



By JONATHAN MARTIN and MEGAN THEE-BRENAN  
Though most Americans agree more with Democratic policy positions, the Republican Party is benefiting from the support of self-described independents.  
February 27, 2014, Thursday

# To Pay for Infrastructure Repairs, Obama Seeks Tax Changes

By MICHAEL D. SHEAR and CORAL DAVENPORT

President Obama will ask Congress to overhaul corporate and business taxes to pay for a four-year, \$302 billion transportation bill.

February 27, 2014, Thursday

[illegible]





# Currently

The NYT Provides Aggregate  
News Pages for Certain Subjects  
(such as the POTUS)



...Later

A Dynamic Timeline for News Articles  
With Interactive Exploration and  
Linking of Related Subjects' Timelines

O

NOV  
2008

O

FEB  
2014

Bounded  
by starting  
node or date

Approximately  
concurrent  
events

Again, what does position  
mean here? Might need to

do API tests  
to see what data can  
be encoded vertically.



NOV  
2008



(Rects are a  
bit big here)

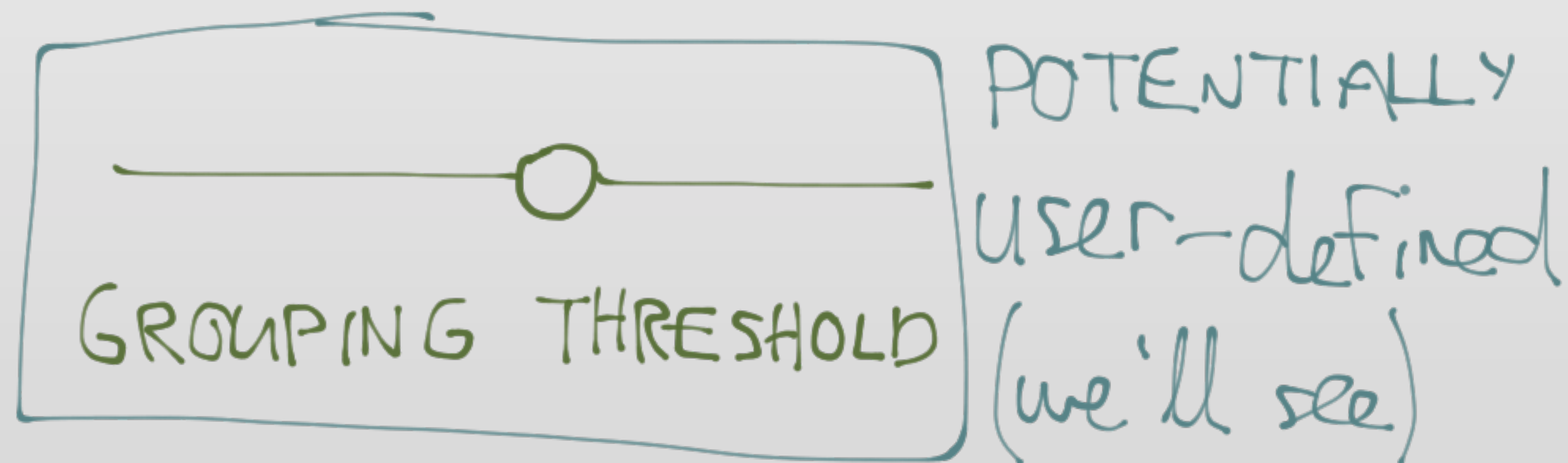
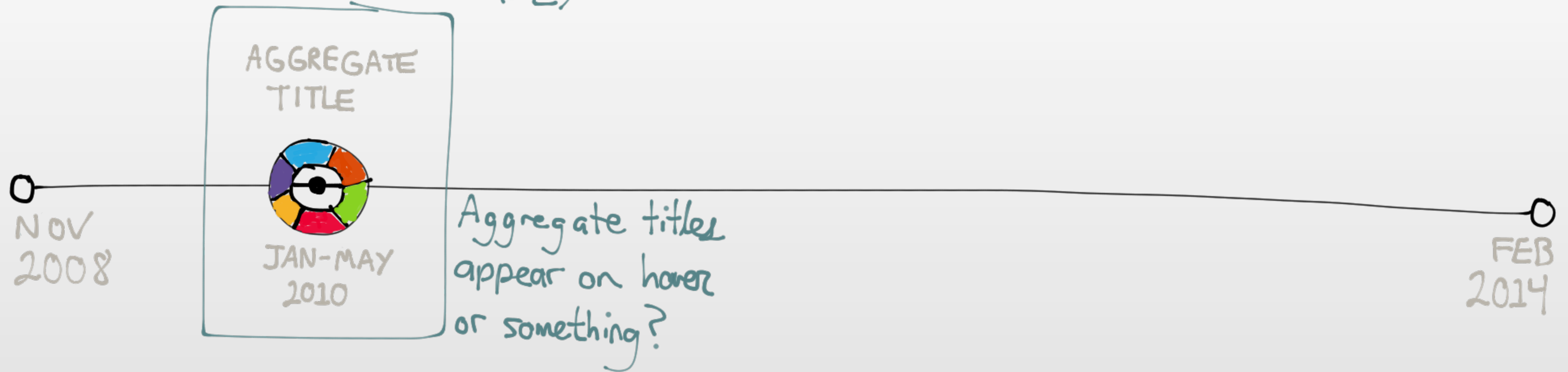
Though: they might be  
top-sized now?



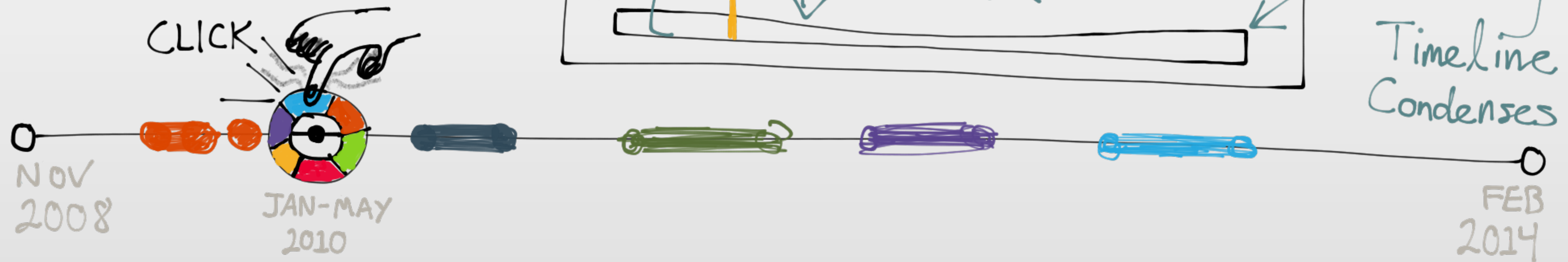
FEB  
2014

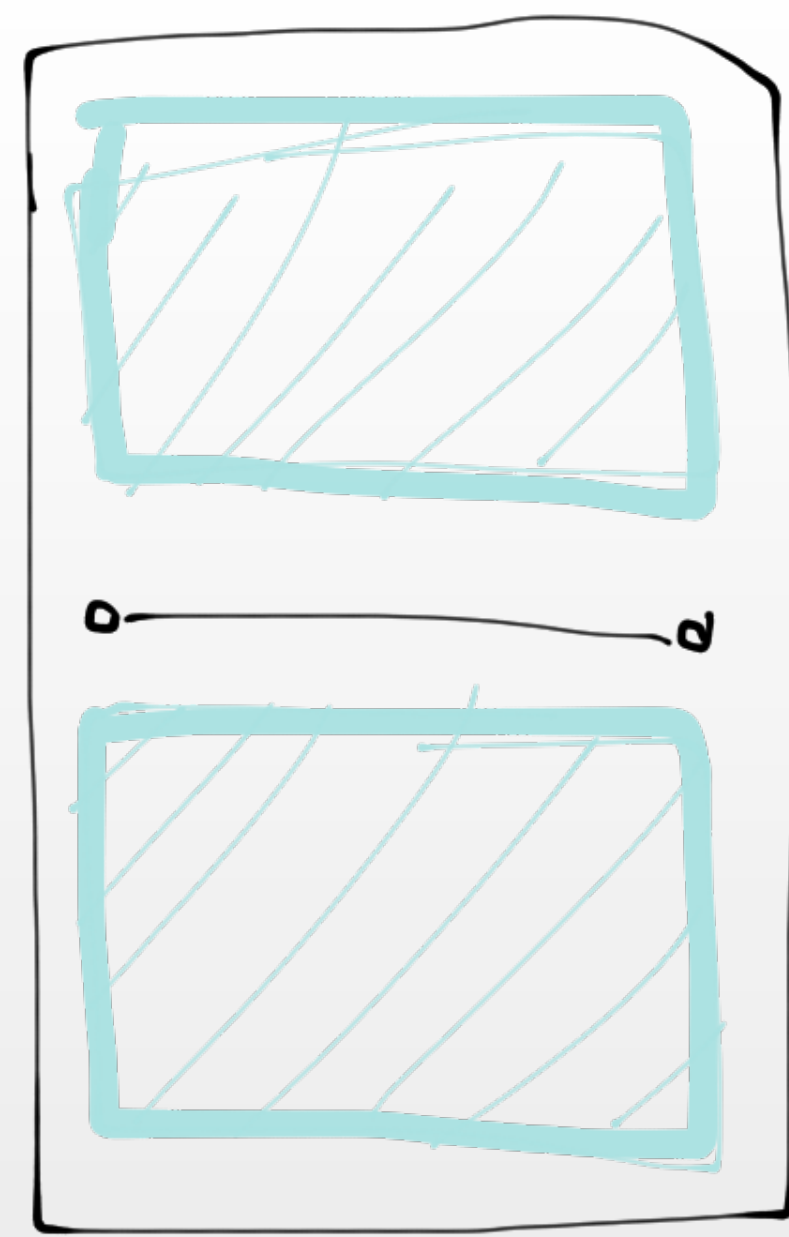
A timeline  
with start/end  
point labels

# MULTIPLE STORIES AGGREGATED CHRONOLOGICALLY



Transition to  
a new, linked  
timeline

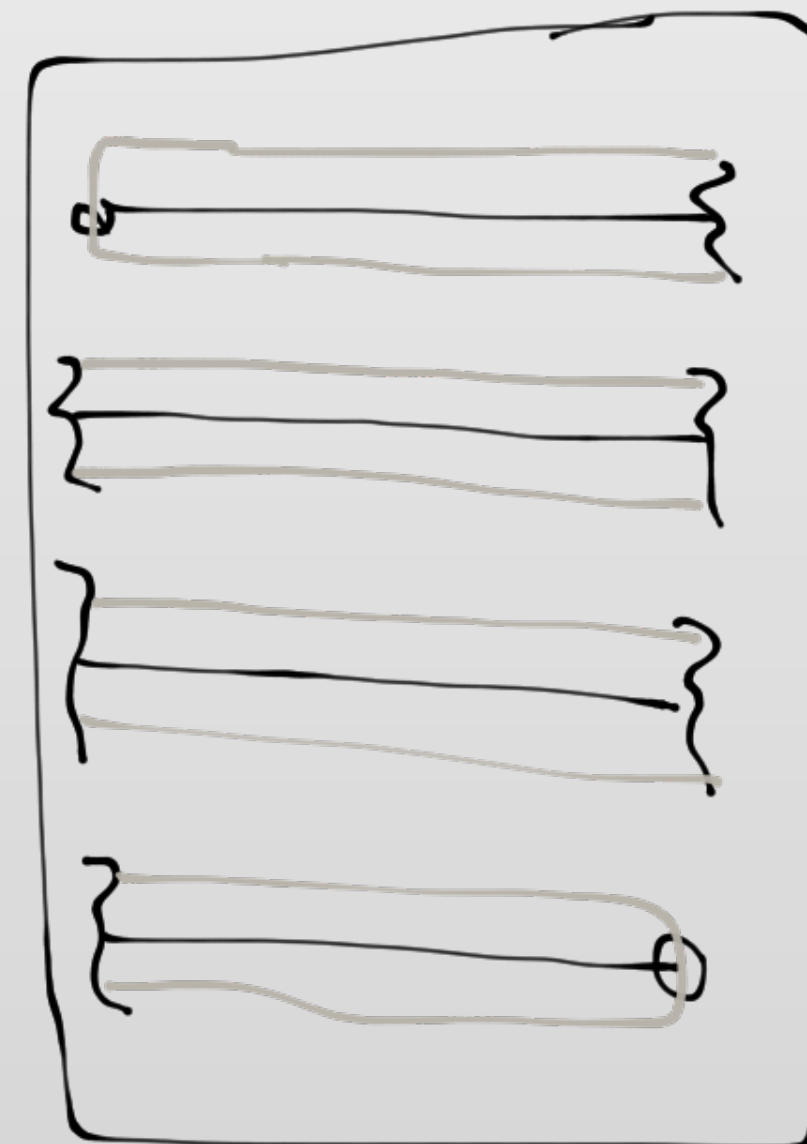




↑ The average  
browser window

← WASTED  
AREA  
←

CAN WE WRAP  
THE TIMELINE?



OBVIOUS  
INSPIRATION  
FROM  
FCPX

Details on Demand

Efficient Use of Area

Lenses

Linked Displays

Narrative Visualization

Summarization

# ThemeRiver Havre et al, 2000

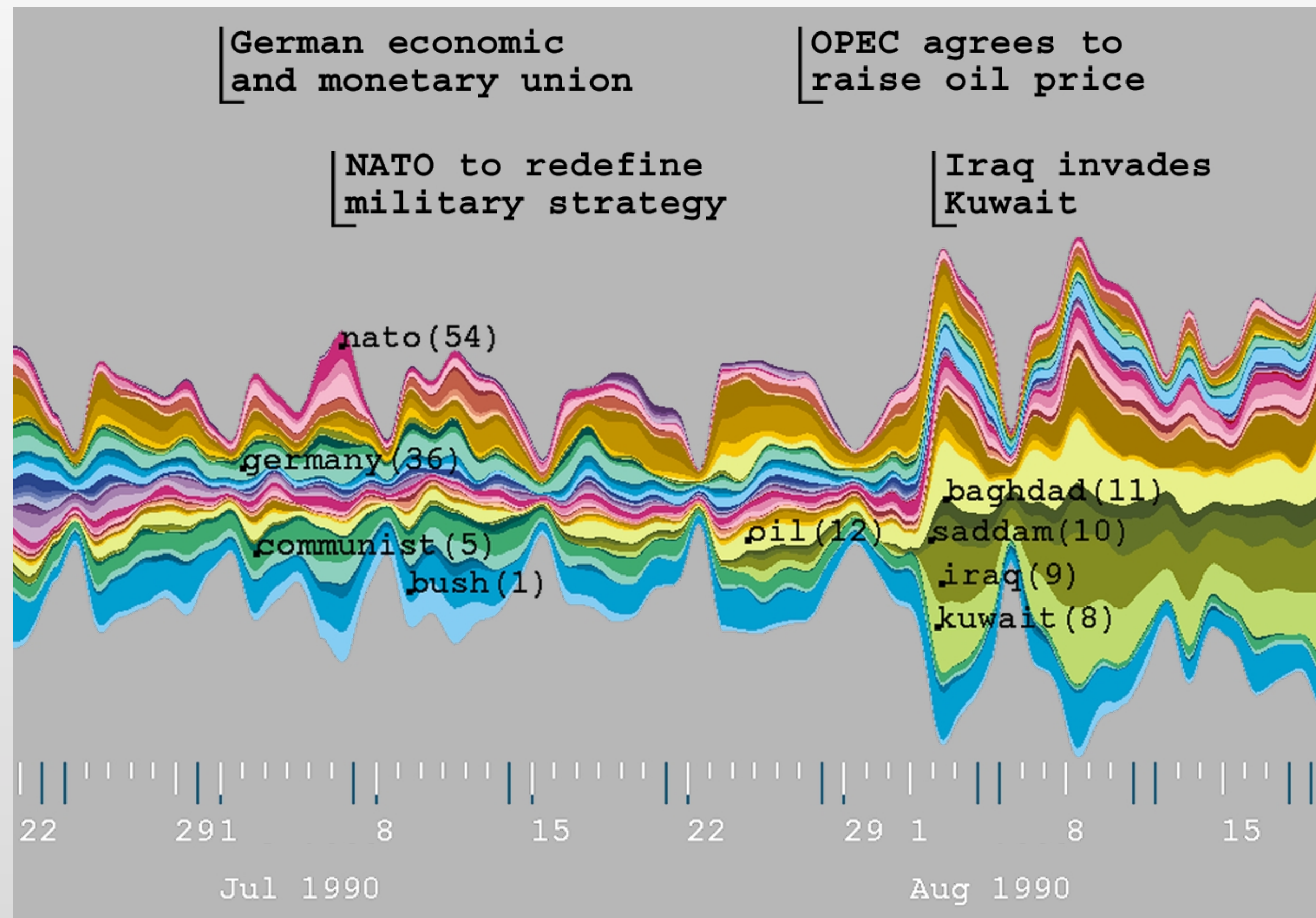


Figure 3: AP data from July - August 1990. A wide current in the river indicates heavy use of a topic, while changes in color distribution correlate to changes in themes.

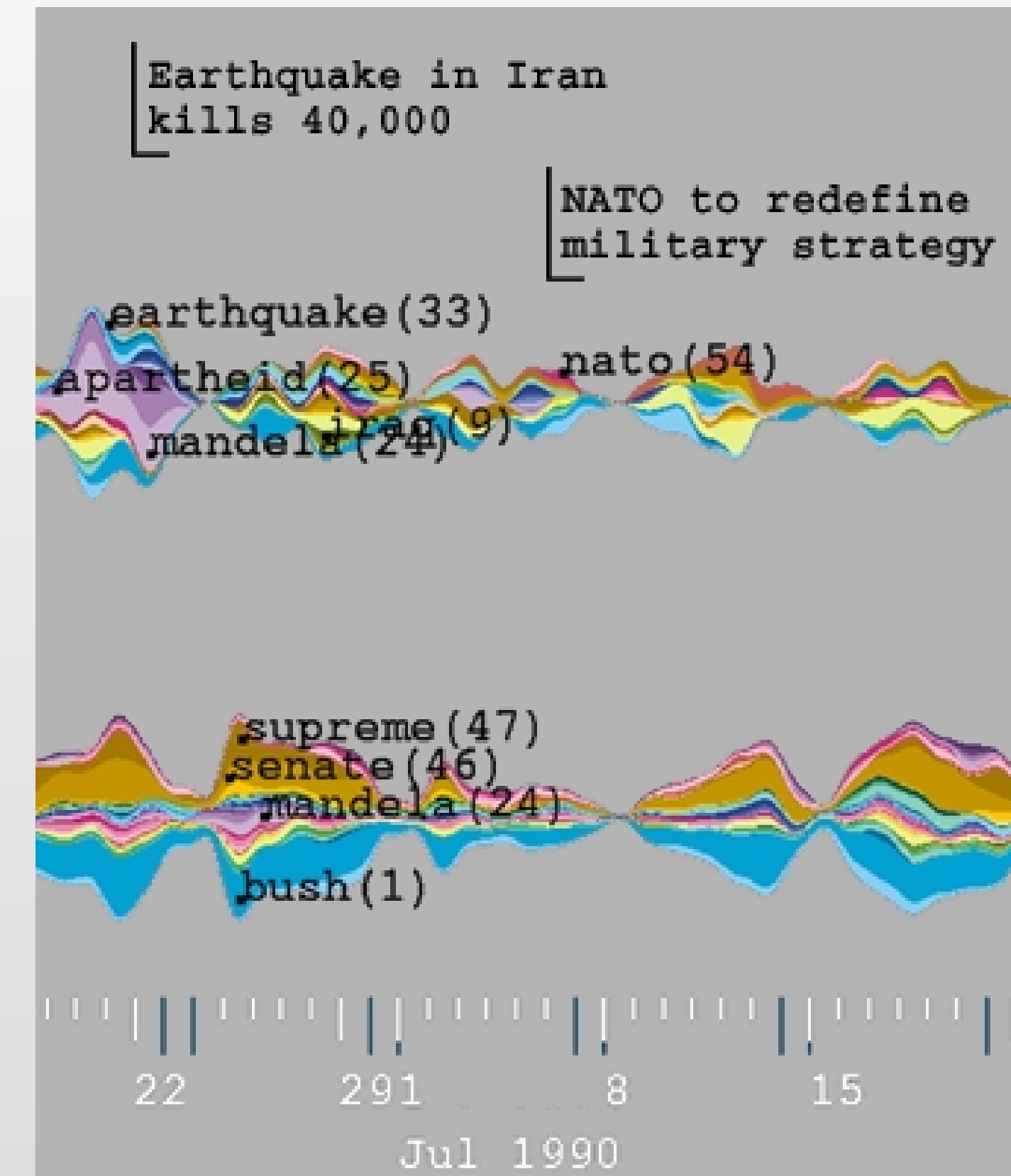
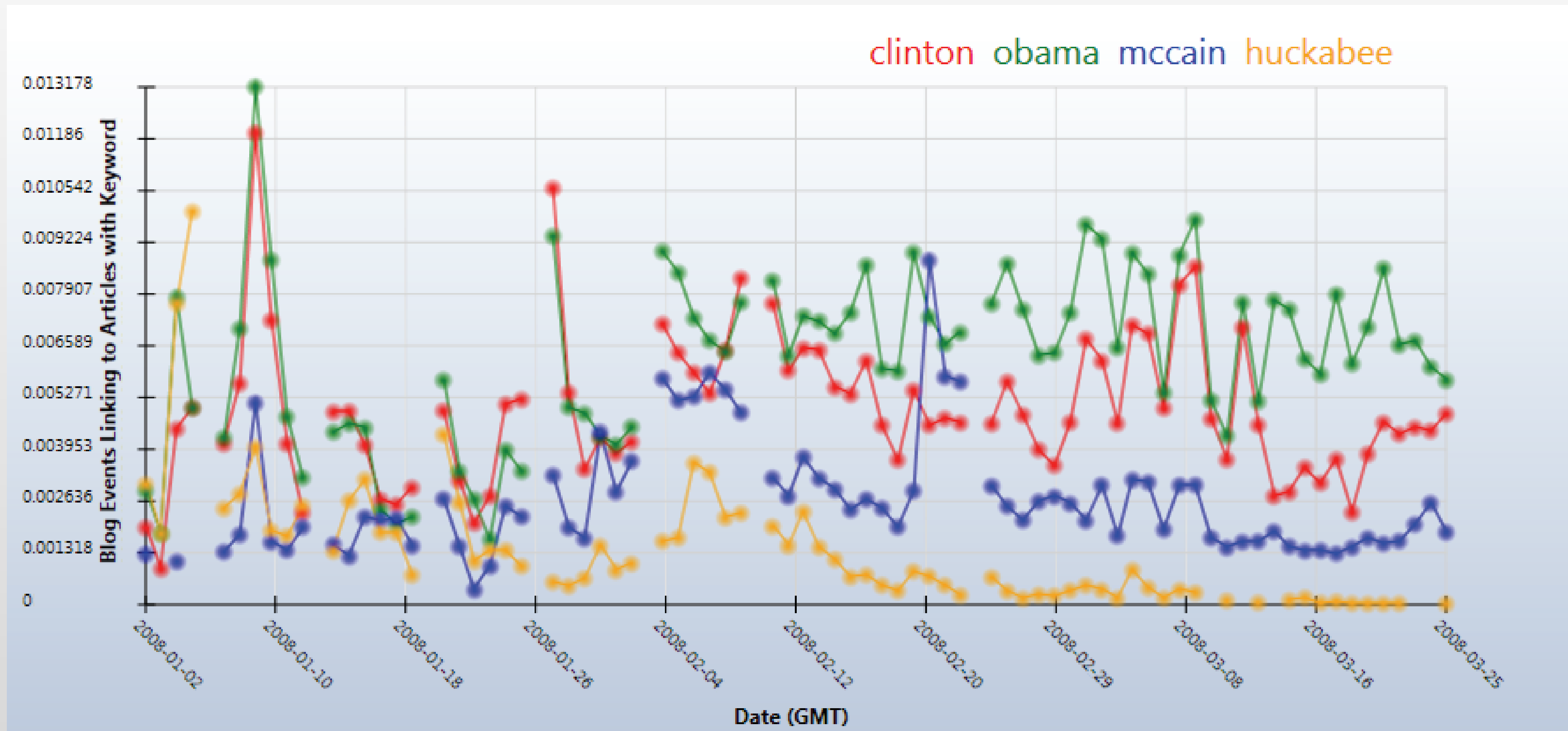


Figure 6: Parallel rivers let users compare AP data from Washington, D.C. and New York from the same time period.

# Narratives Fisher et al, 2008







# Front Row to Fashion Week

By MIKE BOSTOCK, SHAN CARTER, ERIK HINTON, CATHY HORYN and ERIC WILSON | September 12, 2013

Of the more than 300 collections shown during New York Fashion Week, here were the ones that created the most buzz and left the biggest impressions on fashion editors as they headed off to the next round of shows in London, Milan and Paris.

[View Full Screen](#)

## Calvin Klein

A beautiful, innovative collection in which Francisco Costa layered references to urban tribes, '80s art, handcraft and even, seemingly, radical chicks of the 1920s. It added up to a modern expression of fashion.

Read more: [Calvin Klein in Full Color](#)



Sand-colored, orange-accented canvas wrapped into a dress and suit

A large emerald tweed coat with frayed, pronounced seams

A boxy black jacket fringed with multicolored confetti strings

# Project Plan

3/04 Prepare Initial Data Set

Working demonstration of project

3/06 Finish View Experiments

Figure out if we can use the live NYT APIs (or not)

3/11 Finalize Overall Interaction Design

Have a full plan for user interactions and events

**3/13 Final Project Poster Presentation**

Working demonstration of project

**3/20 Final Project Paper and Software Due**

Finish the paper, write documentation, and get the final code on GitHub

# Questions & Feedback

Michael Beswetherick & Joe Kohlmann

# **Visualization on Cosmological Halo Merger Trees**

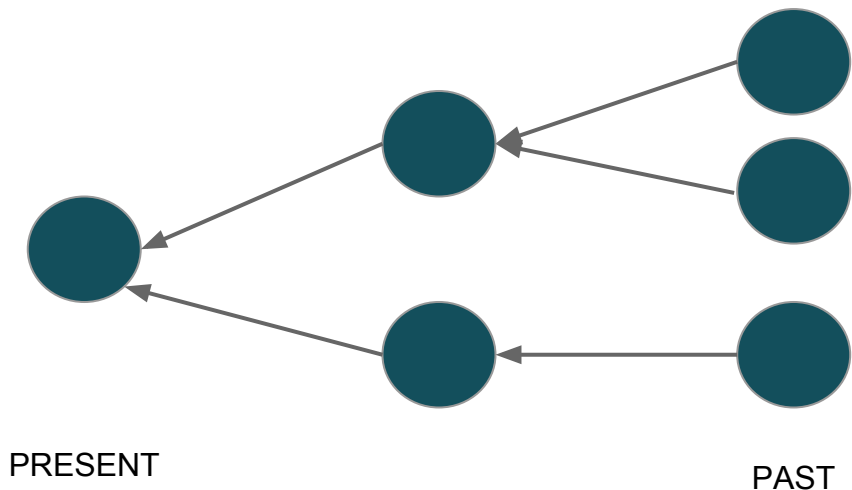
Laurel Orr and Jennifer Ortiz  
CSE 512

# Halo Merger Trees

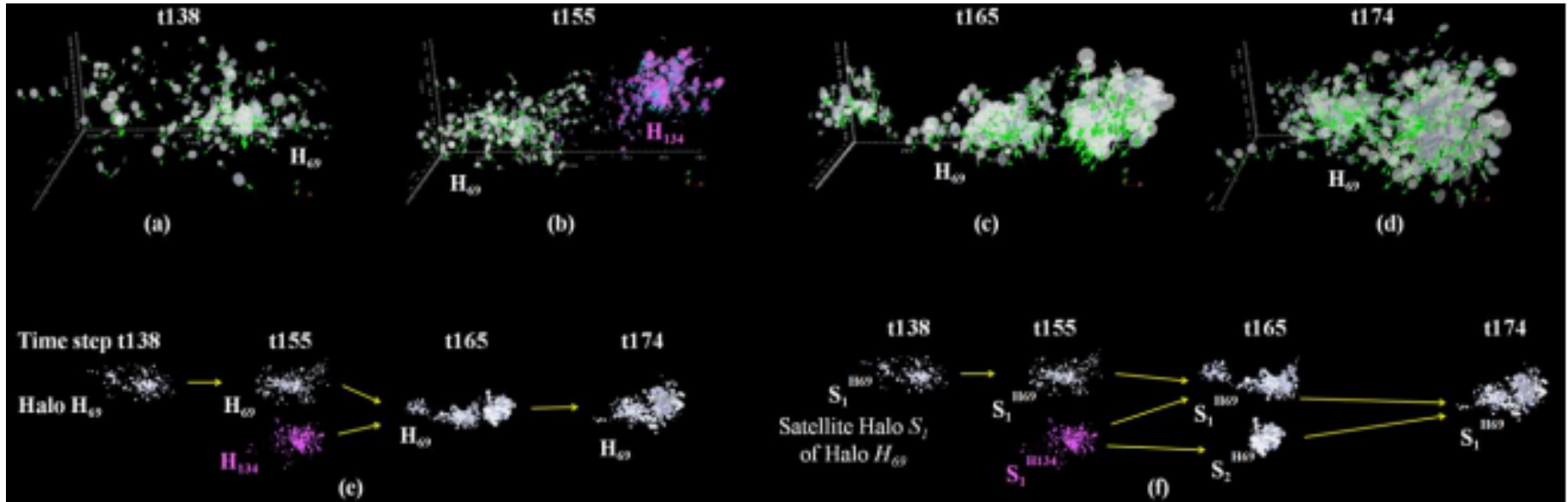
**Halo:** dense clumps of particles that interact and merge

**Merger Tree:** shows the history of joining (merging) and splitting of halos over time

Scientists want to compare and analyze structures of different trees



# Previous Work



Takle, J.; Silver, D.; Heitmann, K., "A case study: Tracking and visualizing the evolution of dark matter halos and groups of satellite halos in cosmology simulations," Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on, pp.243,244, 14-19 Oct. 2012.

# Our Task

320,589 halos with 286 present day roots across 27 timesteps

Create interactive visualization for astronomers to view and analyze halos

- Zooming
- Comparison between merger trees
- Path tracing of important halos
- Filtering
- Node/Path Annotations

Filters

Merger Tree

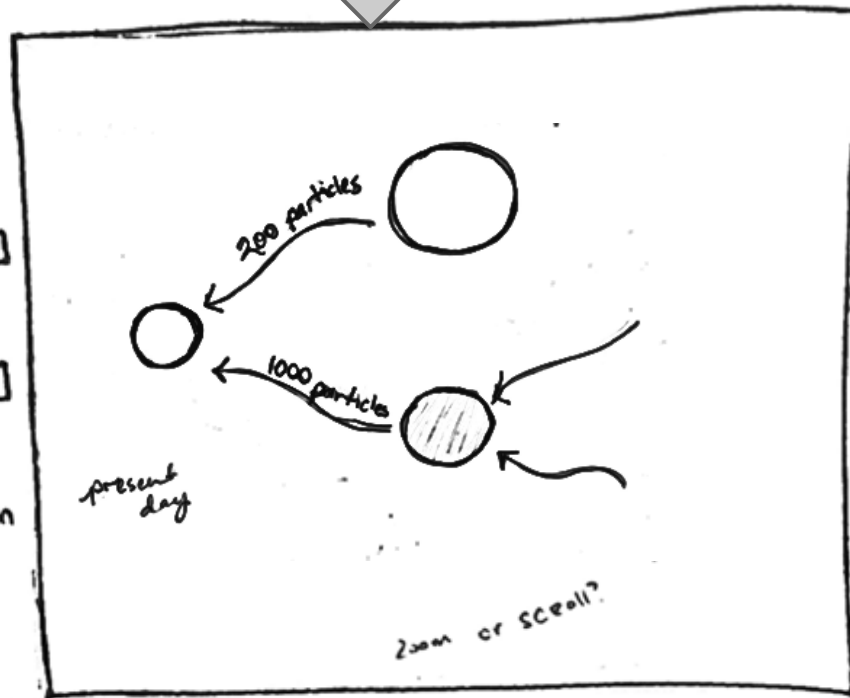
Present Day  
[ 3 ] [ ▾ ]

Mass  
[ 10 ] [   ]

particle count  
[ < ] [   ]  
[ > ] [   ]

☐ dark only  
☐ show stellar formation

GO



structure  
[   ]

Find Similar  
[   ]

Halo stats

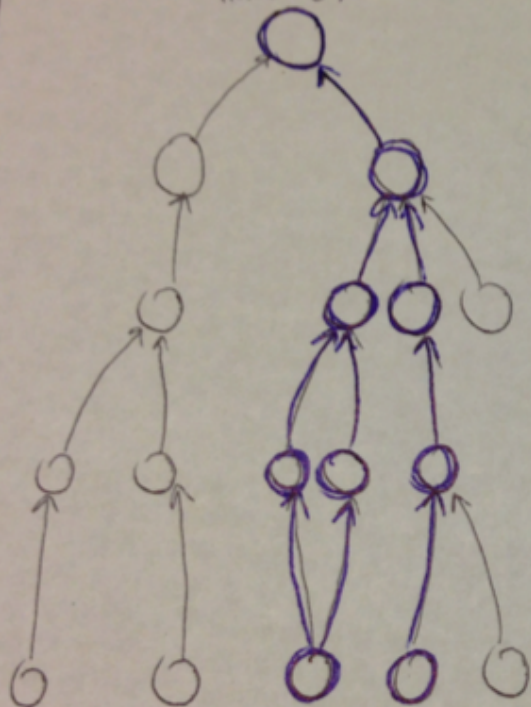
[   ]

← by lower

Halo Stats

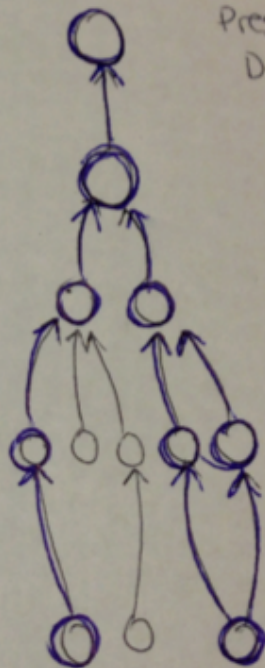
# Comparing Two Halos

Halo 87



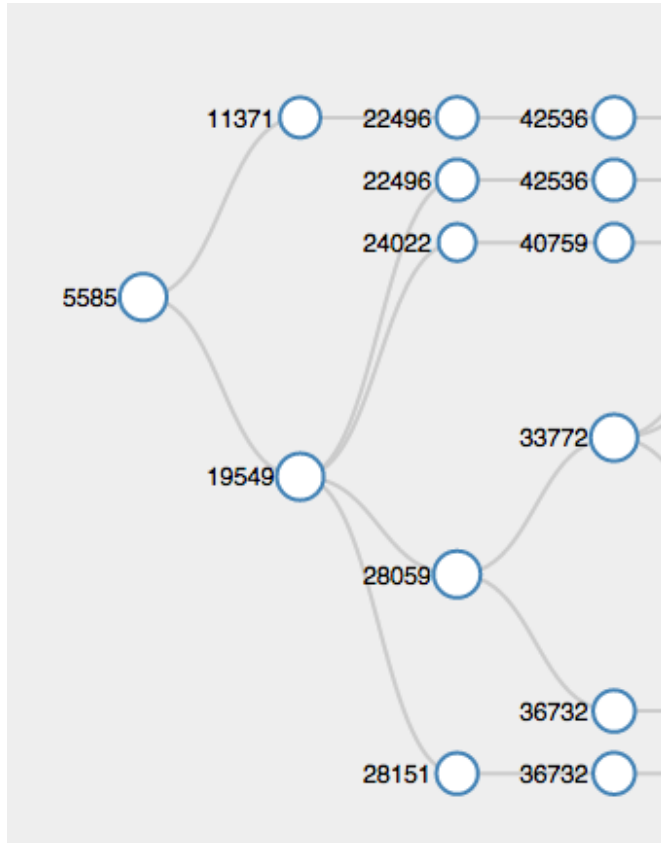
Halo 25

Present Day



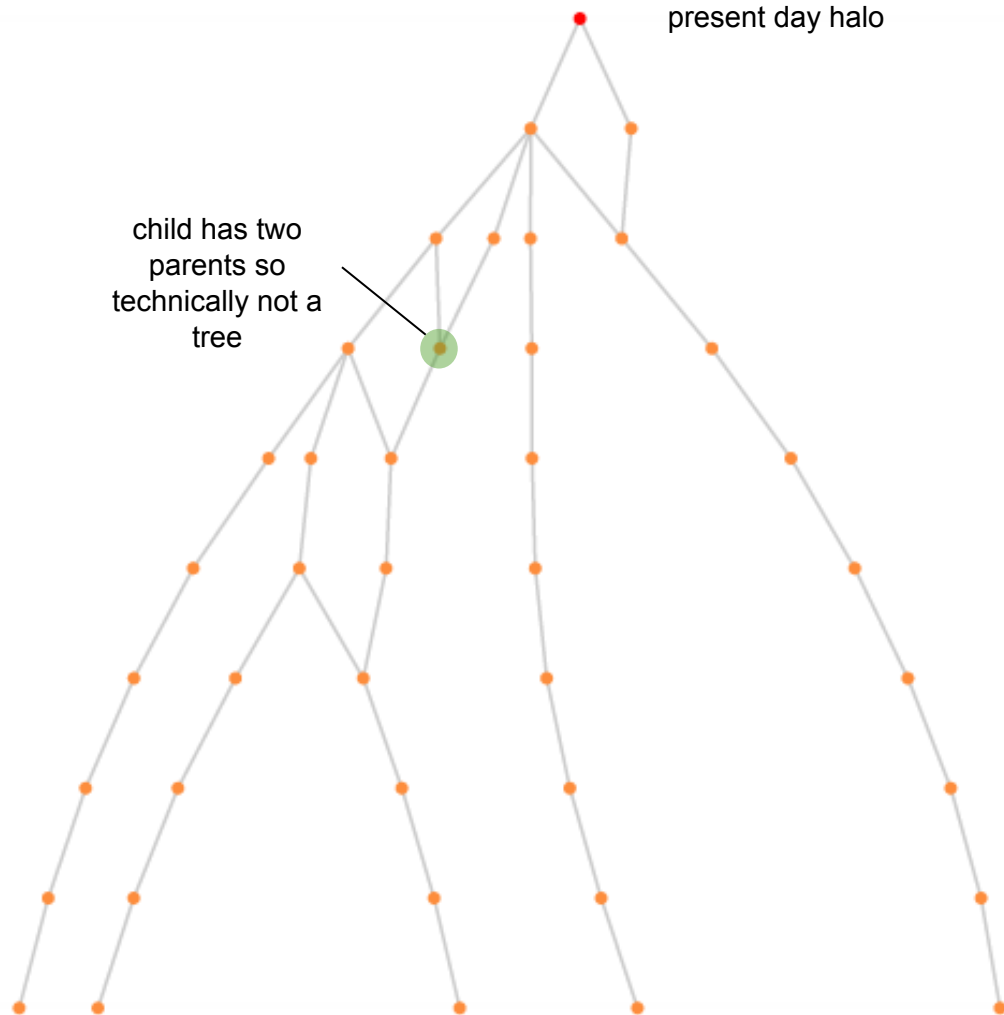
Subgraphs with  
Same Structure

# Collapsable Tree



# Force-Directed Graph

Using D3's force directed layout with constraints to make it look like a tree



# Help!

- How do we show a global view of a tree while allowing for zooming? How can we summarize the graph?
- Should the users be allowed to move the nodes?
- How can we allow for tree comparison?

(Presented by Laurel Orr, Jennifer Ortiz)

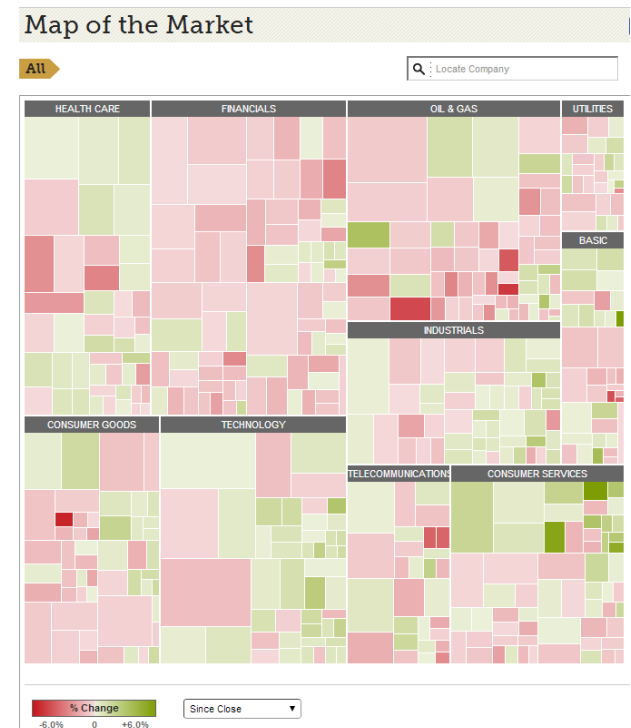
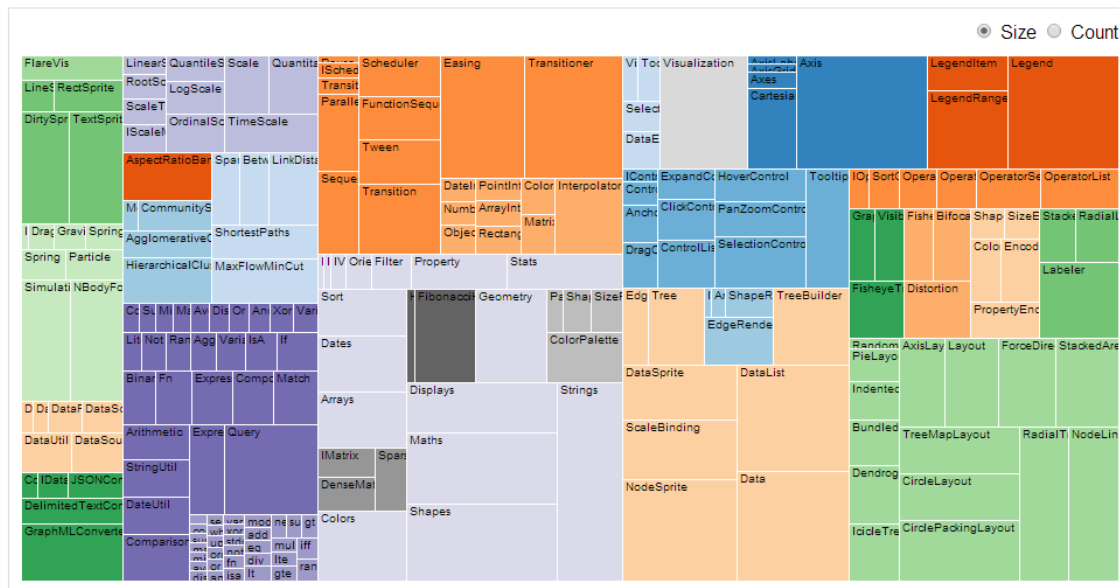
# Voronoi Treemaps in D3

Paul Vines

Peter Henry

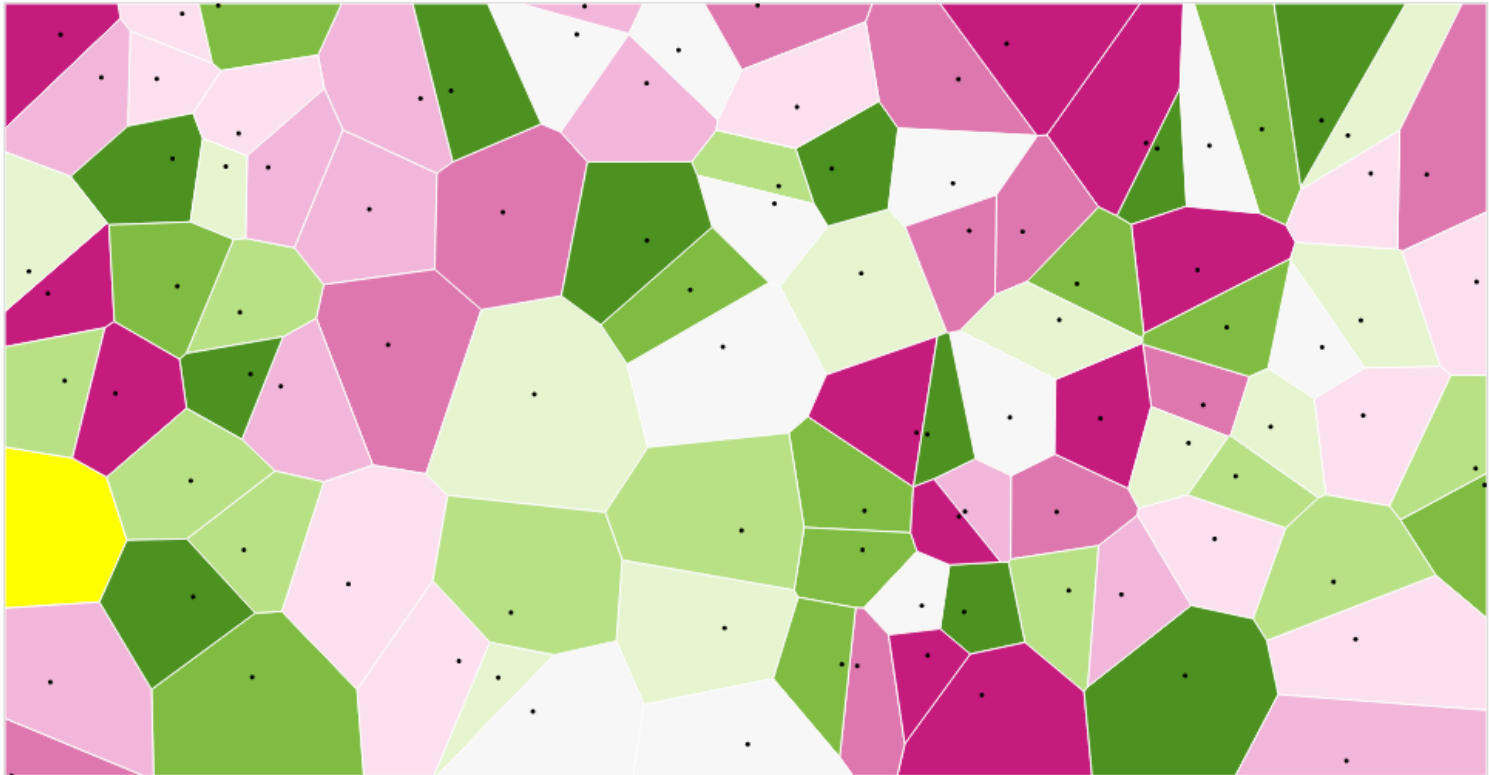
# Treemaps

- *Tree-maps: A space-filling approach to the visualization of hierarchical information structures.* Johnson and Shneiderman, IEEE Vis 1991.
- *Squarified Treemaps.* Bruls et al., Data Visualization 2000.
- Already available in D3
- Real-time, interactive



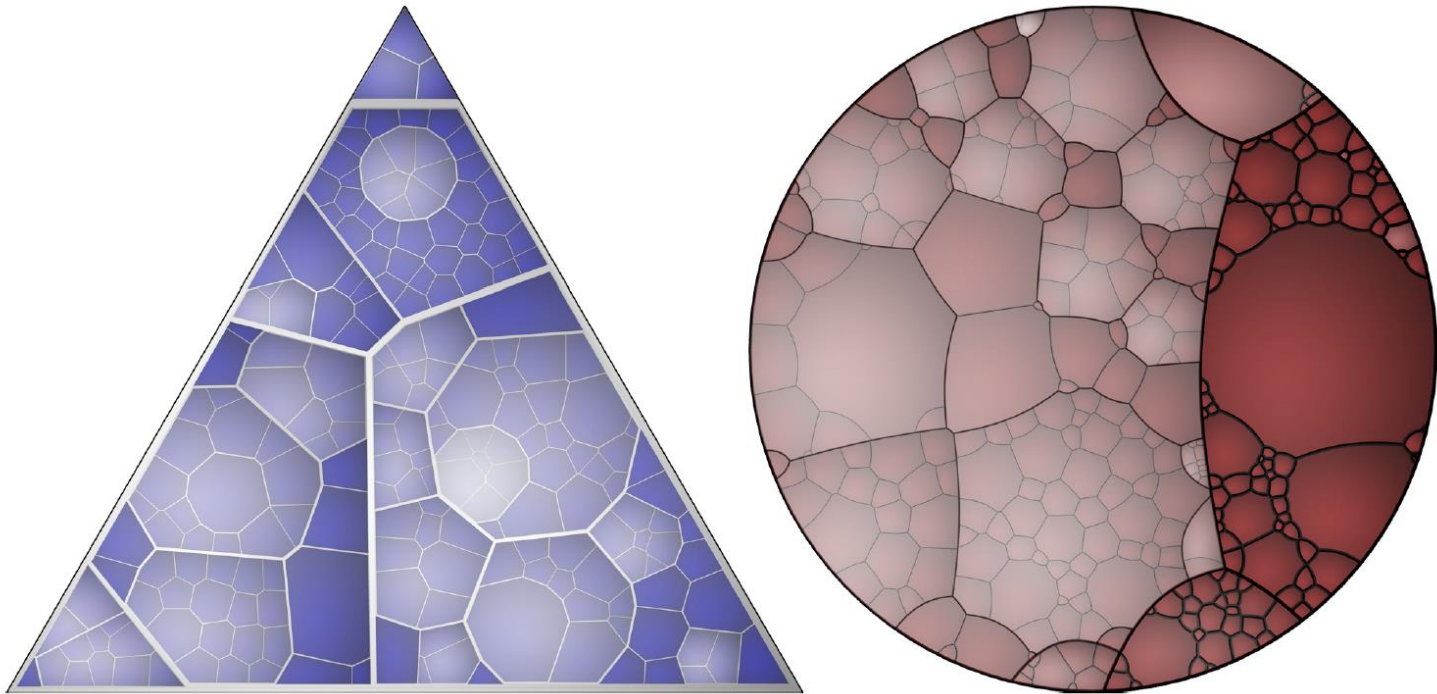
# Voronoi Diagrams

- Partition plane based on closest point
- Available in D3



# Voronoi Treemaps

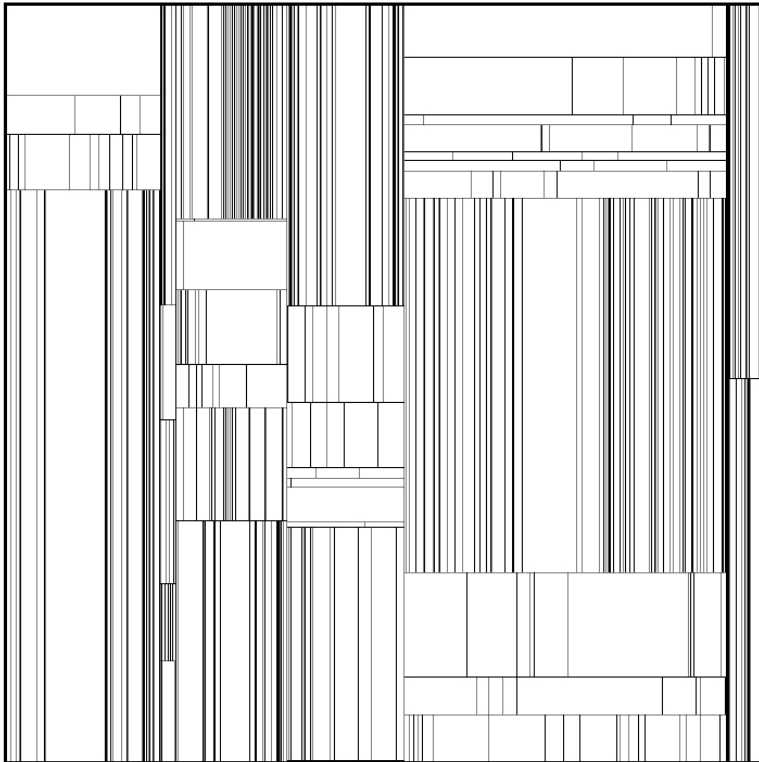
- *Voronoi Treemaps*. Balzer and Deussen, InfoVis 2005.



# Voronoi Treemap Advantages

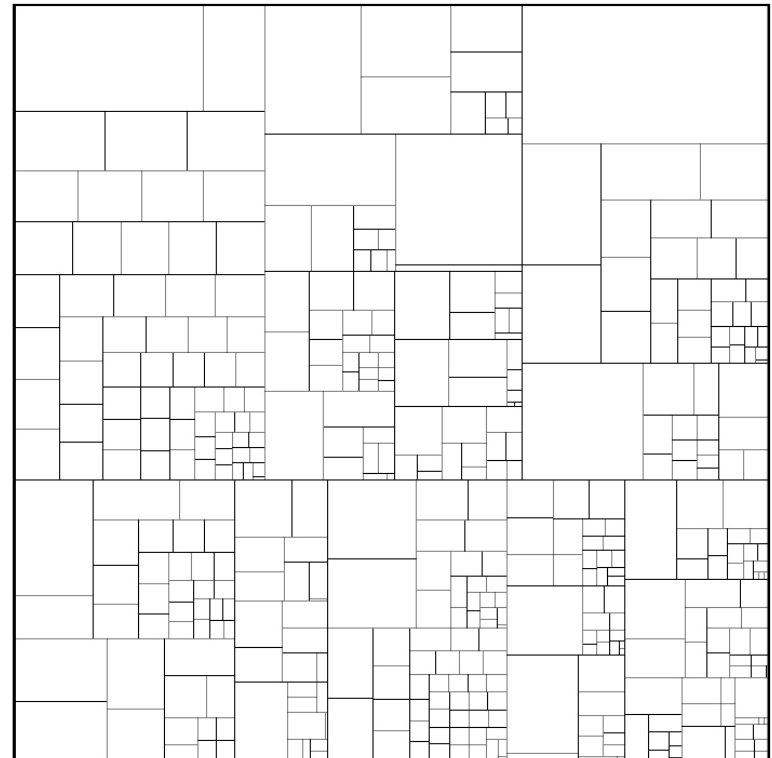
**Original treemap:**

**Unbalanced aspect ratios**



**Squarified treemap:**

**Obscure heirarchy**



# Voronoi Treemaps Are Slow (Even on an 8 core Xeon 2.4GHz)

**~7 minutes:**

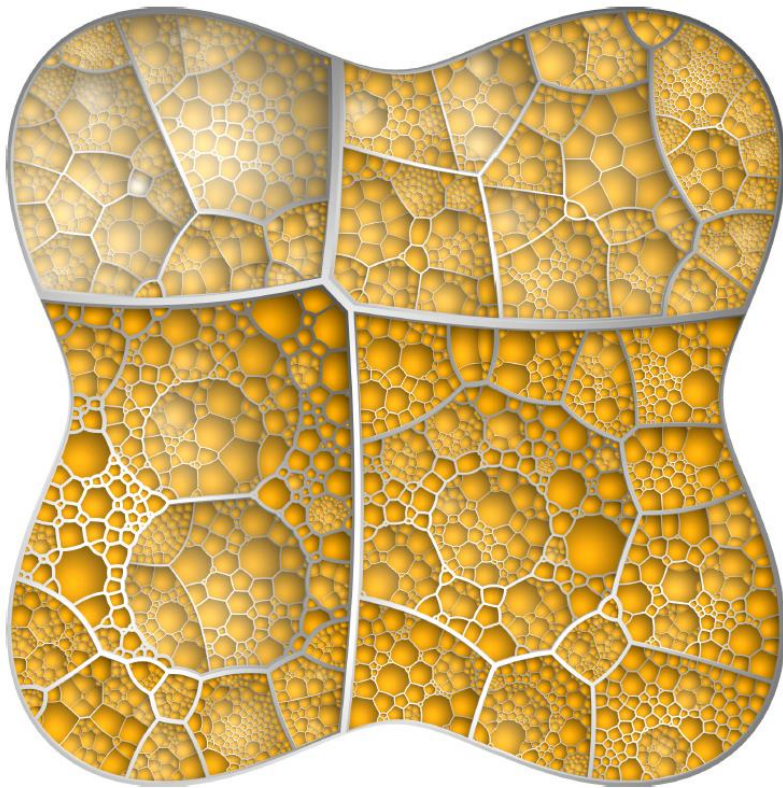


Figure 10: Enhanced AW Voronoi Treemap layout of 4075 nodes at 10 hierarchy levels (a brighter color indicates a lower hierarchy level)

**~6 minutes:**

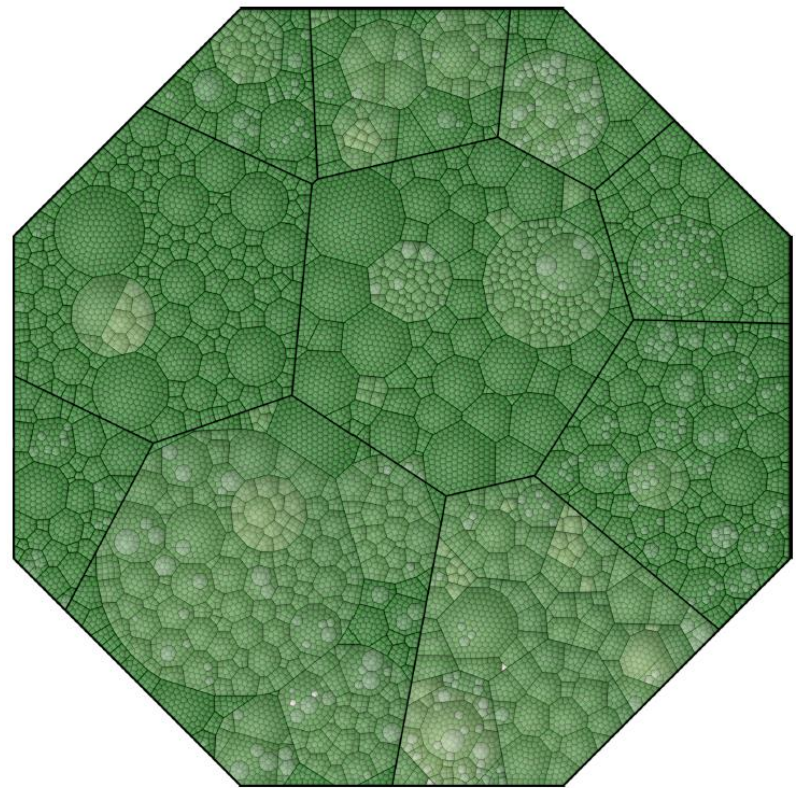


Figure 11: Enhanced PW Voronoi Treemap layout of 16288 nodes at 7 hierarchy levels (a brighter color indicates a lower hierarchy level)

# Tasks

- Literature for faster options (done)
  - *Fast Dynamic Voronoi Treemaps*. Sud et al., 2010.
    - GPU-based, ~50x speedup
  - *Computing Voronoi Treemaps: Faster, Simpler, and Resolution-independent*. Nocaj and Brandes, 2012.
    - CPU-based, core algorithm changes, beats above GPU solution
- Implement original core algorithm in D3
- Investigate Optimizations
  - Most promising is Nocaj and Brandes because it is CPU-based.

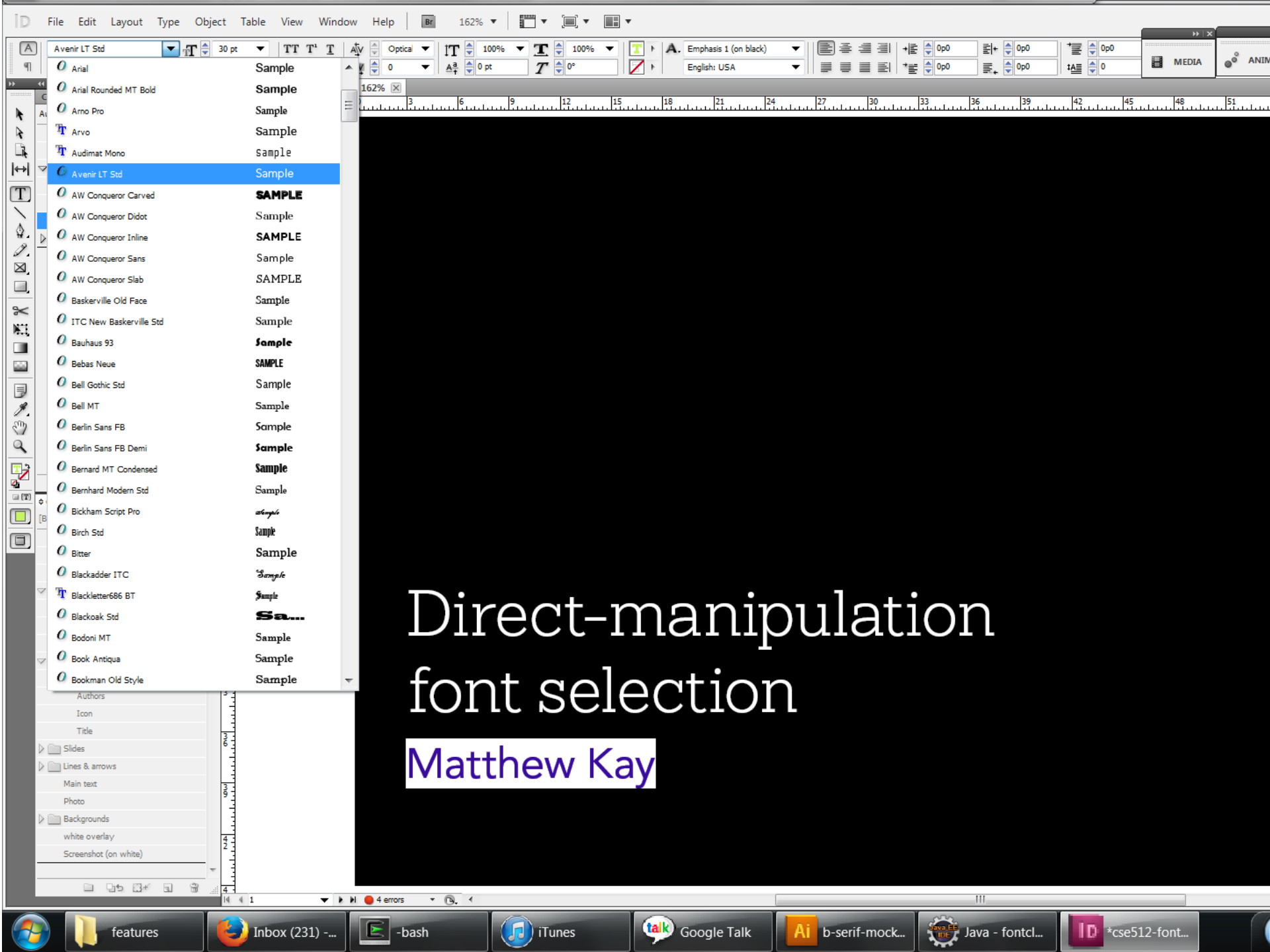
# Questions

- How can we parallelize our tasks?
  - This project is somewhat serial in nature
- How can we parallelize our algorithm (LOL)?
  - Is WebGL or WebCL a good idea?
- Good ideas for optimization?
  - Plot only the first  $n$  levels of the dataset
  - Plot first few levels as a standard treemap, then the remainder as a Voronoi Treemap
- What are some good available sample datasets?
- What interactivity can/should we support?

(Presented by Paul Vines & Peter Henry)

# Direct-manipulation font selection

Matthew Kay



# Direct-manipulation font selection

Matthew Kay

FA 14 11 3

Zoom

3

nature and technology tell us th  
rows out of function (purpose),  
nic or technical), and out of how  
marvellous forms of nature and  
chnology originated.

Elementar —  
a **parametric** type family

Style

?

abc

abc

abc

Parameters

Height

14

Weight

11

Width

3

a

a

a

a

a

a

a

a

a

a

☐ Low Contrast

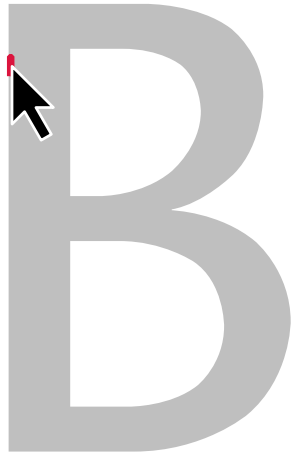
Add

Reset

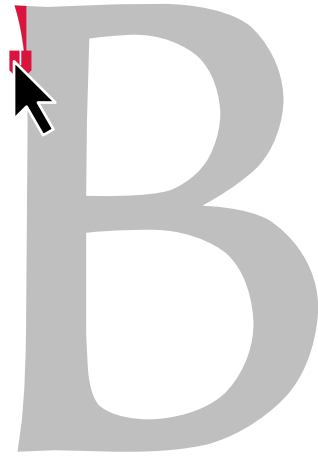
How about  
parametric, **direct**  
**manipulation**  
selection of  
**existing** fonts?

B

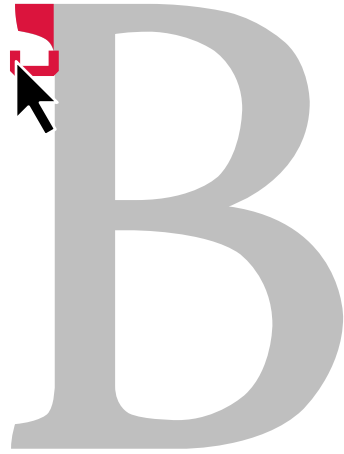
Stone Sans  
Humanist sans-serif



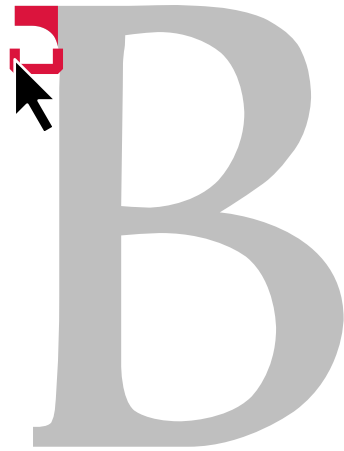
Stone Sans  
Humanist sans-serif



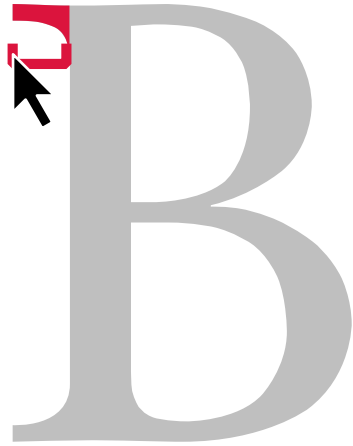
Angie Sans  
Humanist semi-serif



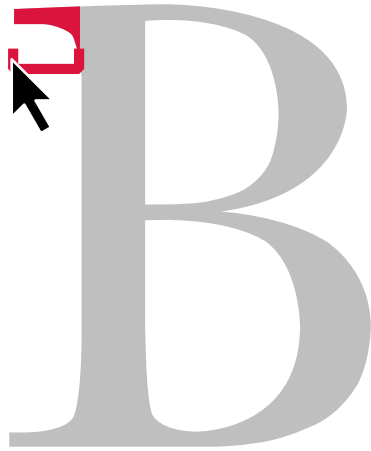
Vollkorn  
Serif



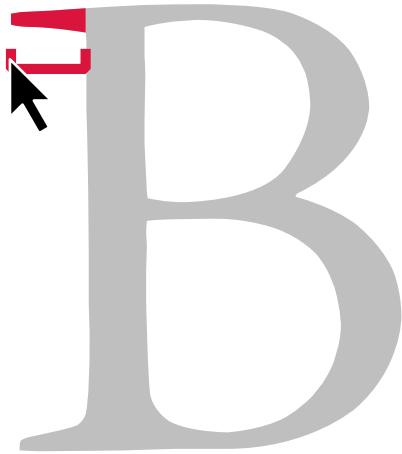
Apolline  
Old-style serif



Sabon  
Old-style serif



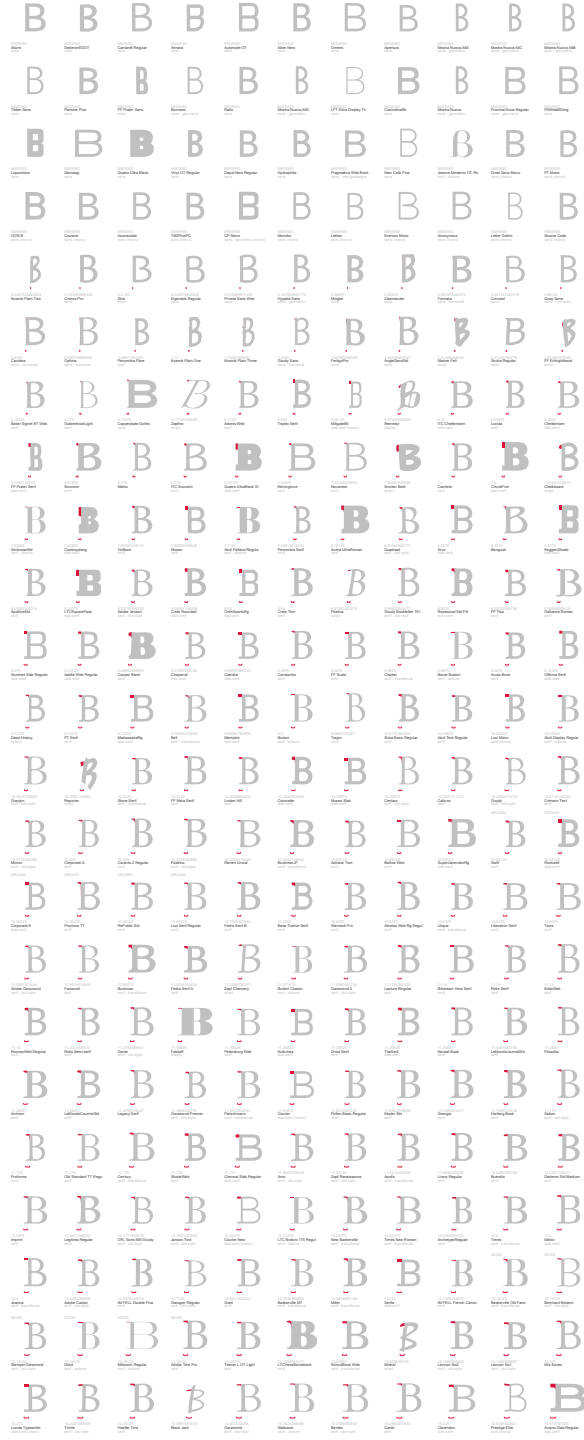
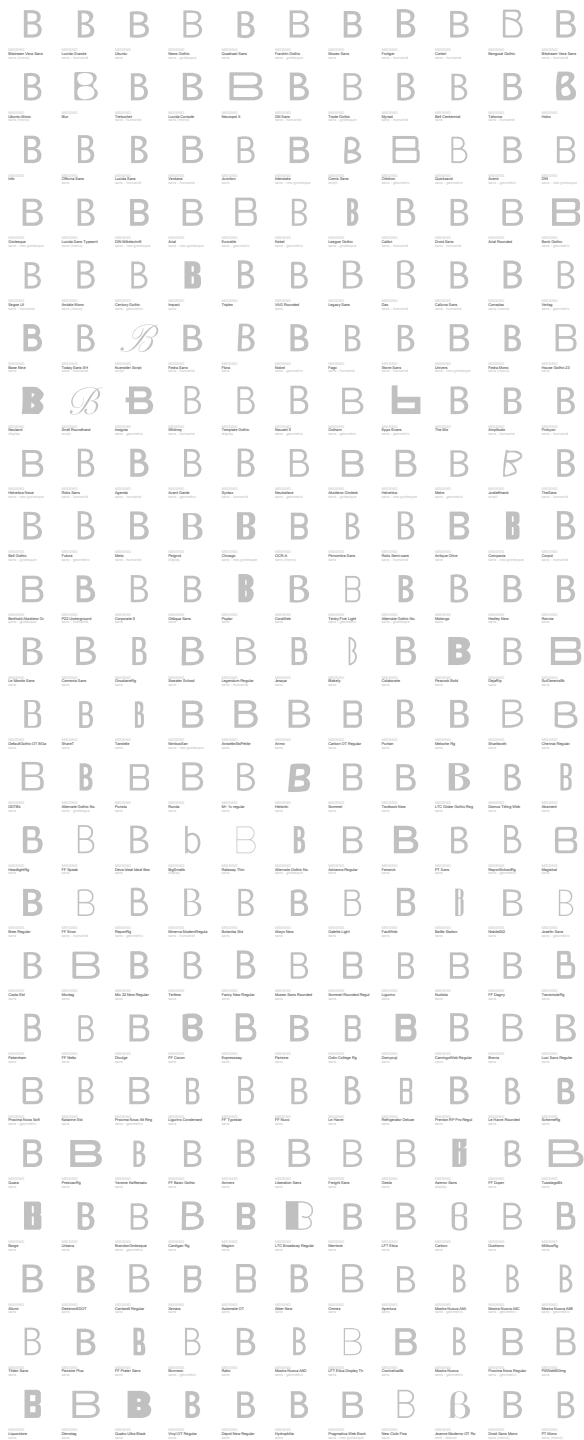
Baskerville  
Transitional serif



Garamond  
Old-style serif

A large, light gray capital letter 'B' in a Garamond Old-style serif font. The letter features a prominent red highlight on the top left serif, which is a characteristic design element of this typeface.

Garamond  
Old-style serif



B

175 Regul  
e

B

12.24375  
New Baskerville  
serif - transitional

B

12.24375  
Times New Roman  
serif - transitional

B

12.2906025552  
ArchetypeRegular  
serif

B

12.4  
Times  
serif - transitional

B

12.4  
Melior  
slab-serif

B

862  
MT  
tional

B

12.9814691148  
Miller  
serif - transitional

B

13.025  
Serifa  
slab-serif

B

13.1209149872  
IM FELL French Canon  
serif

B

MODE

13.18125  
Baskerville Old Face  
serif - transitional

B

MODE

13.18125  
Bernhard Modern  
serif - old-style

B

417  
onoblack

B

13.80625  
SchoolBook Web  
serif - transitional

B

13.925849139  
Mistral  
script

B

14.1090691866  
Lexicon No2  
serif - old-style

B

14.1092809957  
Lexicon No1  
serif - old-style

B

14.275  
Mrs Eaves

B

369  
e

B

15.3559982836  
Bembo  
serif - old-style

B

15.4962875832  
Cardo  
serif

B

15.525  
Clarendon  
slab-serif

B

15.8305820821  
Prestige Elite  
serif (mono)

B

15.975797439  
Aviano Slab Regular  
slab-serif

1.3239416099  
FF Cocon  
sans

1.32738137571  
Bembo  
serif - old-style

1.34180936133  
Dederon Std Medium  
serif

1.34523432432  
OFL Sorts Mill Goudy  
serif - old-style

1.35186002531  
Adobe Jenson  
serif - old-style

1.3630774664  
Sirba Basic Regular  
serif

1.36341348654  
FF Duper  
sans

1.36869613718  
Linden Hill  
serif

1.37180594358  
Zapf Chancery  
script

1.37414419201  
Candara  
sans - humanist

1.50642432828  
Proforma  
serif

1.50768995896  
Cronos Pro  
sans

1.50920372089  
Cantarell Regular  
sans

1.51095085798  
Cambria  
serif

1.51128266365  
Bree Regular  
sans

1.52934052991  
IM FELL French Canon  
serif

1.52961936085  
Corporate E  
slab-serif

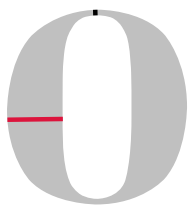
1.52997581472  
Abril Display Regula  
serif - didone

1.53338012329  
Coronette  
slab-serif

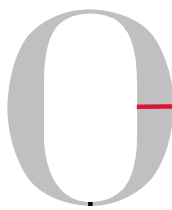
1.53340785951  
Brevia  
sans



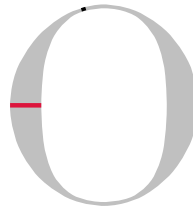
0.1433  
ay Regular



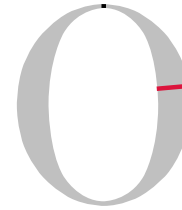
0.100819014473  
Abril Fatface Regula  
serif - didone



0.105085721522  
Jeanne Moderno OT Ro  
serif - didone



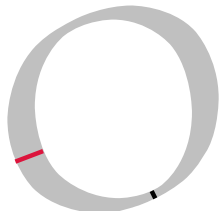
0.113136176665  
Didot  
serif - didone



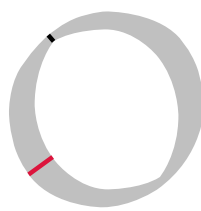
0.12030805281  
Bauer Bodoni  
serif - didone



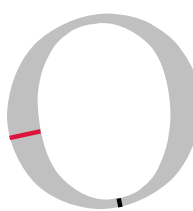
0.346  
le



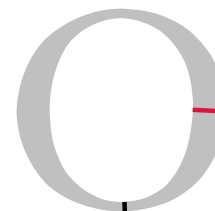
0.275000421729  
Brioso Pro Regular  
serif



0.276429598502  
Rieven Uncial



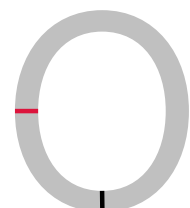
0.279202642021  
Stone Serif  
serif - transitional



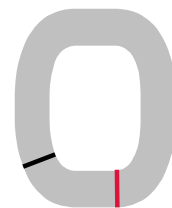
0.280475296642  
Garamond 3  
serif - old-style



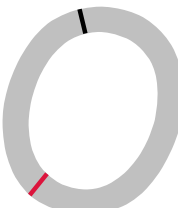
0.539  
ght  
tric



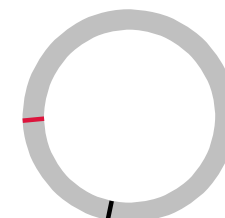
0.909064187305  
Pitch  
slab-serif (mono)



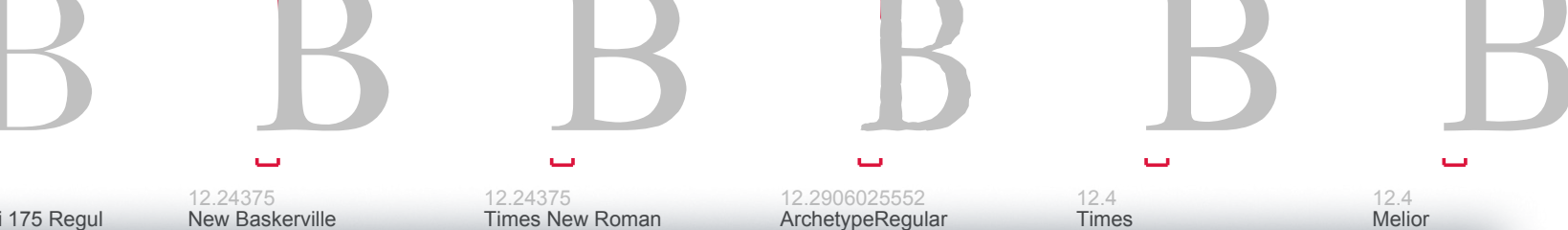
0.90975331595  
Domyouji  
sans



0.909992405507  
Chalkboard  
script



0.911458943686  
ReportRg  
sans - geometric



# Thanks! Discussion—

Preventing **jitter** — optimal path through this space?

Handling **multiple constraints**?

**Feature selection?**

# Narrative Visualization

A story is worth thousand pictures.

Christine(Yenting) Liu, Nina(Zhuxiaona) Wei



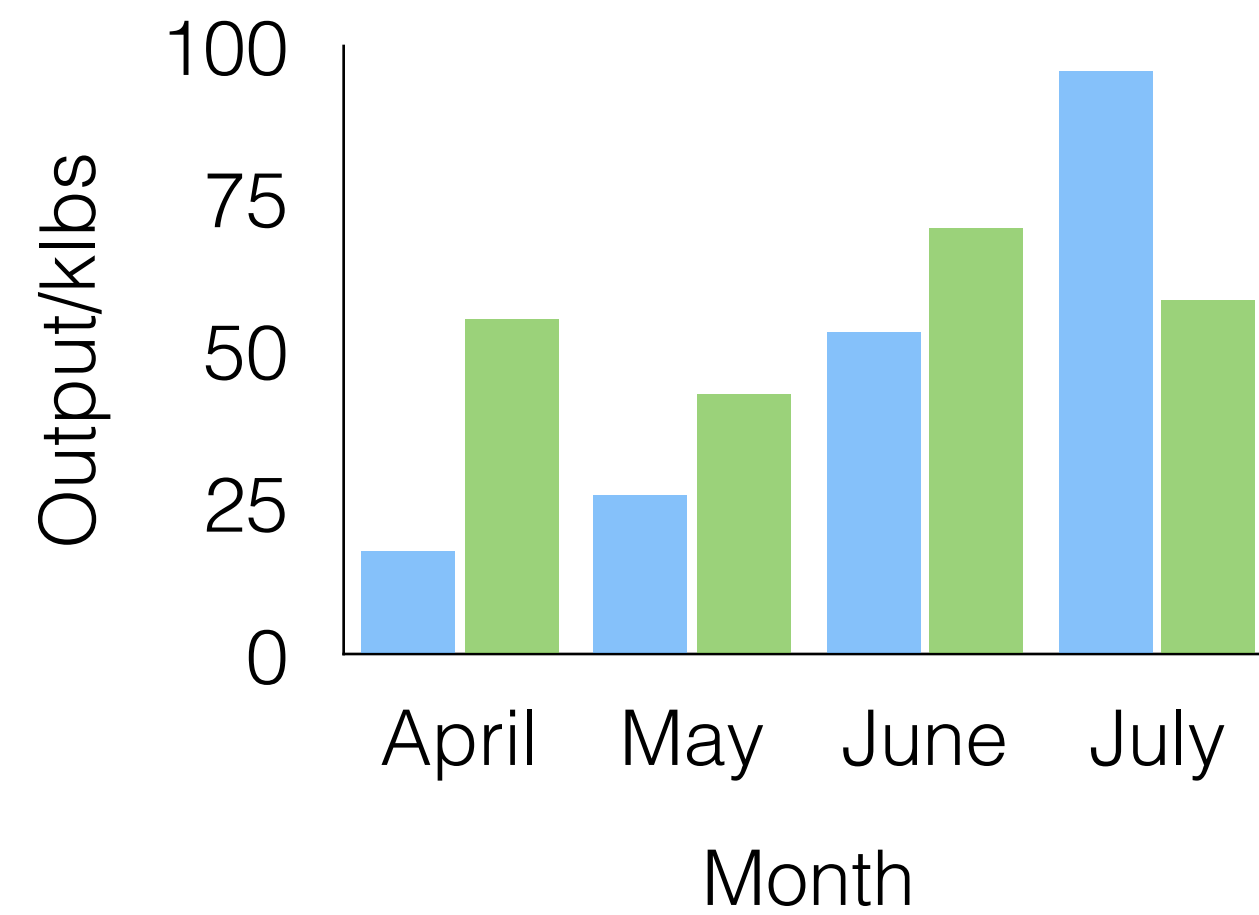
# Question

Which is the more effective way to present statistical data graphics, **synchronously** or **asynchronously**?

# Question

Synchronous

Asynchronous



1

**Synch-**

Cognitive Load

Attention Limitation

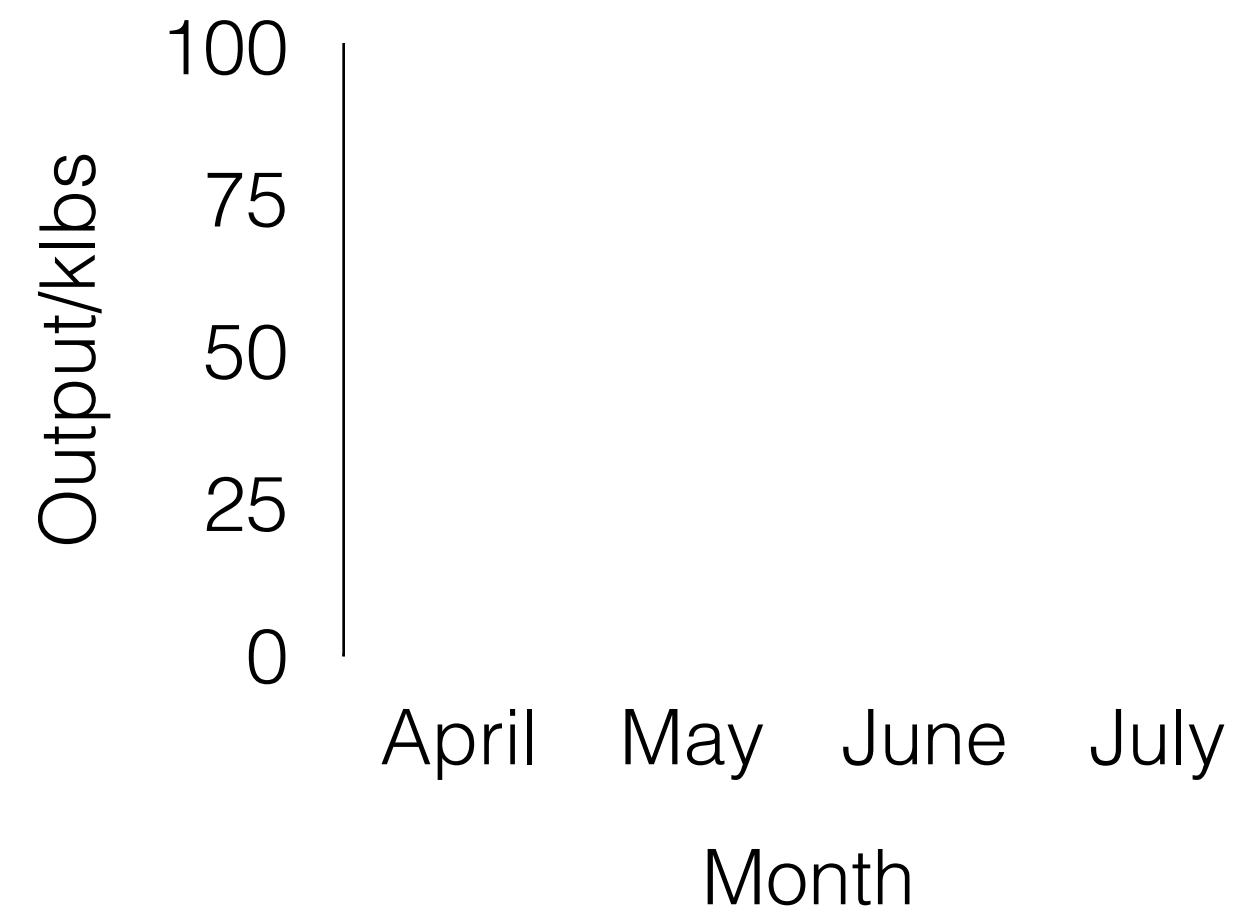
*Top-Down* processing

**Asynch-**

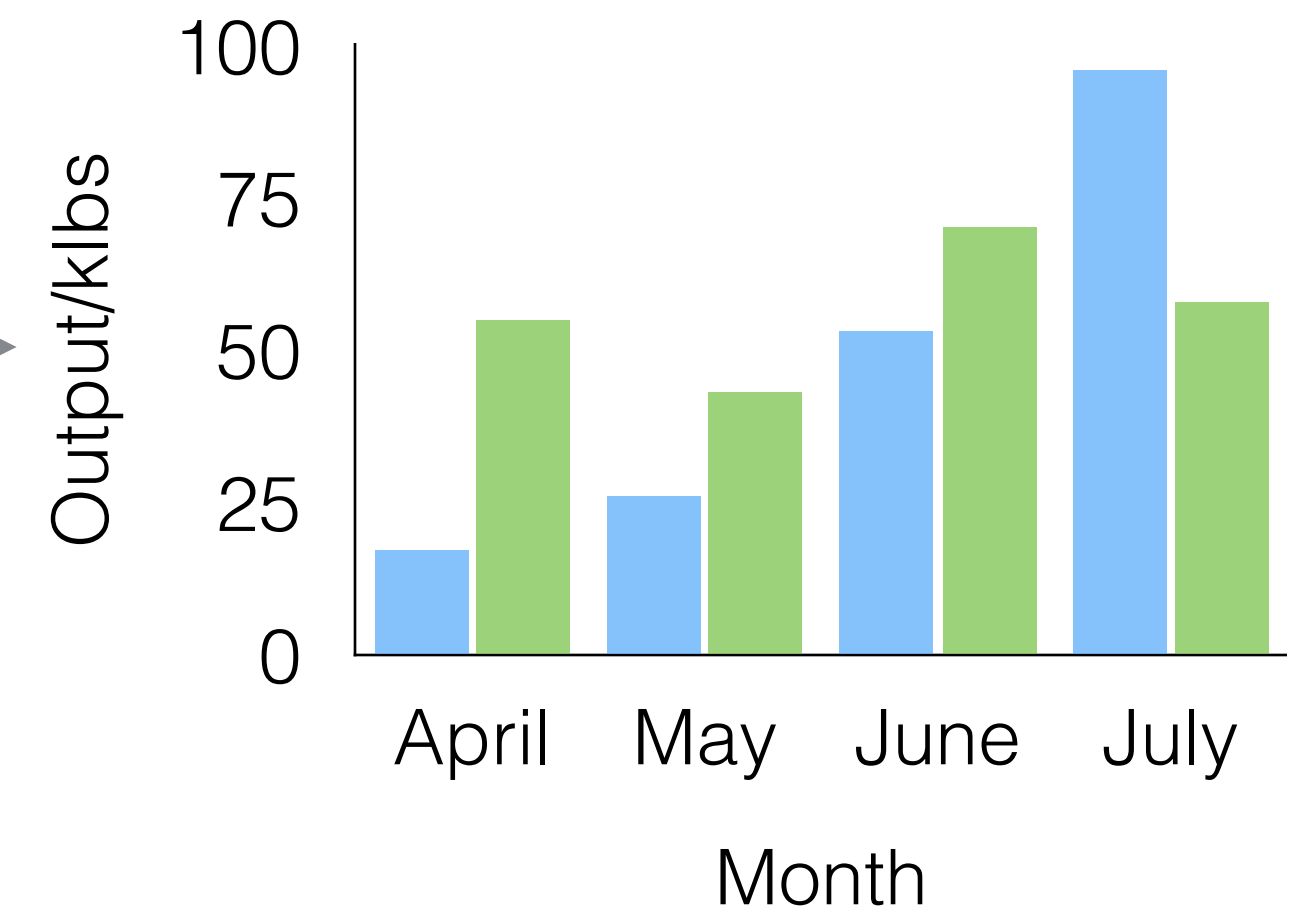
*Bottom-up* processing

April May June July  
Month

1



2



3



## Related-work

The role of Storytelling in visualization, and the importance of narrative visualization.

Gershon & Page, 2001; Wojtkowski & Wojtkowski, 2002; Ma et.al, 2012; Kosara & Mackinlay, 2013;

Design category, design space analysis, case study

Segel & Heer, 2010;

Visualization rhetoric, framing effects

Hullman & Diakopoulos, 2011

Animation transitions, Sequence of narrative visualization

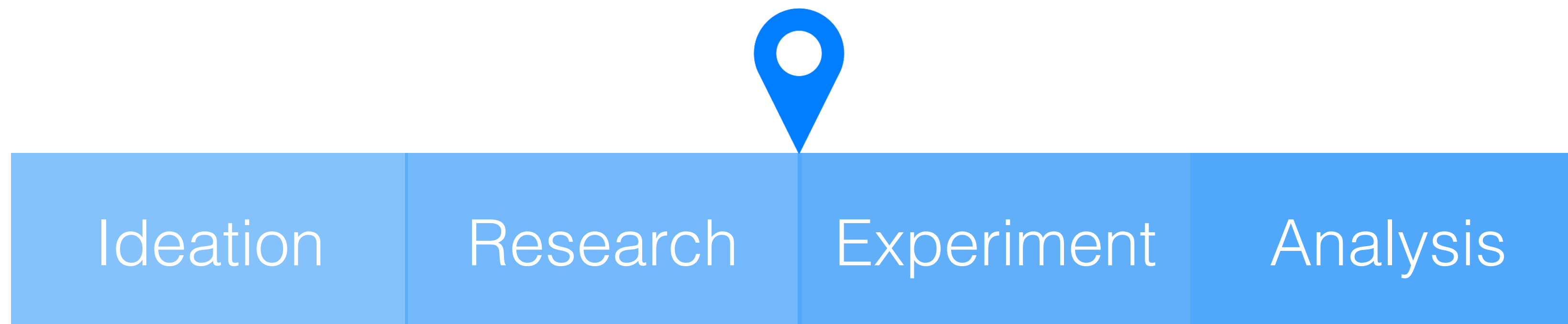
Heer & Robertson, 2007; Hullman et al, 2013

# Progress

**Done:** Literature Review

**Doing:** Experiment Design, Pilot Test

**To do:** Experiment, Data Analysis, Paper Write-up



# Progress



Experiment

## **Within-Subjects Design**

### **Independent Variables:**

1. Synchronous, Asynchronous

### **Dependent Variables:**

1. Accuracy (describe the level of interpretation), objective
2. Likert Attitude Scale (based on the Design Criteria), subjective

### **Experiment Stimulus:**

Basic Charts (bar) or Typical Narrative Visualization

Synchronous

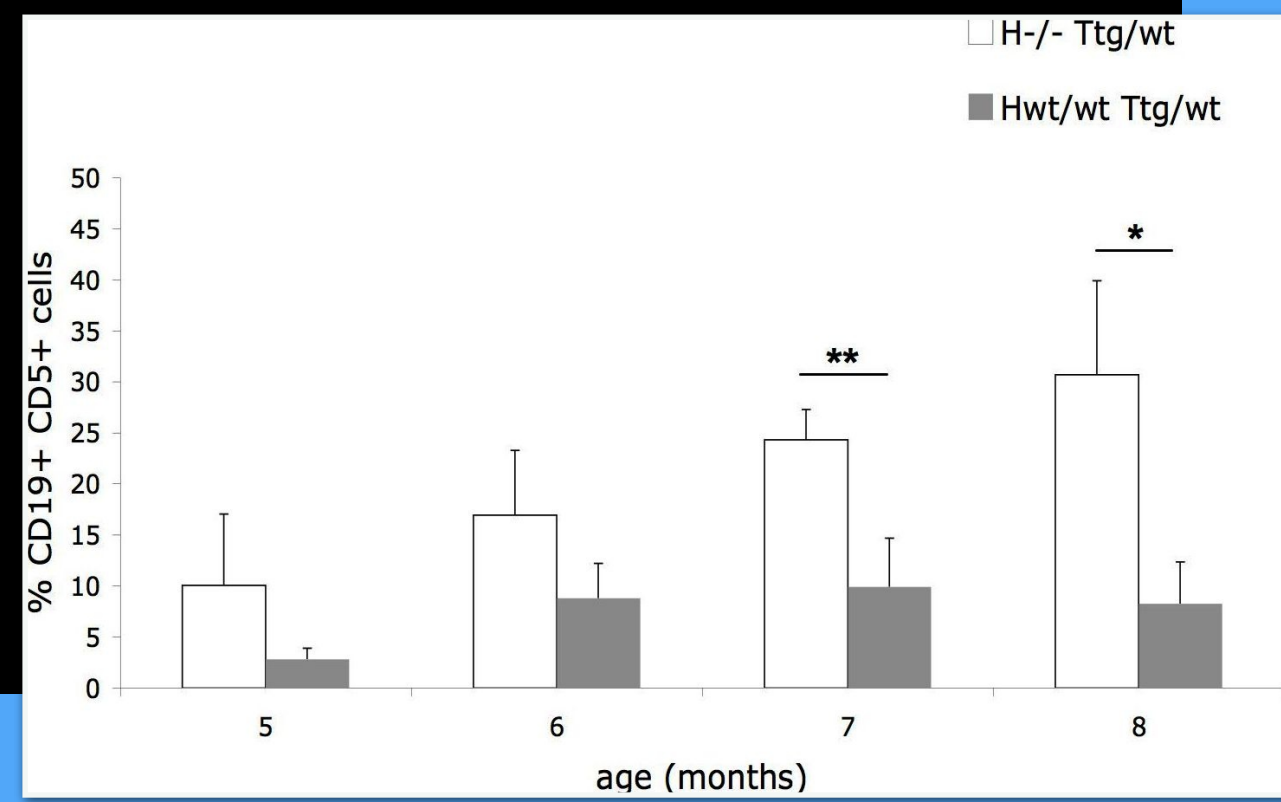
+

500ms

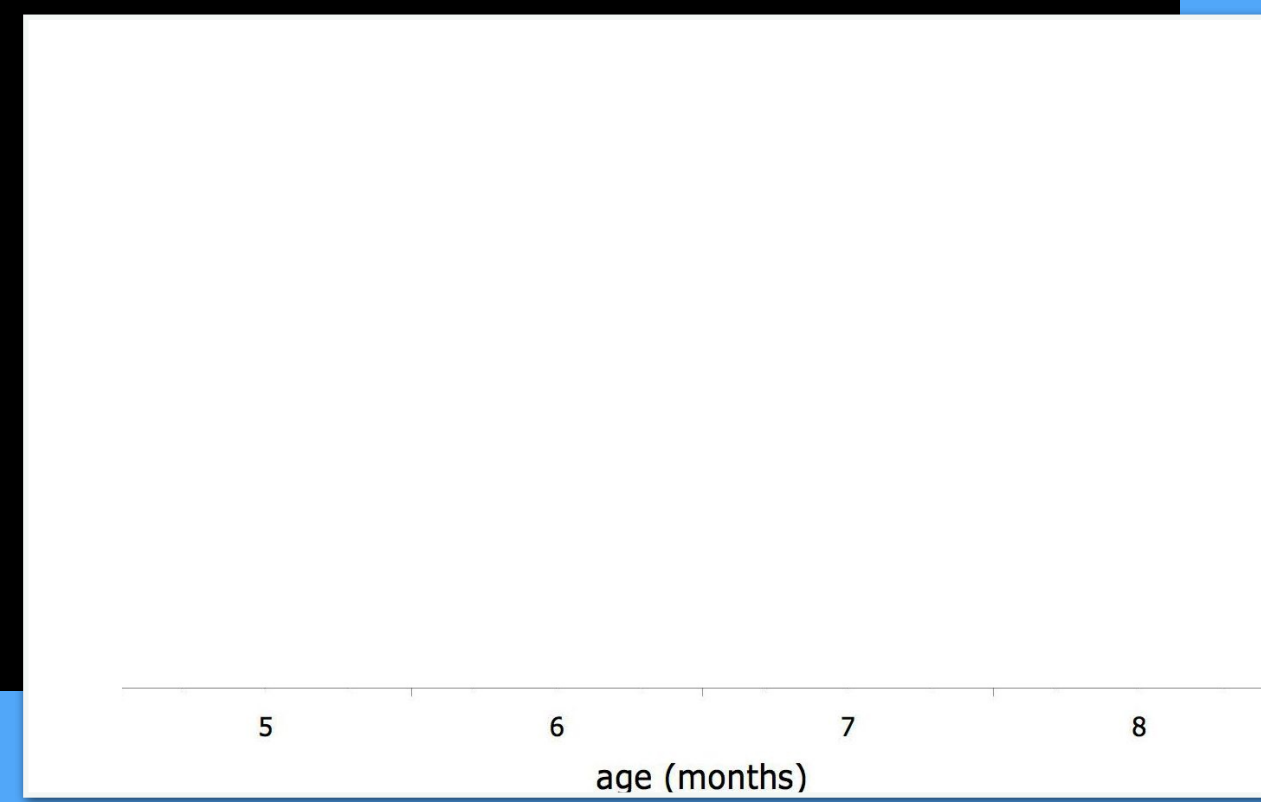
Asynchronous

+

## Synchronous

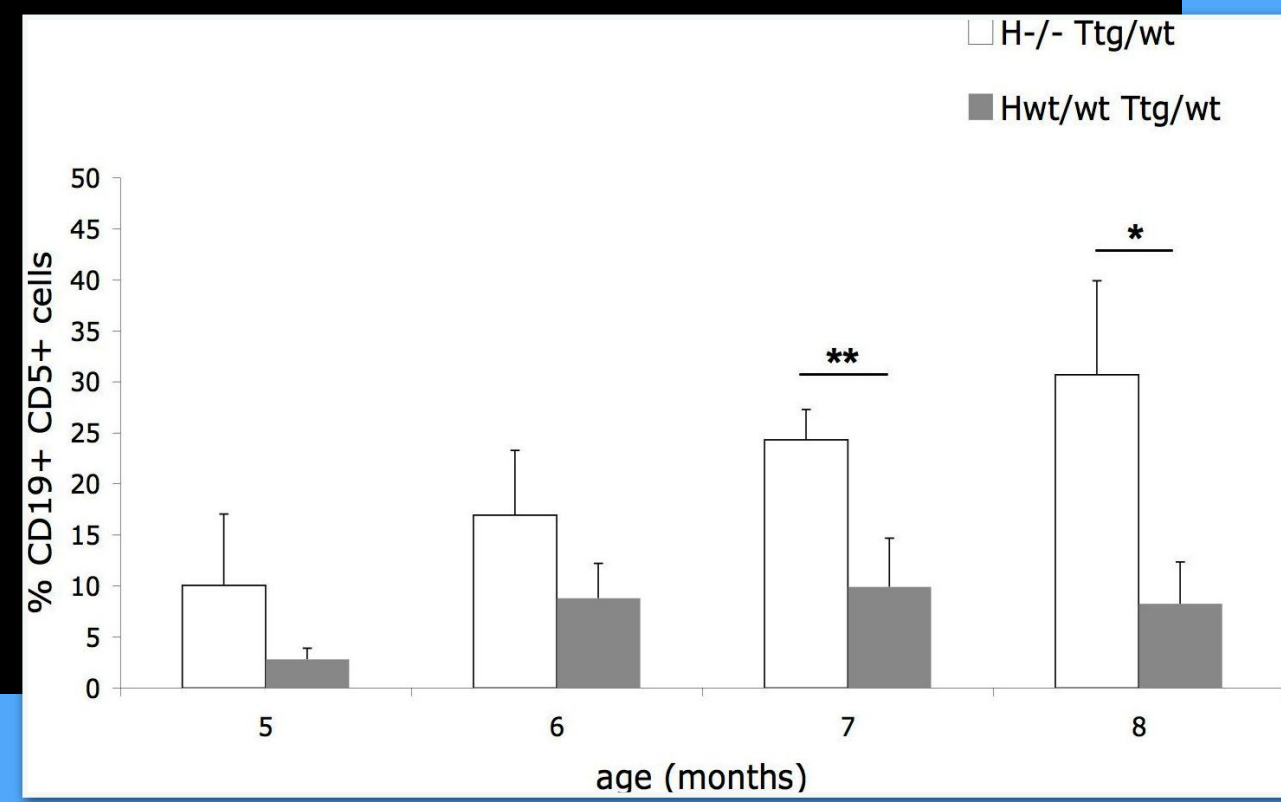


## Asynchronous

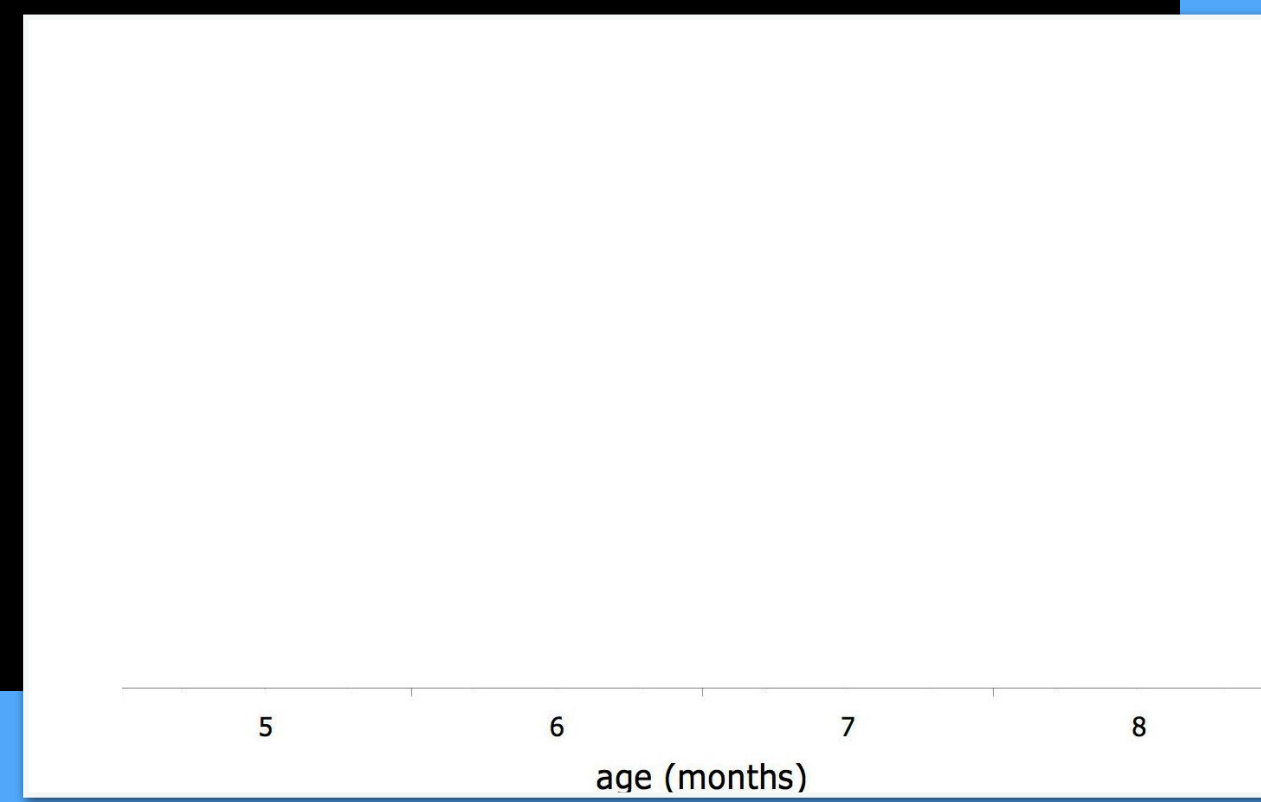


500ms

## Synchronous

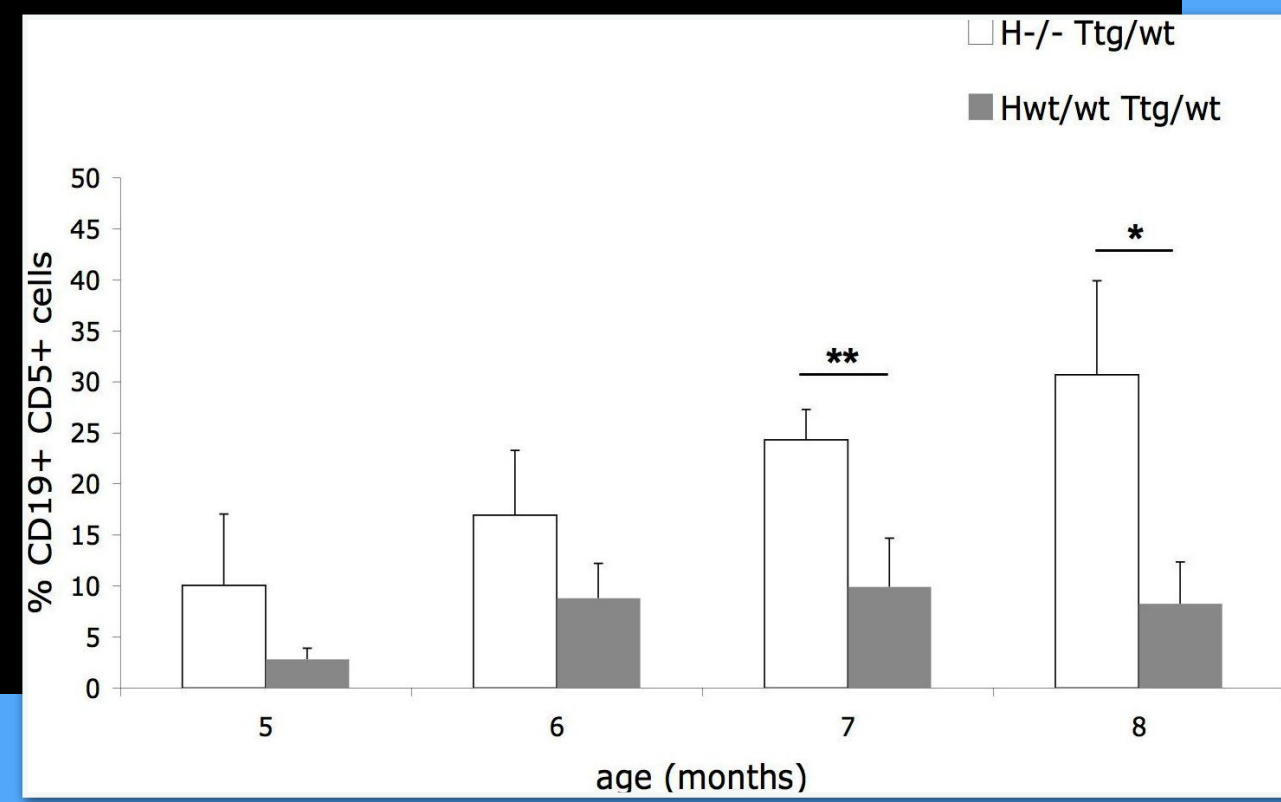


## Asynchronous

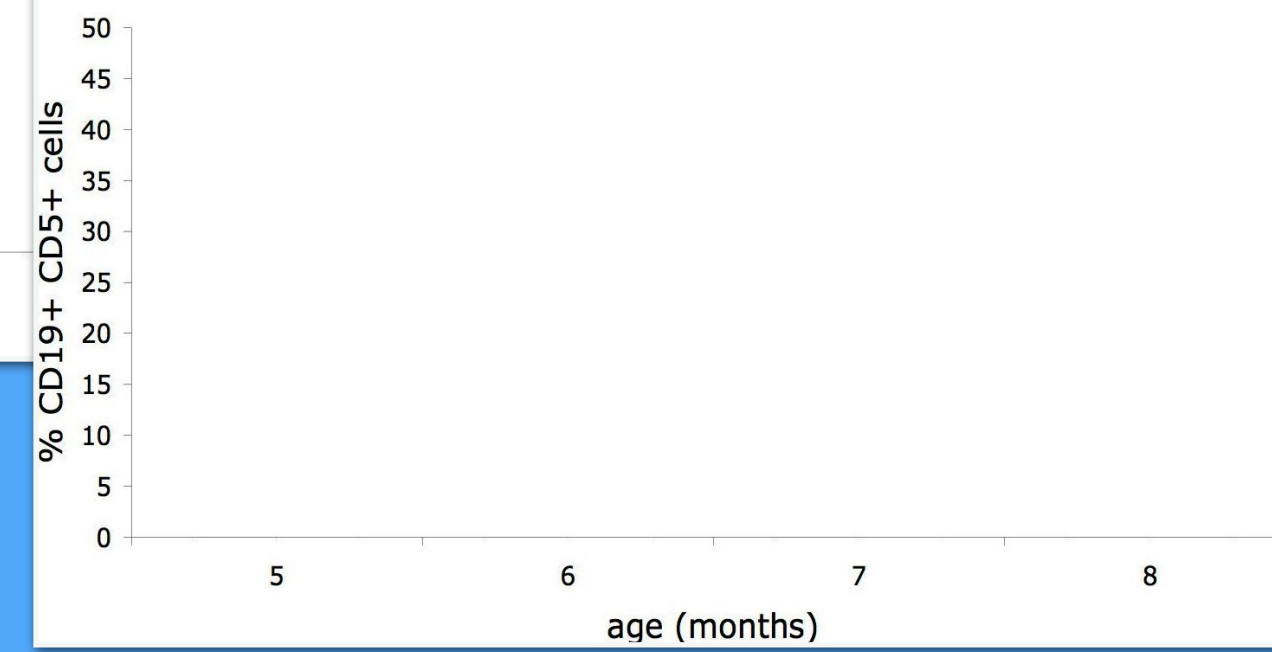


500ms

## Synchronous

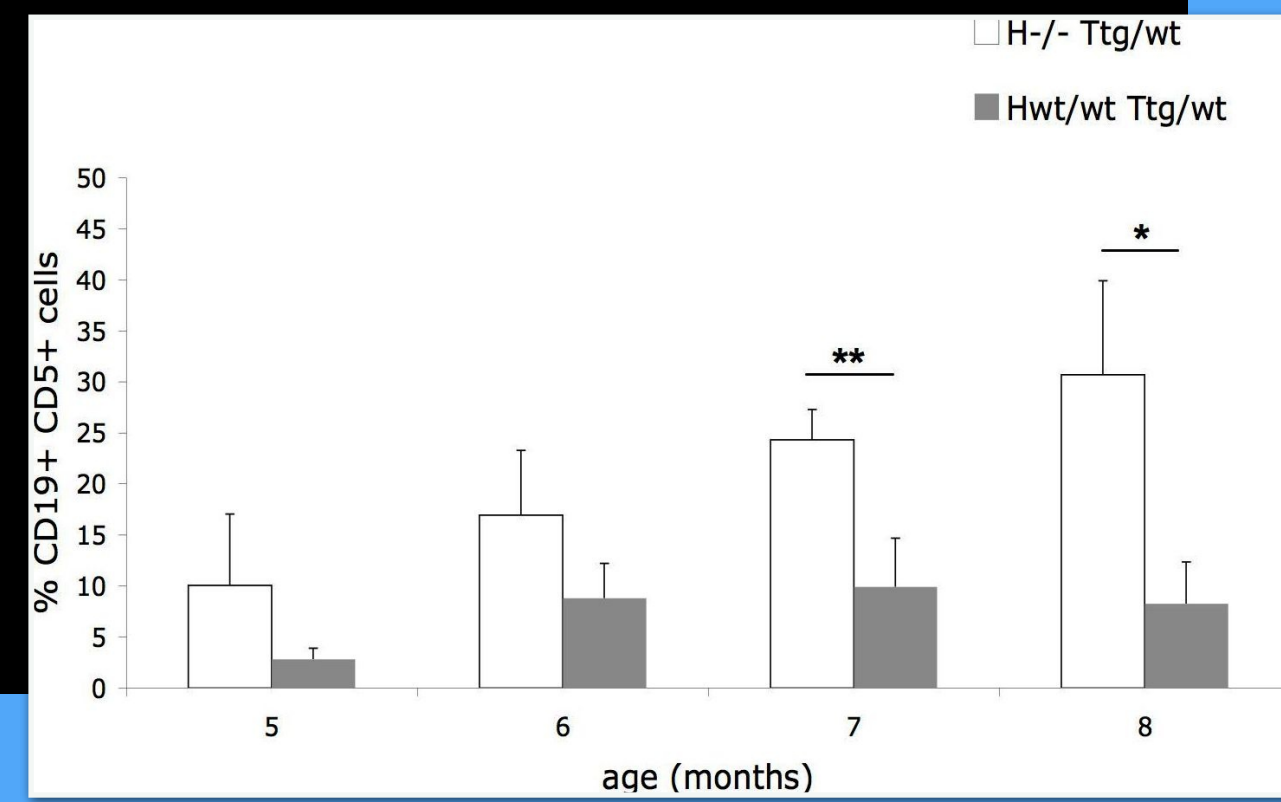


## Asynchronous

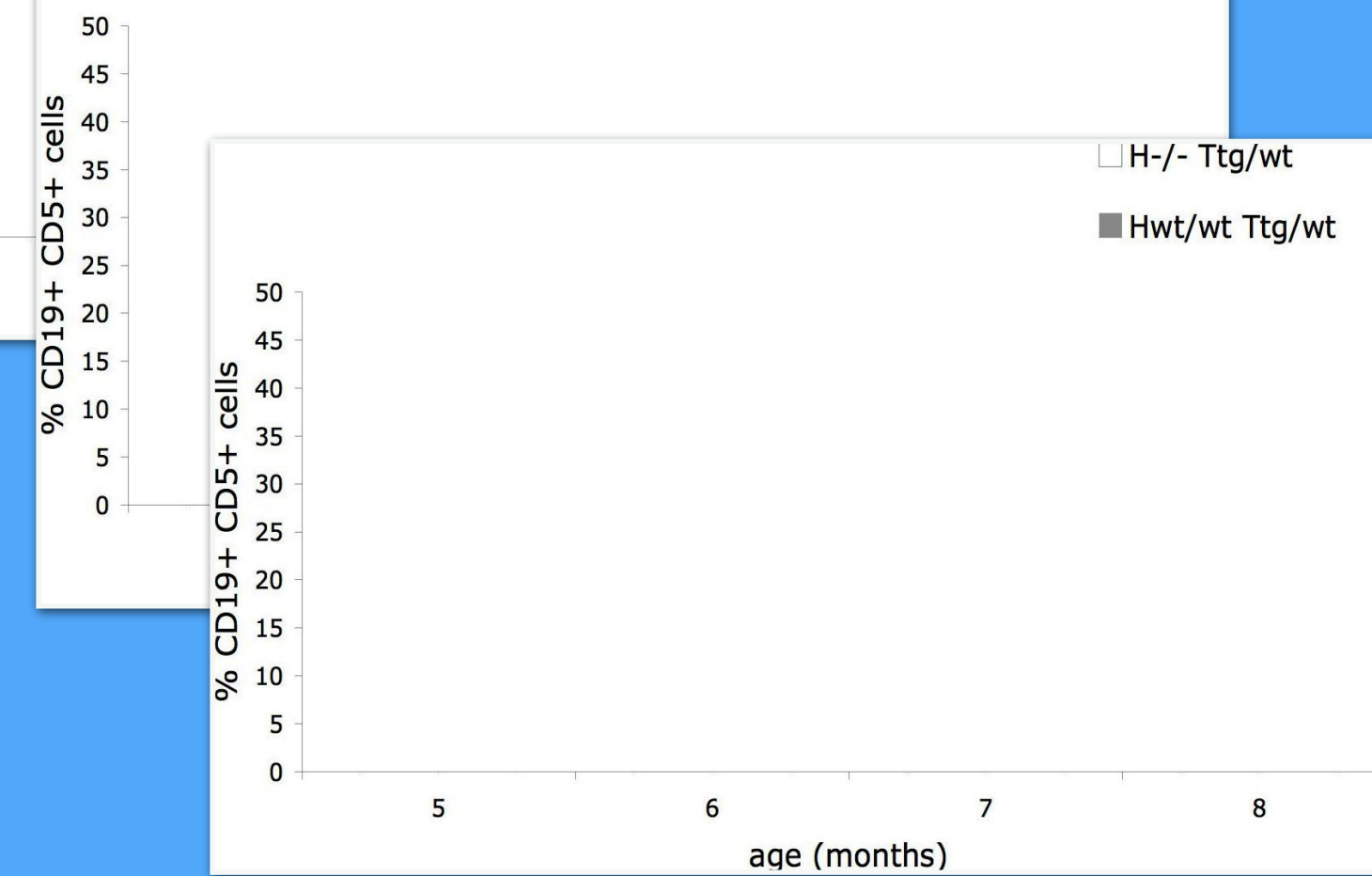


500ms

## Synchronous

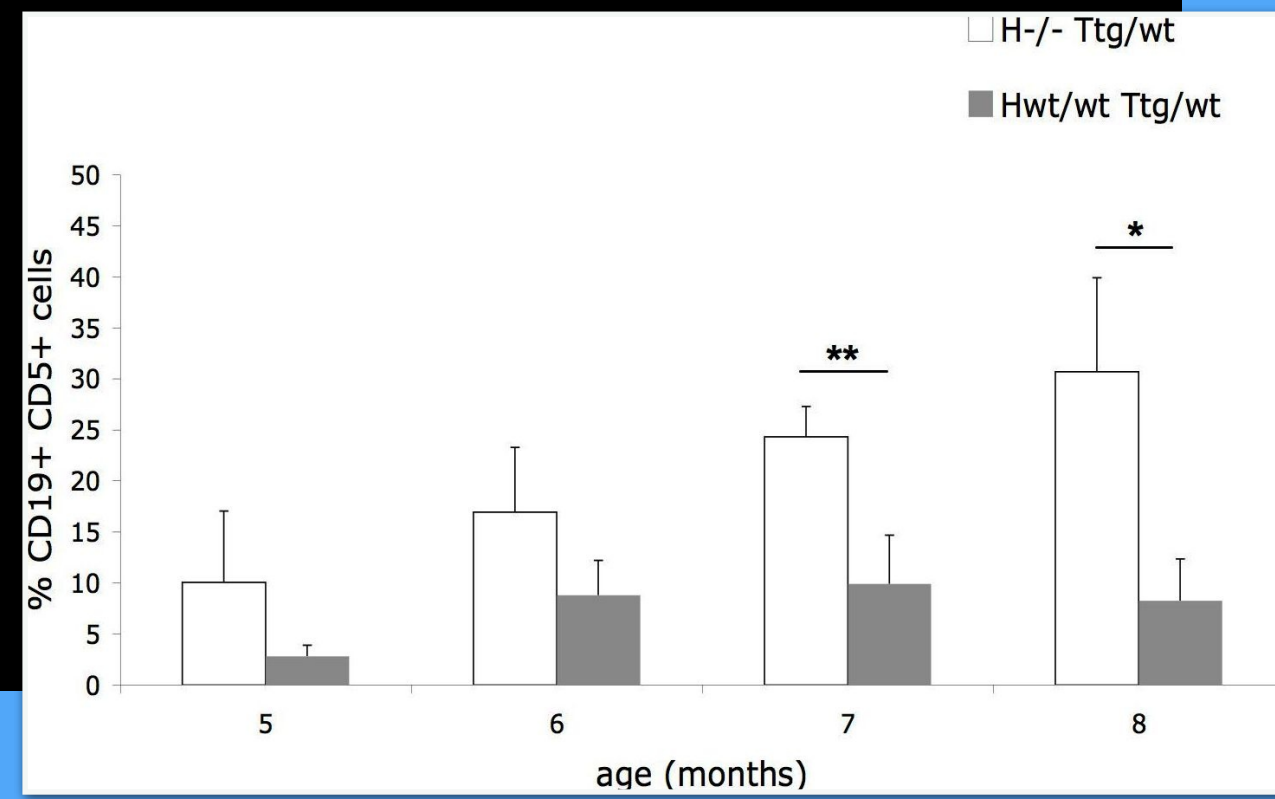


## Asynchronous

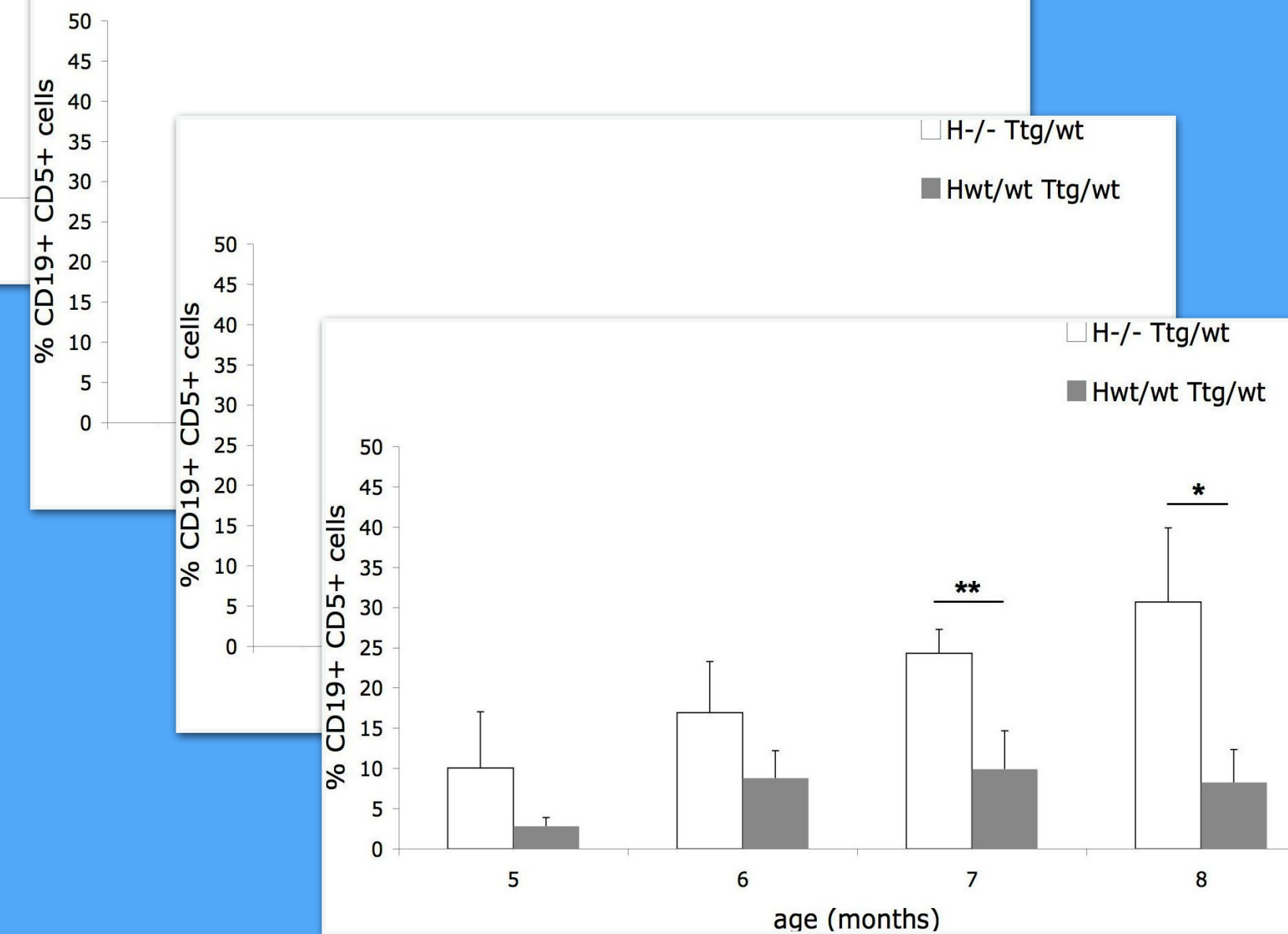


500ms

## Synchronous

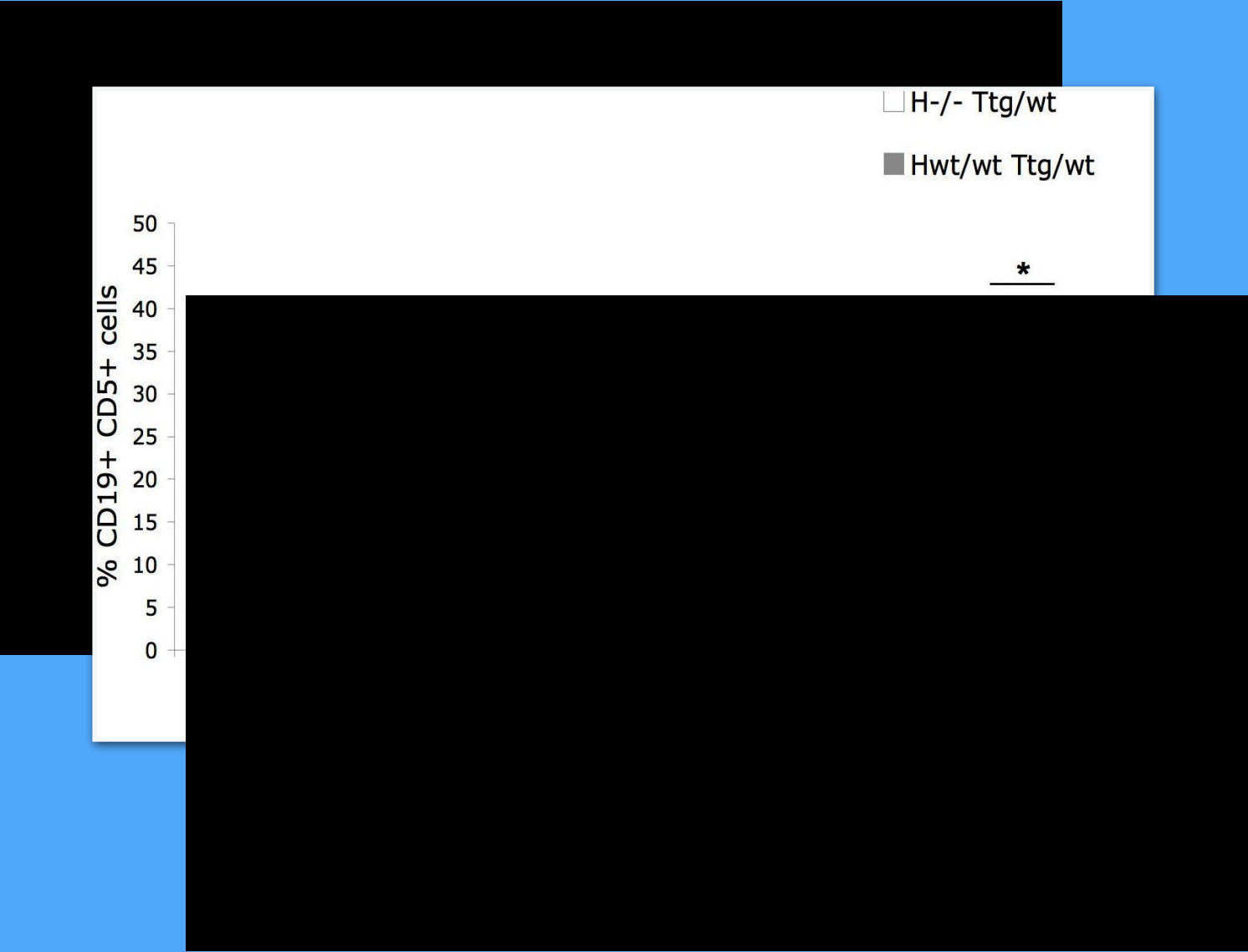


## Asynchronous



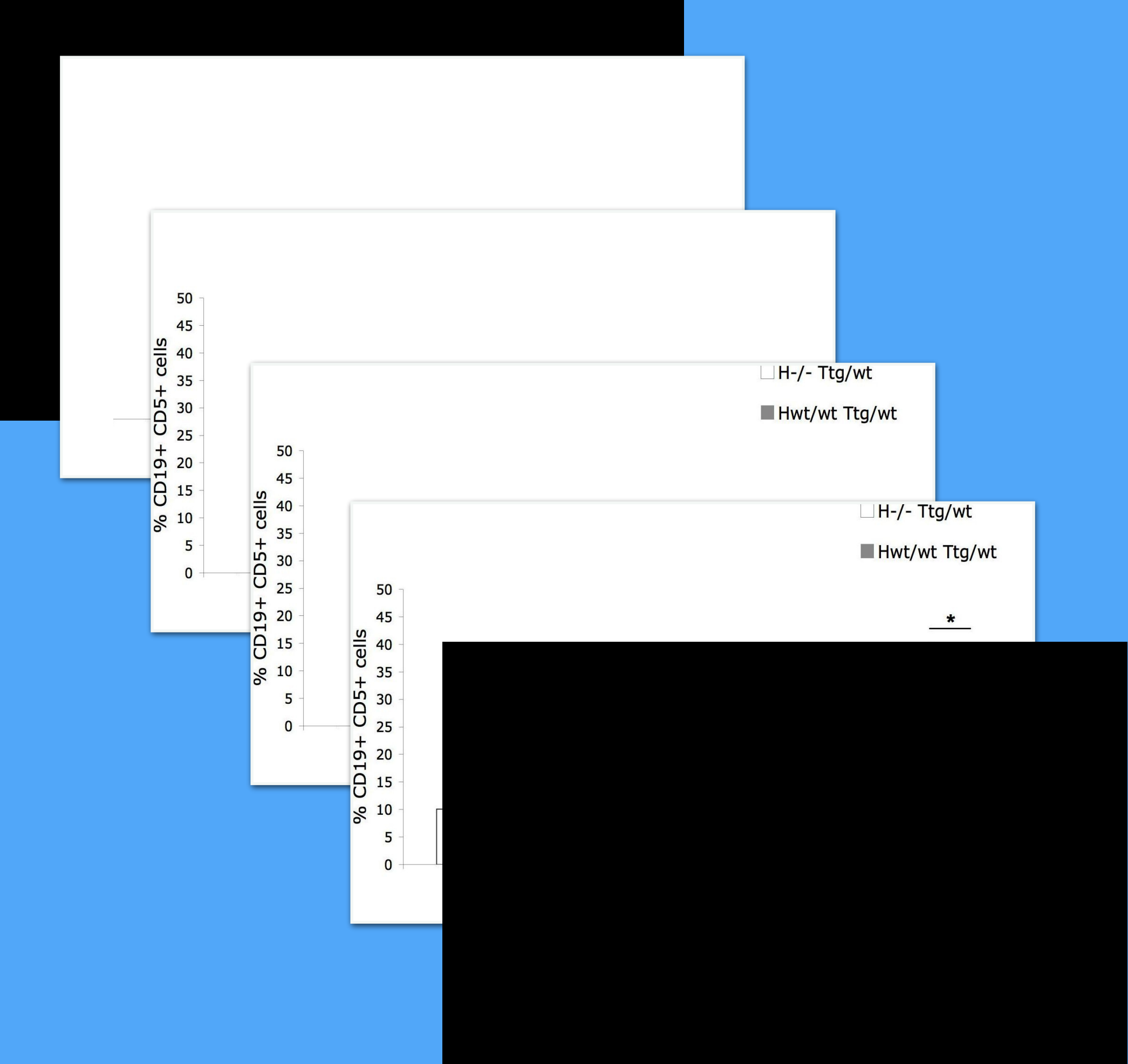
500ms

Synchronous

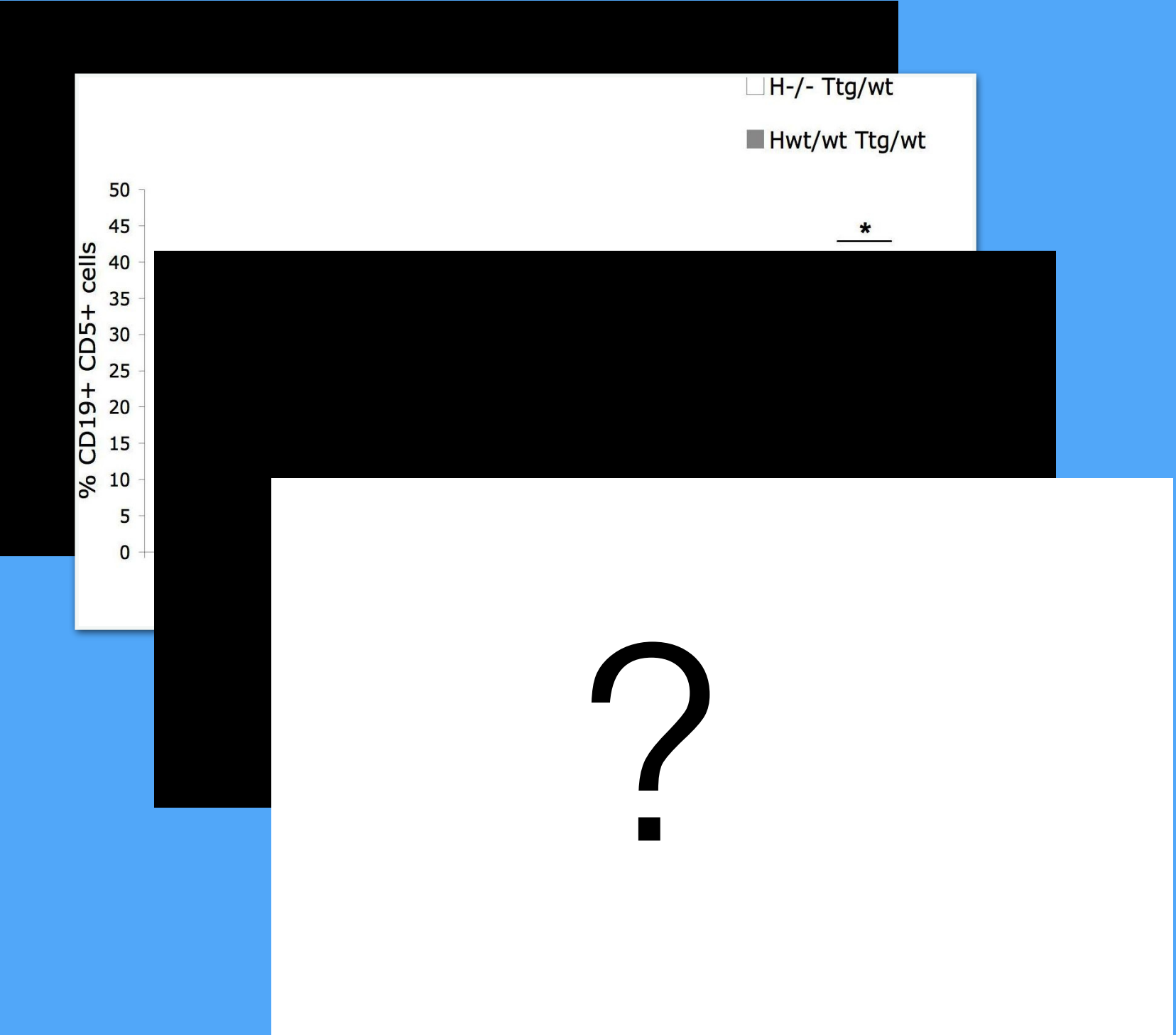


800ms

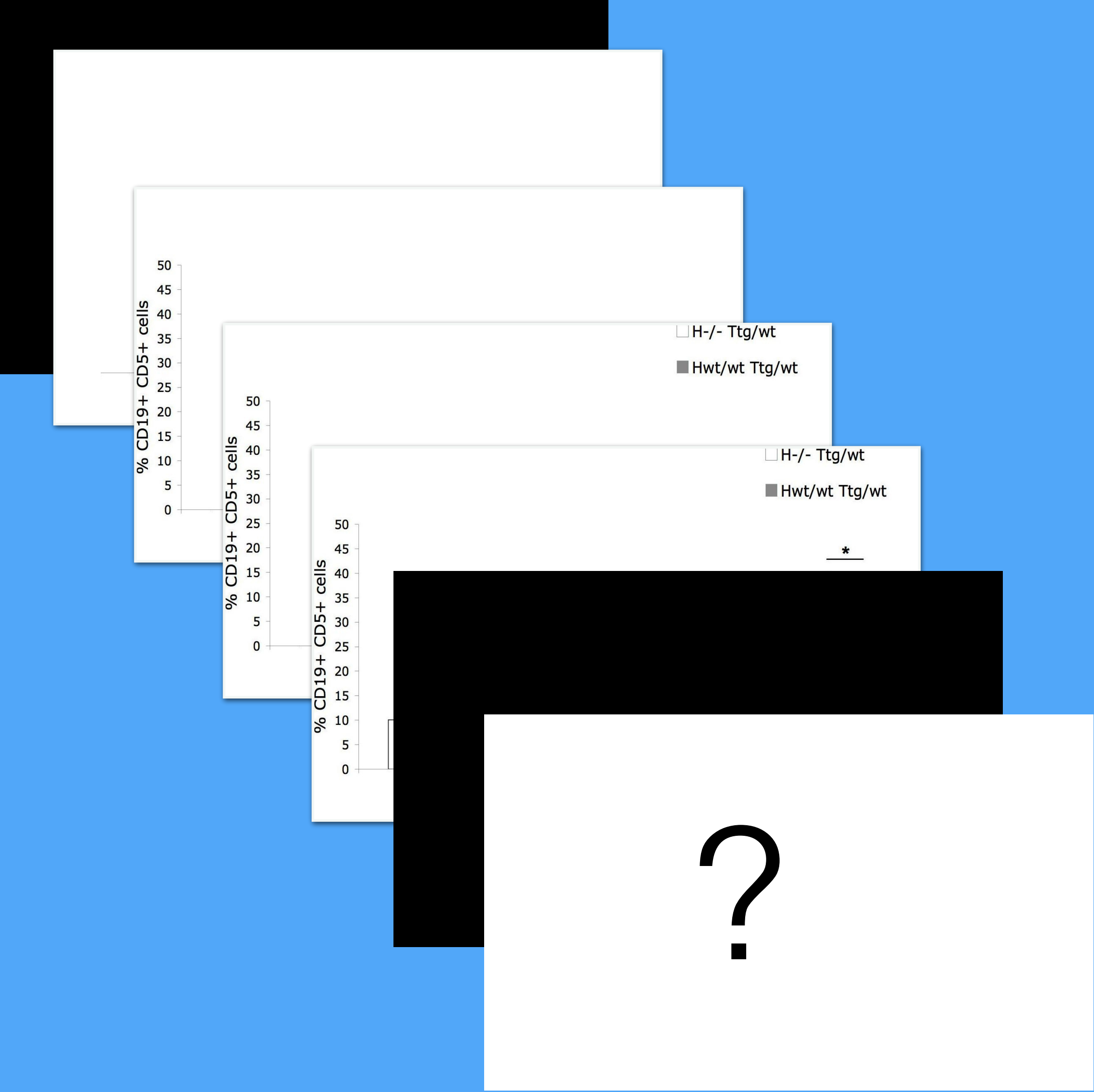
Asynchronous



Synchronous



Asynchronous



# Appreciate any feedback

(Presented by  
Nina(Zhuxiaona) Wei  
and Christine(Yentin) Lin)

Research question specific enough, meaningful?

Suggestions for experiment design?

- Within-Subjects or Between-Subjects design
- Should study different types of chart? Charts with different complexity?
- Independent or Dependent variables?

Suggestions for experiment materials?

- Professional statistical data charts/figures from papers, unfamiliar domain Biology, Chemistry etc.
- Online narrative visualization, Human Development Trend etc.

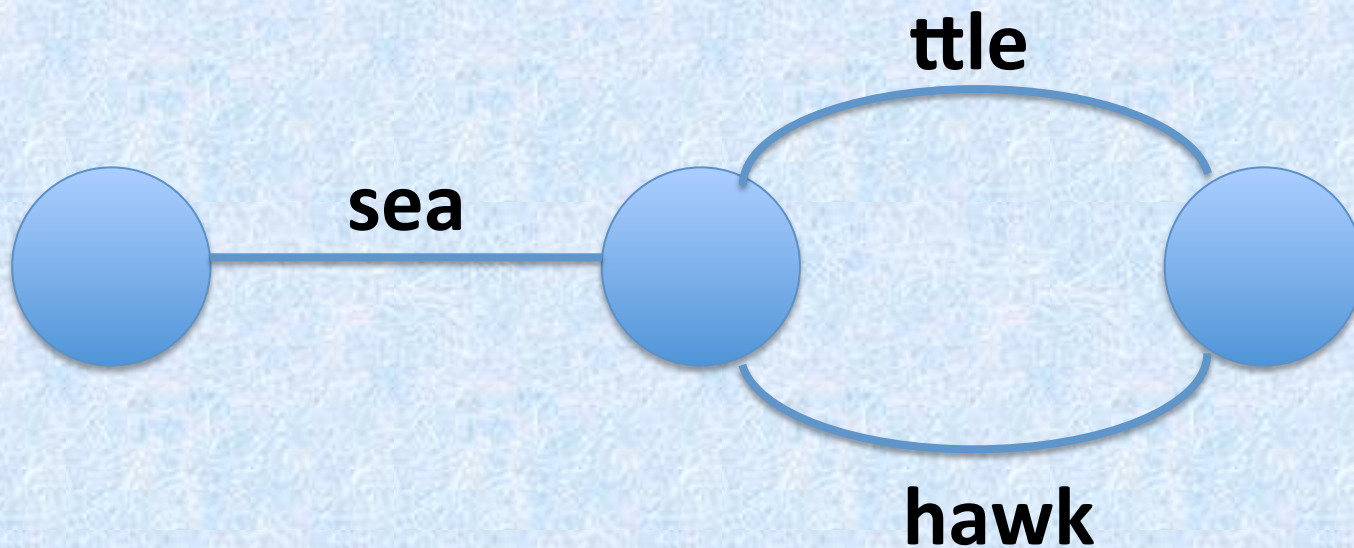
Others?



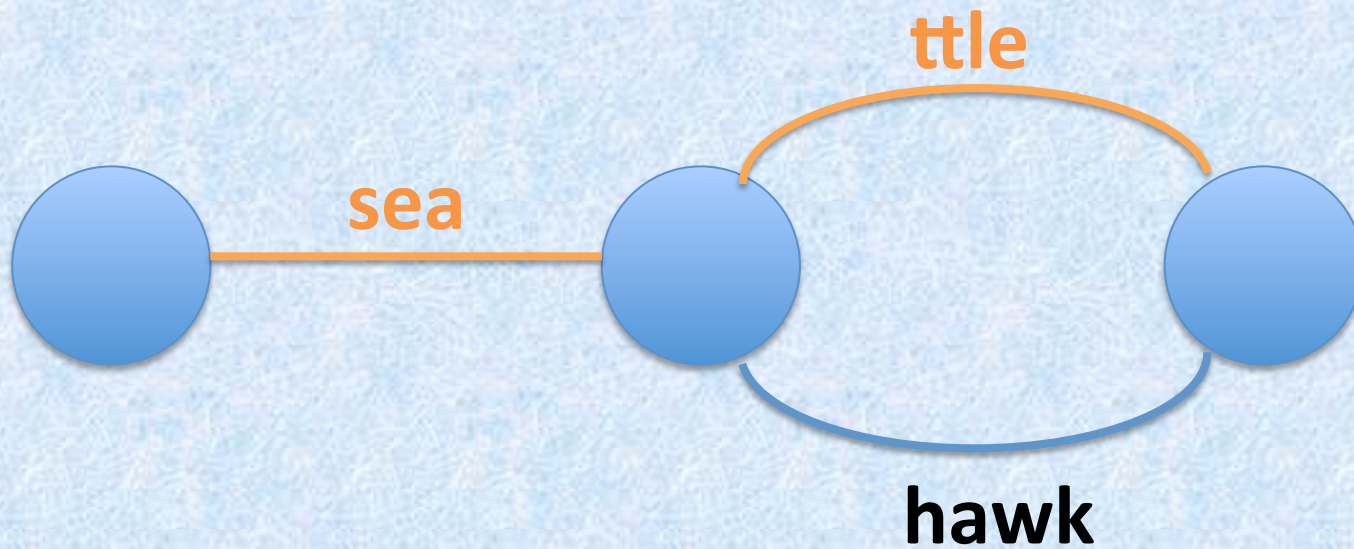
# Visualization of Lattice Structure

Shengjie Wang

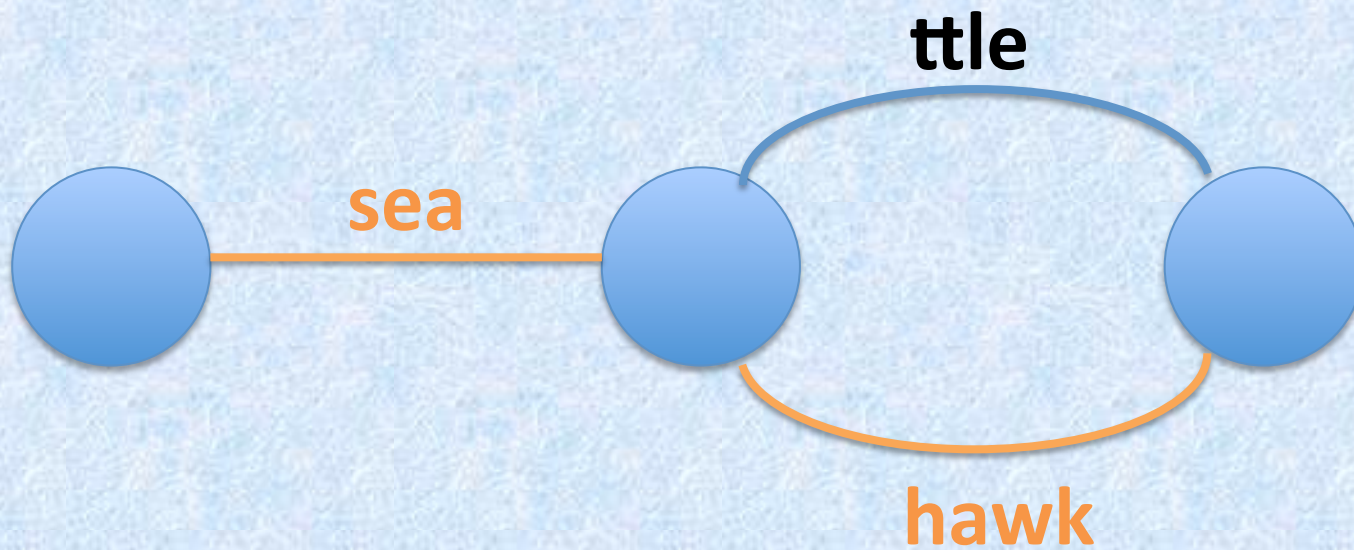
# A Simple Lattice



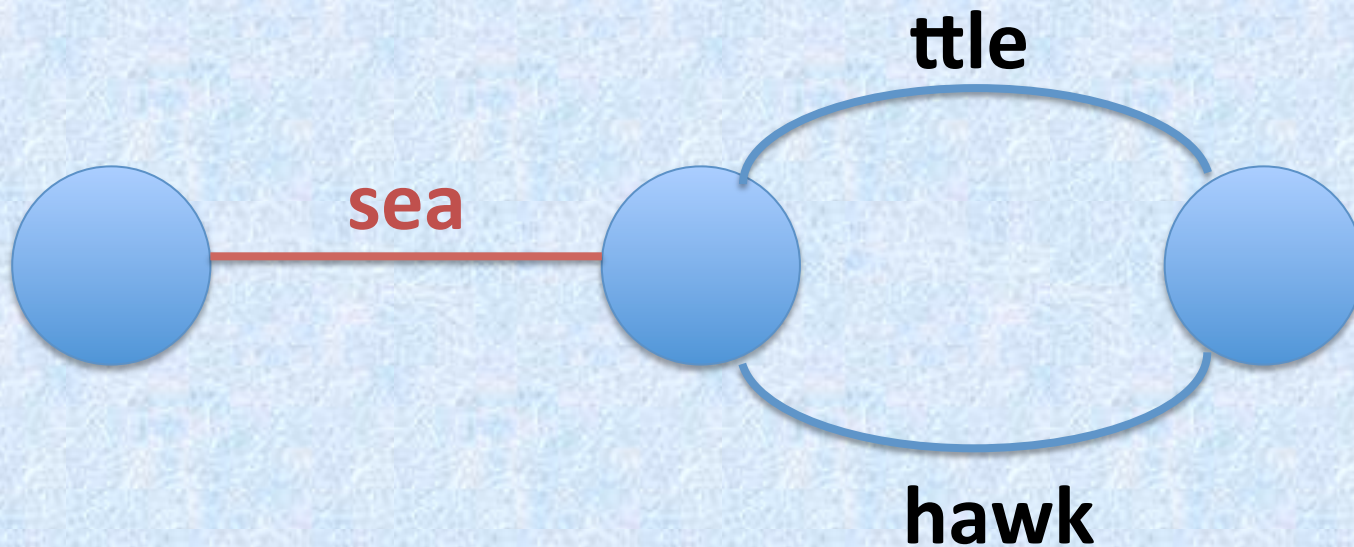
# Seattle



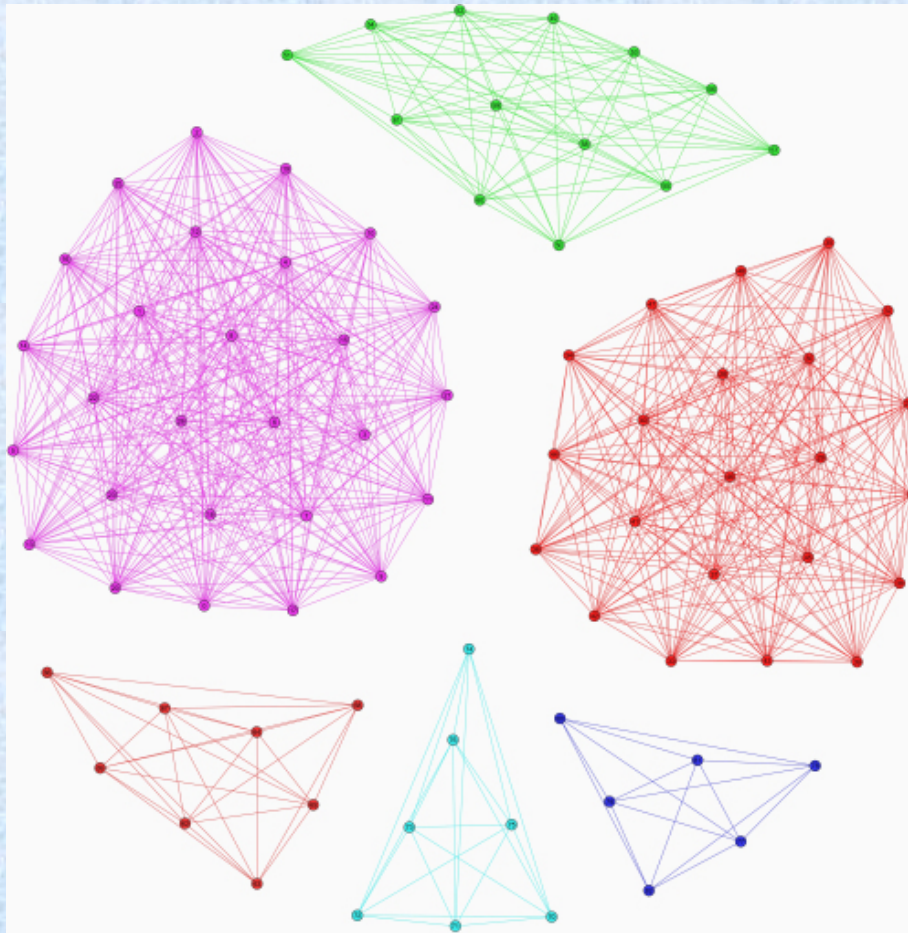
# Seahawk



# Common Edge: “sea”

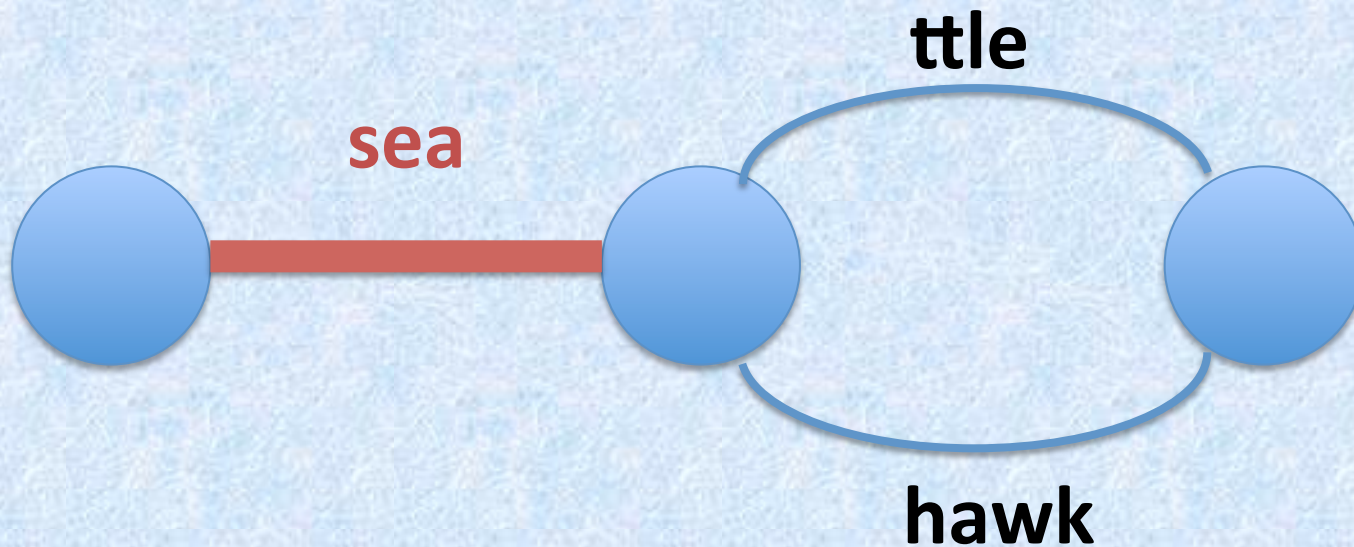


# Visual Encoding: Graph Cluster

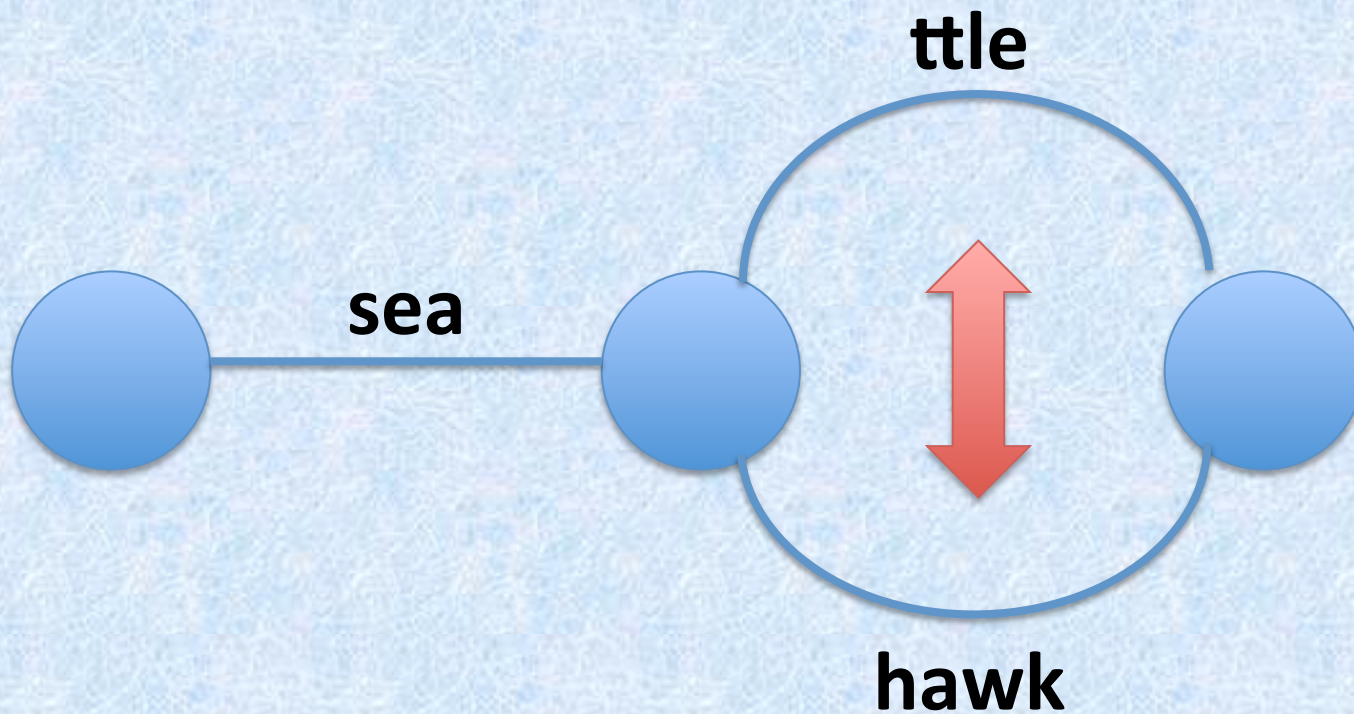


<https://wisonets.wordpress.com/2010/11/22/>

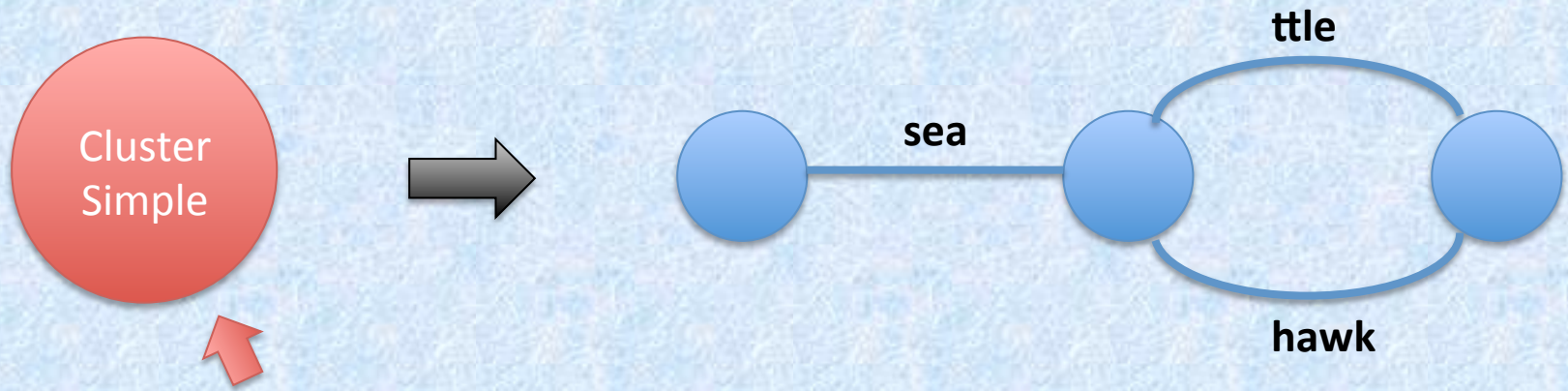
# Visual Encoding: Common Edge



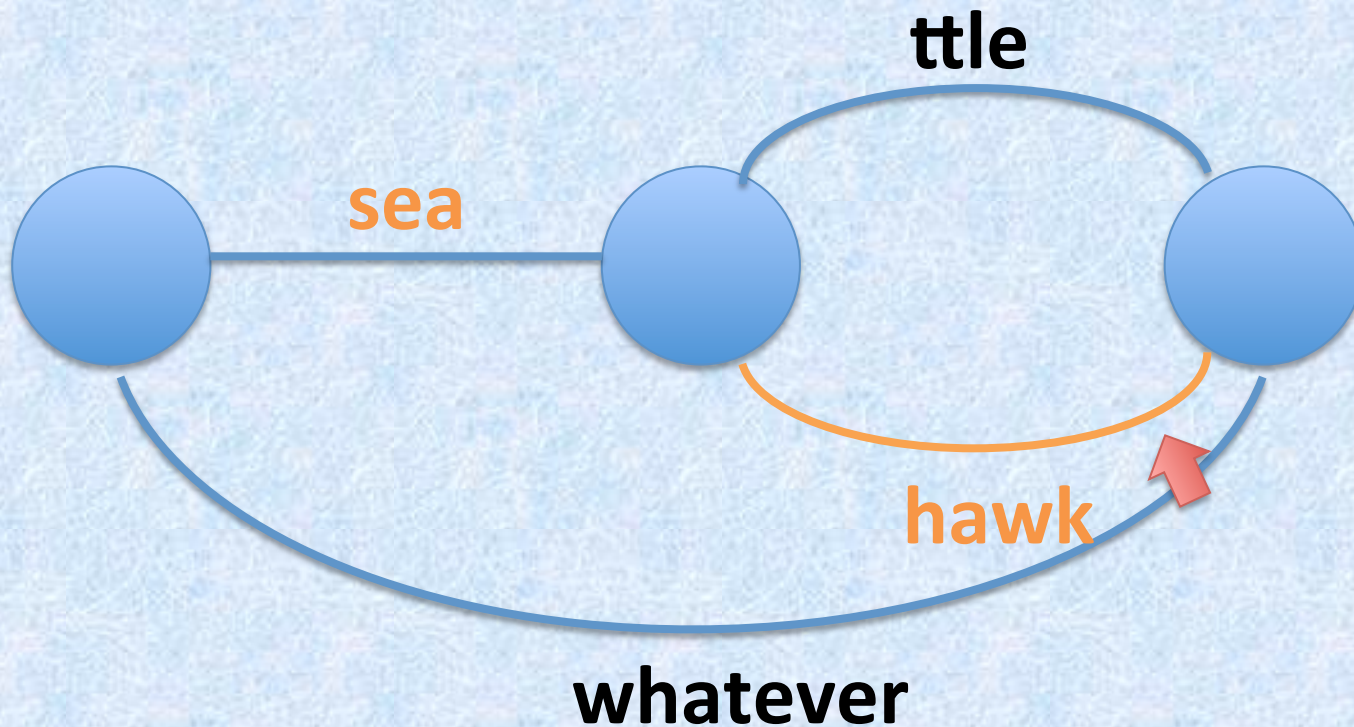
# Visual Encoding: Compelling Edges



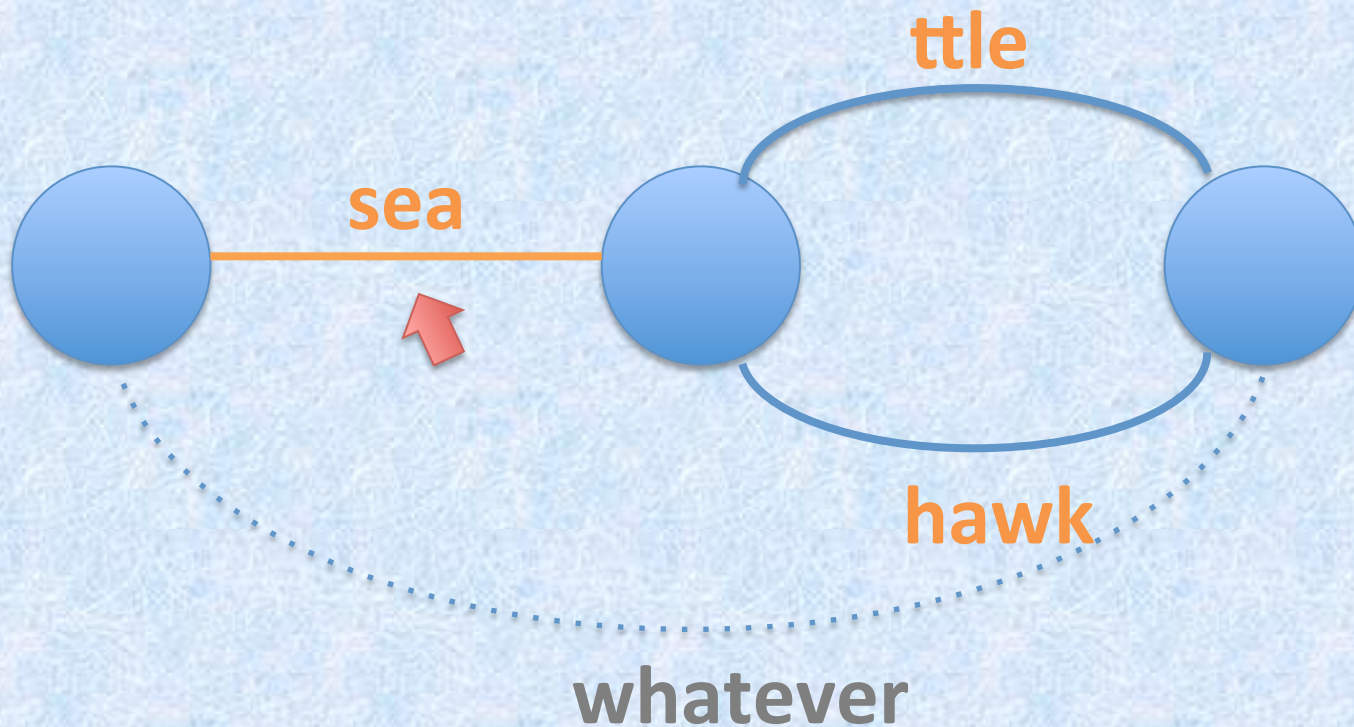
# Interaction: Drill Down Cluster



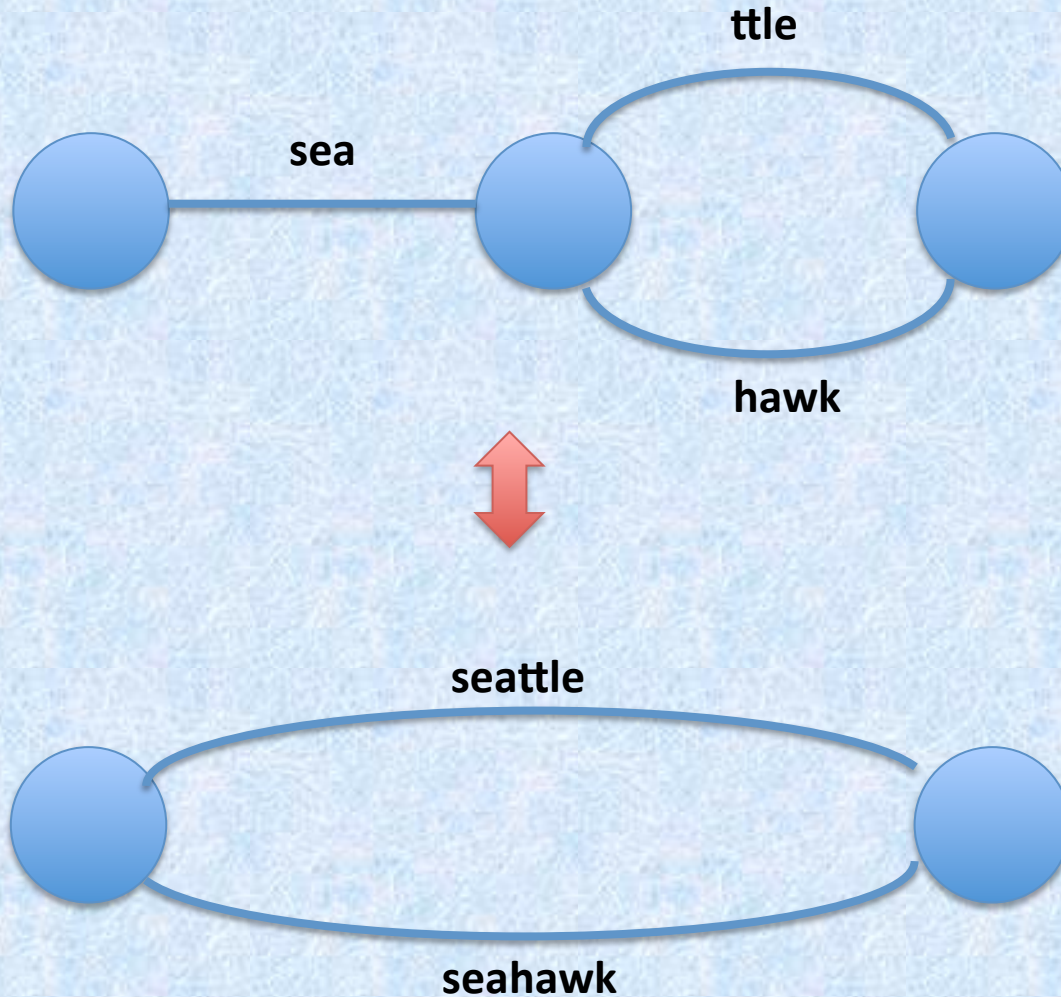
# Interaction: Query Edge



# Interaction: Query Edge



# Interaction: Add/Remove C.E.



# Related Works

- [1] Abello, James, Frank Van Ham, and Neeraj Krishnan. "Ask-graphview: A large scale graph visualization system." Visualization and Computer Graphics, IEEE Transactions on 12.5 (2006): 669-676.
- [2] Holten, Danny, and Jarke J. Van Wijk. "Force-Directed Edge Bundling for Graph Visualization." Computer Graphics Forum. Vol. 28. No. 3. Blackwell Publishing Ltd, 2009.
- [3] Schaeffer, Satu Elisa. "Graph clustering." Computer Science Review 1.1 (2007): 27-64.
- [4] Cui, Weiwei, et al. "Geometry-based edge clustering for graph visualization." Visualization and Computer Graphics, IEEE Transactions on 14.6 (2008): 1277-1284.

# Feedback

- Should we cluster the graph based on nodes?  
Can we effectively show the information of conveyed in each path if we apply the clustering?
- My name: Shengjie Wang



# Hamid Izadinia and Fereshteh Sadeghi

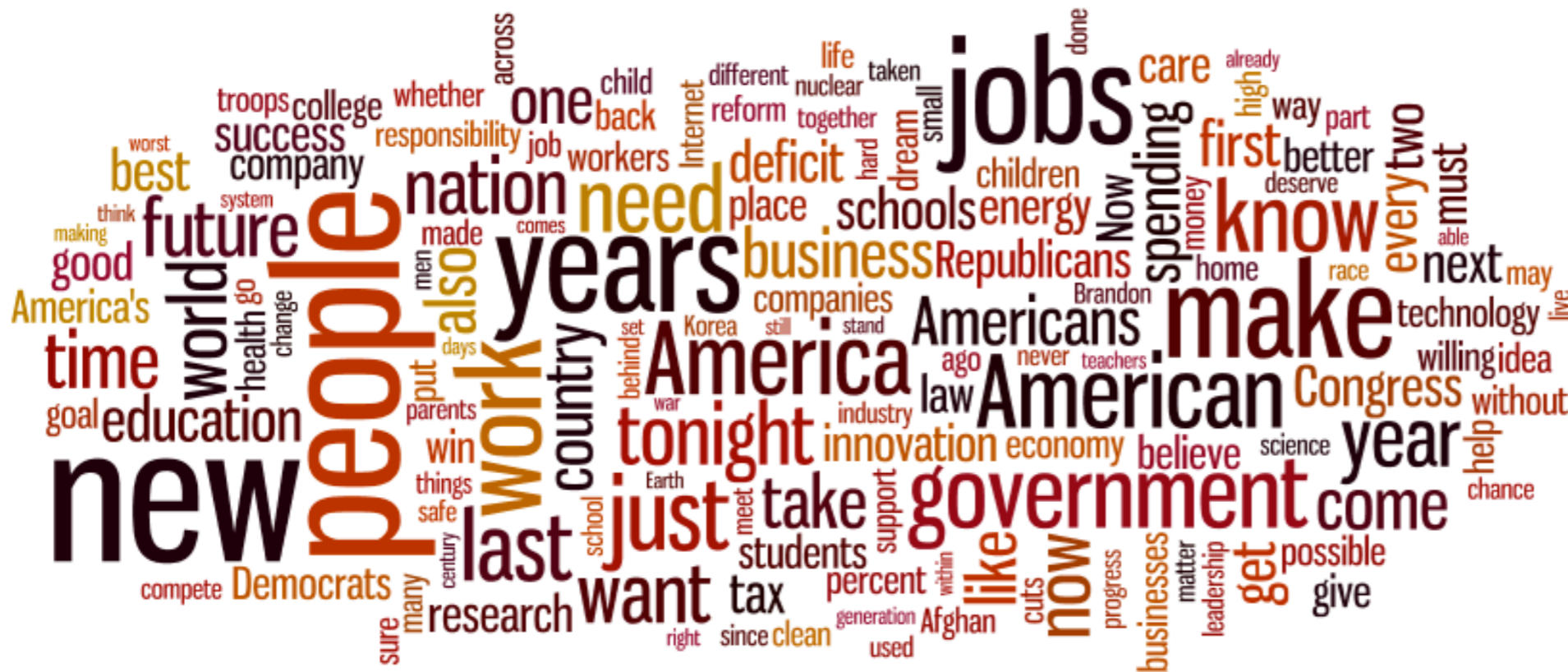
# Final project progress presentation

## Data Visualization

### (CSE512)

# From Text Visualization to Data

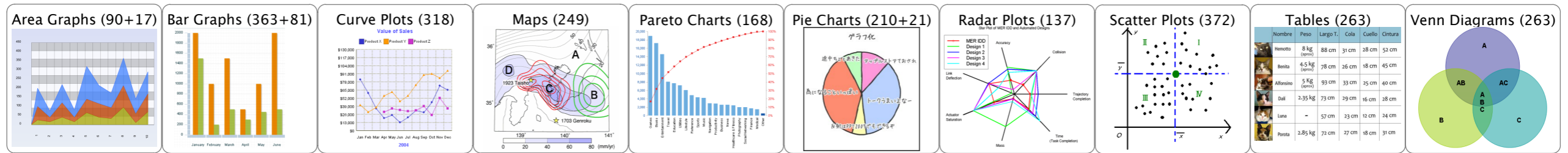
# Input



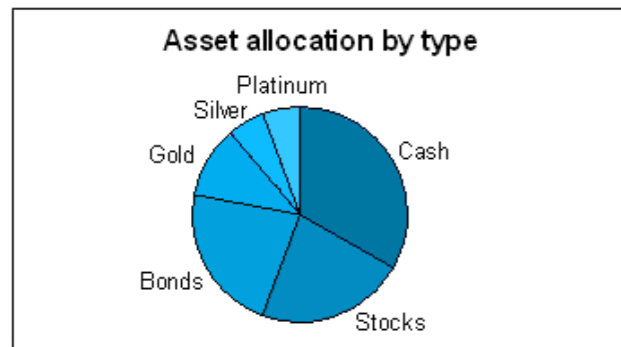
# Output

(People, 100)  
(new , 83 )  
(jobs , 78 )  
(years , 77 )  
(make , 69 )  
.  
.  
.  
(care , 8 )  
(child , 7 )

# Related Works

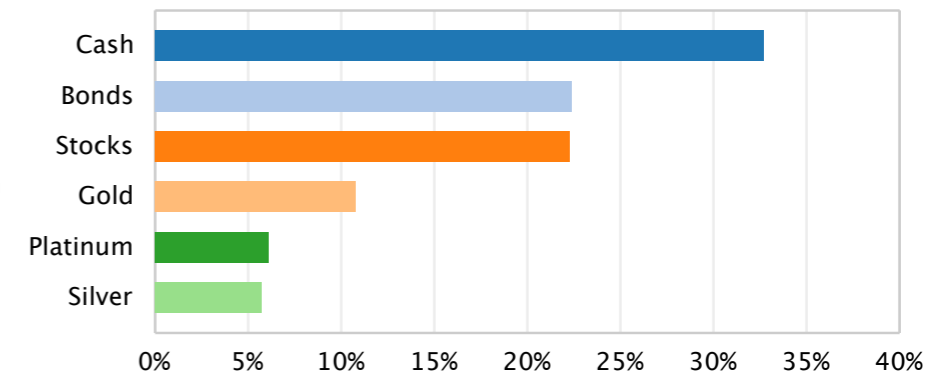


## Input Image (upload)

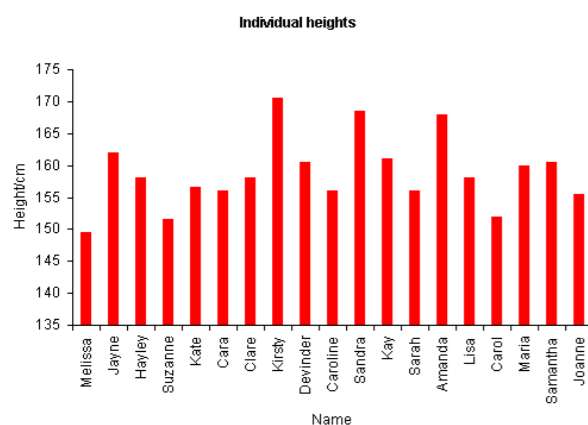


## Data Table (export)

Label	% of Total
Cash	33%
Bonds	22%
Stocks	22%
Gold	11%
Platinum	6%
Silver	6%

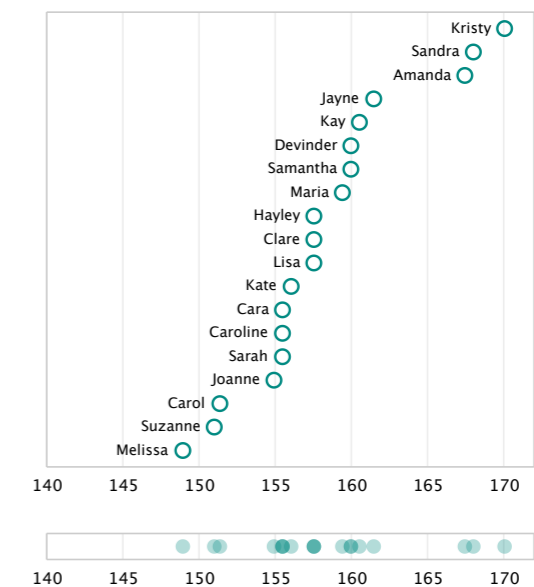


## Input Image (upload)



## Data Table (export)

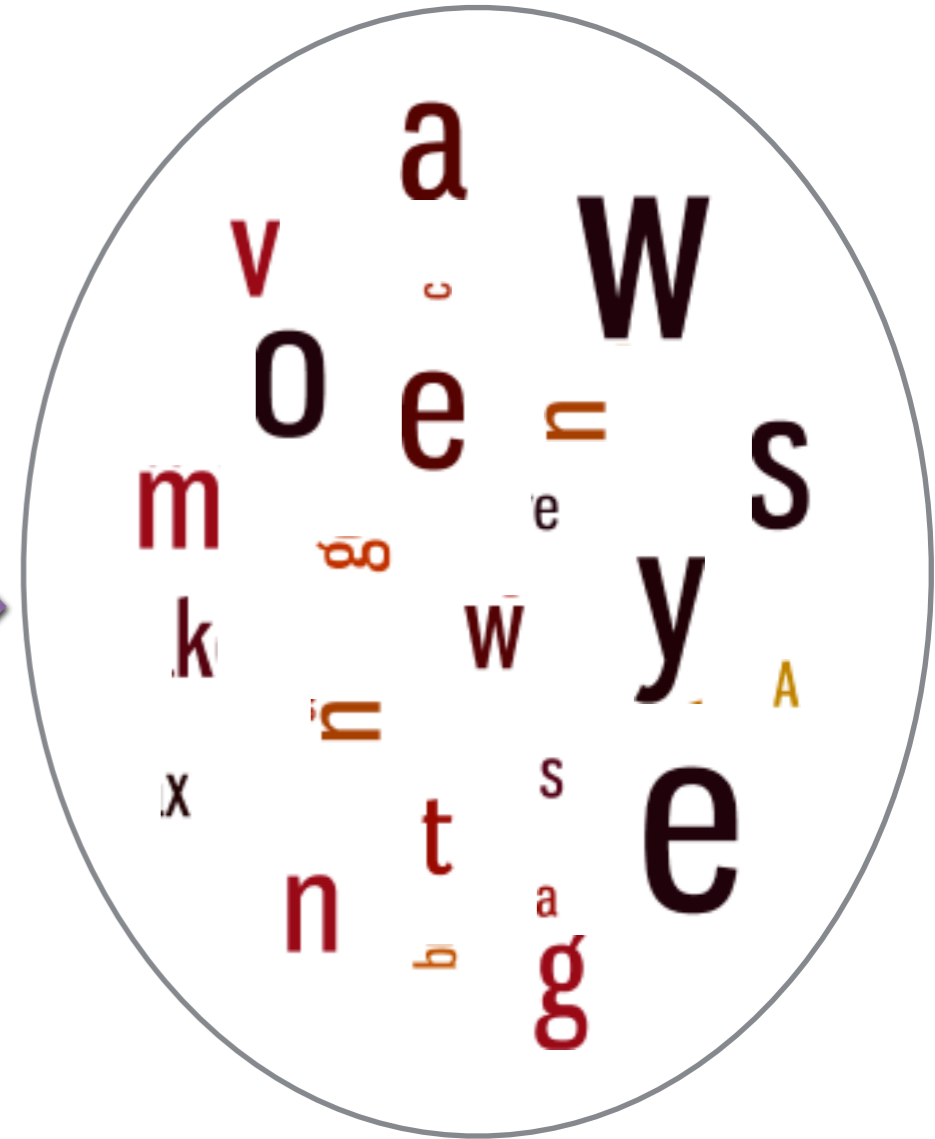
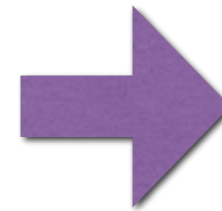
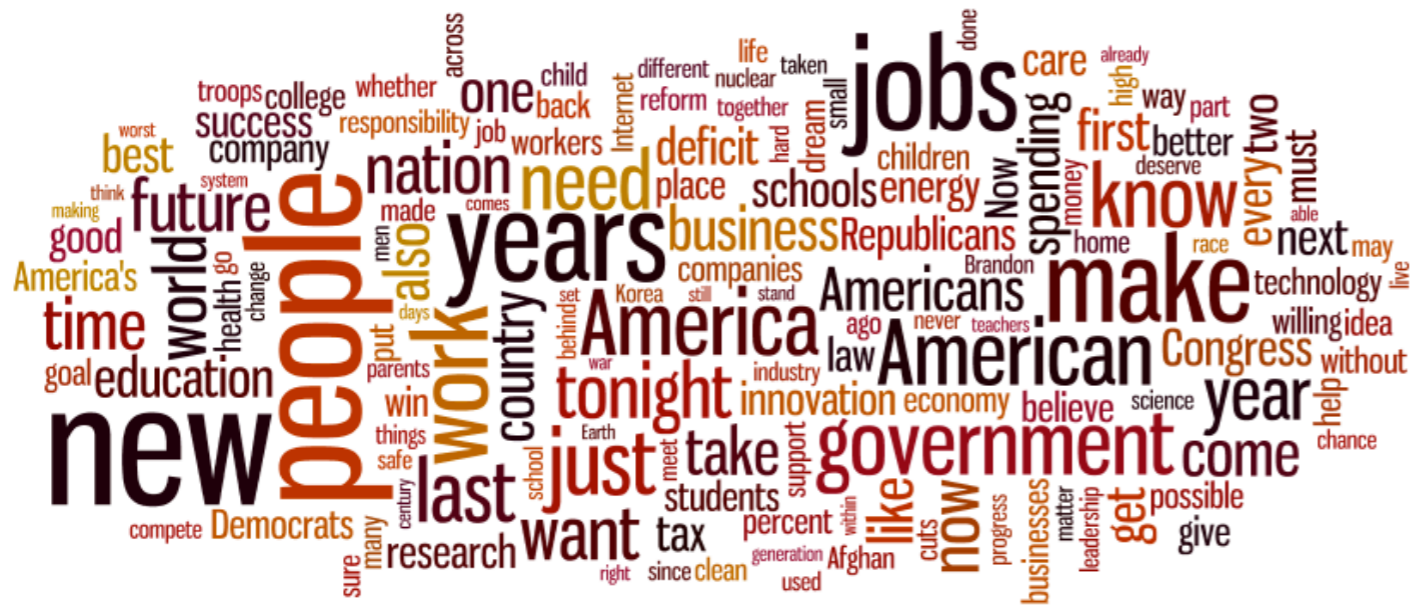
Label	Value
Kristy	170
Sandra	168
Amanda	167
Jayne	161
Kay	161
Devinder	160
Samantha	160
Maria	159
Hayley	158
Clare	158
Lisa	158
Kate	156
Cara	155
Caroline	155
Sarah	155
Joanne	155
Carol	151
Suzanne	151
Melissa	149



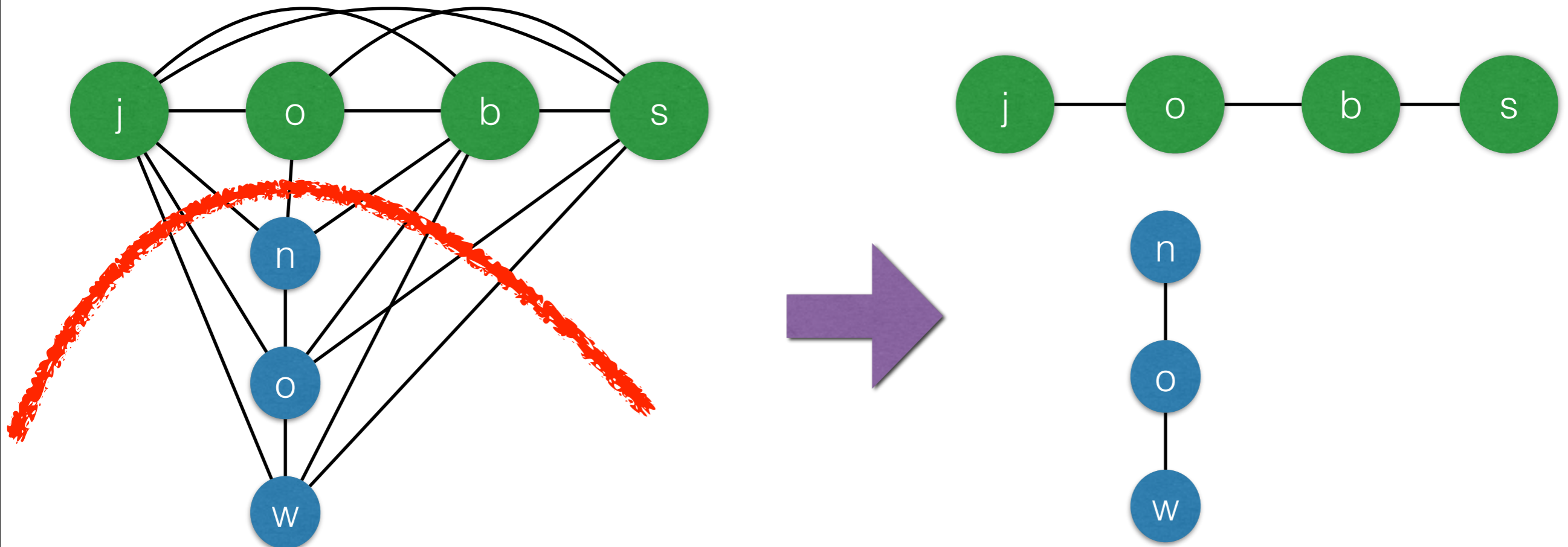
ReVision: automated classification, analysis and redesign of chart images.

Savva, M and Kong, N and Chhajta, A and , Fei-Fei L. and Agrawala, M and Heer, J.

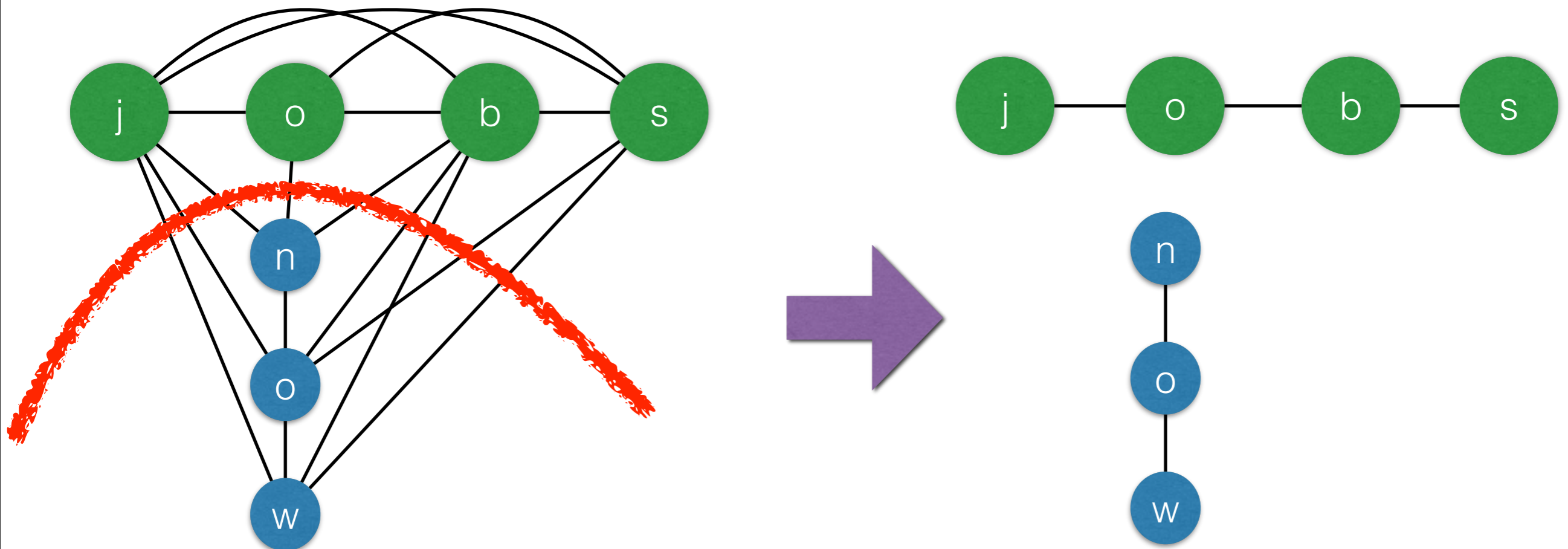
# Letter extraction & OCR



# Word extraction



# Word extraction & Size estimation



new

time

take

one

companies

able

# Challenges

- Dense arrangement of letters and words
- Different Fonts
- No color clue



# Feedback

- Is there any data set available for this task?
- What kind of evaluation is appropriate? Text
- Any idea for redesigning the text visualization
- Do you know of any scaling mechanism in standard text visualization softwares? (linear, log, ranking, ...)

# Visualize NLP Annotations for Crowdsourcing

Hanchuan Li, Haichen Shen, Shengliang Xu and Congle Zhang

# Statistical Nature Language Processing

- NLP: interaction between computer and human languages
- Statistical approaches have made great success



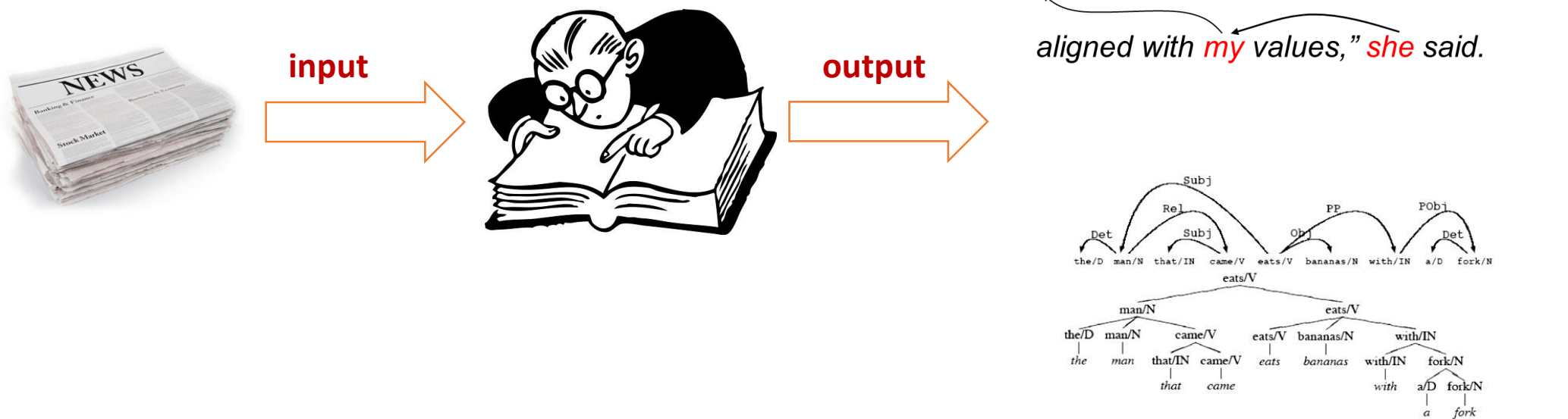
# But ...

- Need labeled training data, which is very expensive



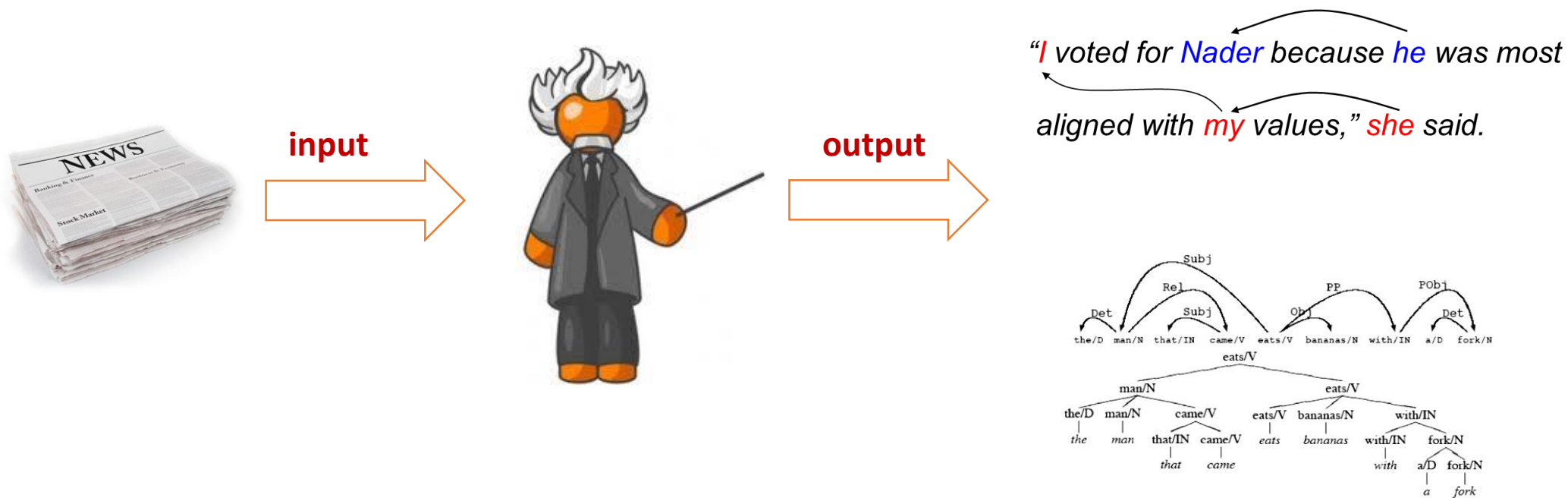
# Even worse...

- Structured prediction, expert only



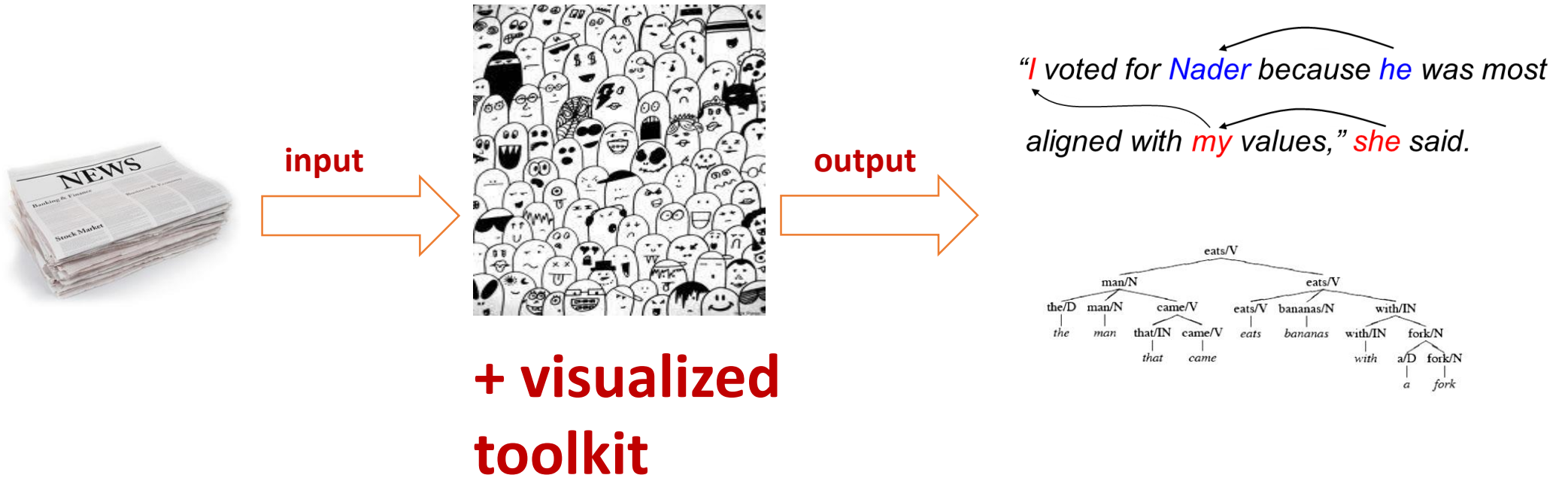
# Even worse...

- Structured prediction (generate trees/graphs), expert only



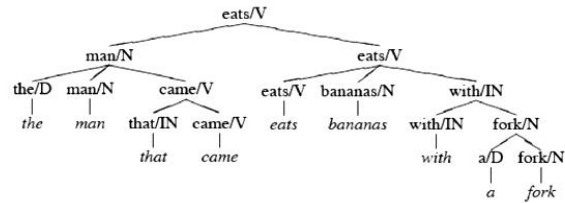
# Let Crowds do the job

- Visualize the NLP annotation

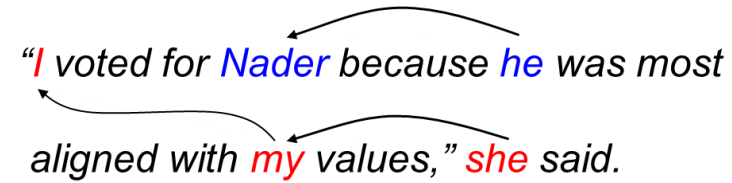


# Goal

- Visualize two categories of annotations



## Tree construction



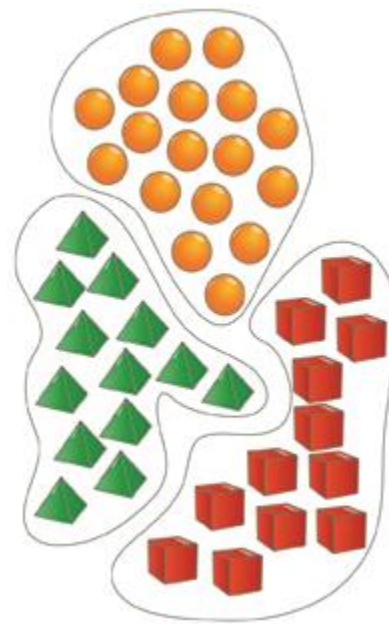
## Clustering

- Undo/redo

# Design

Input

“I vote for Nadar because he most aligned with my values”, she said.



Interaction

Link (he, Nadar)  
Tag (she, “speaker”)  
Link (I, my)

History

# Question

- Flexibility: easily transferred to other tasks
  - Take HTML as input, with target objects tagged.
  - Let NLP experts to design the input visualization
- Interface

Presented by  
Congle Zhang, Shengliang Xu, Haichen Shen, Hanchuan Li

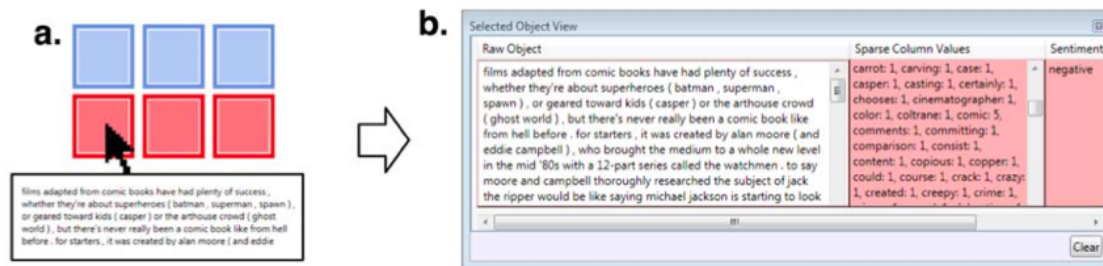
# Using Visualization to Optimize Text Classification

Jeff Snyder, Zorah Fung,  
Brian Walker, Zening Qu

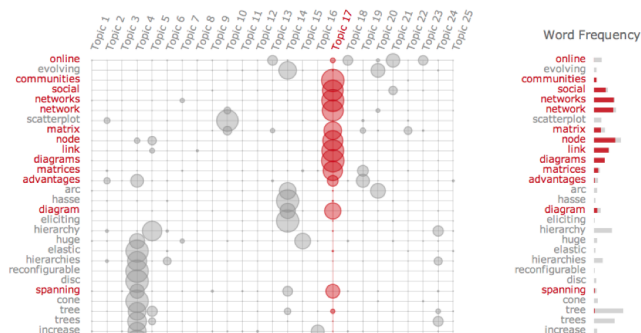
# Problem Description

- Text classification tasks are important in a wide variety of domains
- Create a model that assigns documents (tweets, papers) into classes ([important/not important], [positive/neutral/negative])
- We use visualization to
  - Help analysts to understand and tune the inner workings of their models,
  - Compare the performance of different models, at both a high level and on the level of individual documents, and
  - Understand the essential contours of their data that affect model performance.

# Related Work



Patel, Kayur, et al. "Gestalt: integrated support for implementation and analysis in machine learning." Proceedings of the 23rd annual ACM symposium on User interface software and technology. ACM, 2010.



Chuang, Jason, Christopher D. Manning, and Jeffrey Heer. "Termite: Visualization techniques for assessing textual topic models." Proceedings of the International Working Conference on Advanced Visual Interfaces. ACM, 2012.

# Prototype and Storyboard

## Tweet Classification

		Predicted Label		
		Pos	Neu	Neg
Actual Label	Pos	546	213	401
	Neu	235	497	102
	Neg	89	195	783

Lorem ipsum Lorem ipsum Lorem ipsum Lorem ipsum  
ipsum Lorem ipsum Lorem ipsum Lorem ipsum

Lorem ipsum Lorem ipsum Lorem ipsum Lorem ipsum  
Lorem ipsum Lorem ipsum

Lorem ipsum Lorem ipsum Lorem ipsum Lorem ipsum  
ipsum Lorem ipsum Lorem ipsum Lorem ipsum

Contributed against  
classification



Contributed to  
classification



Regularization



Model for predicted label



Model for actual label

# Prototype and Storyboard

## Tweet Classification

		Predicted Label		
		Pos	Neu	Neg
Actual Label	Pos	546	213	401
	Neu	235	497	102
	Neg	89	195	783

Lorem ipsum Lorem ipsum Lorem ipsum Lorem ipsum  
ipsum Lorem ipsum Lorem ipsum ipsum Lorem ipsum

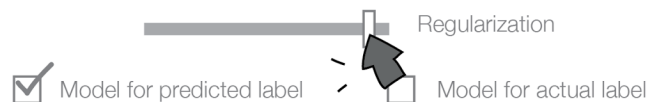
Lorem ipsum Lorem ipsum Lorem ipsum Lorem ipsum  
Lorem ipsum Lorem

Lorem ipsum Lorem ipsum Lorem ipsum Lorem ipsum  
ipsum Lorem ipsum Lorem ipsum

Contributed against  
classification



Contributed to  
classification



# Demo

<http://cse512-14w.github.io/fp-jasnyder-zorahf-bdwalker-zqu/web>

Jeff Snyder, Zorah Fung, Brian Walker, Zening Qu

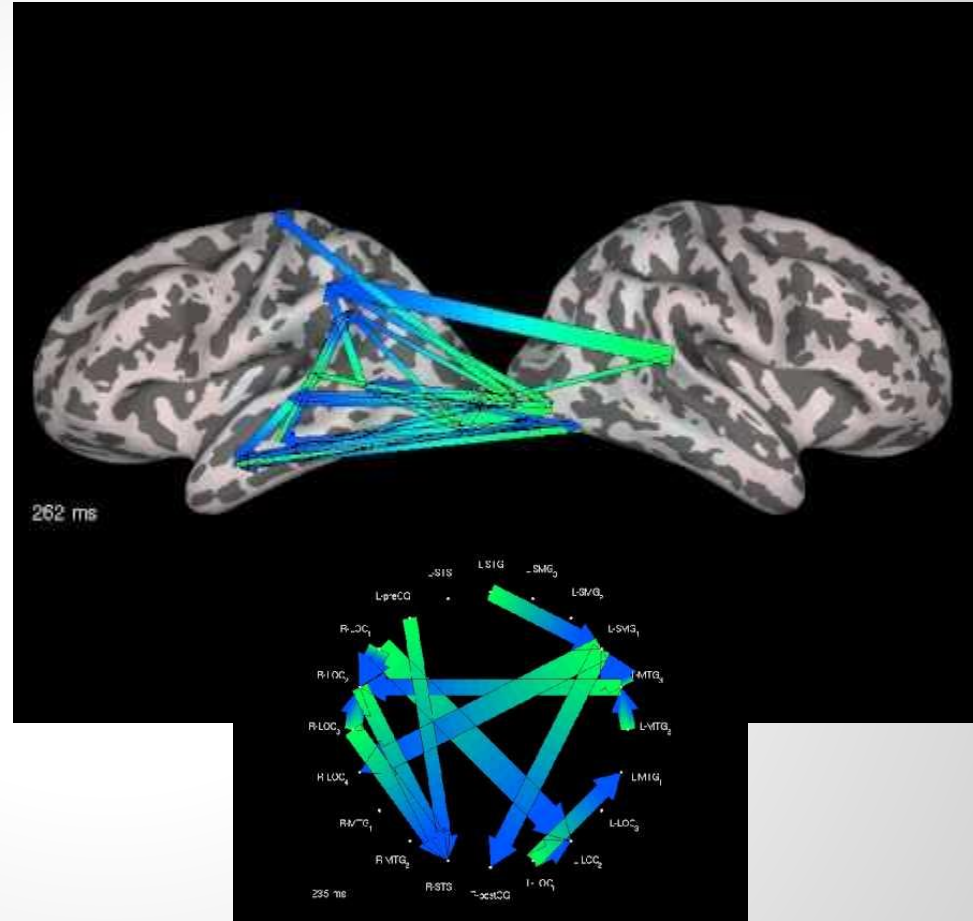
# **CSE512 Final Project**

Yang Chaoyu, Aniket Handa, Gregory Nelson, Alexander Conrad Nied

# Motivation

Networks +  
Time Varying Edge Data

Dangerous amounts of  
information, may not be  
meaningful.



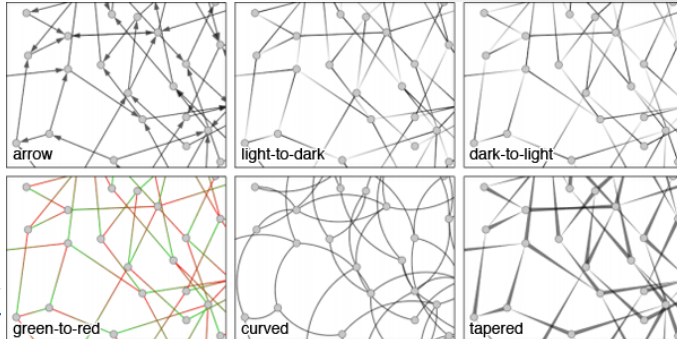
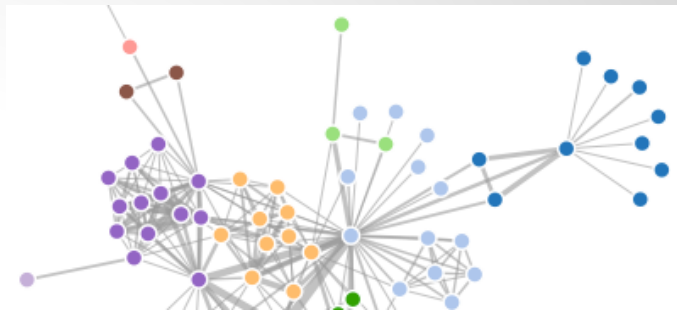
# Workspace

Using D3

Traffic flow is very popular

Not integrated to D3

Time-varying visualizations rare



<http://bl.ocks.org/mbostock/4062045>

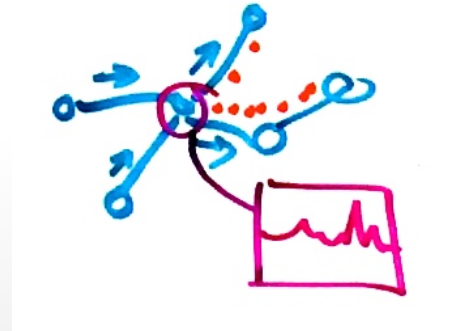
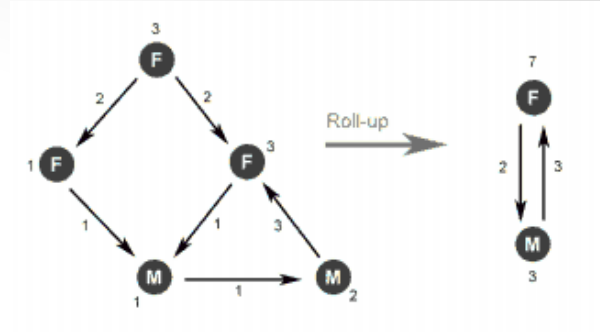
<http://boundary.com/blog/2012/05/14/visualizing-network-flow-data/>

# Methods

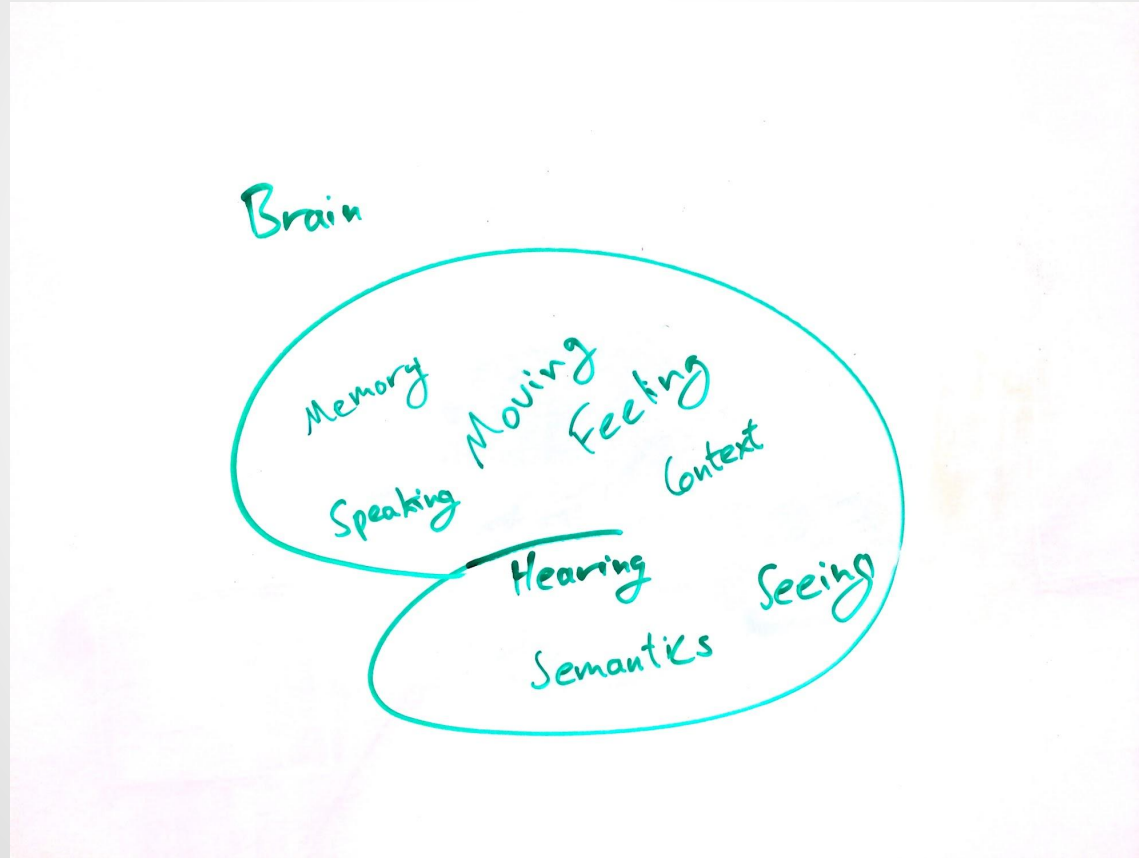
“PivotGraph” & roll-up

Featural Edges or  
Animated Packets

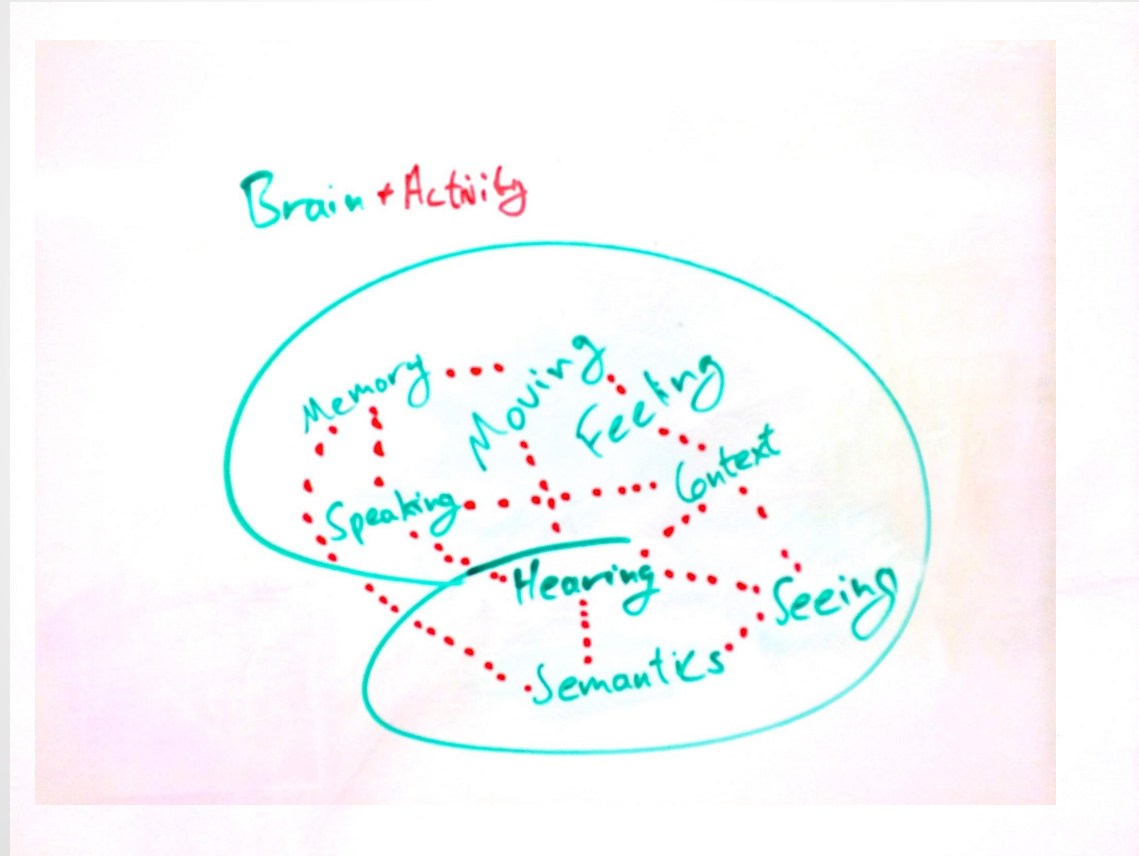
Sensors - Select a node  
Get history/plots and traces



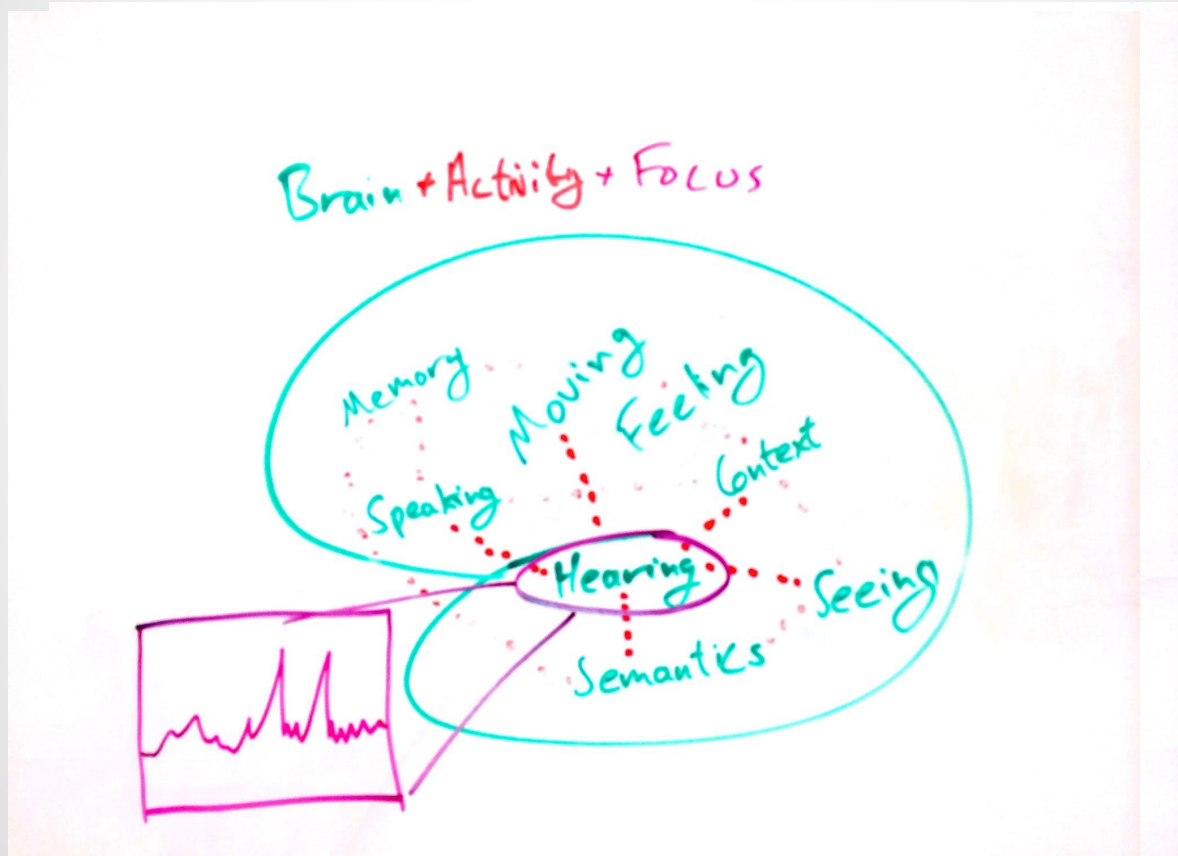
# Storyboard



# Storyboard



# Storyboard



# Questions

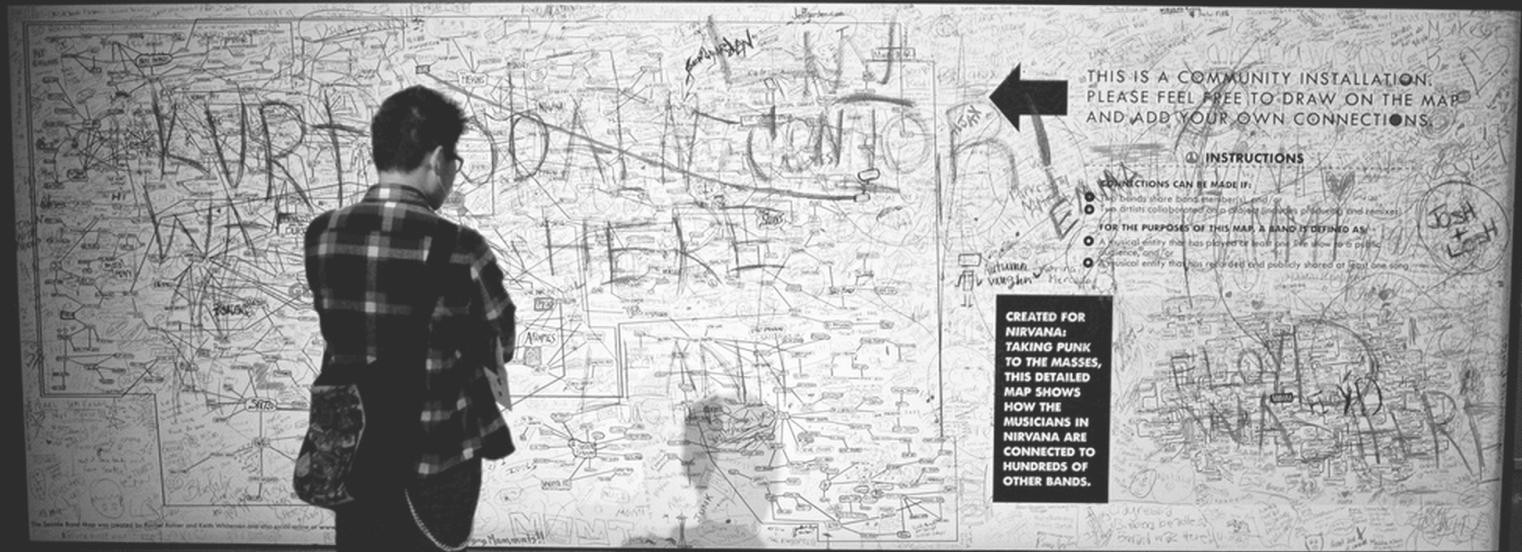
Did we miss any existing work?

Do you have a juicy dataset for this?

Do you have questions?

# SEATTLE BAND MAP

THE SEATTLE BAND MAP SHOWCASES THE NORTHWEST'S VIBRANT MUSIC SCENE BY DOCUMENTING HUNDREDS OF REGIONAL BANDS THAT HAVE PERFORMED THROUGHOUT THE DECADES. IT ALSO EXPLORES HOW THESE BANDS INTERCONNECT THROUGH PERSONAL RELATIONSHIPS AND COLLABORATIONS.



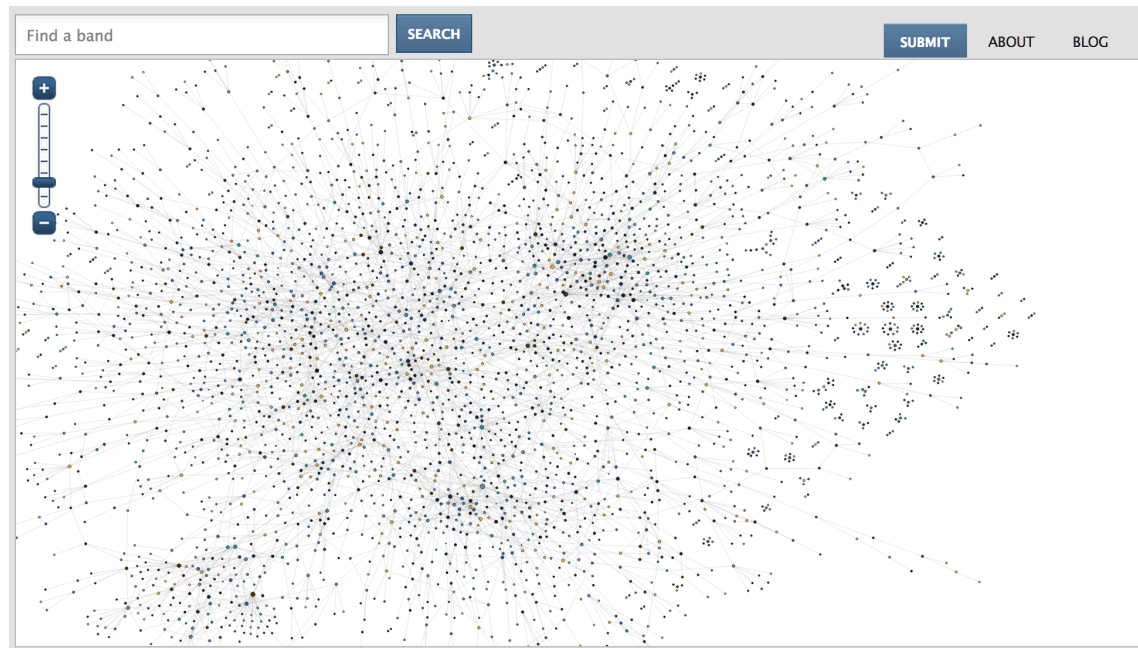
## Seattle Band Map

CSE 512 Final Project

Chase Wu | Susanne Hsu

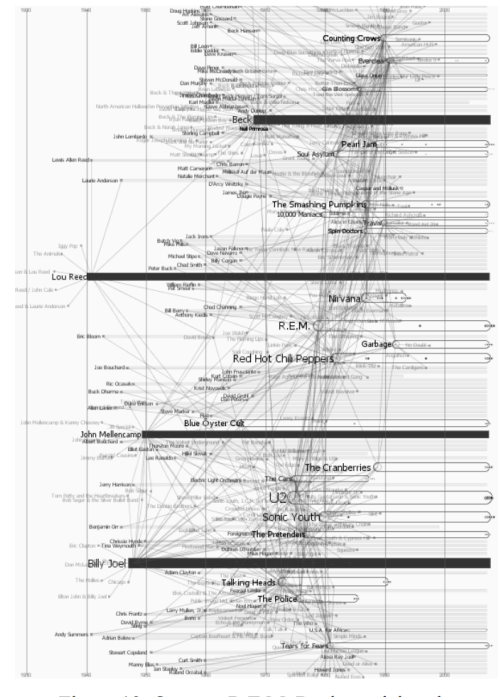
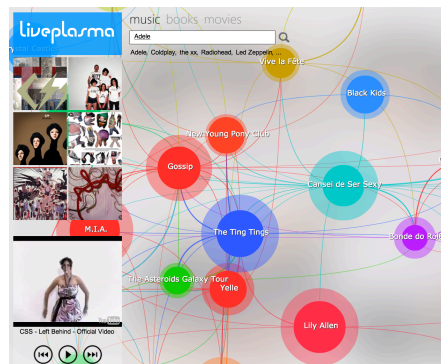
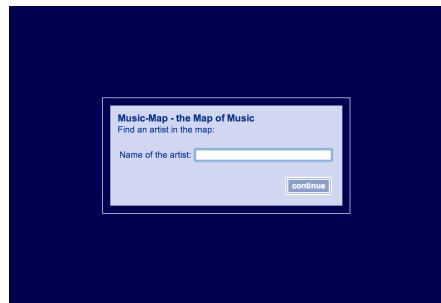
# Seattle Band Map

The current visualization is difficult to navigate and is not organized in a way that supports exploration or music discovery



## Relevant Work

Existing visualizations force users to define a starting point, rather than using the network structure to reveal areas of interest and guide discovery



Metal

Alternative  
Rock

Death Cab For Cutie

Death Cab For Cutie



1997 - Present  
Bellingham, WA

...more

Grunge Rock

Electronic Pop

## SEATTLE BAND MAP

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed velit neque, hendrerit vel quam quis, commodo lacinia ligula. Nullam fermentum egestas tellus in hendrerit. Vivamus et adipiscing dui, quis placerat massa. Fusce mollis urna et neque eleifend, in auctor felis euismod..

### Filter

1950  
TIME 2014

detail  
SCALE overview

VIEW

GENRE TIME PLACE

## Challenges

How to best support exploration and discovery: what's the most appropriate clustering and layout algorithm?

What information would you want to see/hear on the map and what would you want to do with it?

What information should be available in the global view vs. zoomed in view?

Susanne Hsu, Chase Wu