CourseRatings

CSE 512 Final Project

The problem

- UW’s course evaluation catalog has an unintuitive interface
- Only data for the past 2 - 3 quarters is available
- Difficult to compare classes

Course Evaluation Catalog GAMEDES - GWSS

Information on this site is intended for use by members of the UW academic community. Publication or redistribution of this information to any individual without a current UW NetId is prohibited.

UW Bothell GAMEDES 110 Lee Hammock Instructor AU14
UW Bothell GAMEDES 210 Jay Schneider Instructor WI15
UW Bothell GAMEDES 220 A Daniel Smith Instructor SP14
Executive Master Bus Admin Group GEMBA 514 G Patrick Bettin Lecturer SP14
Genome Sciences GENOME 351 A Leo Pallanck Professor SP14
Genome Sciences GENOME 351 A Evan Eichler Professor SP14
Genome Sciences GENOME 361 A Joshua Akey Professor WI15
Genome Sciences GENOME 361 A Colin Manoil Professor WI15
Genome Sciences GENOME 361 A Frances Cheong Lecturer SU14
Genome Sciences GENOME 361 A Keisha Carlson Lecturer AU14
Genome Sciences GENOME 361 A Frances Cheong Lecturer AU14
Genome Sciences GENOME 361 A Willie Swanson Professor SP14
Genome Sciences GENOME 361 A Maitreya Dunham Assistant Professor SP14
Genome Sciences GENOME 371 A M Raghuraman Associate Professor AU14
Genome Sciences GENOME 372 A John Stamatoyannopoulos Associate Professor AU14

Genome Sciences GENOME 351 A

Leo Pallanck Professor SP14

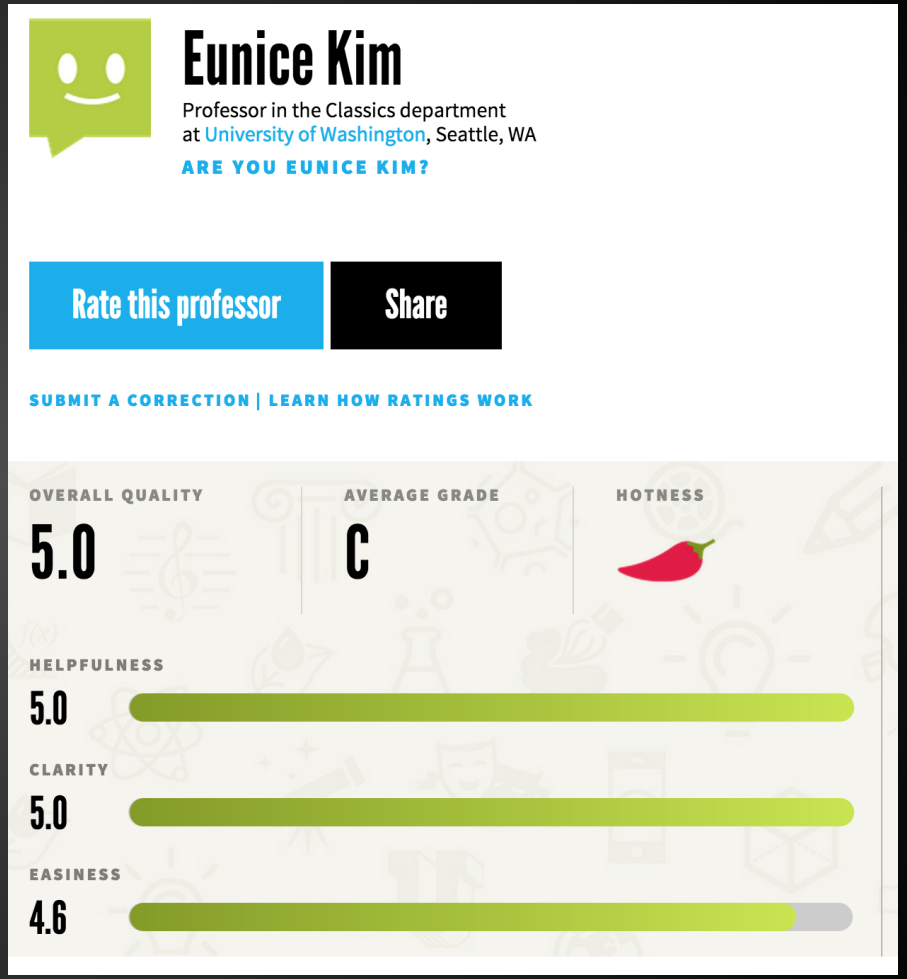
Form B: Large Lecture "40" surveyed "56" enrolled

Question	Excellent	Very Good	Good	Fair	Poor	Very Poor	Median
The course as a whole:	15%	32%	32%	15%	5%	0%	3.42
The course content:	18%	32%	35%	15%	0%	0%	3.50
Instructor's contribution:	28%	22%	30%	20%	0%	0%	3.50
Instructor's effectiveness:	25%	22%	25%	20%	8%	0%	3.40
Instuctor's interest:	30%	30%	22%	12%	2%	2%	3.83
Amount learned:	18%	30%	28%	18%	8%	0%	3.41
Grading techniques:	15%	33%	41%	8%	0%	3%	3.47

For median calculation: 5 = Excellent 4 = Very Good 3 = Good 2 = Fair 1 = Poor 0 = Very Poor

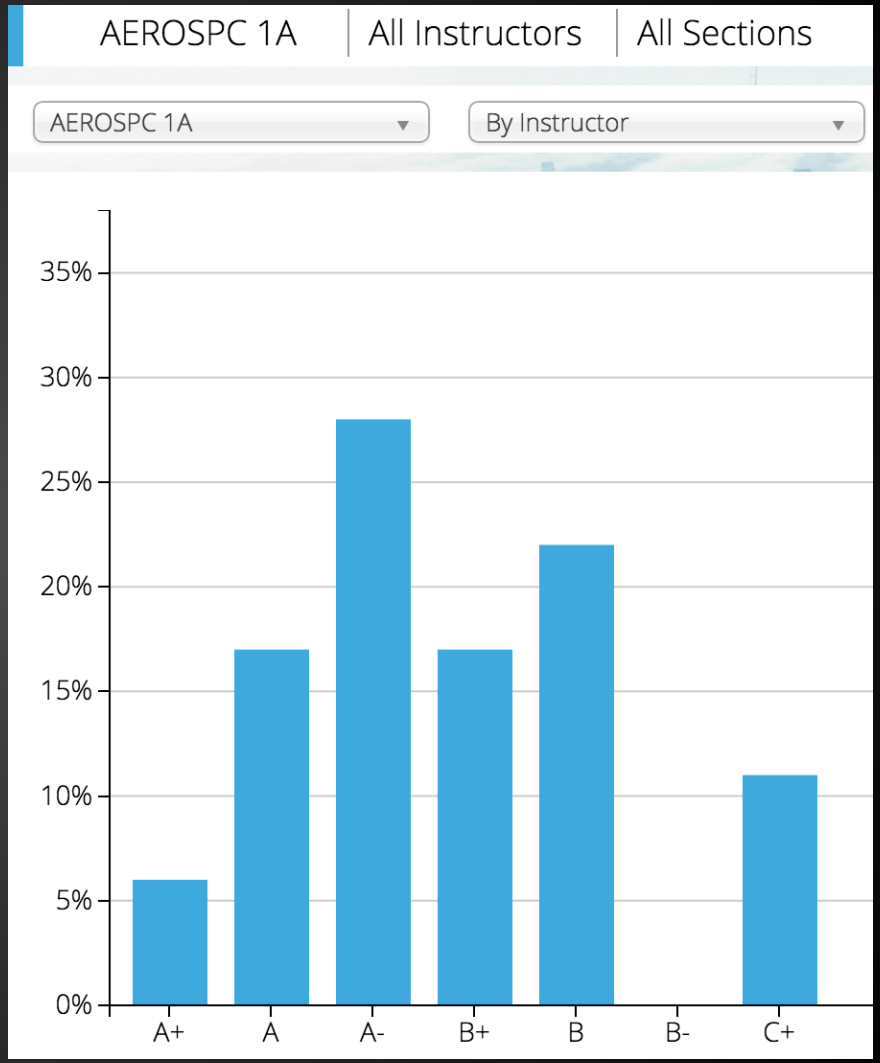
Prior work

- RateMyProfessors
  - Website where students can rate professors
  - Selection bias leads to unreliable ratings
  - Few entries per professor



Prior work

- BerkeleyTime
  - Website that allows students to view grade distributions for classes
  - Allows comparison between classes



How we’re different

- We display evaluations of all students, not just students who wanted to leave a review
- We do not display grade distributions

Recap

- Website that allows students to search through and sort information on courses
- Two years worth of data

http://students.washington.edu/drapeau/course_ratings

Instructor ↓	Overall	Content	Amount Learned	Teaching	Grading
Allison Obourn	3.97	4.16	4.21	3.73	4
Allison Obourn	4.3	4.4	4.45	4.26	4.06
Benson Limketkai	3.95	3.99	4.18	4.16	3.75
Helene Martin	4.66	4.62	4.67	4.72	4.28
Helene Martin	4.66	4.66	4.81	4.84	4.68
Helene Martin	4.84	4.84	4.85	4.85	4.67
Scott Petersen	3	3.33	3.25	3	2.13
Stuart Reges	4.63	4.56	4.64	4.74	4.17
Stuart Reges	4.7	4.71	4.76	4.84	4.45
Tyler Rigsby	4.3	4.37	4.64	4.66	4.01
Whitaker Brand	4.16	4.27	4.35	4.32	3.84
William Paul Osborne	4.28	4.21	4.59	4.25	4.2
Zorah Fung	4.29	4.31	4.61	4.59	4.06

New features

- A column indicating % of students in class who filled out the survey
- Course descriptions on each course page
- Collapsing identical courses
- Search redesign
- Department visualizations

COURSERATINGS

Department ▽ | Course code ▽ | Instructor ▽

CSE 142 : Computer Programming 1

Basic programming abilities and concepts including procedural programming, basic control structures, arrays, and objects. Intended for students without prior programming experience. Offered AWSpS.

Instructor ↓	Overall	Content	Amount Learned	Grading	Teaching	% filled out
Allison Obourn	~	~	~	~	~	~
Benson Limketkai	~	~	~	~	~	~
Helene Martin	~	~	~	~	~	~
Scott Petersen	~	~	~	~	~	~
Stuart Reges	~	~	~	~	~	~
Tyler Rigsby	~	~	~	~	~	~
Whitaker Brand	~	~	~	~	~	~
William Osborne	~	~	~	~	~	~

Feedback

- What kind of visualizations would you like to see for each department?
- What other features would you like to see?

HIP HOP COLLABORATION IN THE UNITED STATES

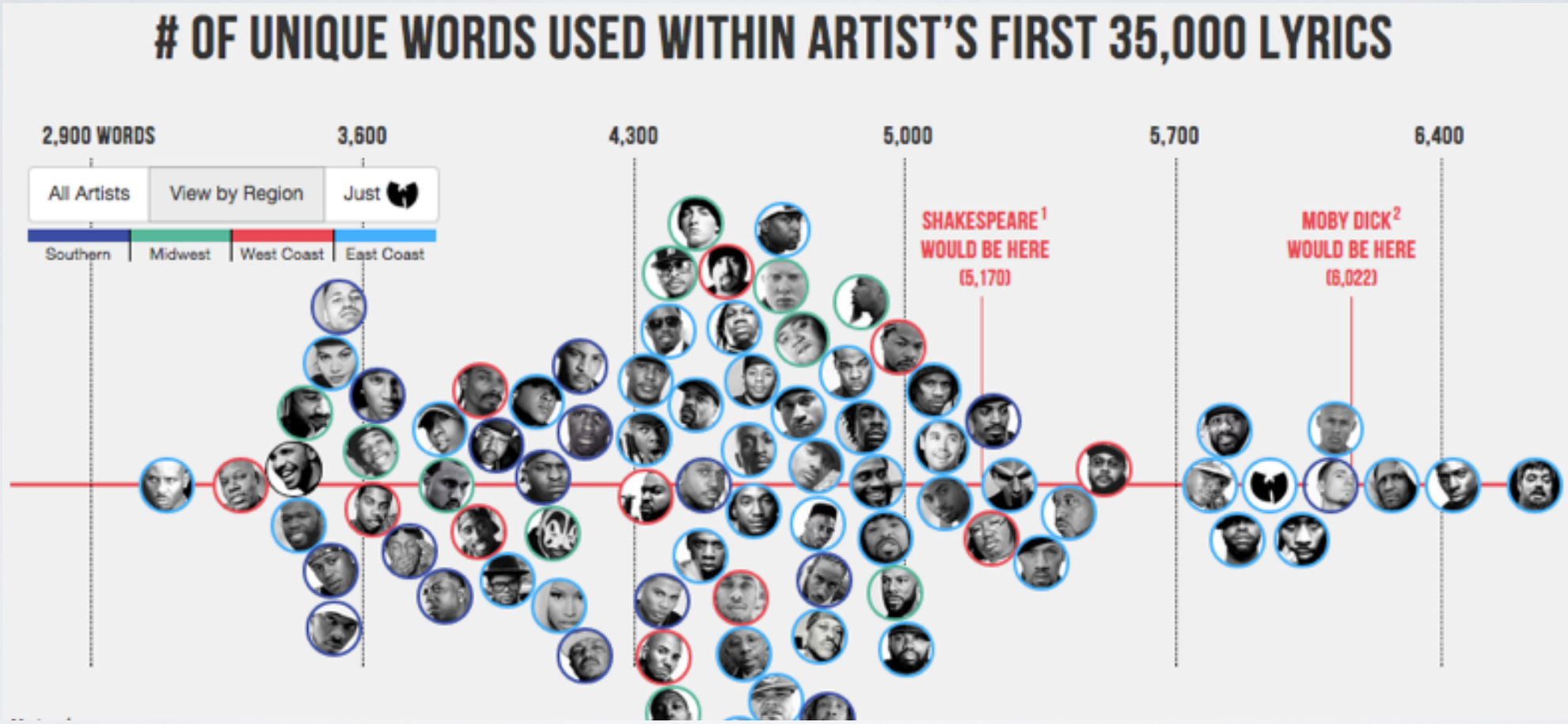
Kevin Quinn, Vinod Rathnam, Sonja Khan, and Riley Porter

OUR QUESTION

How can we show interesting regional and temporal patterns in collaborations between Hip Hop artists in the United States?

INSPIRATION:

HIP HOP VOCABULARIES



LINK

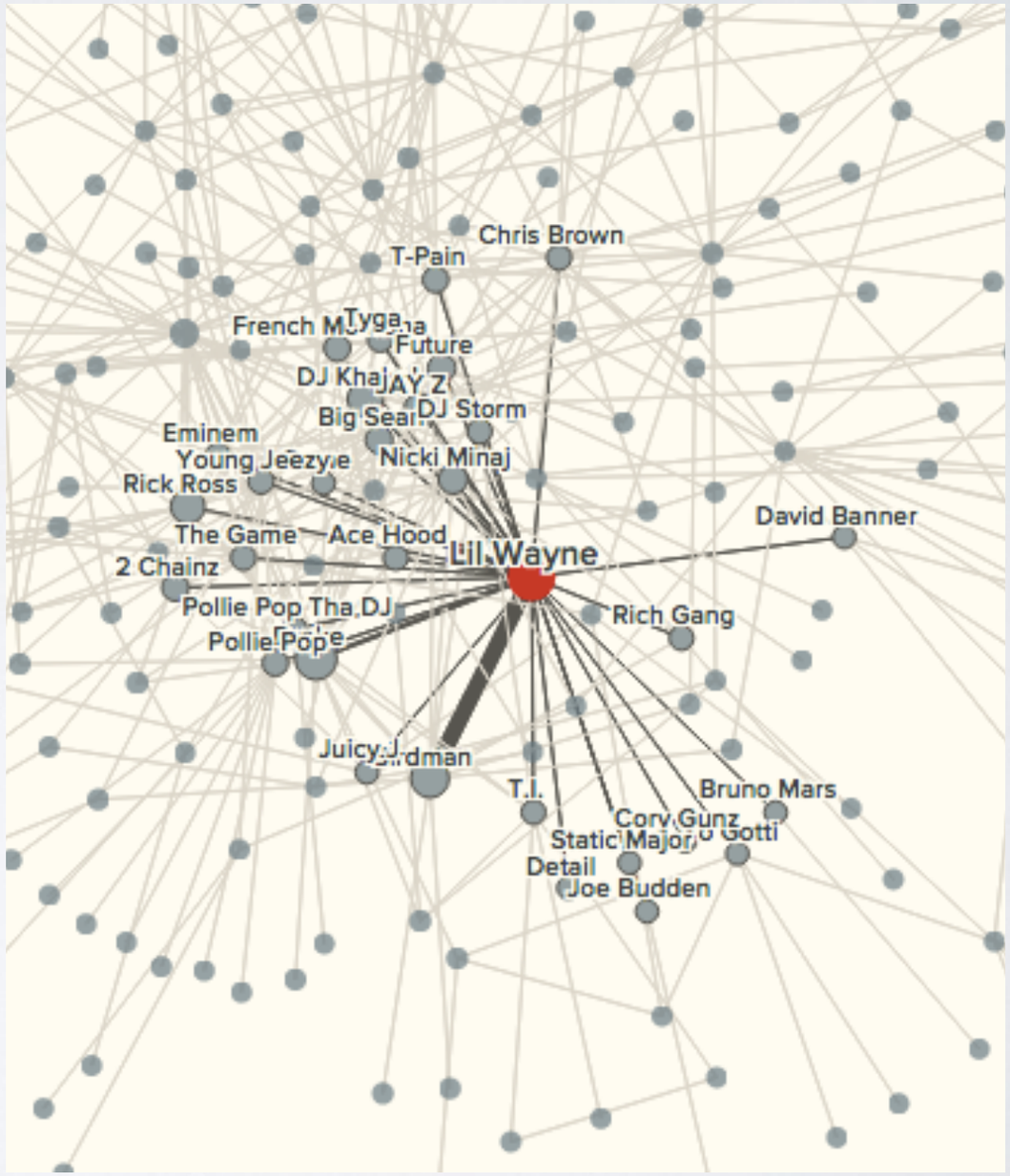
- showed size of vocabulary for various artists and groups
- incorporated regional data for artists, but displayed with no geographical context

RELATED WORK:

COLLABORATION NETWORK

LINK

- interactive network of collaborations
- limited data: 2002-2014
- based on Billboard top 40 artists, not all hip hop
- no regional or time data

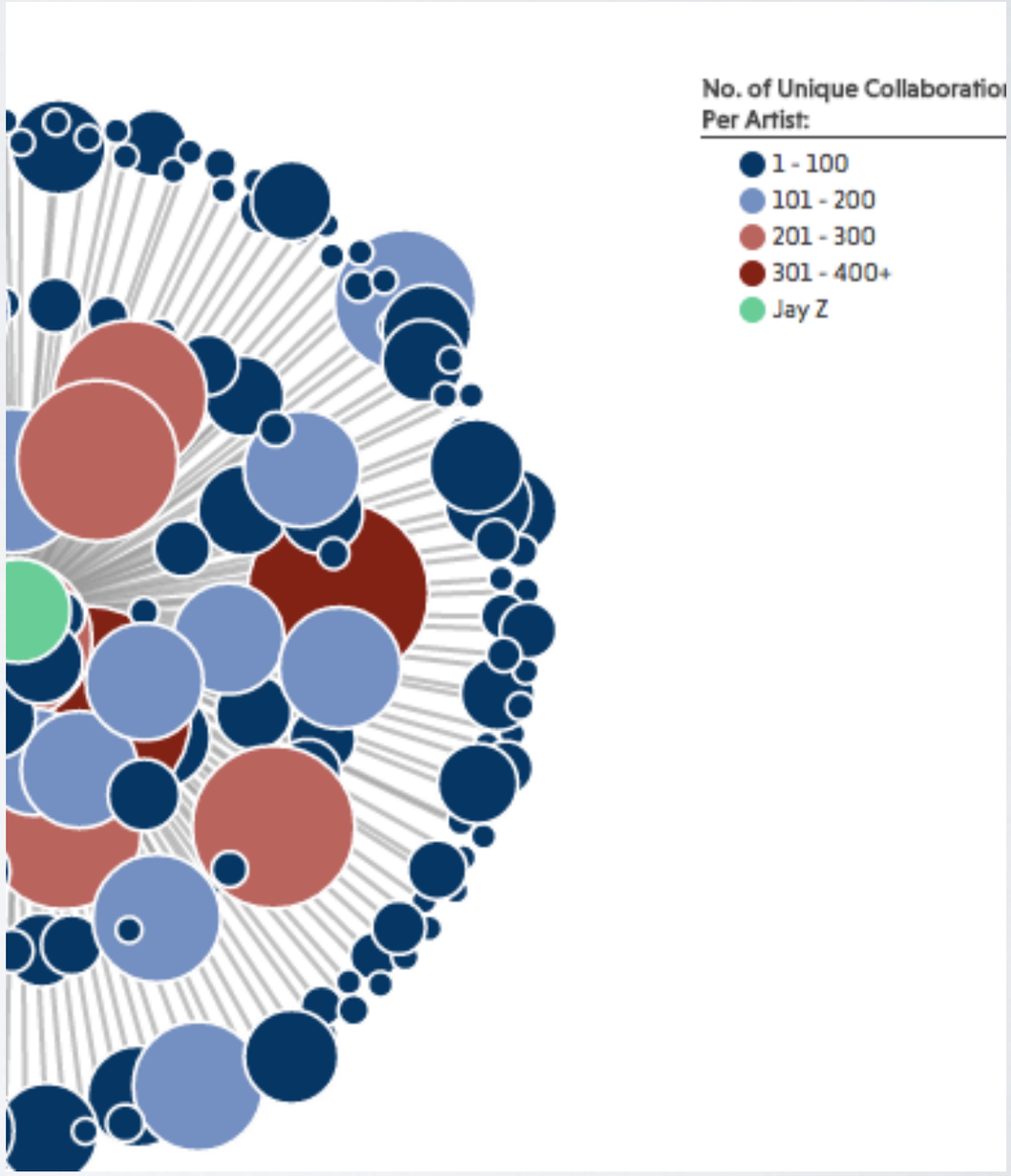


RELATED WORK:

COLLABORATION NETWORK FOR A SINGLE ARTIST

LINK

- more detailed information for the given network
- hard to interact with and gain more data from
- still no regional data



OUR IDEA

- incorporate regional and temporal data to collaboration data

- have more extensive data (1970s - present)

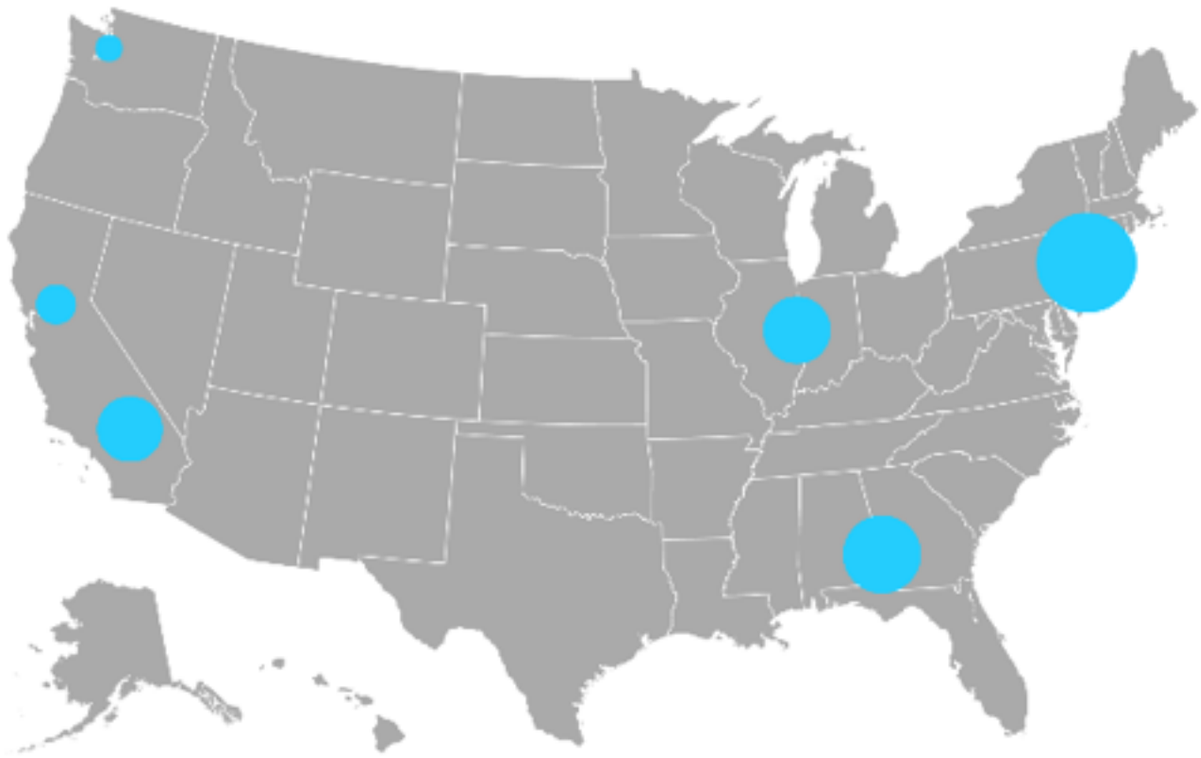
OUR DATA

- Compiled a list of 134 popular artists and metadata spanning 1972 - present day

- Queried Musicbrainz API to gather collaboration song data for the artists

Storyboard: overall view

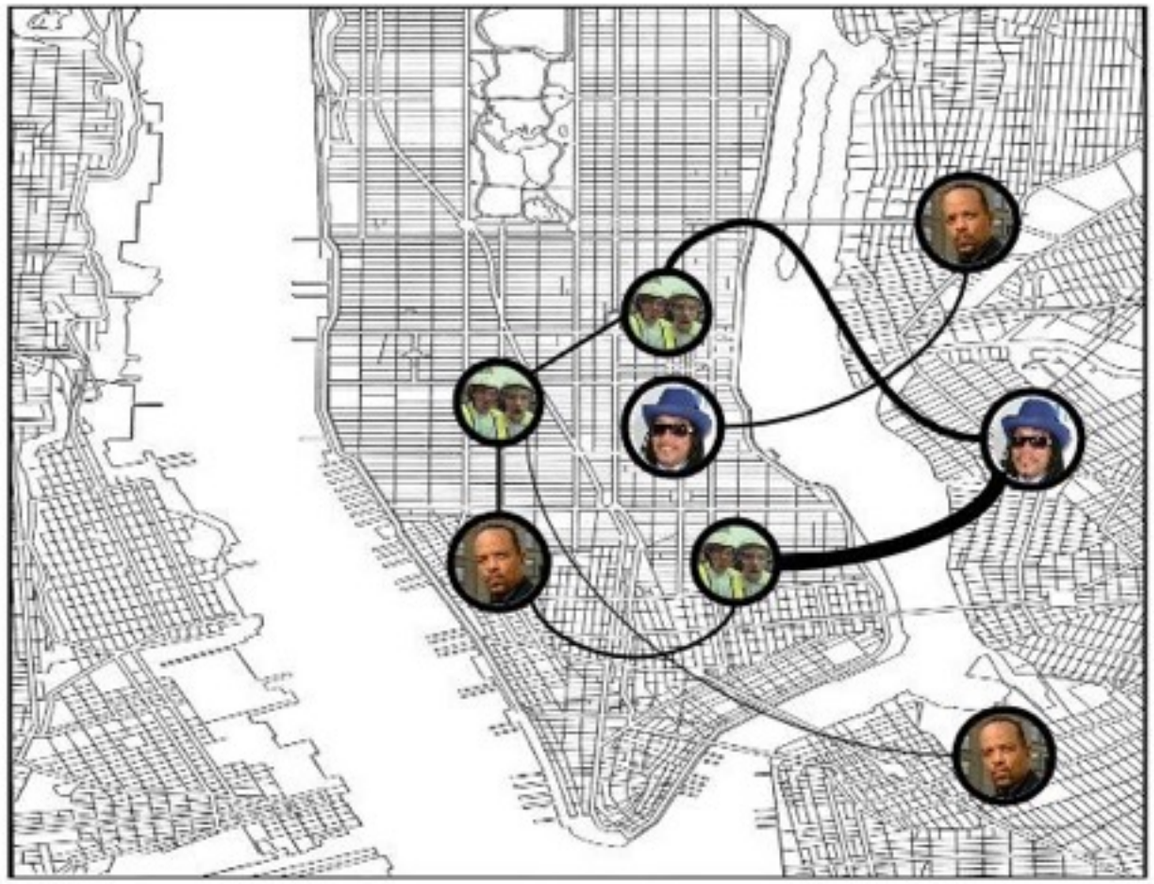
1996



1990 1991 1992 1993 1994 1995 1996

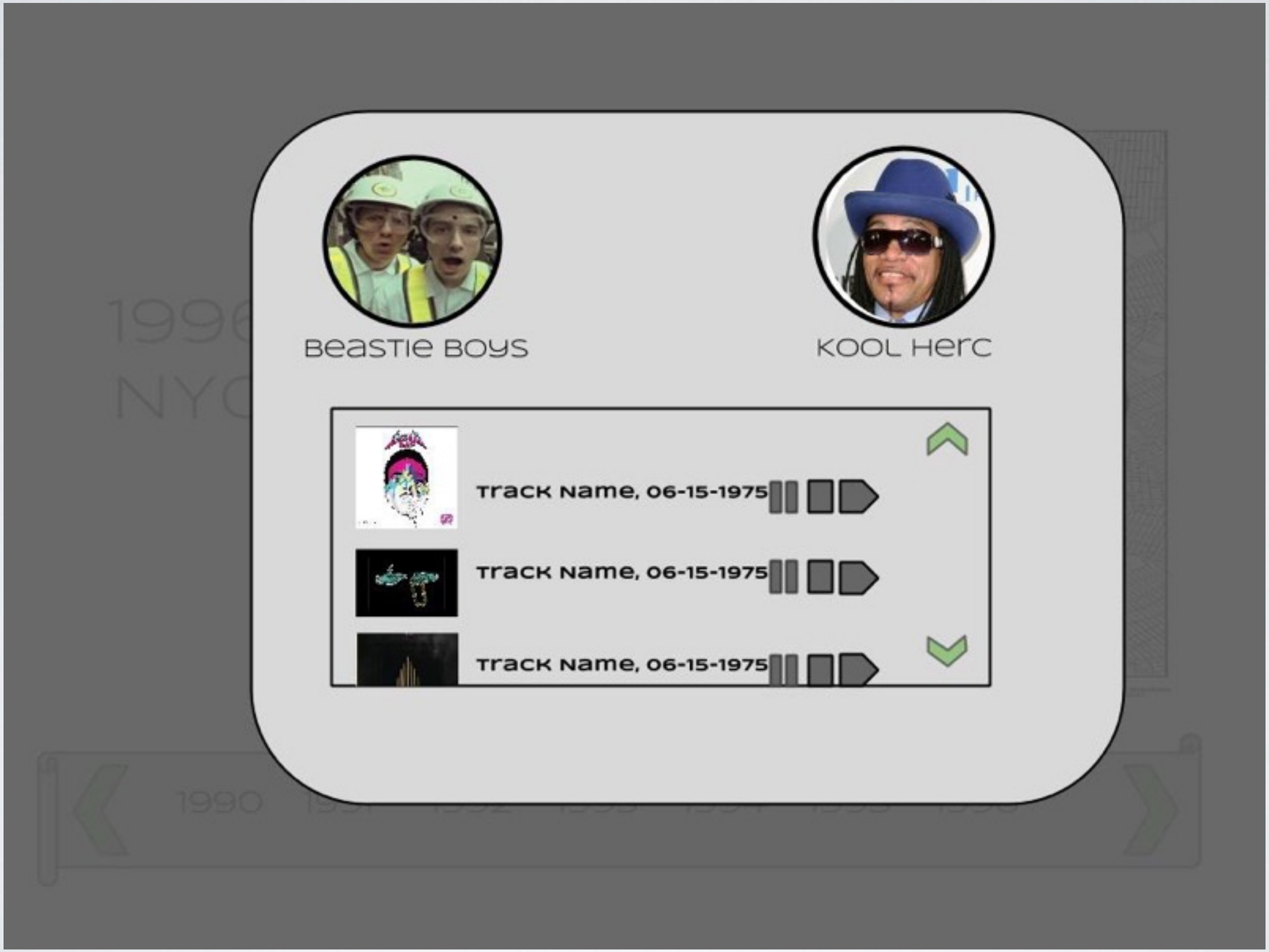
Storyboard: regional view

1996
NYC



1990 1991 1992 1993 1994 1995 1996

Storyboard: collaboration view



QUESTIONS

- What views of this data are we not considering?
- What potential interaction problems does this design have?
- What other interesting questions could we be answering with this data?
- Anything else?

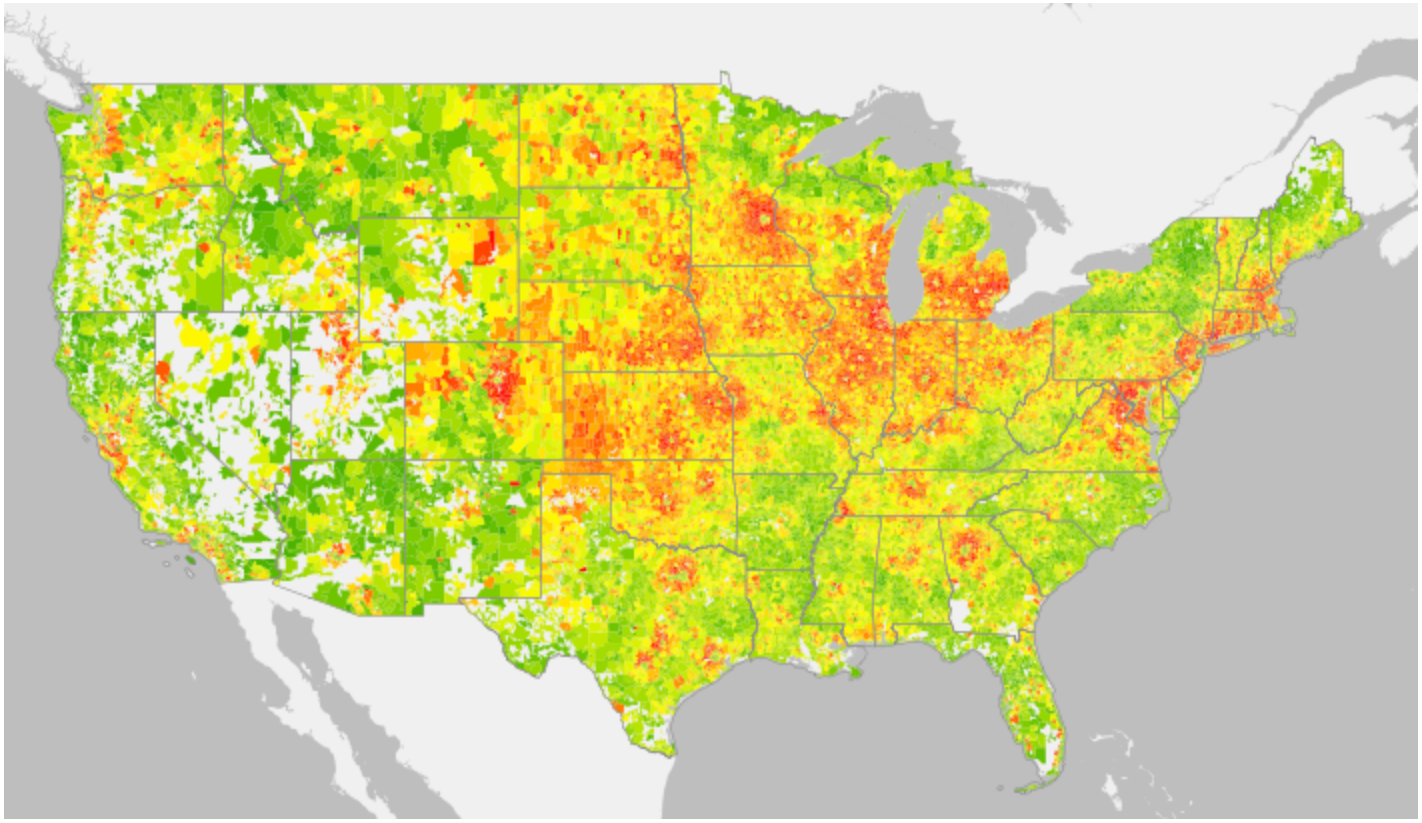
Visualizing the Statewide Impact of a Carbon Tax

Justin Bare, Aditya Kaul, Nandita Anand, Richard Li

Problem

- Revenue-neutral carbon tax in Washington
  - Added tax on fossil fuels
  - Added low income tax rebate
  - Reduced sales tax, business taxes
- How can we visualize the statewide impacts of this policy?
  - High-dimensional
  - Lots of uncertainty
  - Several partially complete, disparate datasets

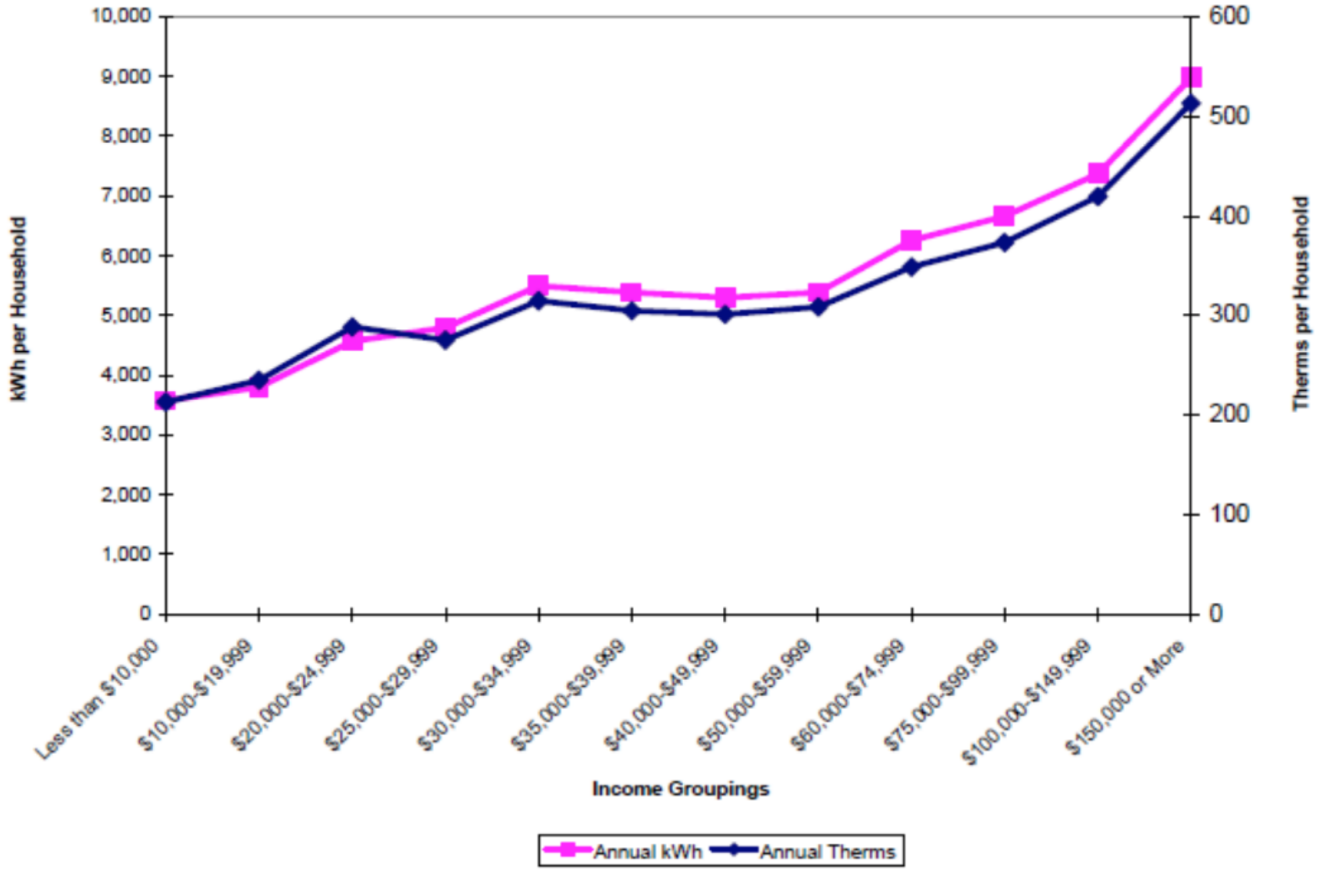
Prior Work



http://coolclimate.berkeley.edu/maps

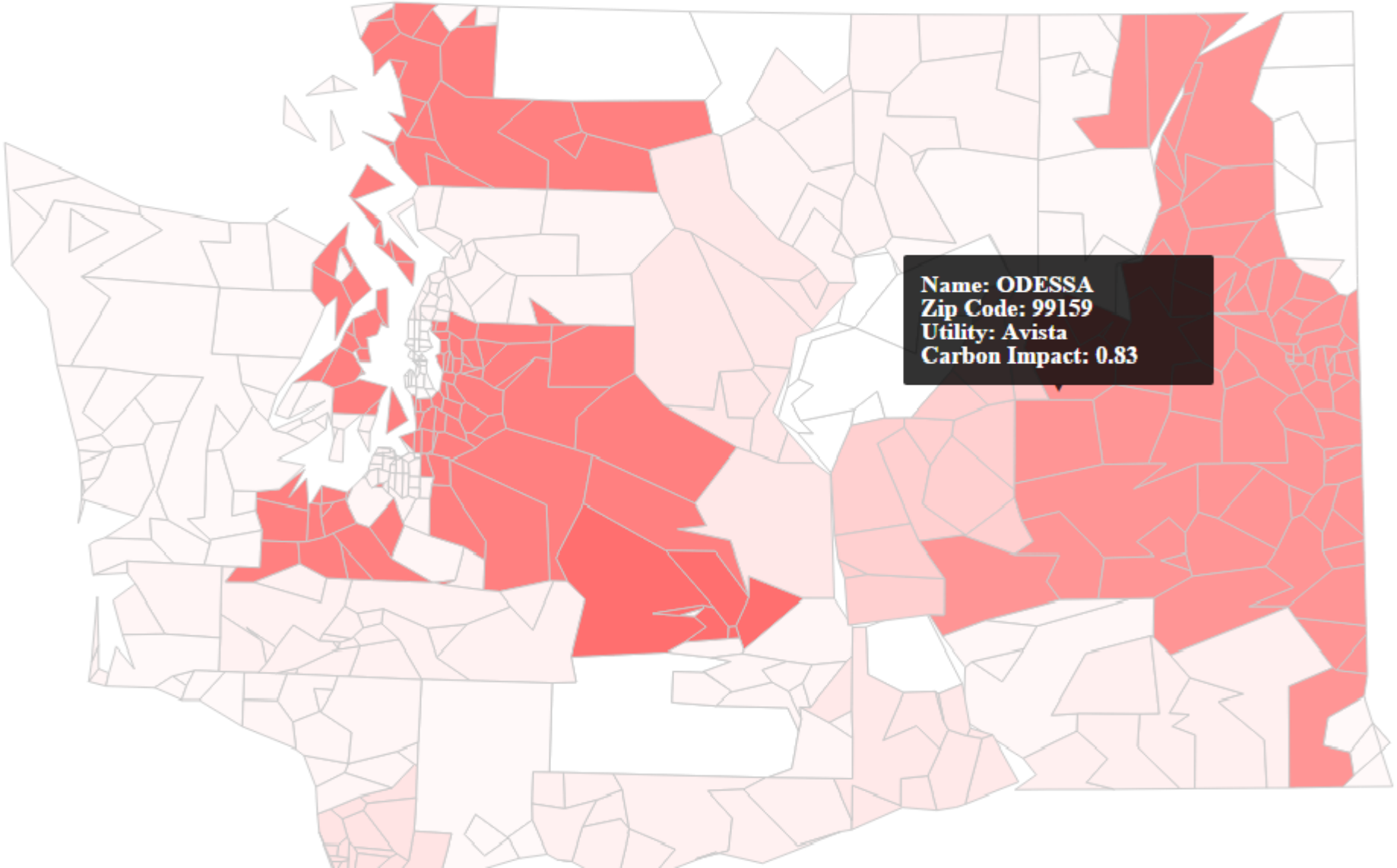
Prior Work

Figure I. Average Electricity and Natural Gas Consumption by Income

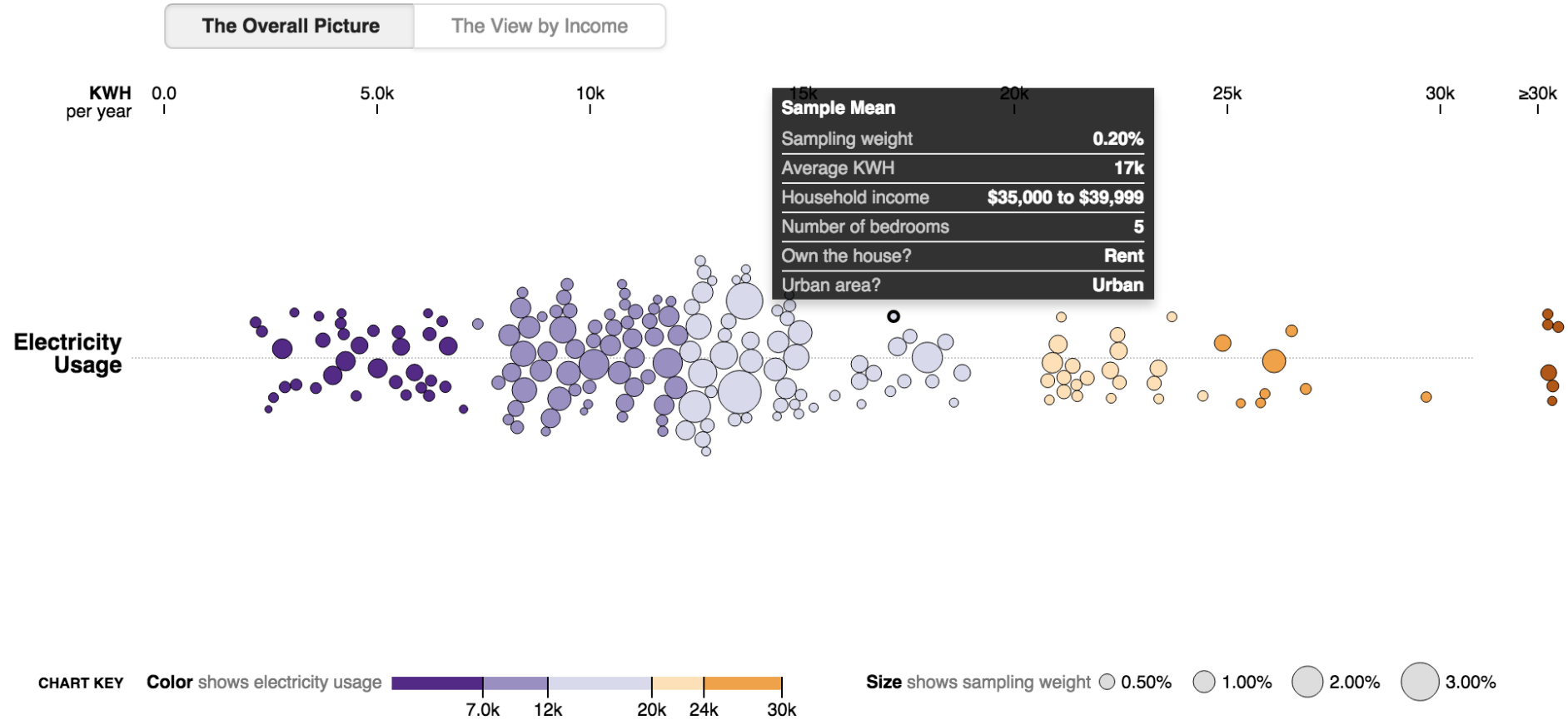


Source: KEMA (2010), p. 32.

Current Progress



Current Progress



Next Steps

- Filters
- Zoom feature on map
  - More details on zoom: consumption based on utilities in that zipcode, household details, etc.
- Interactive bar/bubble charts to appear when hovering on map
- Animated transitions between views of different datasets
- Polishing the encodings according to the latest perceptual theories of color, graphics, space, etc.

Displaying CO_2 Emissions Worldwide

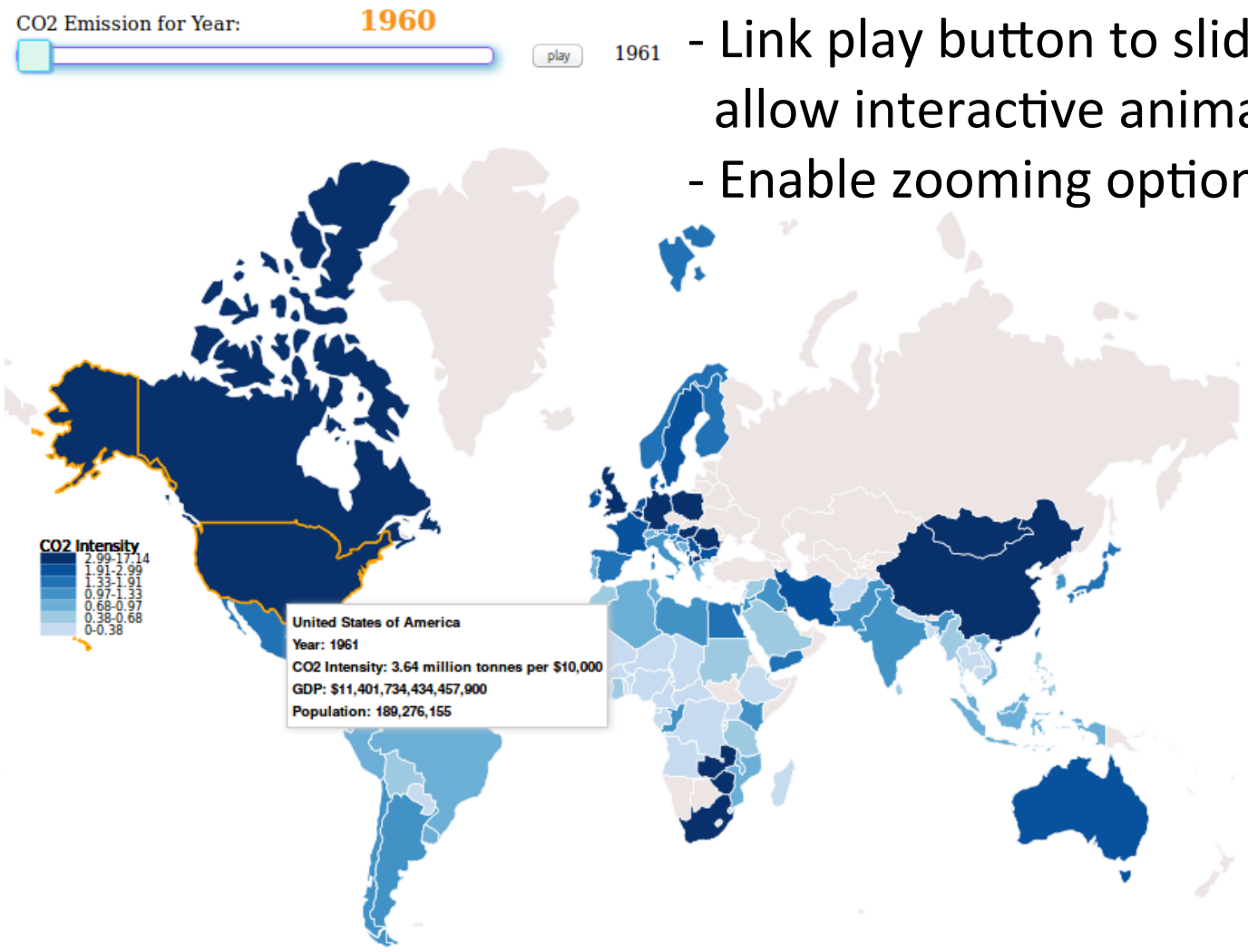
Alec Zimmer (azimmer8)
Hoiyi Ng (hoiyi)

Problem

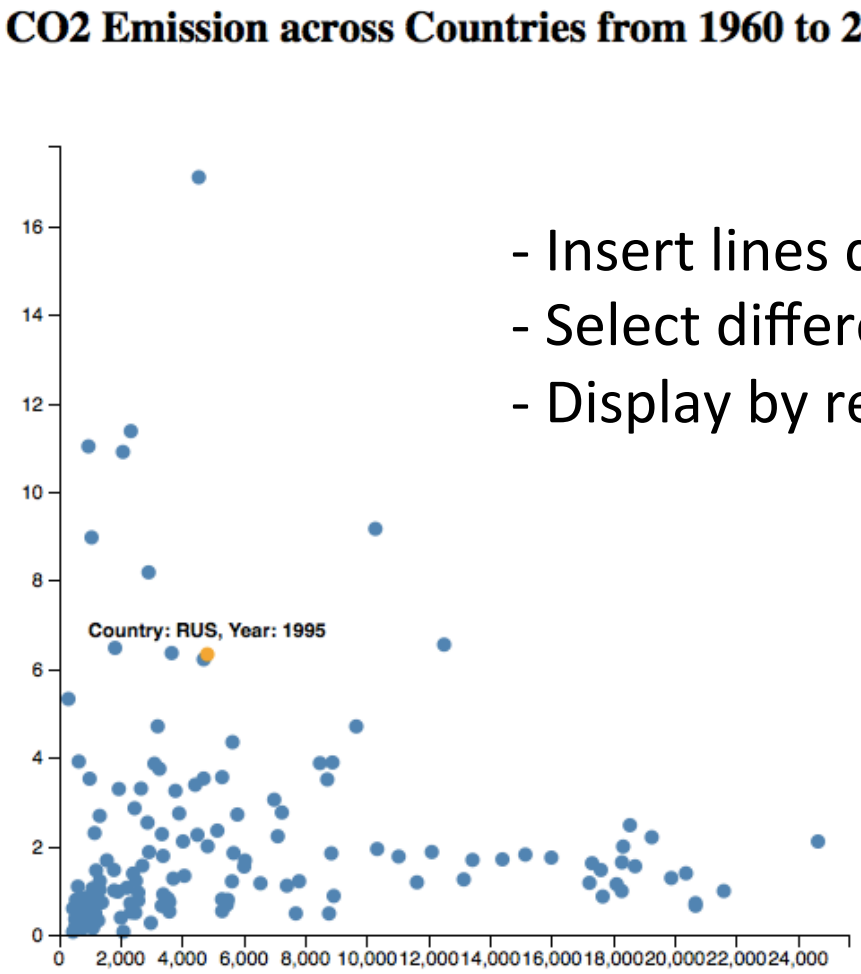
- Display CO_2 emissions by country and year
- Enable analysis and exploration
- Compare with relevant variables (population, GDP, etc.)
- Motivated by work with Adrian Raftery (UW, statistics) and Dargan Frierson (UW, atmospheric sciences)

Interactive map

- Scatter plot in pop-up frame
- Link play button to slider to allow interactive animation
- Enable zooming options



Scatterplots/Trendlines



- Insert lines displaying country data by year
- Select different variables
- Display by region

Questions

- How to make the two approaches interact?
- What would be useful information to display when clicking and hovering?
- Should we allow zooming for the map? Would it be more helpful or more distracting?
- Any creative additions?

Feedback?

1. Suggestions for other visualizations?
2. Any experience in combining information from different surveys?
3. What features are more interesting in describing a household?

Exploration of a Multi-Dimensional Astronomy Dataset

CSE 512 Project Feedback

Nicole Atherly, Mahir Kothary, Grace Telford

Motivation

Dataset: ~tens of properties for ~100,000 galaxies

Issues:

- Parameters derived from modeling light distribution
- Basically all galaxy parameters are correlated with each other → Potential biases!

Goal: Interactive, deep search for anomalous correlations

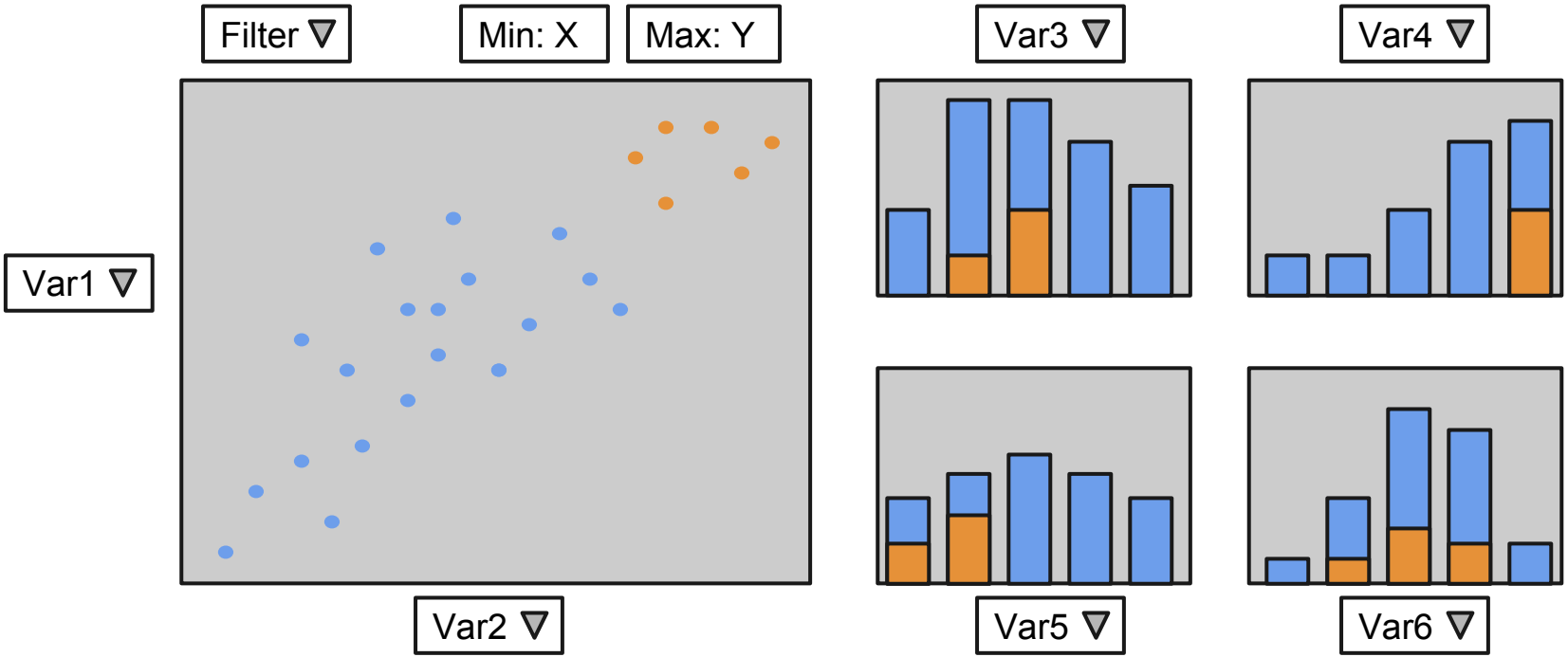
Relevant Prior Work

- Ggobi
- Parallel coordinates

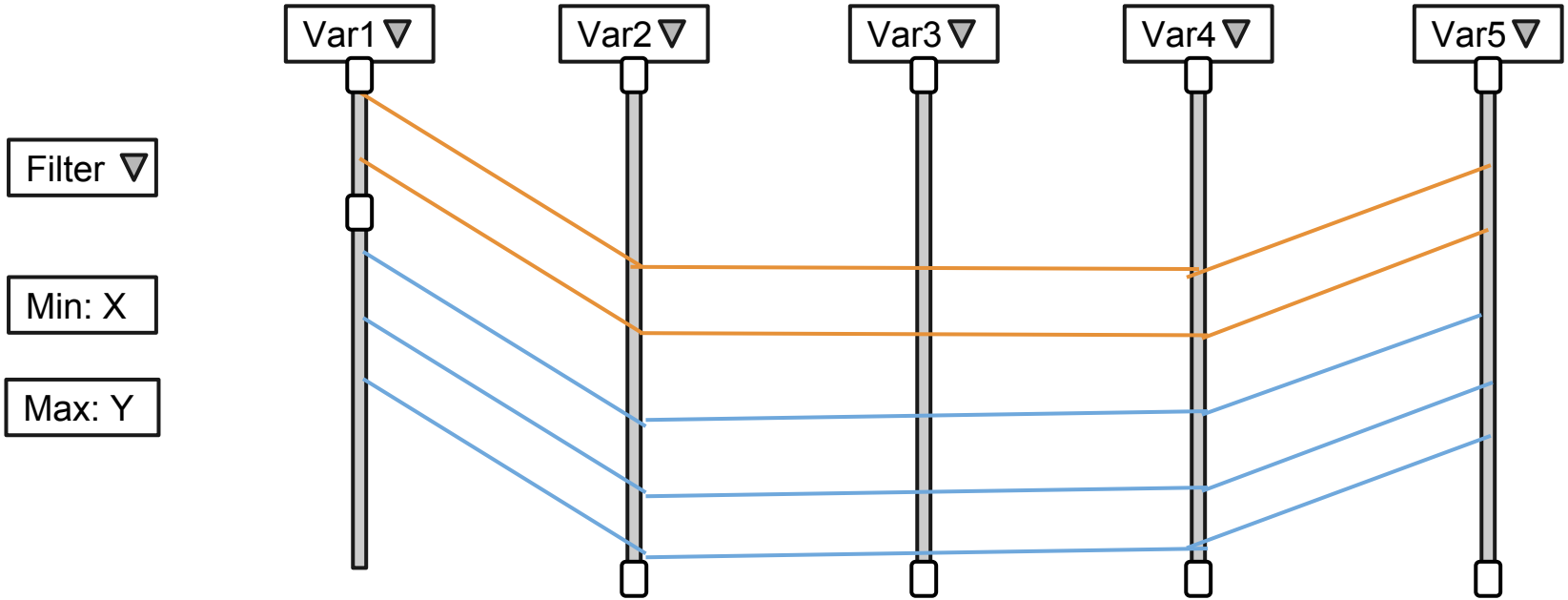
Our project will be tailored to this galaxy dataset and will cater to users who want to drill down into the details

- Dynamically filter data and change variables displayed
- Inspect distributions of subsamples vs. whole

Storyboard



Storyboard



Feedback Questions

- Are there other capabilities you would want if you had to look for biases in a complex, high-dimensional dataset?

- Thoughts on customizability vs. clutter?
  - E.g., ability to change type of plot, number of bins in histograms, axis ranges, …

S.I.E.V.E.

Statistical Interactive Explorer of Vaccine Efficacy

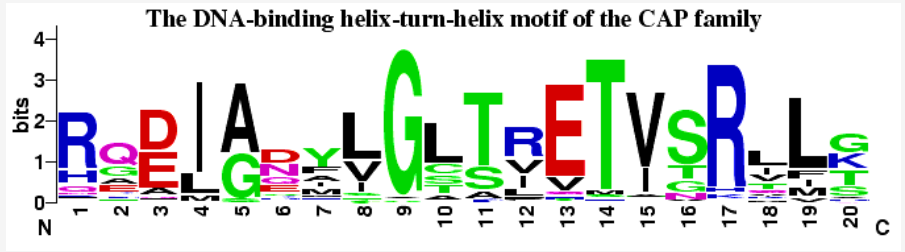
KULLMAN, CLENAGHAN + WANG

SIEVE ANALYSIS

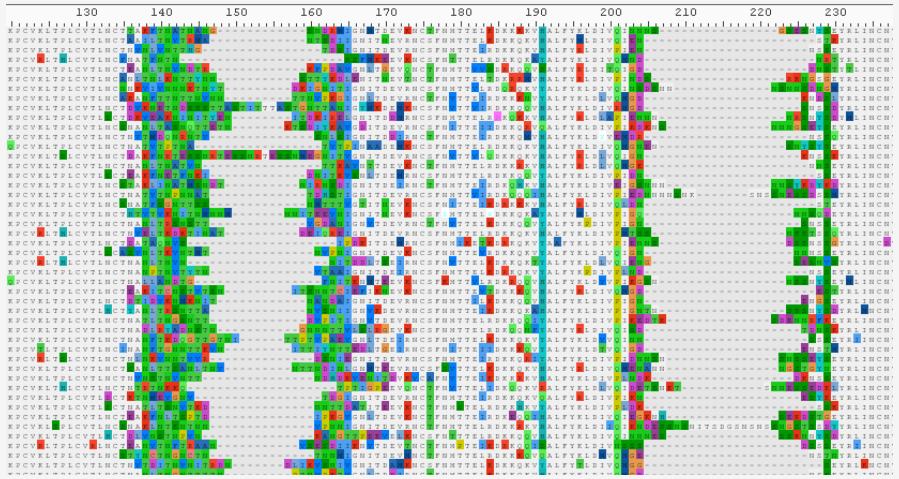
- Vaccines as a whole are ineffective for HIV.
- However, new techniques show they can be effective on certain strains
    - Which ones???
- Sieve analysis studies the similarities of the genomes the HIV in patients to the vaccine administered.
- OUR TASK: create an interactive visualization to aid in exploring this data.
- Working with Andrew Gartland and Allan DeCamp at Fred Hutch

PRIOR WORK

GENERAL SEQUENCE VIEWERS

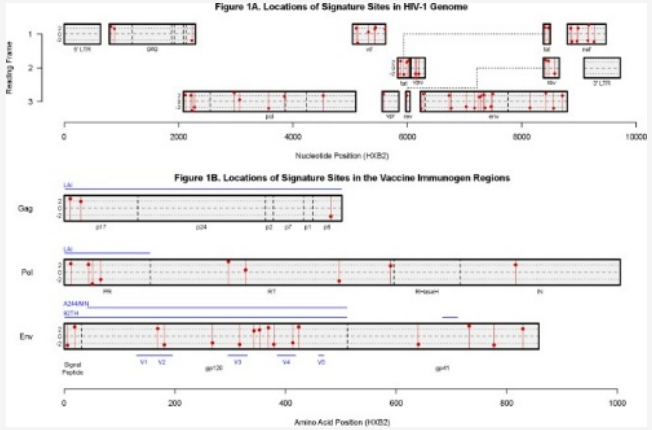


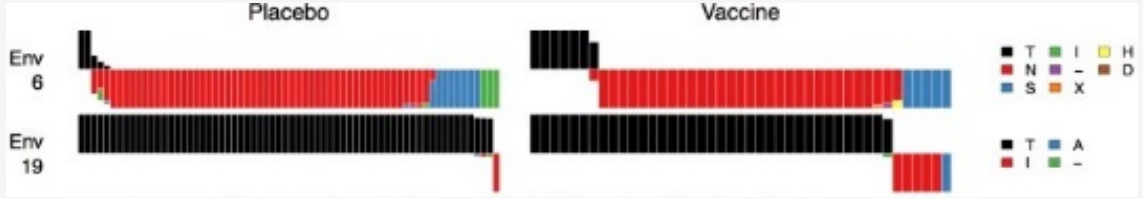
WebLogo



Alignment Viewers

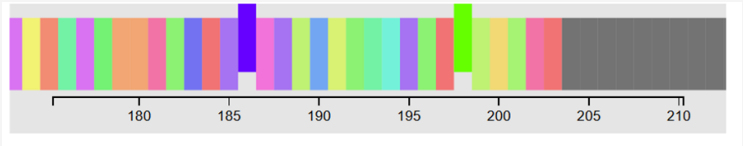
SIEVE ANALYSIS GRAPHICS



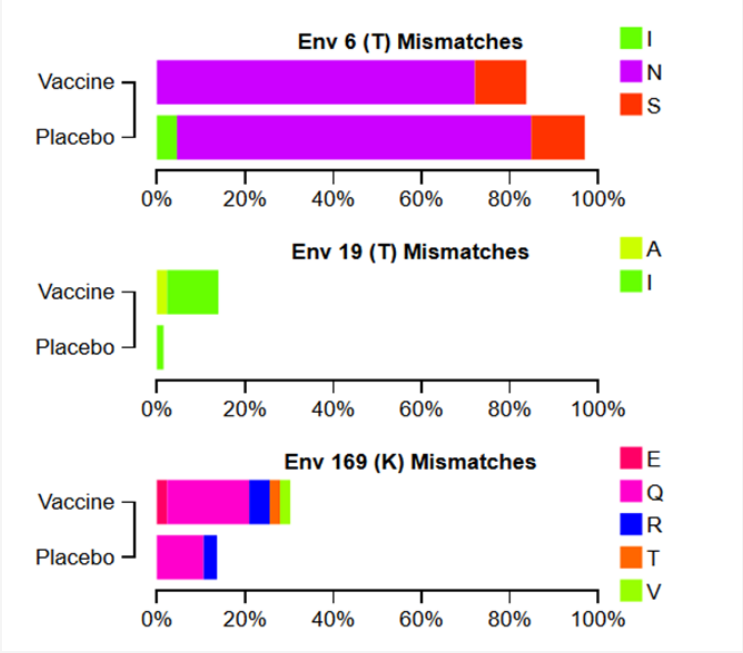


S.I.E.V.E.: THREE PARTS SO FAR

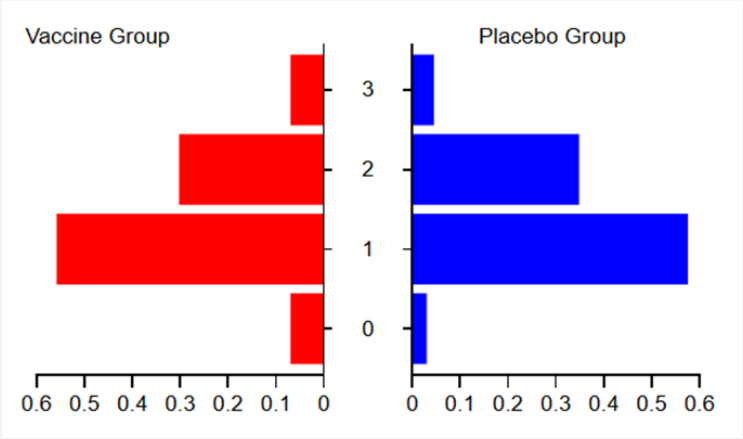
OVERVIEW GRAPHIC / SITE SELECTION



INDIVIDUAL SITE CHARTS



GROUP STATISTICS



WORKS IN PROGRESS

- Better selection
- More statistics shown: visually encode statistically interesting sites to view
- Allow researchers to upload their own data for other vaccine studies
- Implement customization of colors, annotation of graphics, and export of images for use in sieve analysis papers

FEEDBACK SOLICITATION

What selection mechanism would work best for this type of data?

- Data is in a very long sequence, but researchers need fine grained control
- Ideally easily select multiple disconnected sequences across the genome, with ability to quickly select sequences of a specified length

Best way to export a D3 graphic to PDF format for inclusion in papers?

- SVG is inconvenient for the researchers

Visualizing HIV Mutation

Introduction

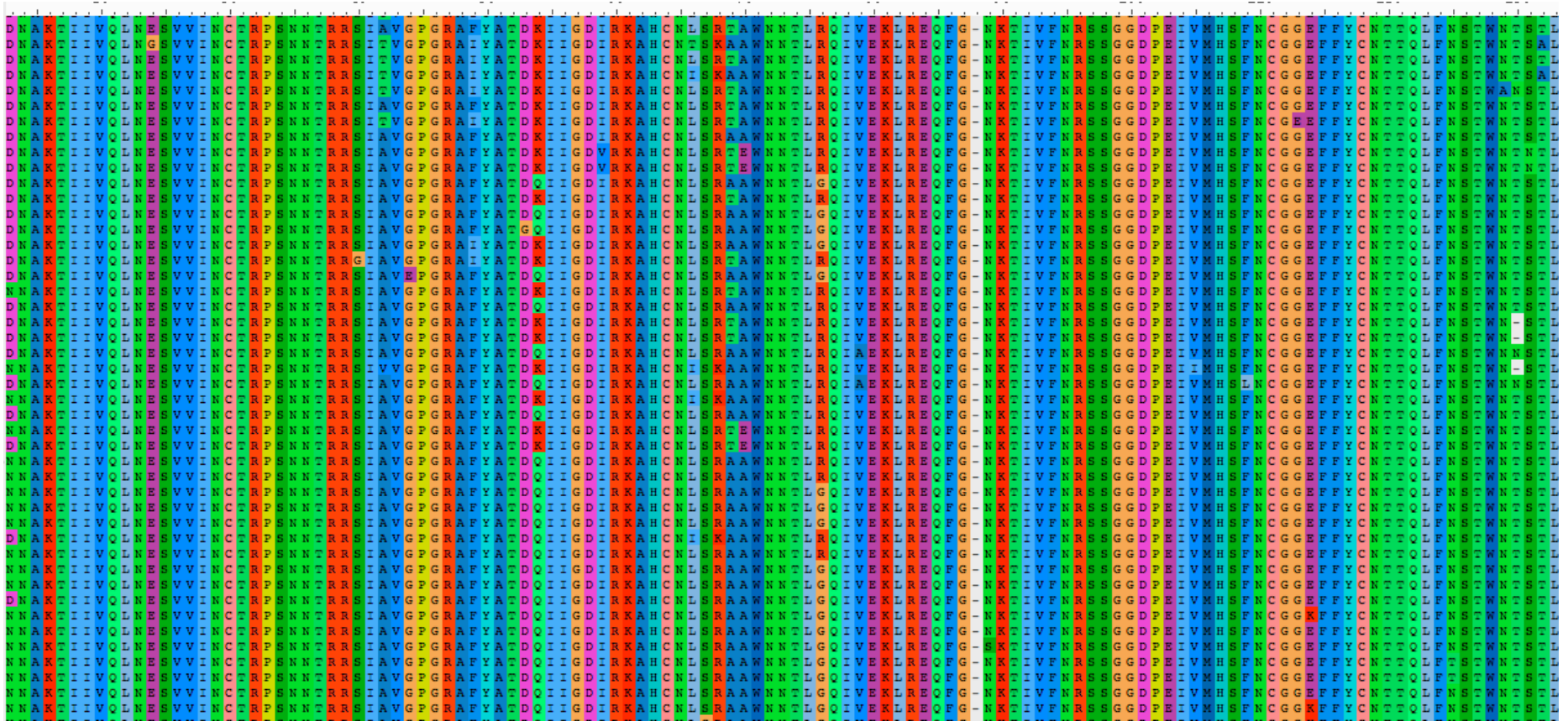
A F R T O P R

------>

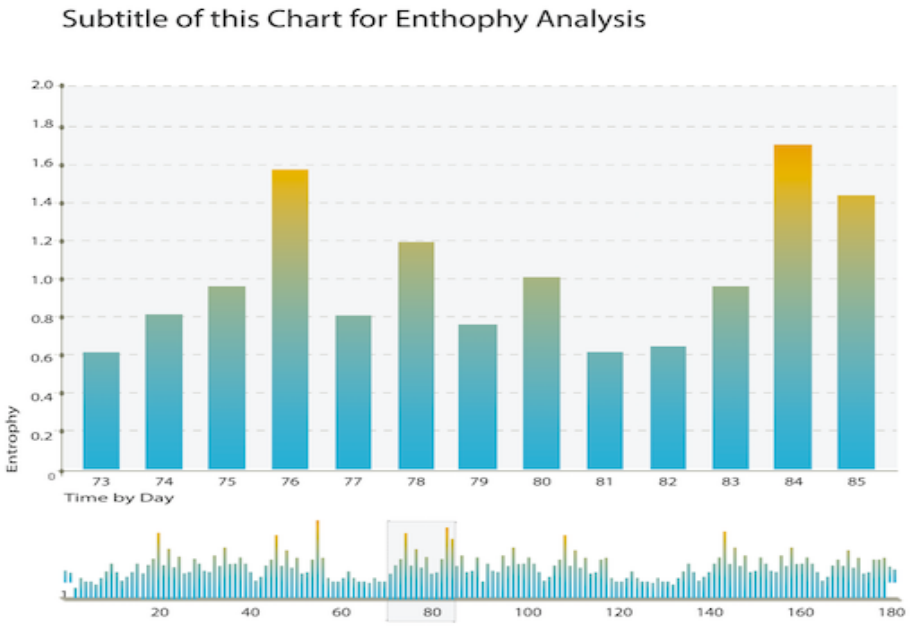
A F R ***A*** O P R

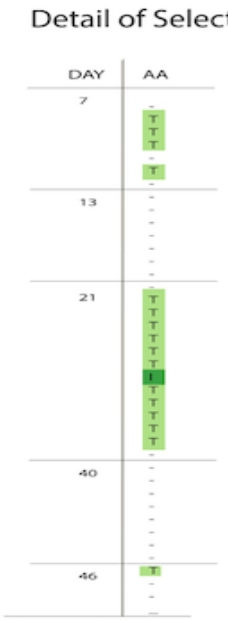
Mutation!

Challenge



Design





More details to be designed

based on further requirement

...

* Dataset: HIV protein sequence from infected individuals; Known T-cell HIV epitopes mapping

Questions / Feedback

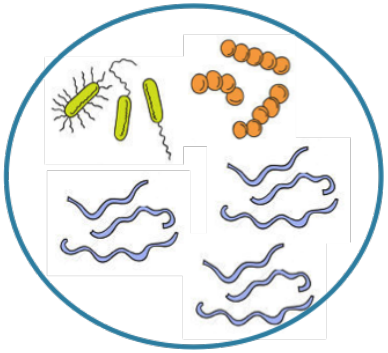
1. We have tight space, how to make good use of it?

Visualizing joint taxonomic and functional data from microbial communities

Cecilia Noecker
Alex Eng
Colin McNally
Will Gagne-Maynard

Motivation: Characterizing microbial communities

Species Composition

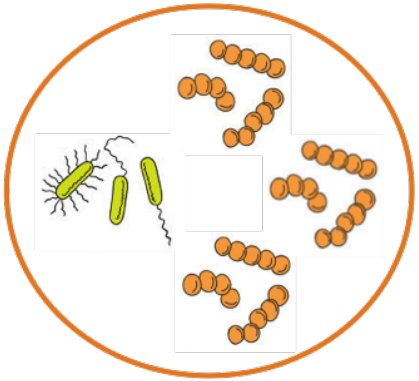


Sample 1

A B C
A C D
B C D
B C D
B C D

Gene Composition

2A 4B 5C 4D

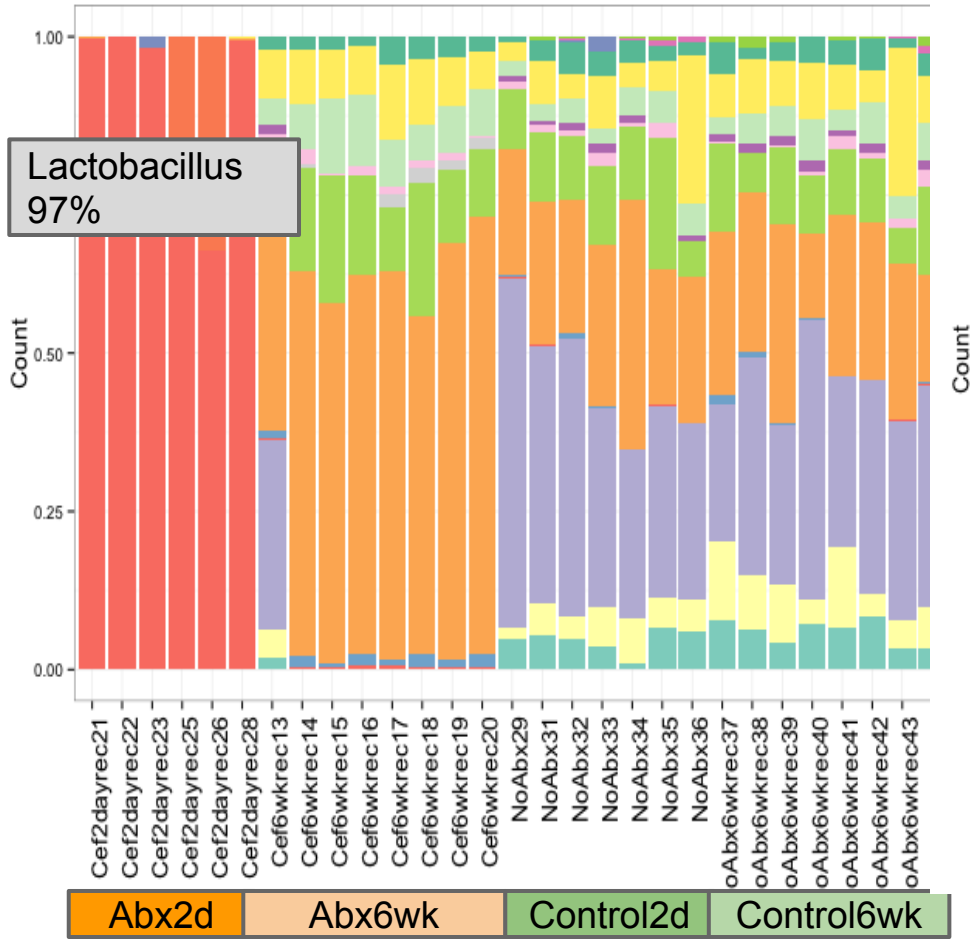


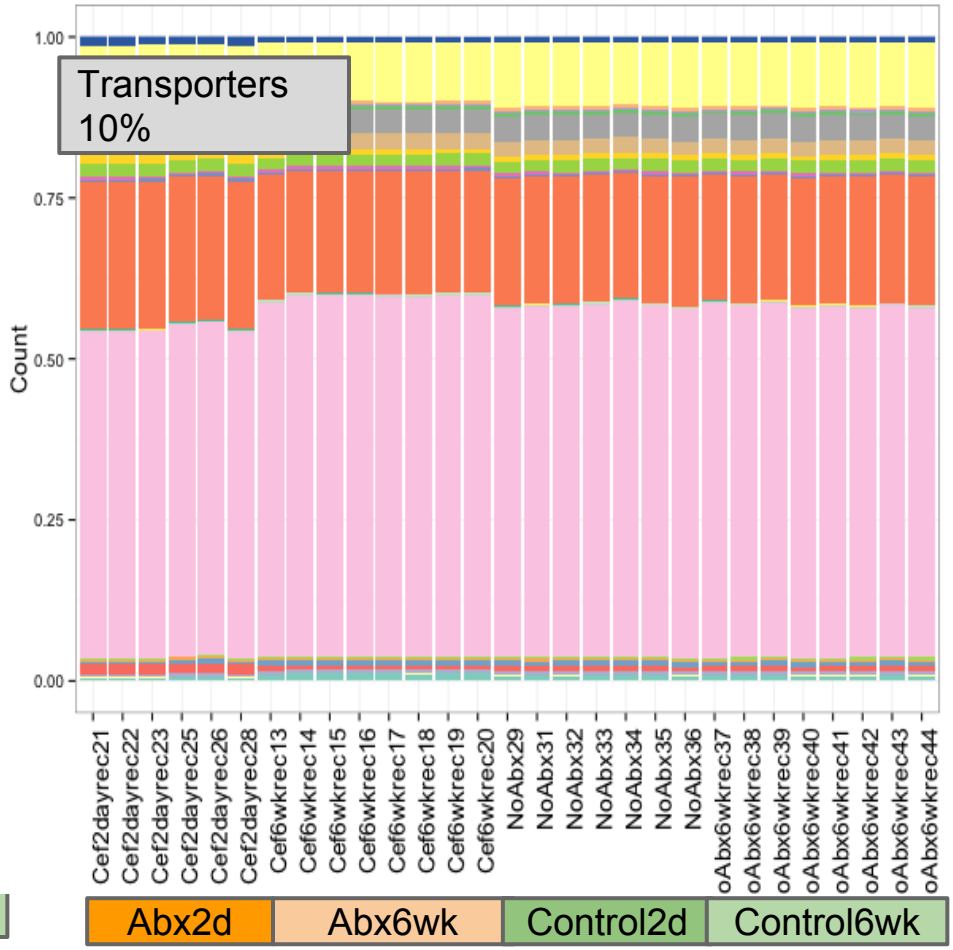
Sample 2

A B C
A B C
A B C
A C D

4A 3B 4C 1D

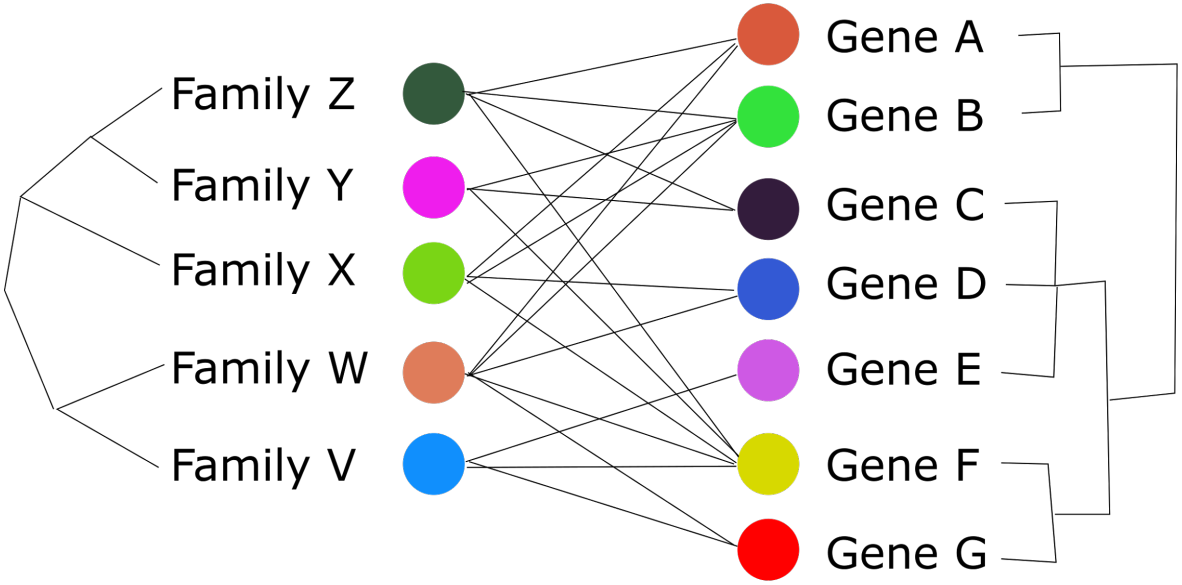
Design plan: Start with traditional visualizations of these data



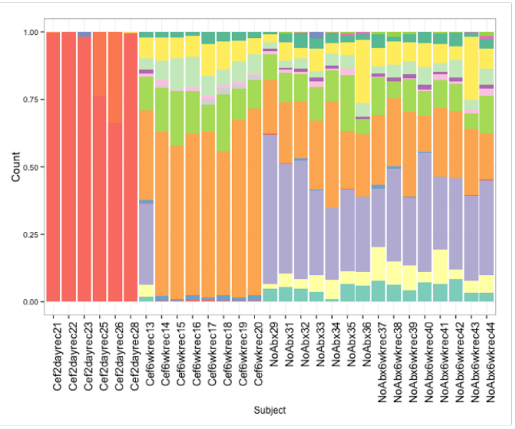


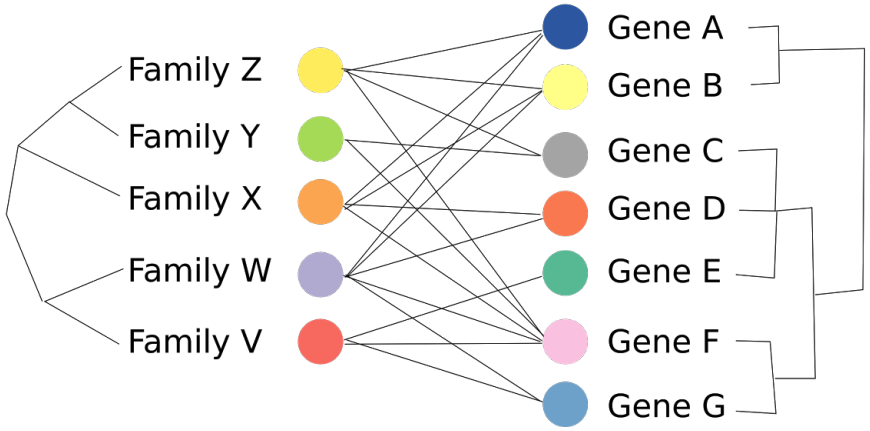
Theriot et al *Nature Comm* 2014

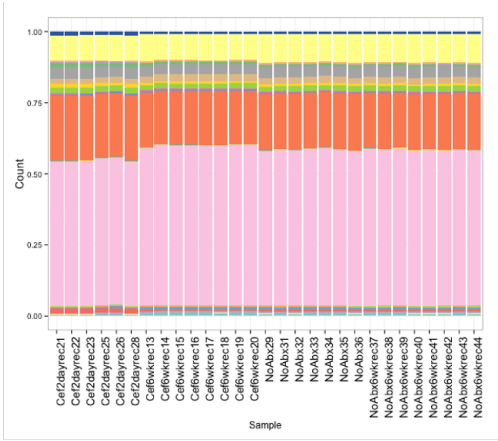
Design plan: Visualize links between taxa and functions



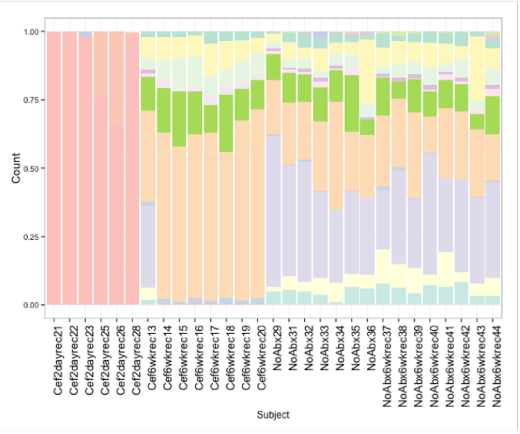
Overall Interface

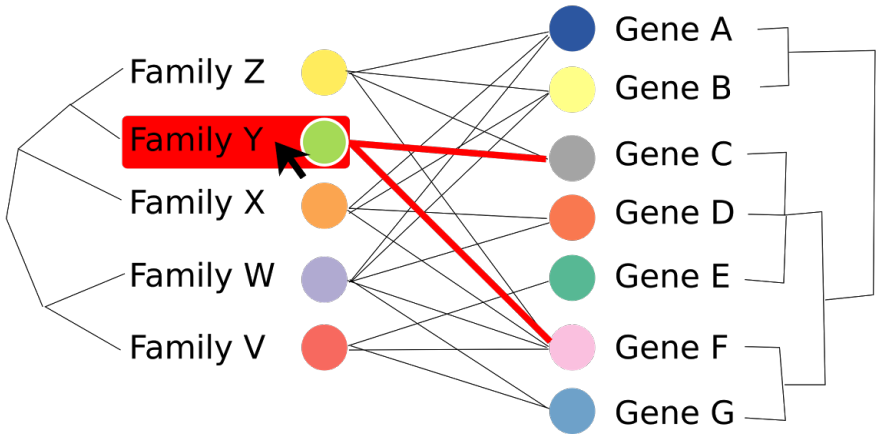


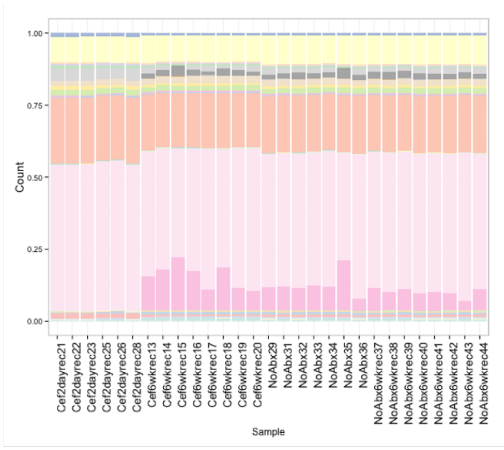




Overall Interface







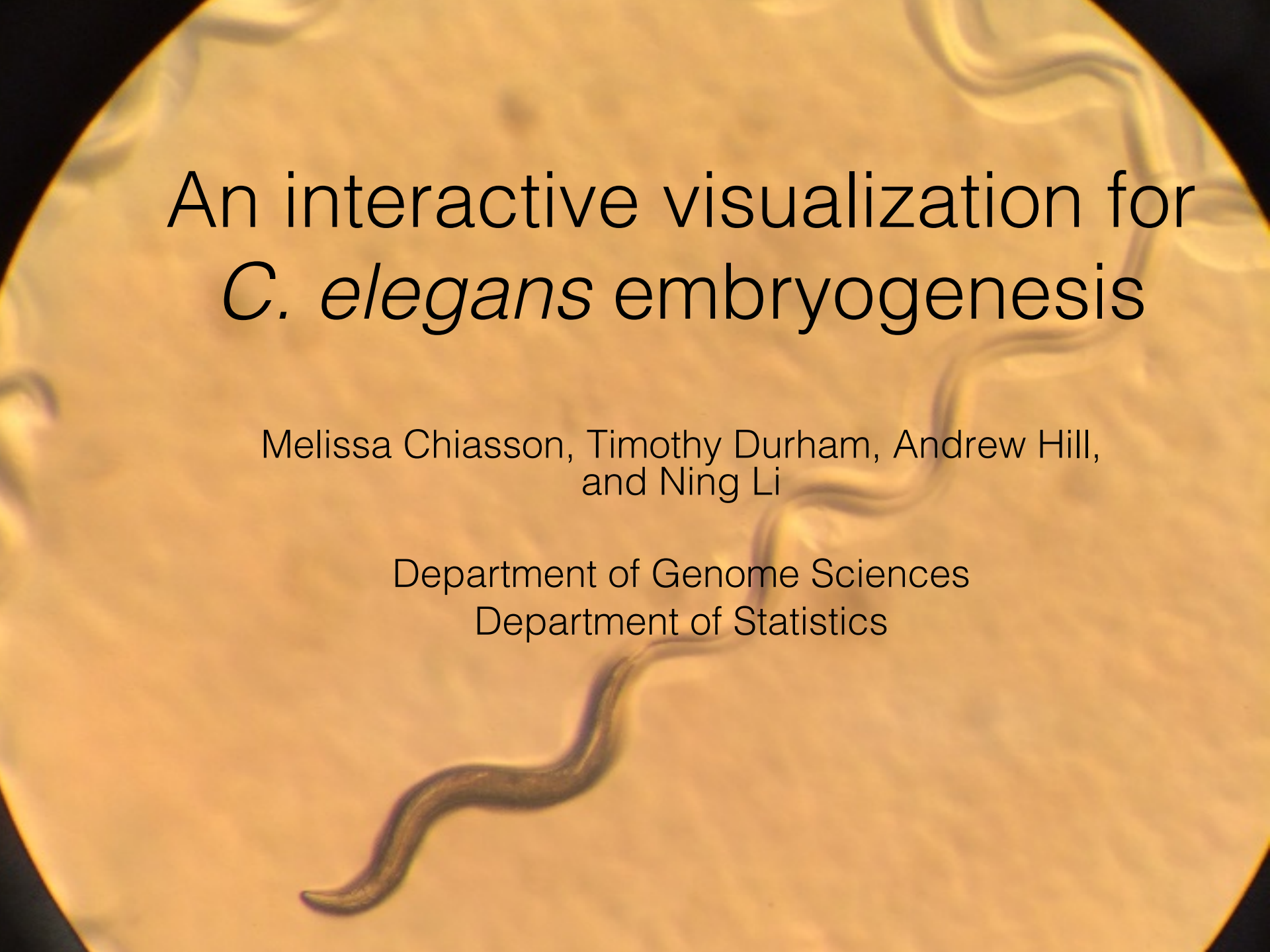
Questions

- Should we consider an adjacency matrix heatmap instead of bipartite graph?

- Should the graph visualization include variable edge widths encoding relative contributions or just show binary links for clarity?

- Any point in visualizing all (100k+) links between genes and species at once, or should we constrain to higher-level categories except when viewing specific subsets?

- Is this the right approach for comparing separate but linked data?



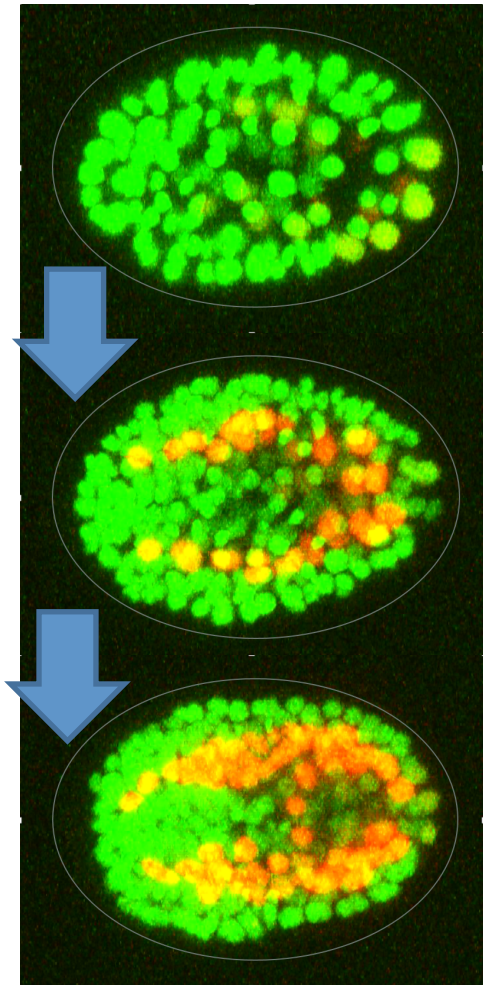
An interactive visualization for *C. elegans* embryogenesis

Melissa Chiasson, Timothy Durham, Andrew Hill, and Ning Li

Department of Genome Sciences
Department of Statistics

What gene regulatory relationships exist in *C. elegans* embryogenesis?

- *C. elegans* is a round worm used widely as a model organism in the genetics community.
- Currently no tool exists to interactively visualize embryo development through time and interrogate how gene expression patterns change with time, lineage, or cell type.

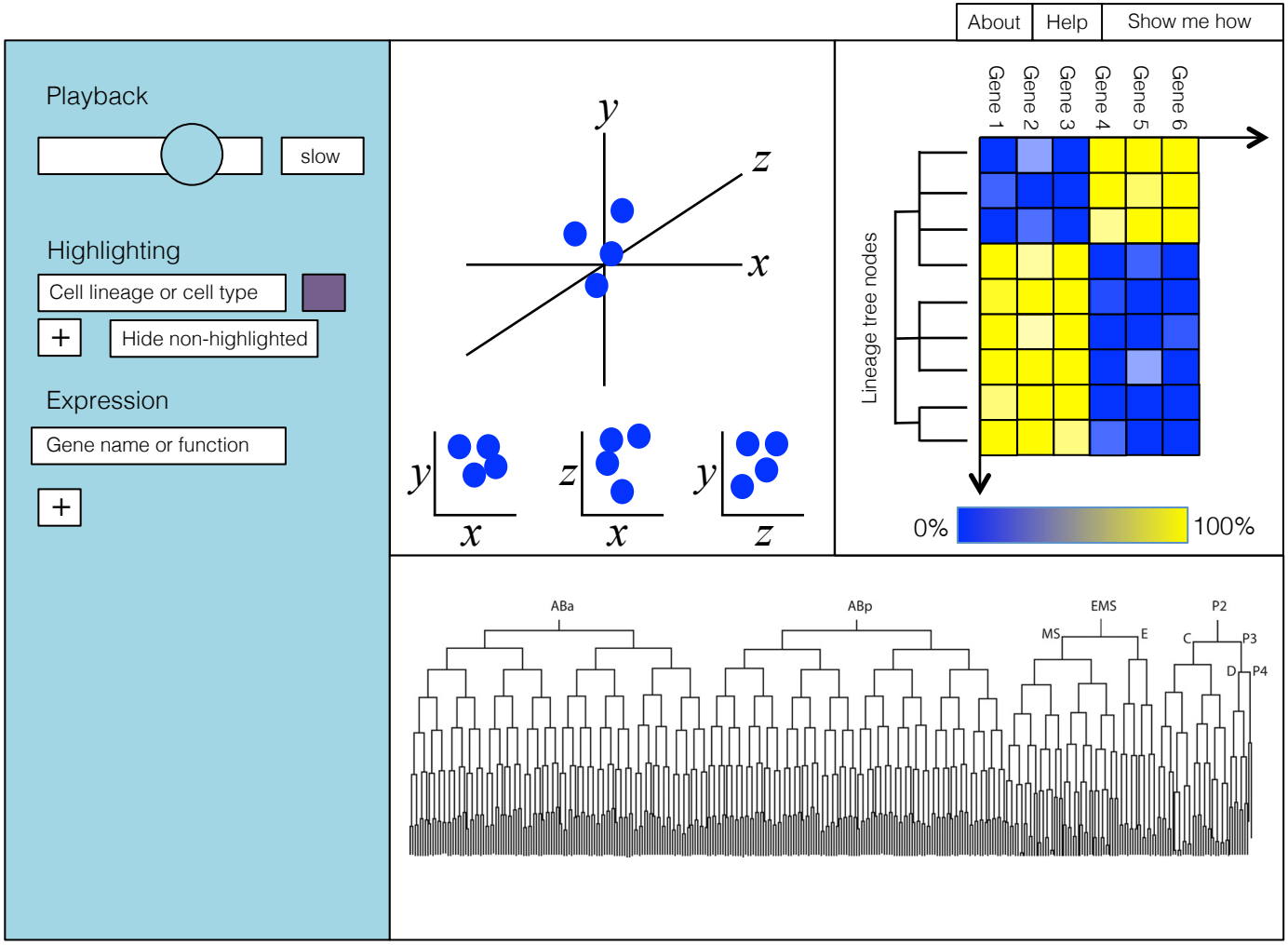


What gene regulatory relationships exist in *C. elegans* embryogenesis?

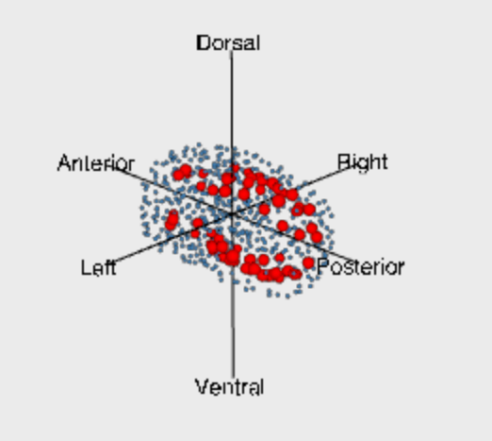
- *C. elegans* is a round worm used widely as a model organism in the genetics community.
- Currently no tool exists to interactively visualize embryo development through time and interrogate how gene expression patterns change with time, lineage, or cell type.

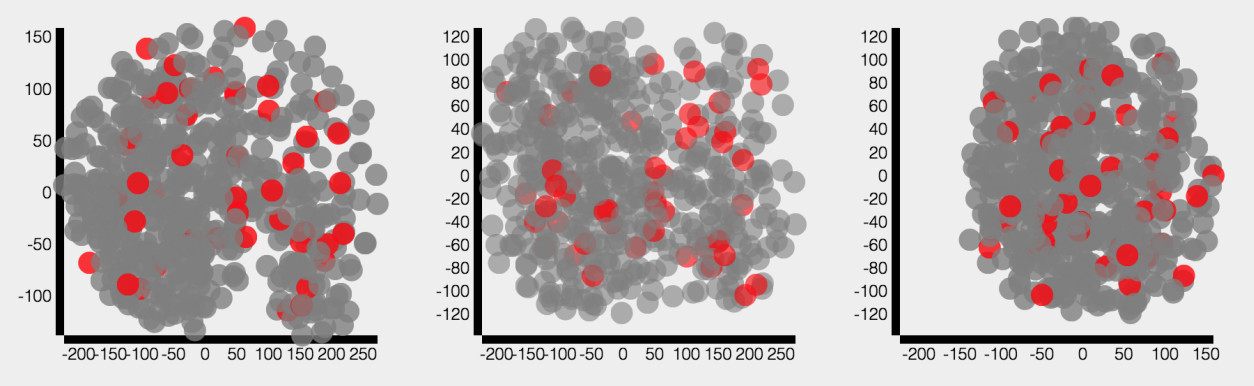
Types of data
Spatial orientation and diameter of cells
Cell types and lineages
Gene expression for ~250 genes in each cell

General layout of visualization



3D plot vs. small multiples





Questions for feedback

- Would it be useful to have the plots on different tabs? Or to select which two plots you prioritize?
- Should we include 2D projections of the 3D plot?
- What would be the most useful visualization for zooming in on gene expression?

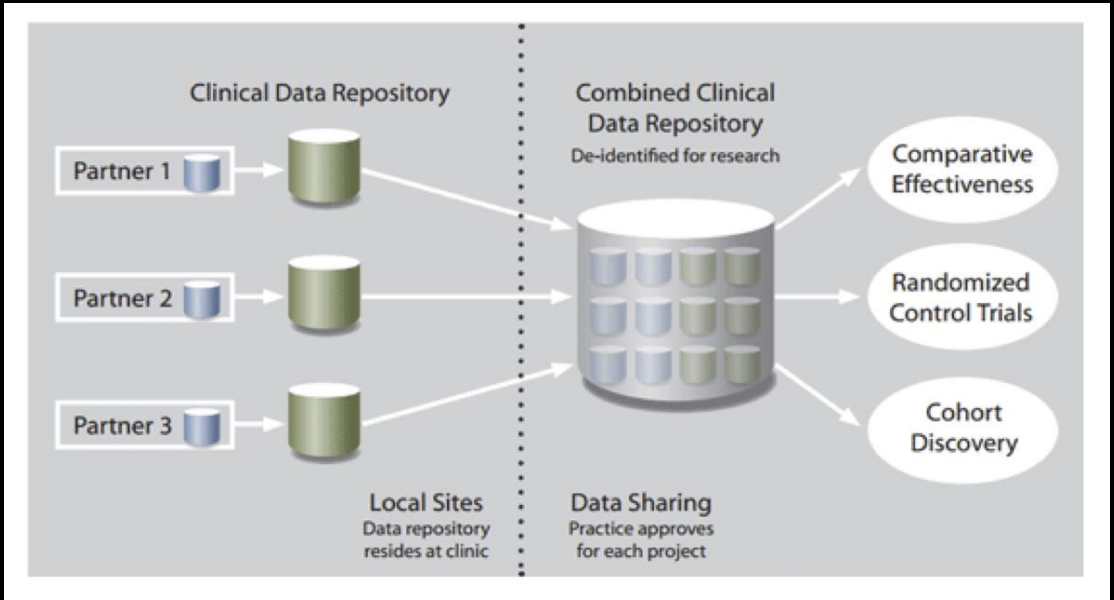
Visualizing Electronic Health Record Data to Aid Exploration

Ahmad Aljadaan, Ross Lordon, Jacob Olsufka

Problem

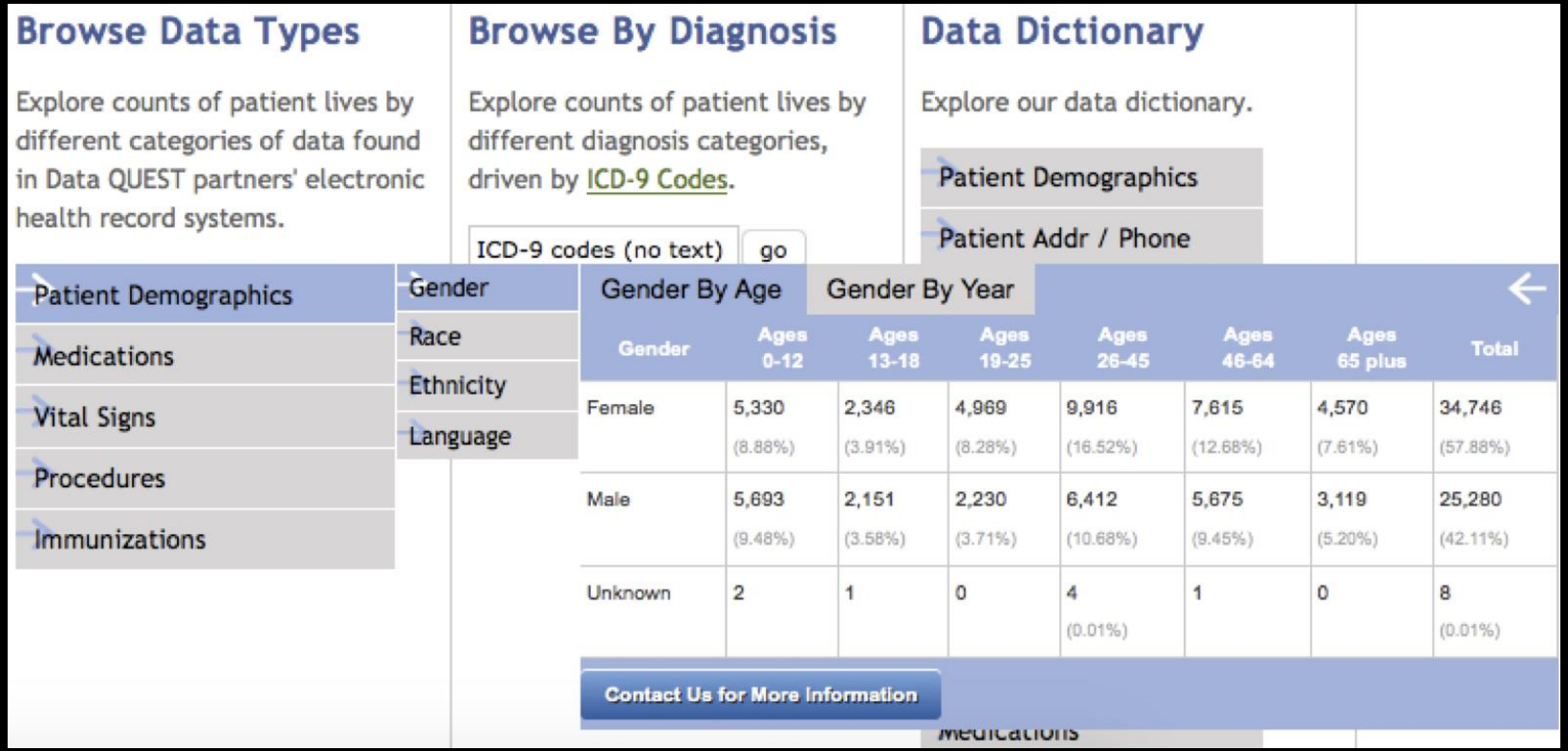
- Data QUEST
  - Supported by ITHS at UW
  - Promotes translation of scientific discovery to practice by fostering innovative research
  - An electronic health data-sharing architecture across community-based primary care practices in Washington and Idaho.

Problem



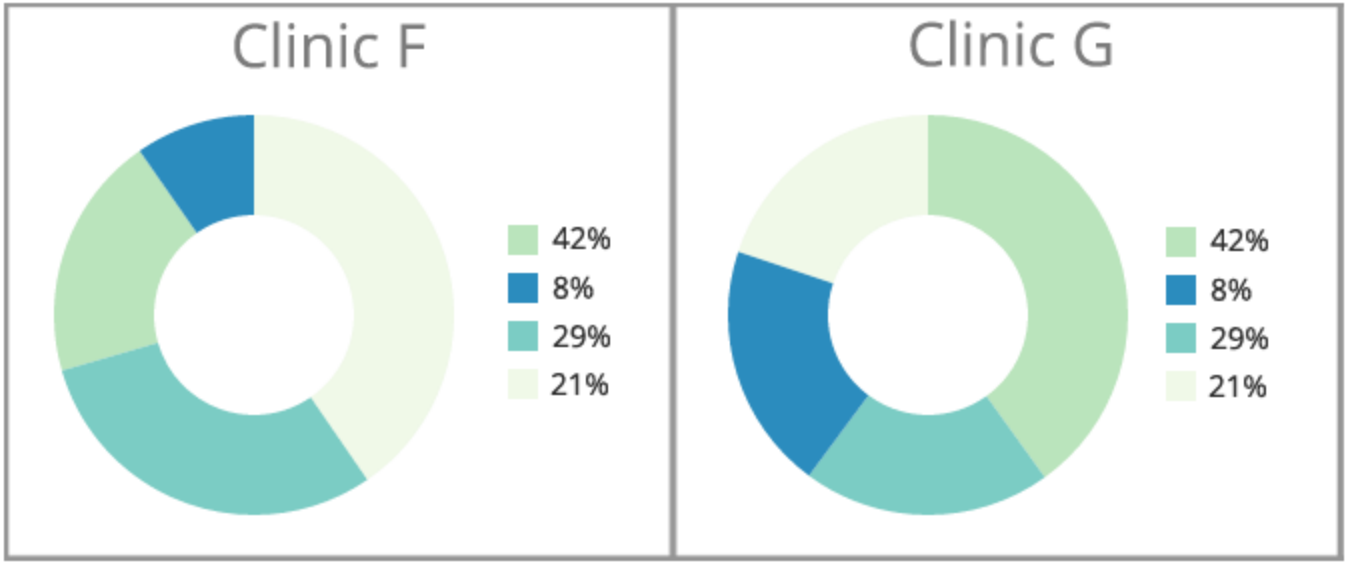
- Types of data
  - Patient demographics
  - Medications
  - Vital Signs
  - Procedures
  - Immunizations
  - Diagnoses
  - Lab Tests

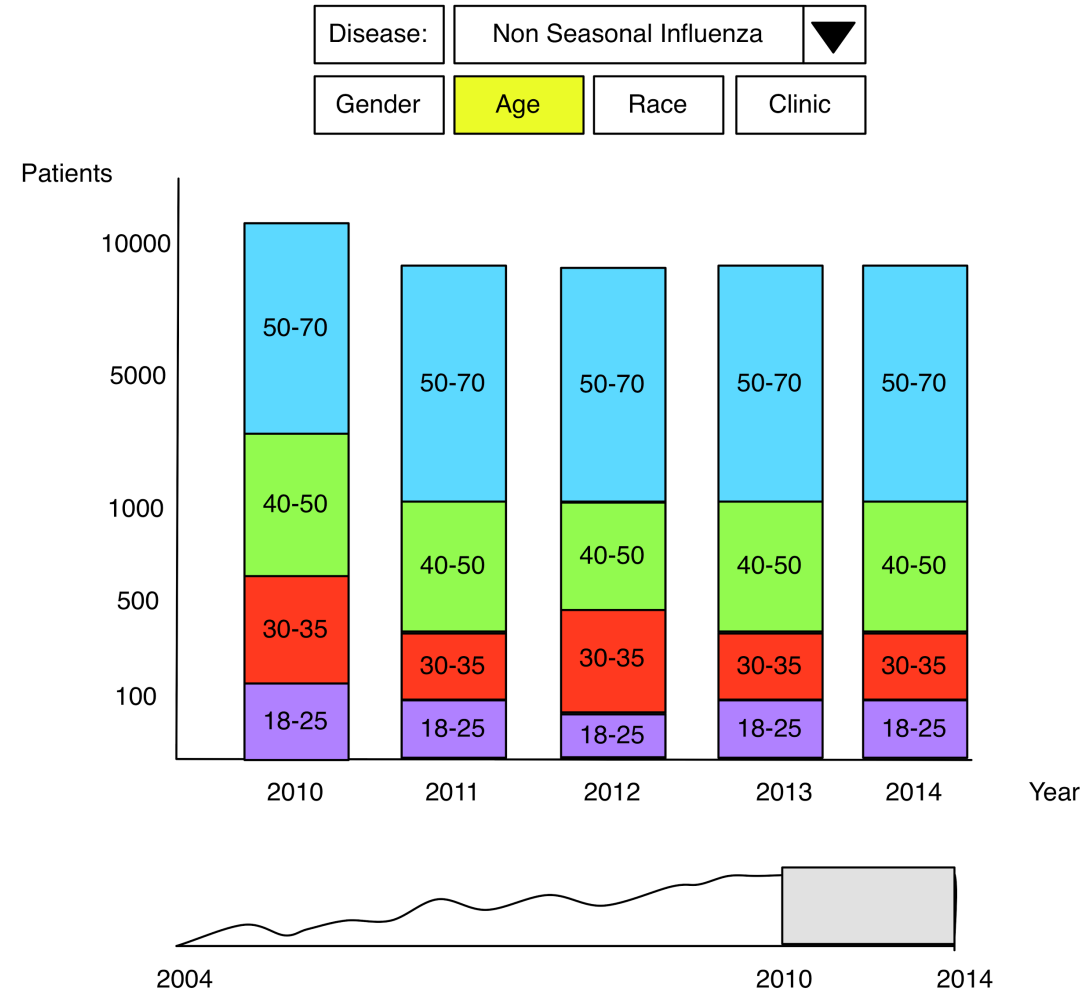
Problem


Browse Data Types
Explore counts of patient lives by different categories of data found in Data QUEST partners' electronic health record systems.
Patient Demographics
Medications
Vital Signs
Procedures
Immunizations
Gender
Race
Ethnicity
Language
Browse By Diagnosis
Explore counts of patient lives by different diagnosis categories, driven by ICD-9 Codes.
ICD-9 codes (no text)
go
Data Dictionary
Explore our data dictionary.
Patient Demographics
Patient Addr / Phone
Gender By Age
Gender By Year
Gender
Ages 0-12
Ages 13-18
Ages 19-25
Ages 26-45
Ages 46-64
Ages 65 plus
Total
Female
5,330 (8.88%)
2,346 (3.91%)
4,969 (8.28%)
9,916 (16.52%)
7,615 (12.68%)
4,570 (7.61%)
34,746 (57.88%)
Male
5,693 (9.48%)
2,151 (3.58%)
2,230 (3.71%)
6,412 (10.68%)
5,675 (9.45%)
3,119 (5.20%)
25,280 (42.11%)
Unknown
2
1
0
4 (0.01%)
1
0
8 (0.01%)
Contact Us for More Information

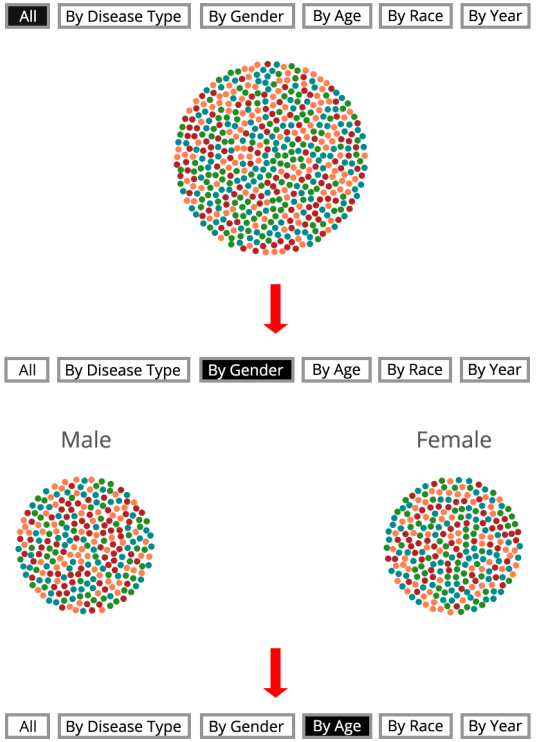

Clinic A
42%
8%
29%
21%
Clinic B
42%
8%
29%
21%
Clinic C
42%
8%
29%
21%
Clinic D
42%
8%
29%
21%
Clinic E
42%
8%
29%
21%
Clinic F
42%
8%
29%
21%
Clinic G
42%
8%
29%
21%
Clinic H
42%
8%
29%
21%
Clinic I
42%
8%
29%
21%
Clinic J
42%
8%
29%
21%
Clinic K
42%
8%
29%
21%
Clinic L
42%
8%
29%
21%
Clinic M
42%
8%
29%
21%
Clinic N
42%
8%
29%
21%
Disease A
Disease B
Disease C
Disease D
Select Variable
Disease Type
Race
Age
Ethnicity
Gender
Procedure

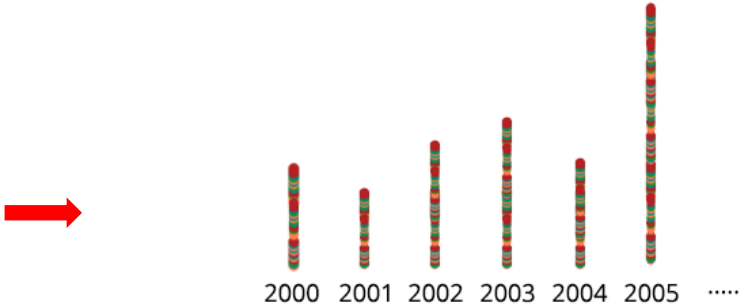
Comparison

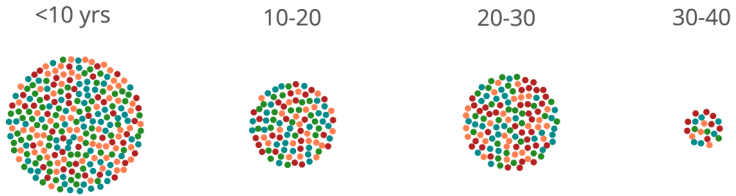



Disease:
Non Seasonal Influenza
Gender
Age
Race
Clinic
Patients
10000
5000
1000
500
100
50-70
40-50
30-35
18-25
2010
2011
2012
2013
2014
Year
2004
2010
2014

Force Bubble Charts with transitions







Timeline

5/19 - Data Request Sent

5/21 - Present ideas to class and refine

5/22 - Present ideas to ITHS team and refine

5/29 - Basic prototypes coded

6/5 - Finalize visualizations

6/7 - Finish Poster

6/8 - Poster Presentation

6/10 - GitHub Repo finalized

Questions for the class

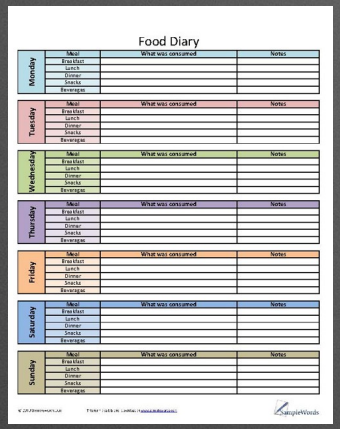
- *What do you think the best comparison technique (bar chart, donut chart, etc) would be best for the small multiples?*
- *Would it be beneficial to offer the viewer more than one small multiple comparison techniques as opposed to only one?*
- *Can anyone think of additional methodologies to demonstrate both the breadth and depth of this data available to researchers?*
- *Which method would be more powerful to allow researchers in this domain to have narrative story about the data?*

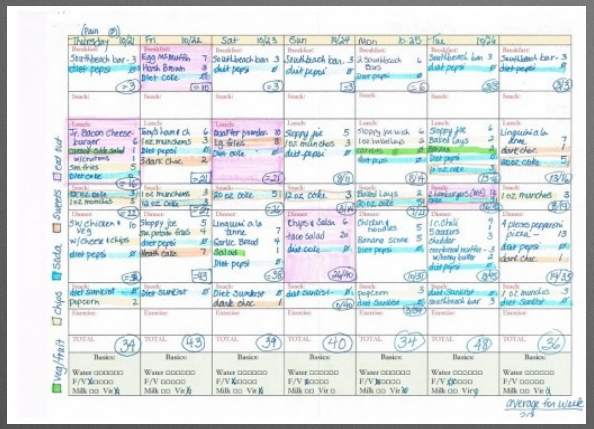
DECAFData / Foodprint

Christina Chung, Rick Huang

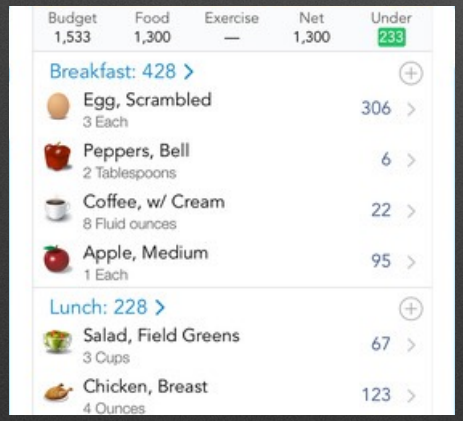
Problem

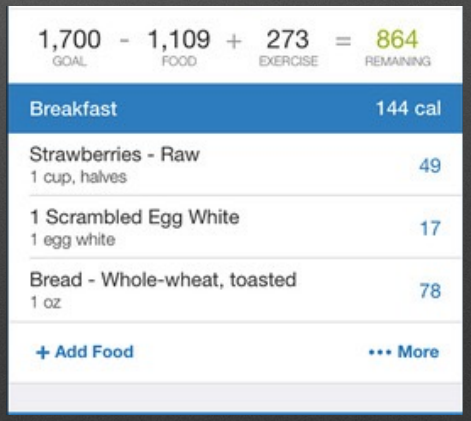
If you type Food Diary into Google image search

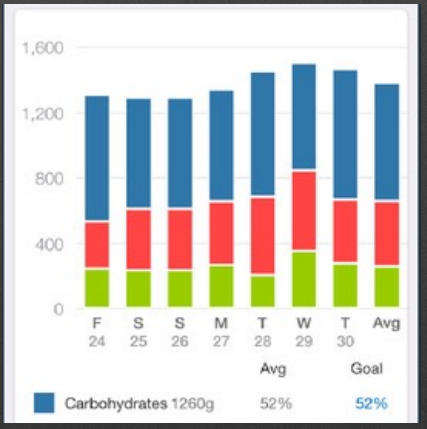




If you type Food Diary into app store







Prior Work

Providers and patients have various goals when reviewing and sharing self-tracking data

Providers (Chung et. al, under review)

Supporting Diagnosis
Personalizing treatment
Learning about patients
Increasing motivation and accountability
Facilitating discussion and managing visits

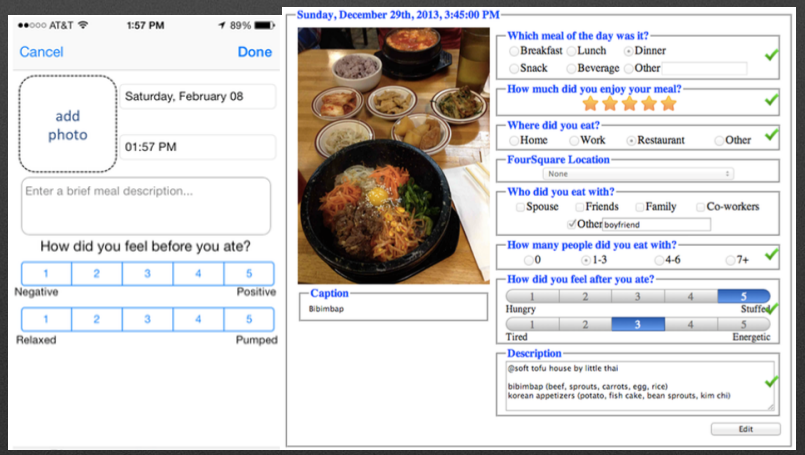
Patients (Chung et. al, under review)

Making sense of data
Showing a Complete picture of daily life
Personalized and actionable care plan
Support motivation and accountability
Seeking recognition and emotional support

People can not estimate calories correctly (Cordeiro et. al, 2015)

Why we are different

Support goal-oriented visualization and summary



Cordeiro et. al, 2015

Find the anomalies

Potential Flow

What is the purpose of this review?

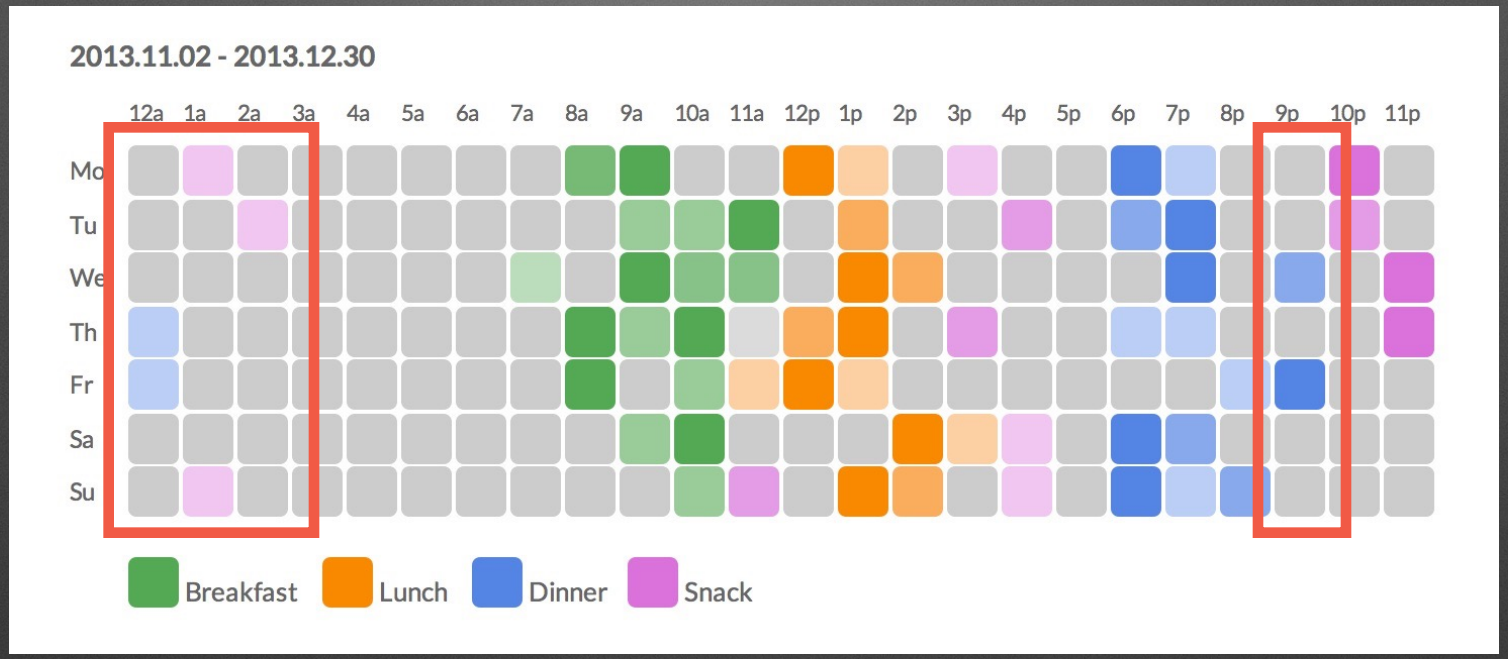
- Ensure regular eating time
- Suggest possible diet plan based on location or social setting
- Help with emotional eating

Overview based on goals

Interaction to support detail

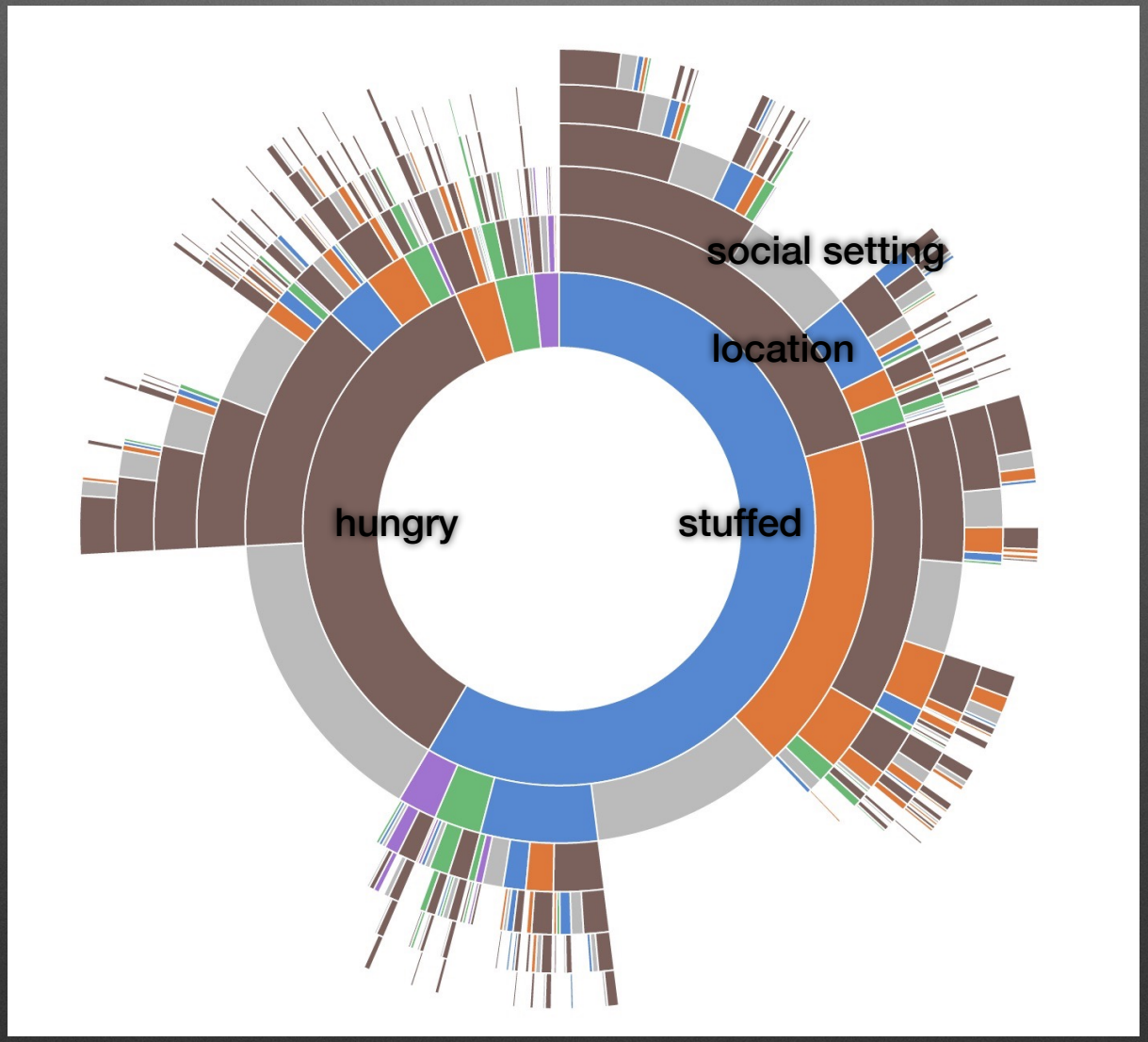
Potential ideas for variables related to time

Observe abnormal meal time



How to show overlapping meals?

Potential ideas for other contextual variables



Dynamically change tiers according to user's preference

Thank You

Mapping the social networks of global health organizations

Katie Doroschak and Jessica Schroeder

The GRO Foundation Lesotho

http://www.grofoundation.org

Organisation Type:
Community-based Organisation, Local NGO

Target Group:
Care-givers, Elderly, Orphans and Vulnerable Children, People living with HIV, Women

Domain:
Training and Skills

Sector:
Agriculture, Education, Environment, Food, Healthcare, HIV/AIDS, Home-based Care, Livelihoods, Power Supplies, Training

District:
Leribe

Core Activities:
Family Scholarship Program: comprehensive scholarship program for OVC focusing on academic tutorials, life skills, career guidance, and post secondary support.
Grandmothers Support Group: OVC community care through Grannies poultry project initiative and home-based care program.
G.R.O. Artisans Collective: Fashion jewelry collective for HIV positive mothers with an emphasis on health care, busines training, and skills development.

Future Plans:
- To establish an eco-sustainable demonstrational poultry housing model for Lesotho.
- To expand G.R.O. Artisans and develop similar collectives in Lesotho.
- To establish a career cdenter, career curiculum, and a post secondary education resource booklet.
- To support community care facilities for OVCs.

Funding:
The G.R.O. Foundation Canada/USA

This organisation's info updated on: Thu, Jun 18, 2009. Please disregard the date below.

Contact Details:
James White, Chairperson
t: +266 5914 7445
e: director@grofoundation.org

The GRO Foundation Canada
Jean Margaritis, Co-director/program manager
t: +266 5943 3797
e: jmargaritis@grofoundation.org

Postal Address:
Postal Address:
Private Bag C0020
Leribe 300
Lesotho

Physical Address:
Physical Address:
Motebang, Leribe

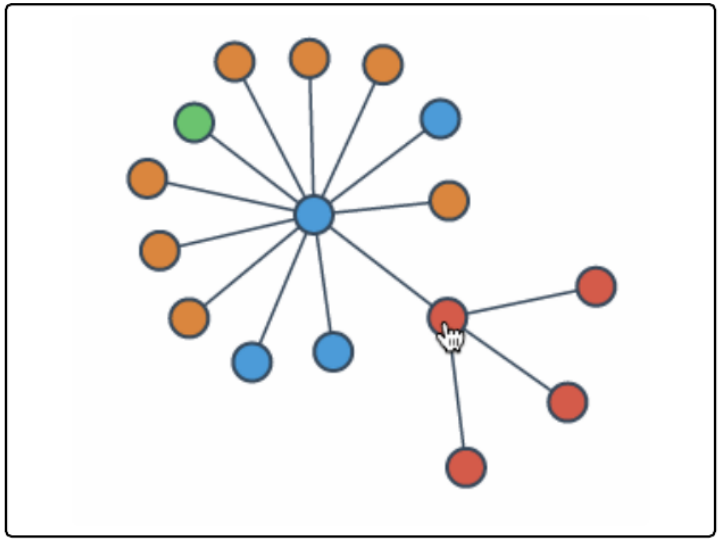
Our Focus

- What are the funding hierarchies of NGO's in the HIV/AIDS response of 2010-2011?

- How peripheral or well connected are smaller community-based organizations?

Title

http://website.com



View Funding Sources and Sinks for Selected Organization

Selected Organization Name

Organization Type:	International NGO
Aid Focus:	Care
Other Aid Given:	Education, Food and Nutrition
Home District:	Berea
Districts Served:	Berea, Butha-Buthe, Maseru

Organization Type

- [x] Community-based Organization
- [ ] International NGO
- [ ] Local NGO

Select all | Select none

Types of Aid Provided

- [x] Care
- [ ] Education
- [ ] Food and Nutrition
- [ ] Infrastructure Services
- [ ] Training and Skills

Select all | Select none

- [x] HIV-focused
- [ ] Religious
- [ ] Advocacy
- [ ] Service

Select all | Select none

Feedback

- What should we show initially?
  - The full network is too overwhelming
    - OK to start with pre-selected filters to view something interesting?
- We hoped to reduce clutter by having names as mouseovers; however, this means we sacrifice the ability to see multiple names at a time.
  - Do you have any suggestions on how/when/if we should include labels?
- Any general suggestions on how to keep it from being overwhelming?

Visualizing Rumor Spread on Twitter

Cynthia Andrews
Graeme Britz
Rama Gokhale
Zeno Koller

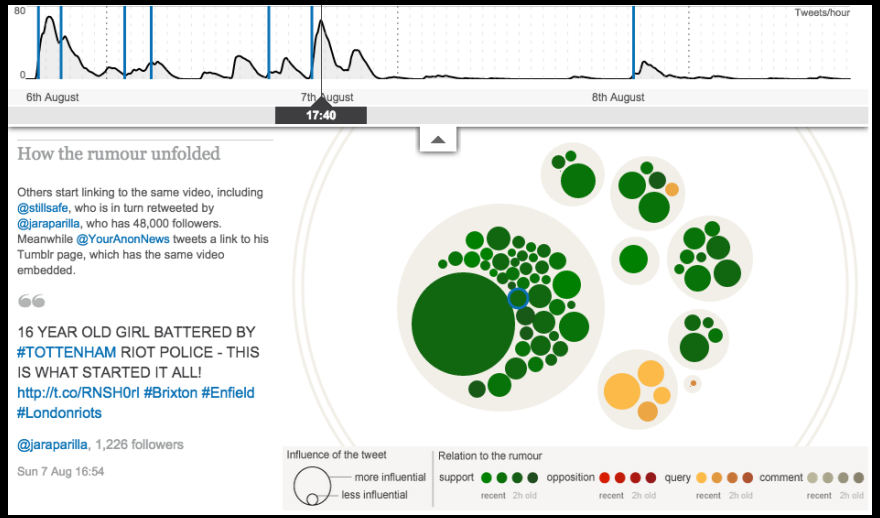
Goal

To visualize the effects of major players in the communication landscape of disaster events on Twitter

In Collaboration With...

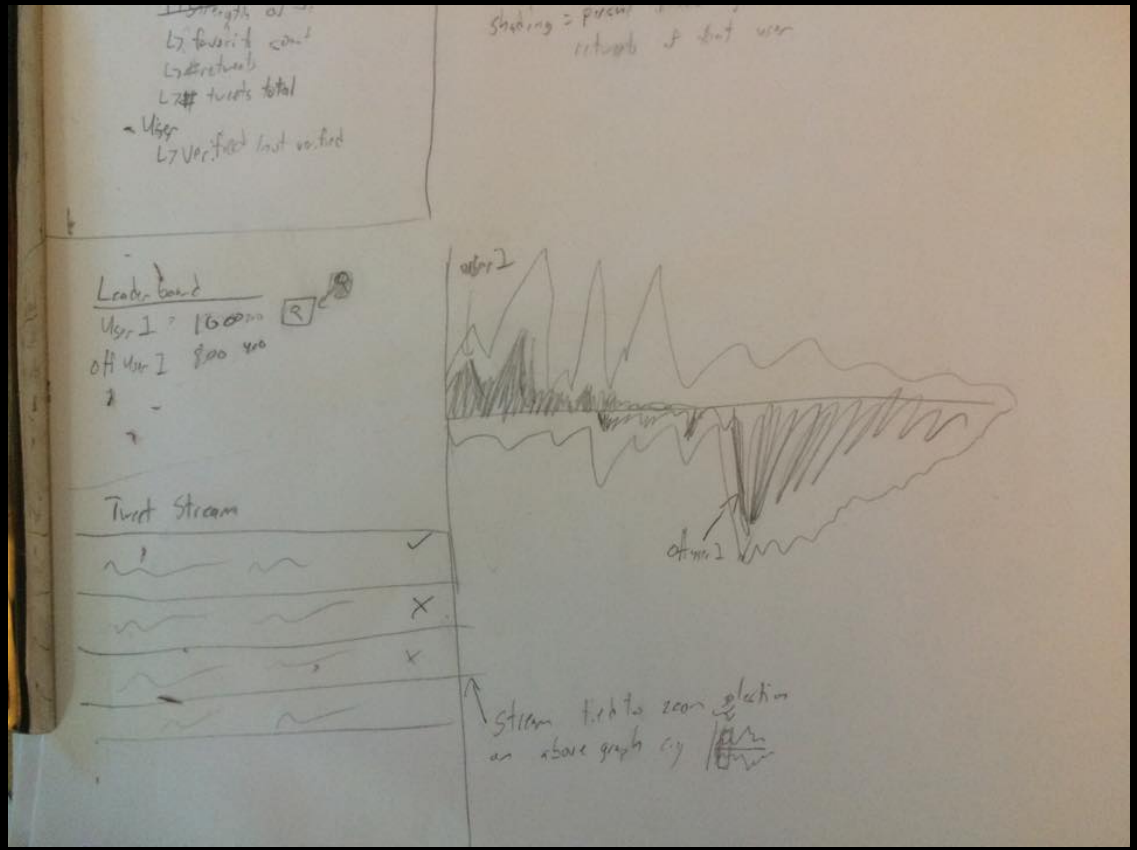
Research on the spread of misinformation through Twitter networks

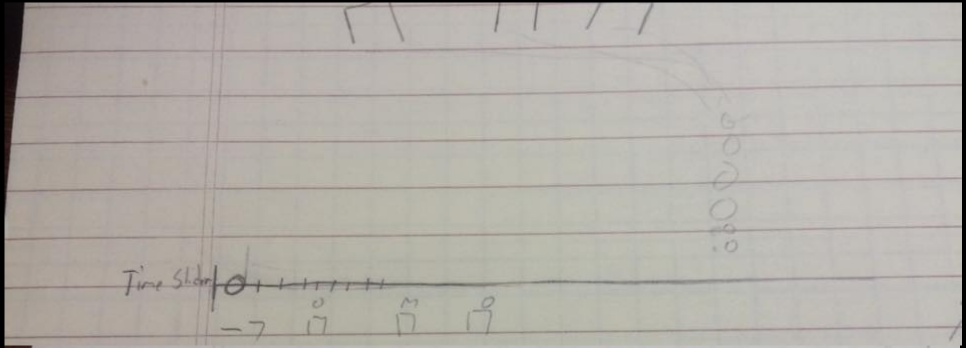
Relevant Work

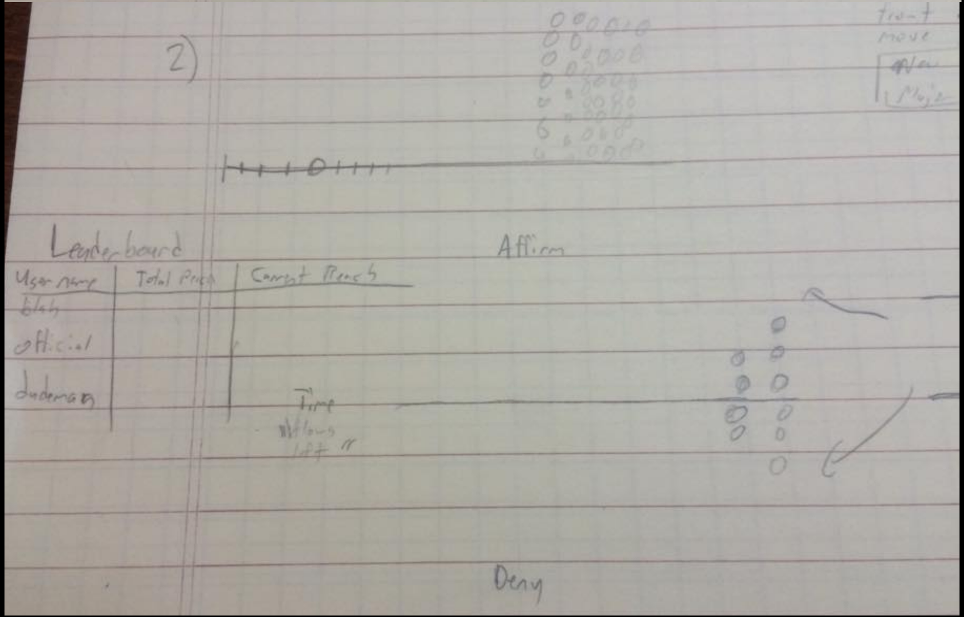


We are focusing more on major players and their influence on popular belief

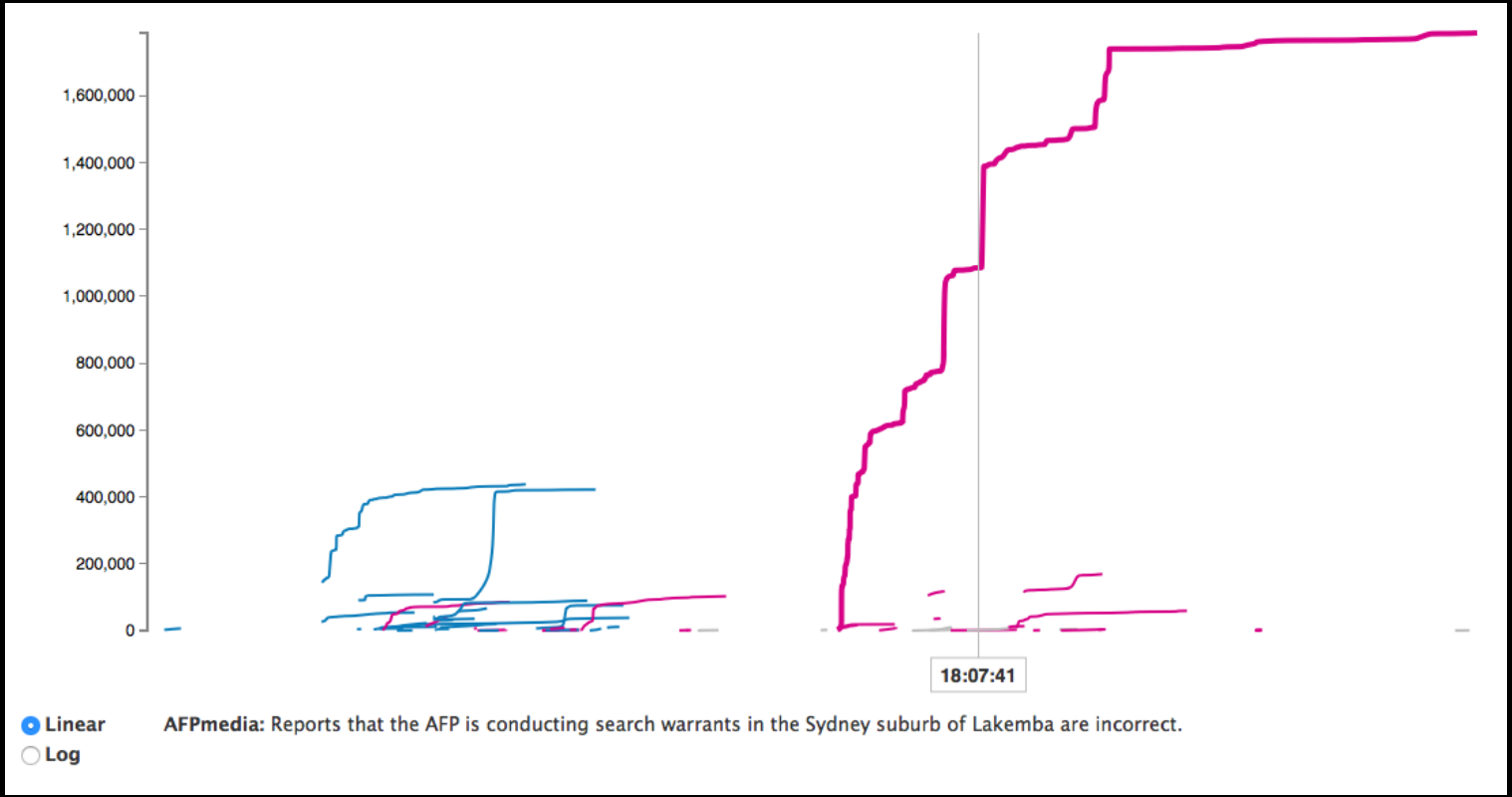
Idea #1: Stream

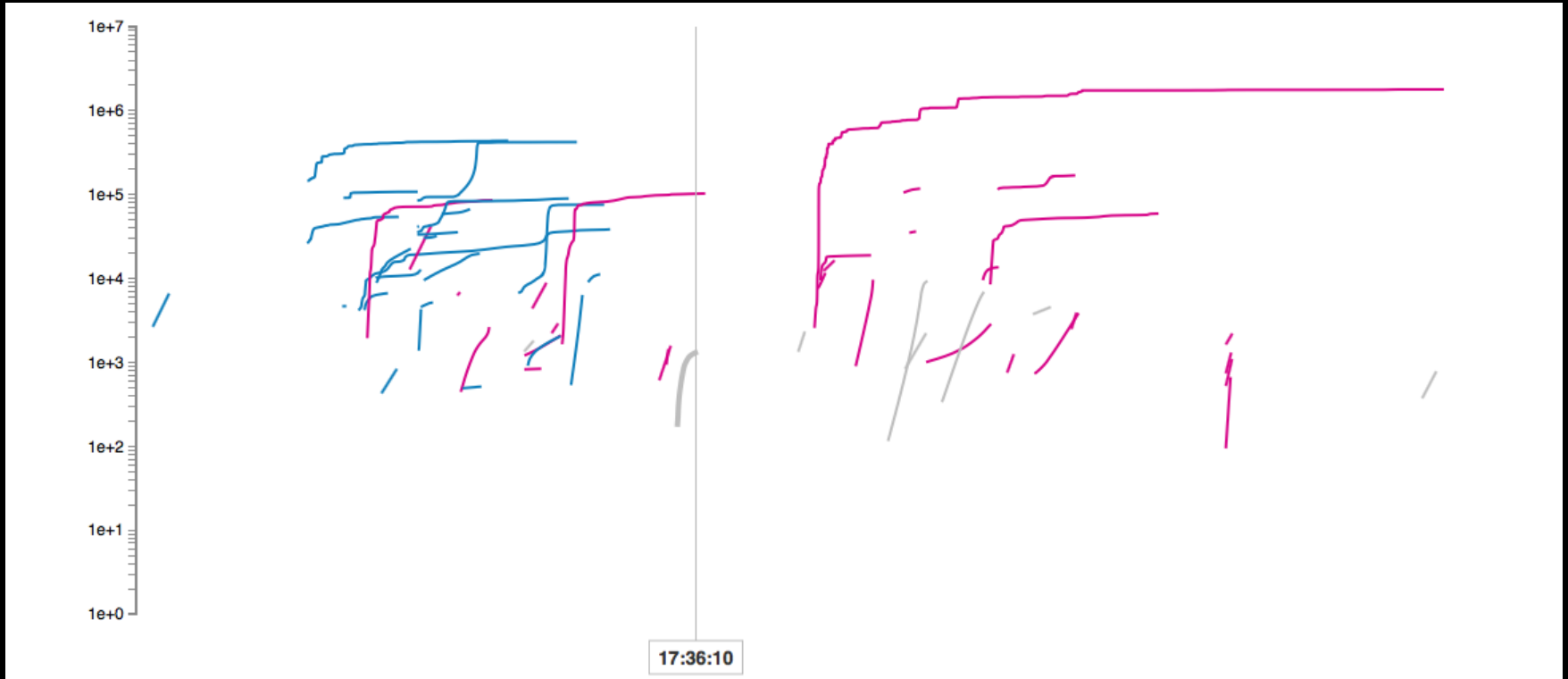



Time Slider


2)
Leaderboard
User name
Total Reach
Current Reach
blah
official
dudeman
Time
Affirm
Deny

Idea #2: Spaghetti

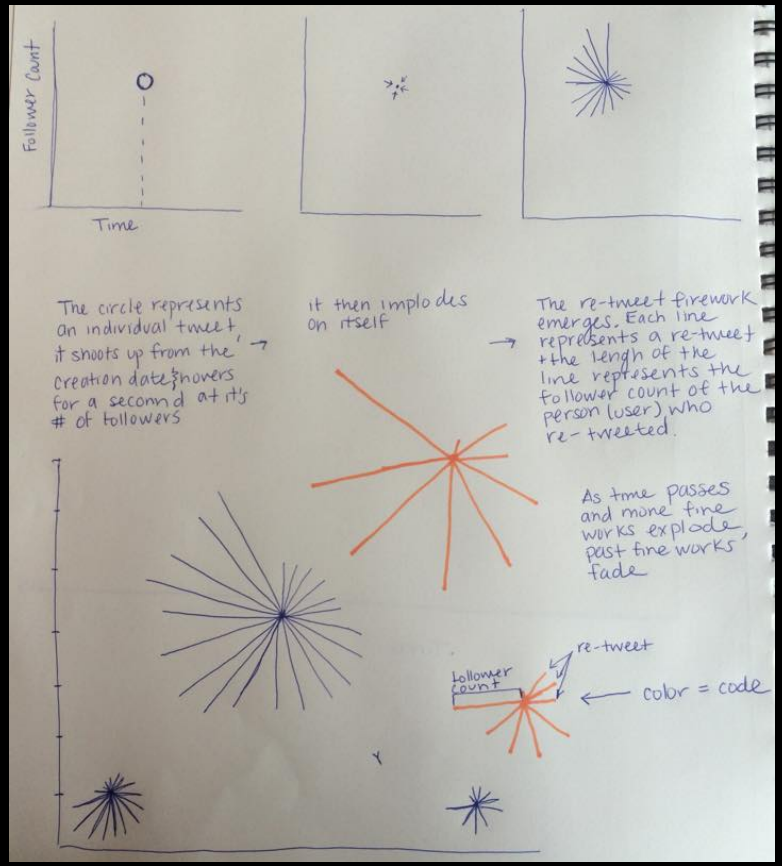




Linear
Log

Lina_Martinezz: So apparently similar attacks are happening in Lakemba and possibly other parts of Sydney, channel 7 idk if it's true but I'm shit scared

Idea #3: Fireworks



Help!

How can we more effectively highlight the most influential individuals?

How can we best present overall tweet volume alongside with tweet influence?

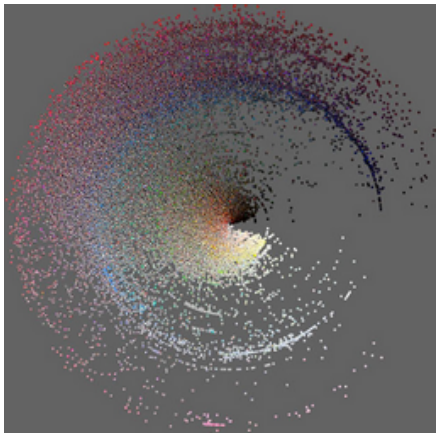
Best Filter?

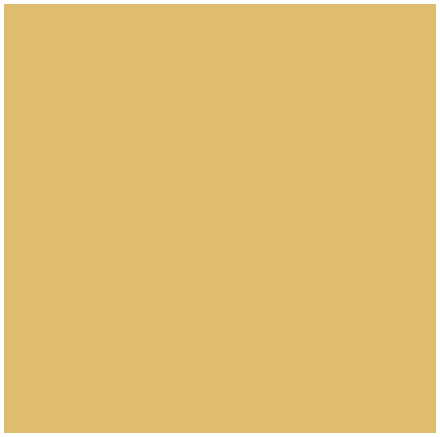
Instagram Data Analysis

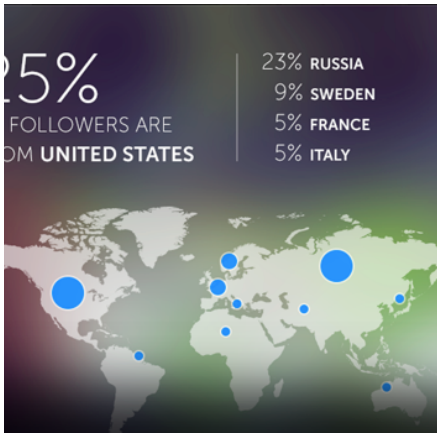
-Yanling He -Xiaoyi Zhang -Xin Yang

Instagram

Prior Work







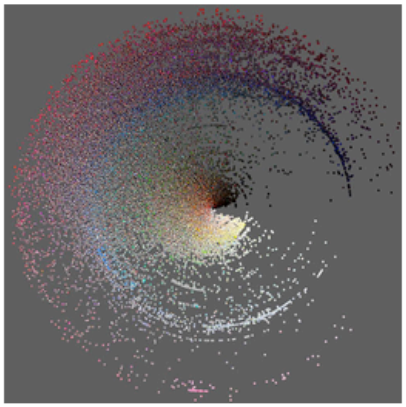
Prior Analysis Work on Instagram Data

Phototrails
Iconosquare

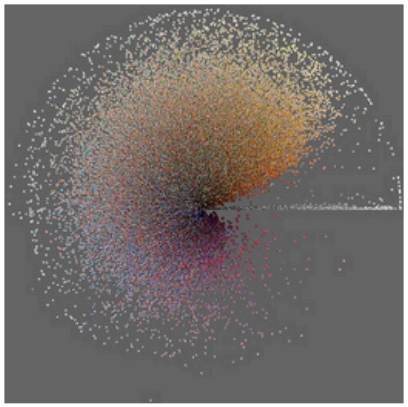
Phototrails

- Instagram Data V.S. Cities
  - Culture differences and social science
  - Show how visual social media can be analyzed at multiple spatial and temporal scales.
  - Analysis social and cultural dynamics in specific places and times

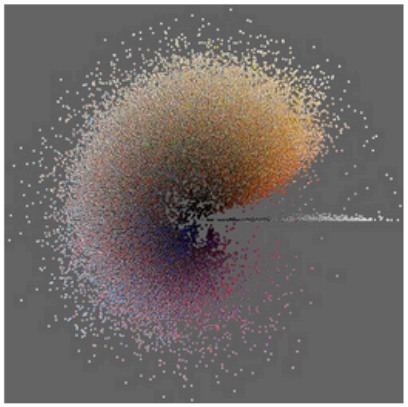
New York



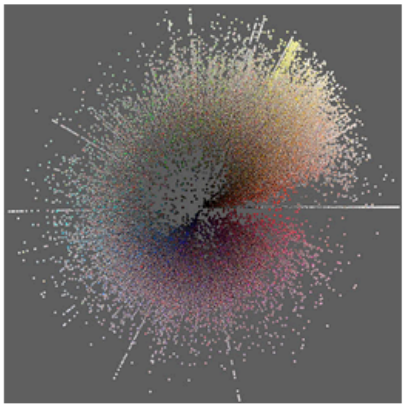
San Francisco



Tokyo



Bangkok

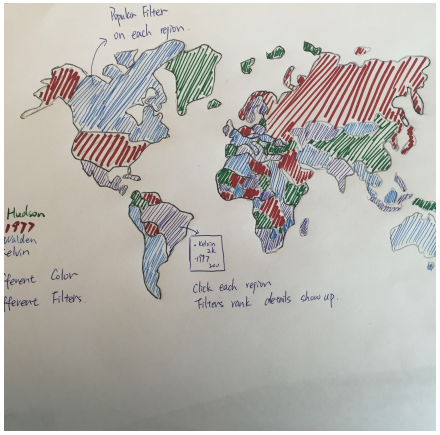


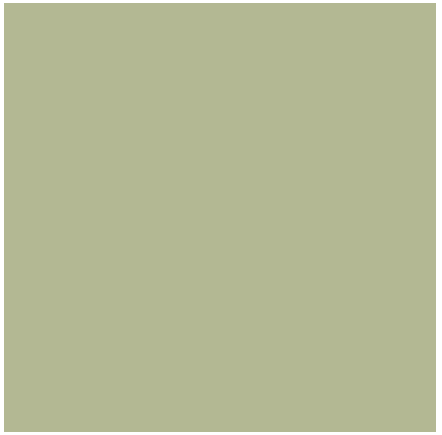
Iconosquare

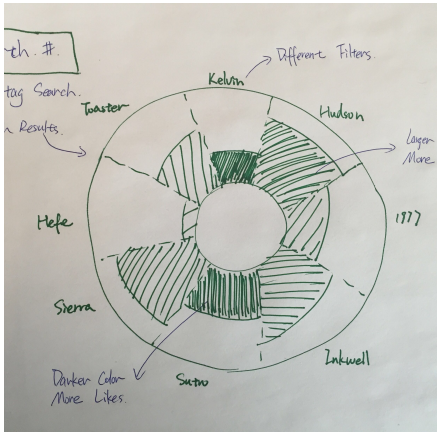
- Statistics about Instagram
  - Likes and comments
  - Followers growing chart
  - Hashtags
  - Promote Instagram account

Our Work

What's Best Filter







Visualize Filter Relationship

Locations
Likes
Tags

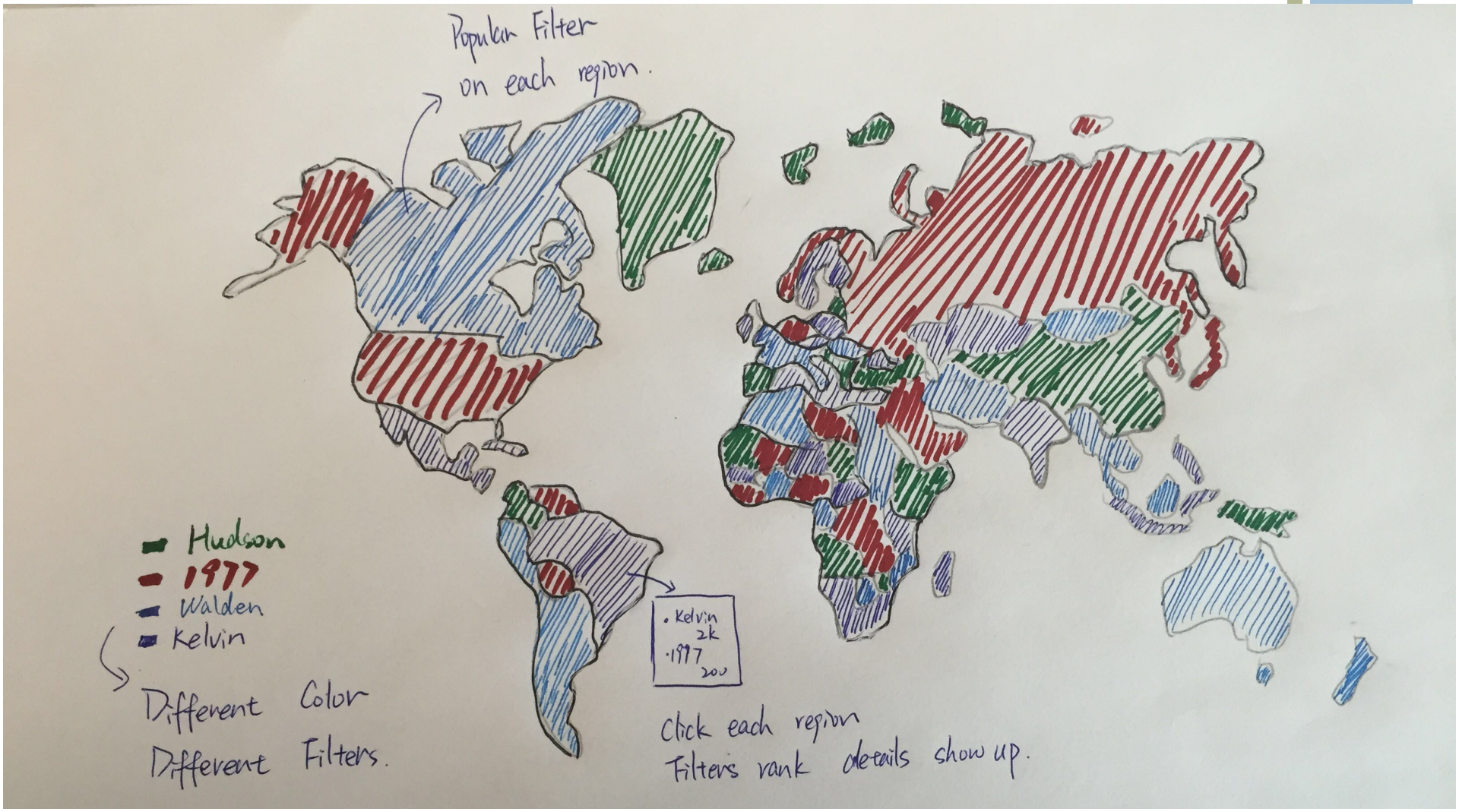
Data Resource

- Instagram API:
  - https://instagram.com/developer/endpoints/

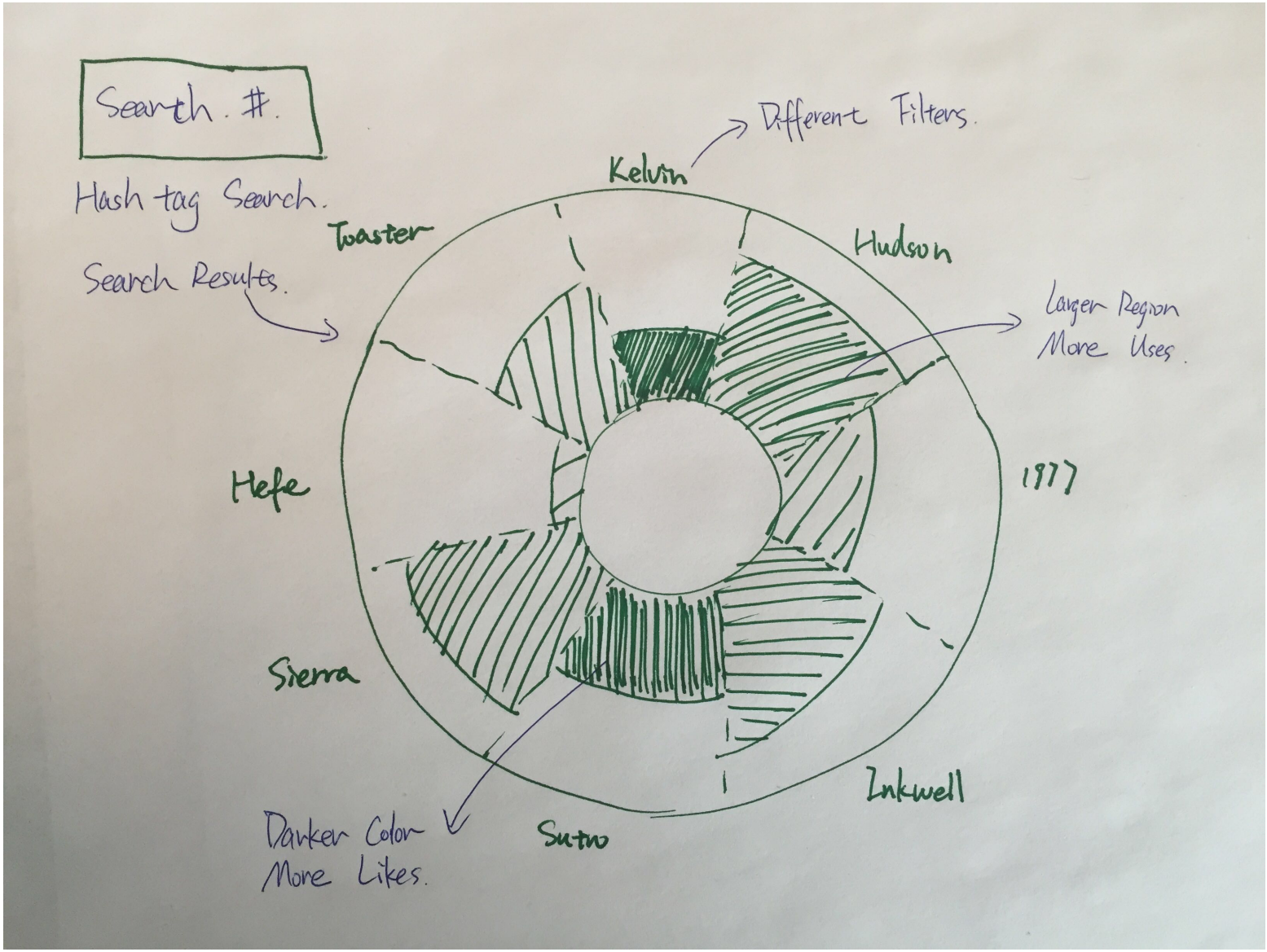

```
"data": [{
    "type": "image",
    "users_in_photo": [],
    "filter": "Earlybird",
    "tags": ["expobar"],
    "comments": {
        data: [],
        count: 0
    },
    "caption": {
        "created_time": "1296532028",
        "text": "@mikeyk pulls a shot on our #Expobar",
        "from": {
            "username": "josh",
            "full_name": "Josh Riedel",
            "type": "user",
            "id": "33"
        },
        "id": "25663923"
    },
    "likes": {
        "count": 35,
```


- Scripting Data:
  - http://www.rldw.de/2015/04/using-the-instagram-api-with-python/
  - https://github.com/rldw/instastats/

Filters - Locations



Filters – Tags and Likes



Open Questions

- How to put different features into one visualization instead of three different ones?
- Major purpose of this visualization?
  - filter suggestions?
  - Promotion?
  - Social or culture analysis?

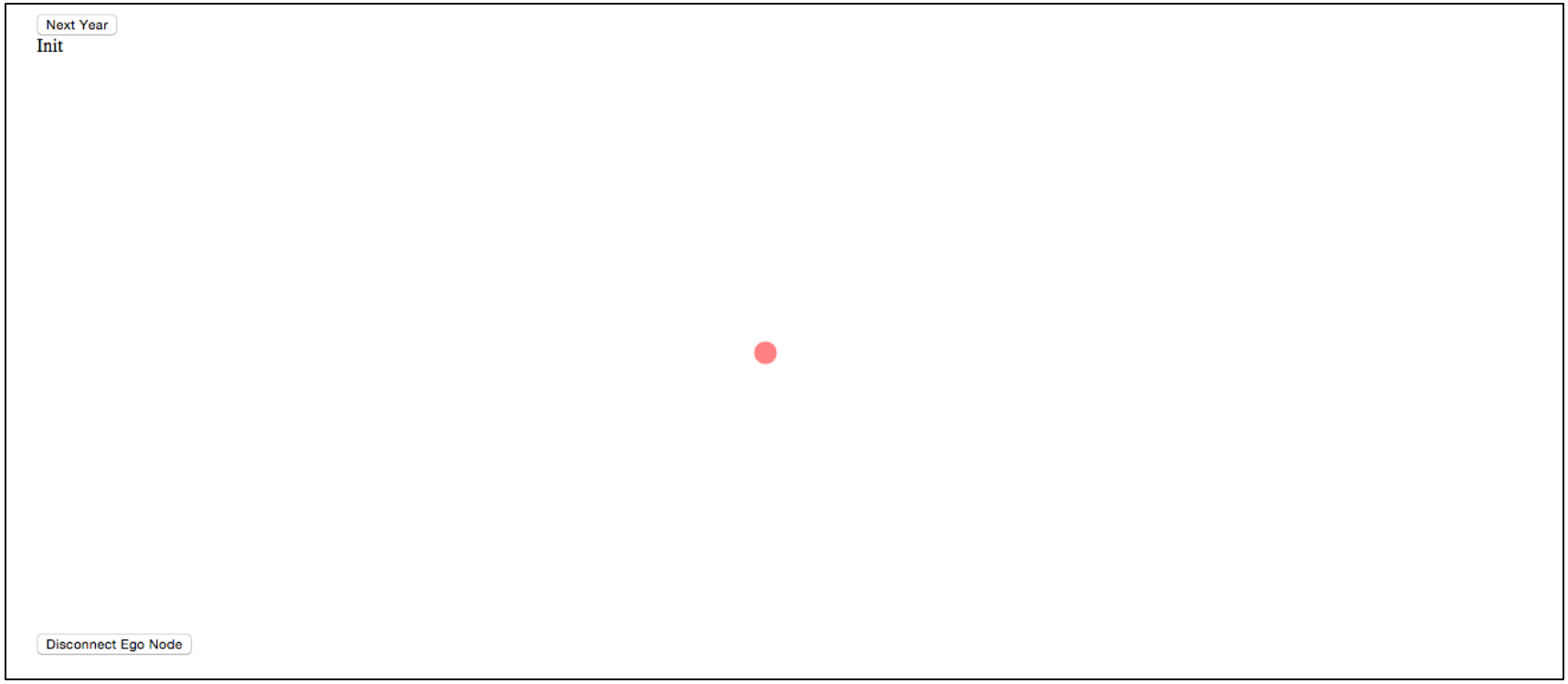
Visualizing Scholarly Impact over Time

Jason Portenoy and Muhammad Raza Khan

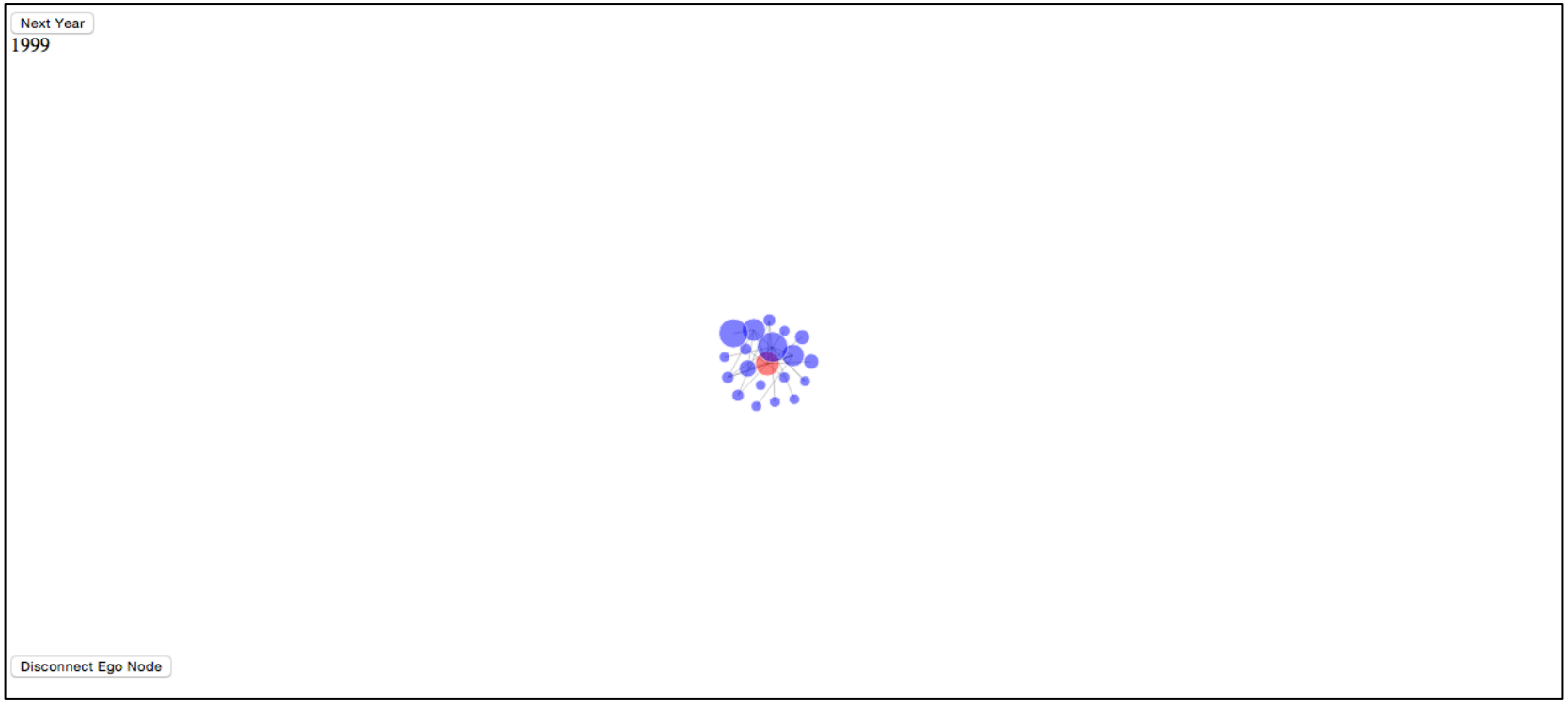
Goal

- Create a compelling narrative showing the influence that a particular researcher or research paper has had in scholarship.
- Use a node-link diagram:
  - Nodes = papers
  - Edges = citations between papers
- Watch the network develop over time.

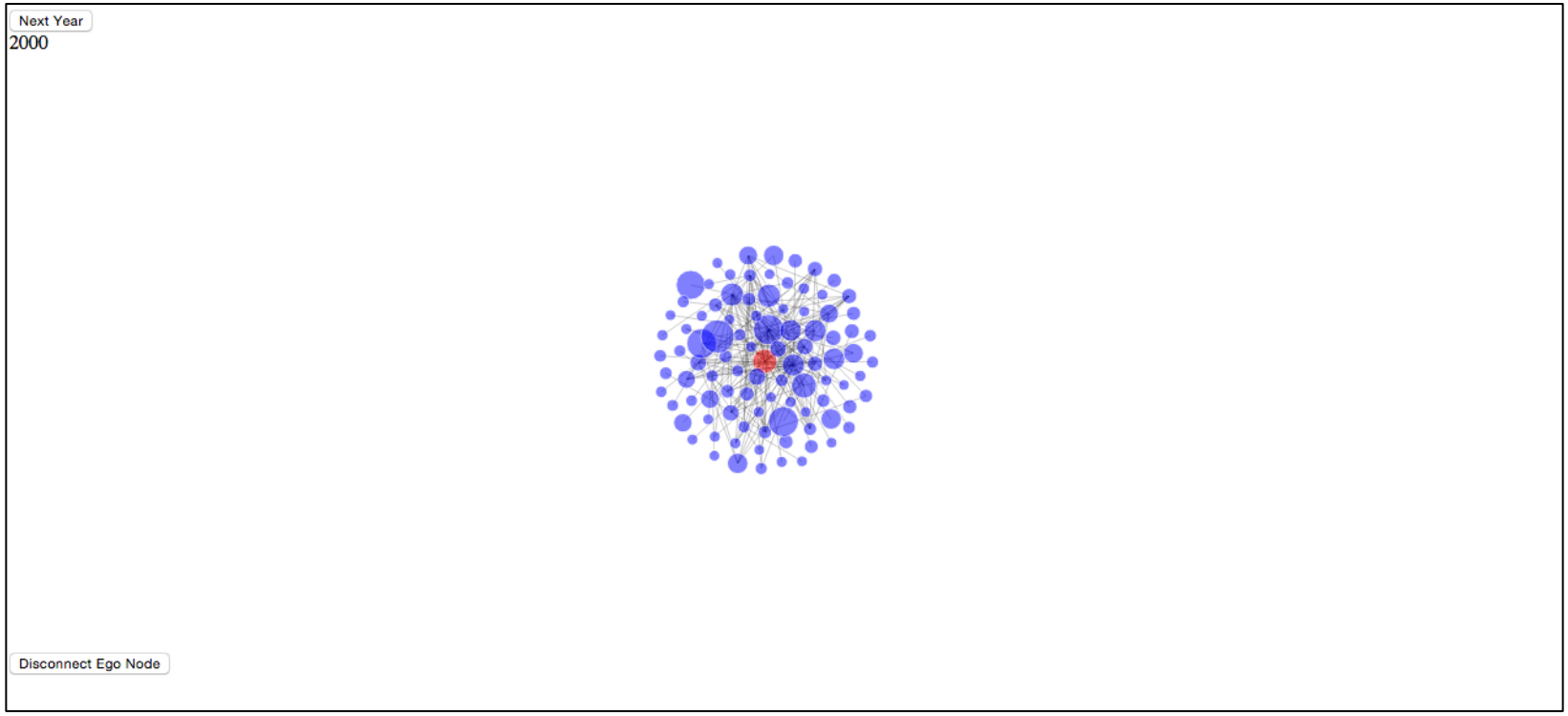
Initial Progress



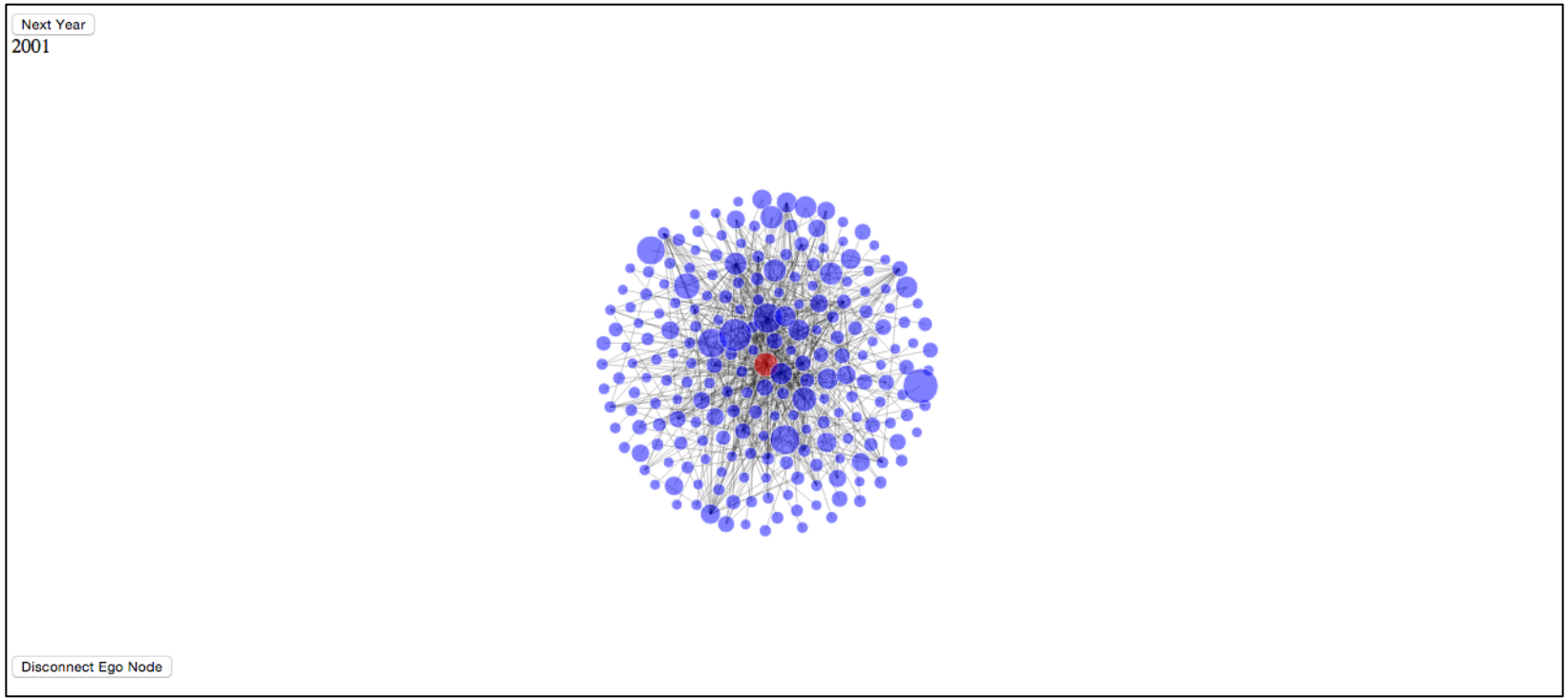
Initial Progress



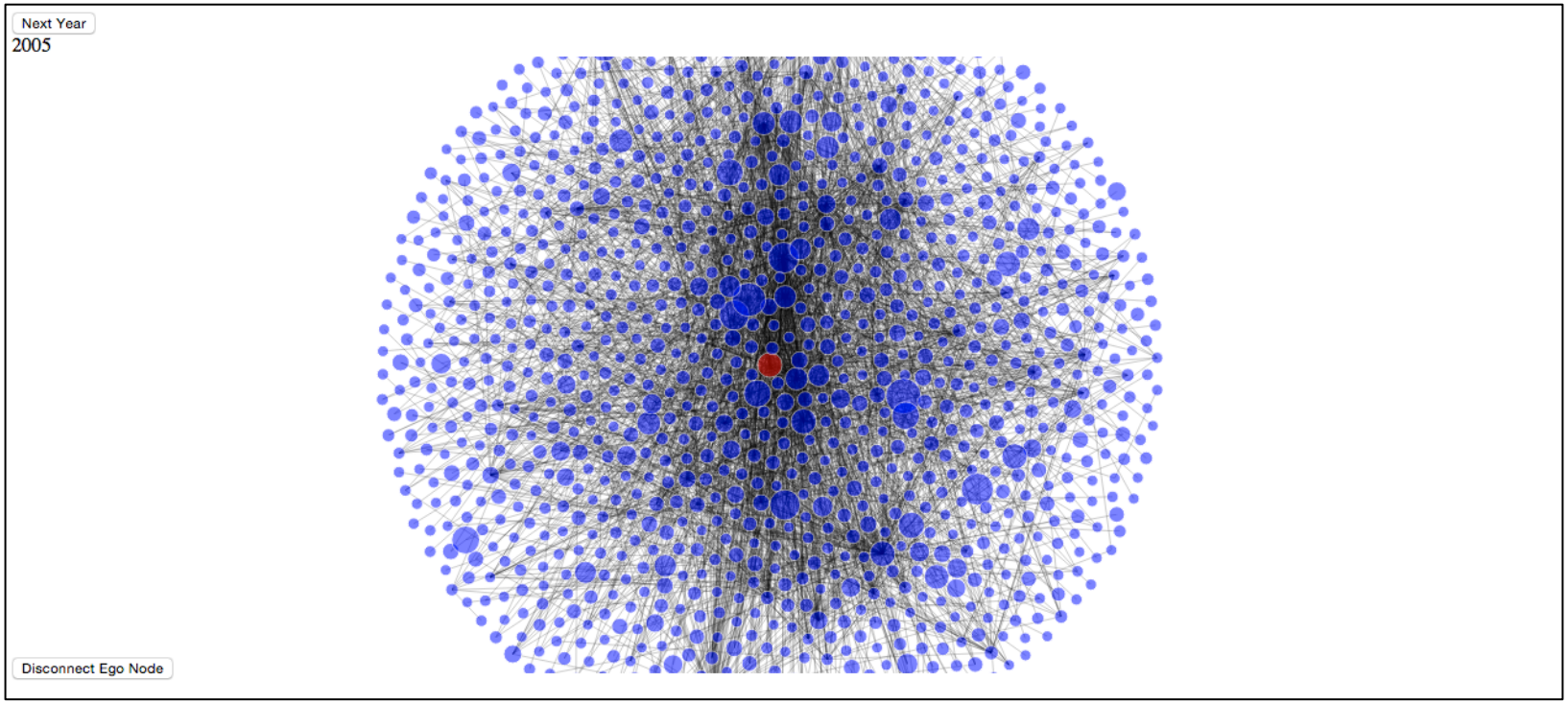
Initial Progress



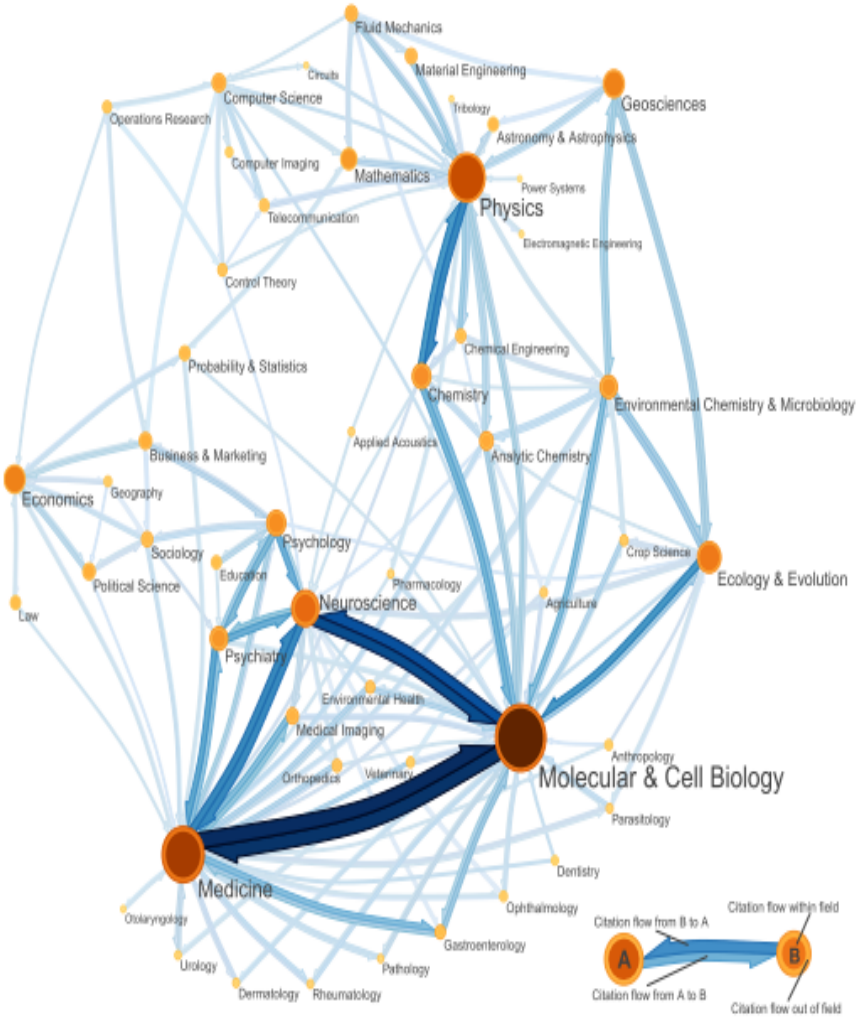
Initial Progress



Initial Progress

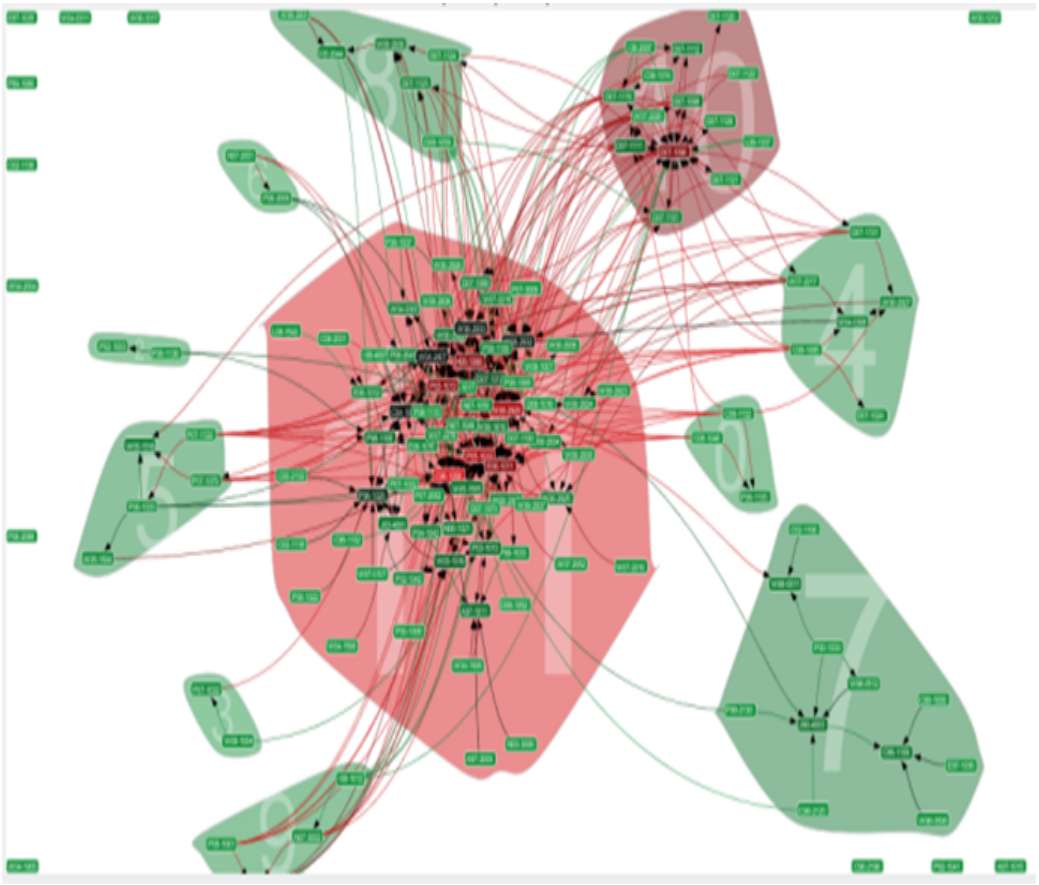


Related Work:
Fields Level Citation Networks

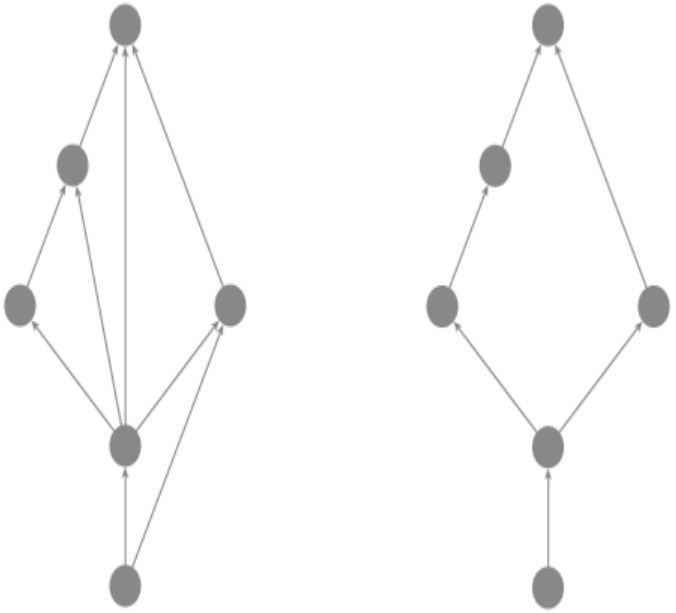


M. Rosvall and C. T. Bergstrom, “Maps of random walks on complex networks reveal community structure,” Proceedings of the National Academy of Sciences, vol. 105, no. 4, pp. 1118–1123, Jan. 2008.

Related Work: Paper-Level Citation Networks



N. J. van Eck and L. Waltman, "CitNetExplorer: A new software tool for analyzing and visualizing citation networks," arXiv:1404.5322 [physics], Apr. 2014.



Transitive Reduction in Networks

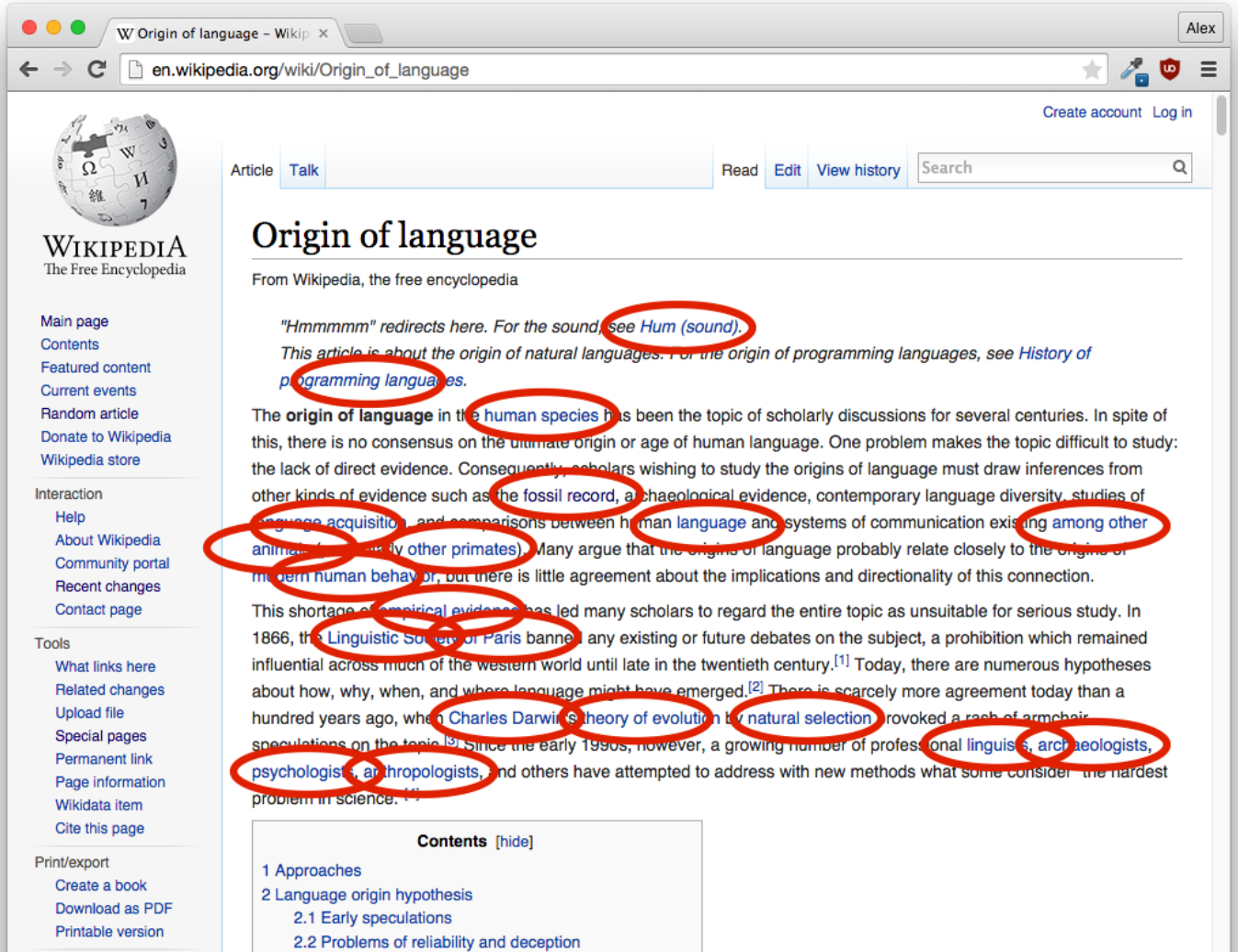
Questions for You

- How to best deal with the hairball problem?
  - We have data on clusters (research fields) and influence (based on network structure).
- Most effective ways to provide context?
- Alternatives to building the network by year?

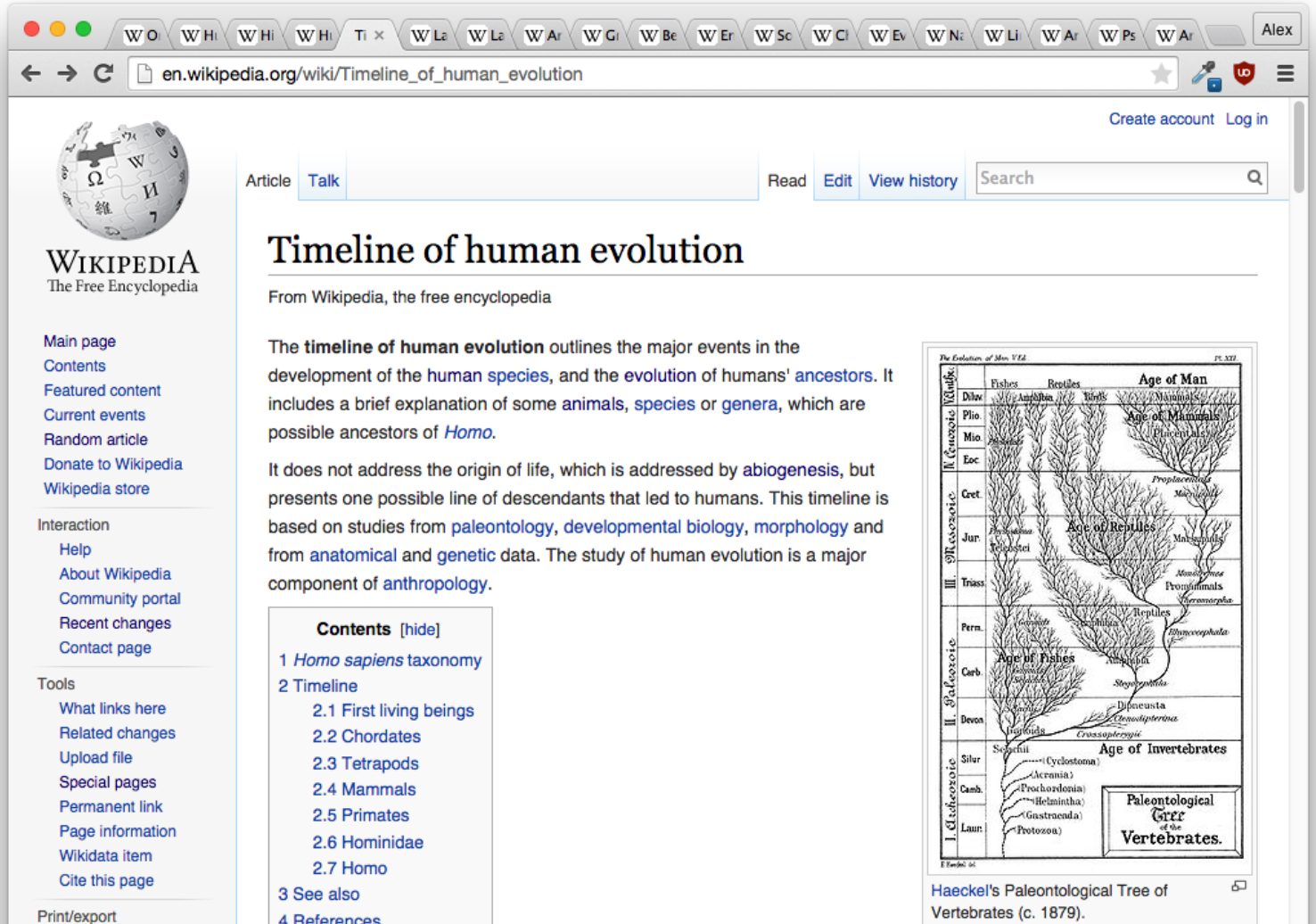
Wikitree

https://wikitree.website

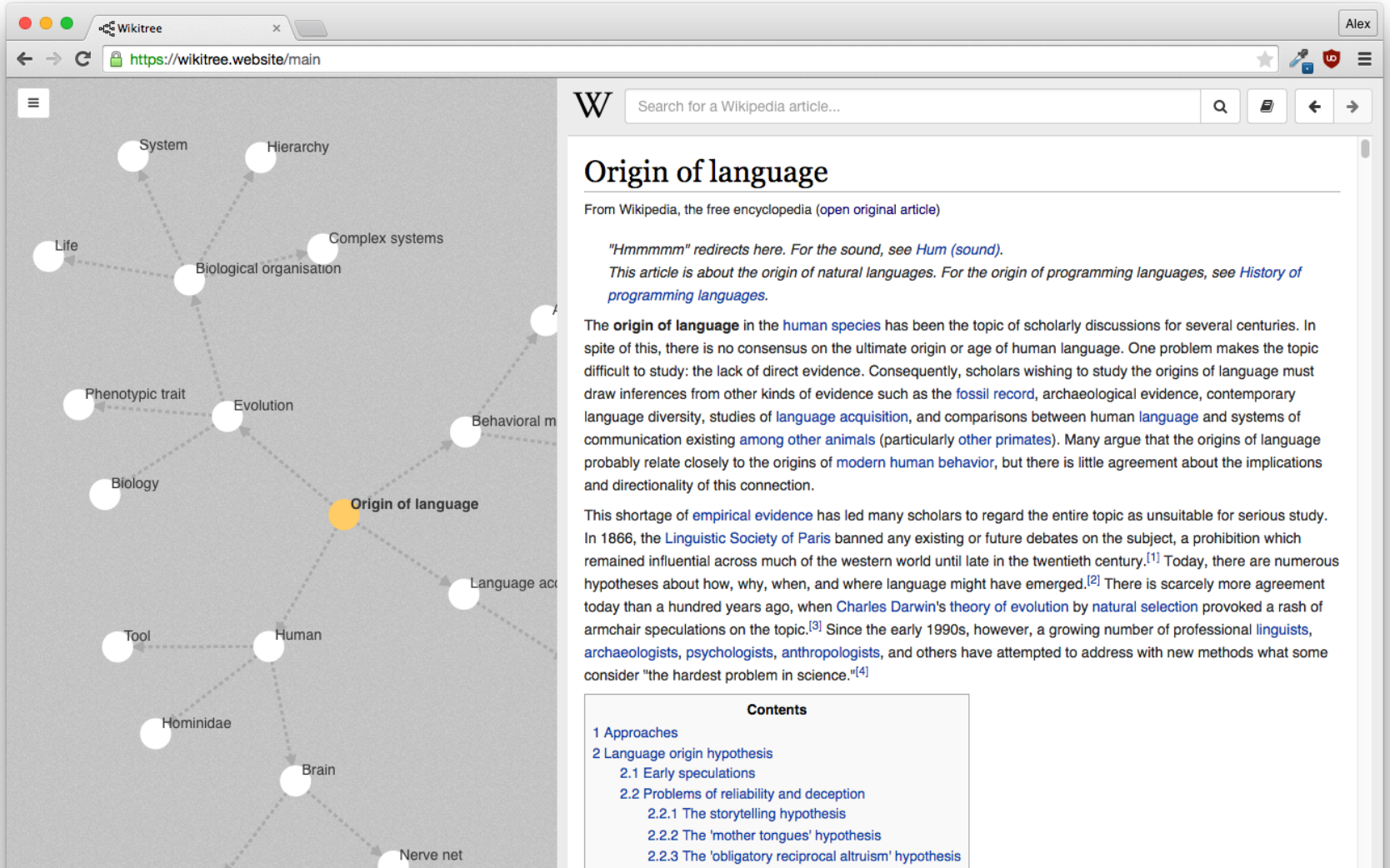
Wikipedia articles contain links to many other Wikipedia articles



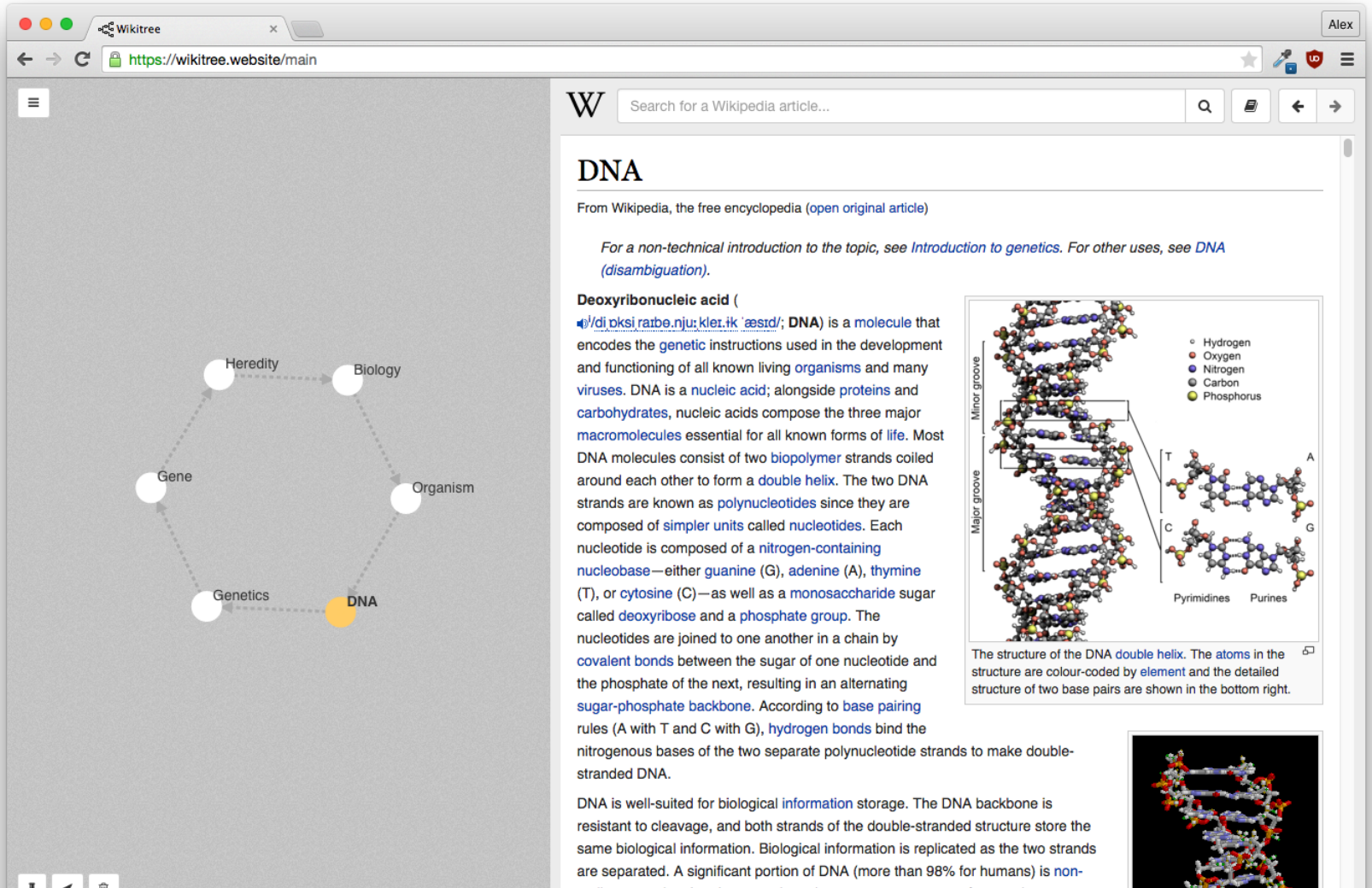
Browser window tabs and navigation histories are serial / linear paradigms



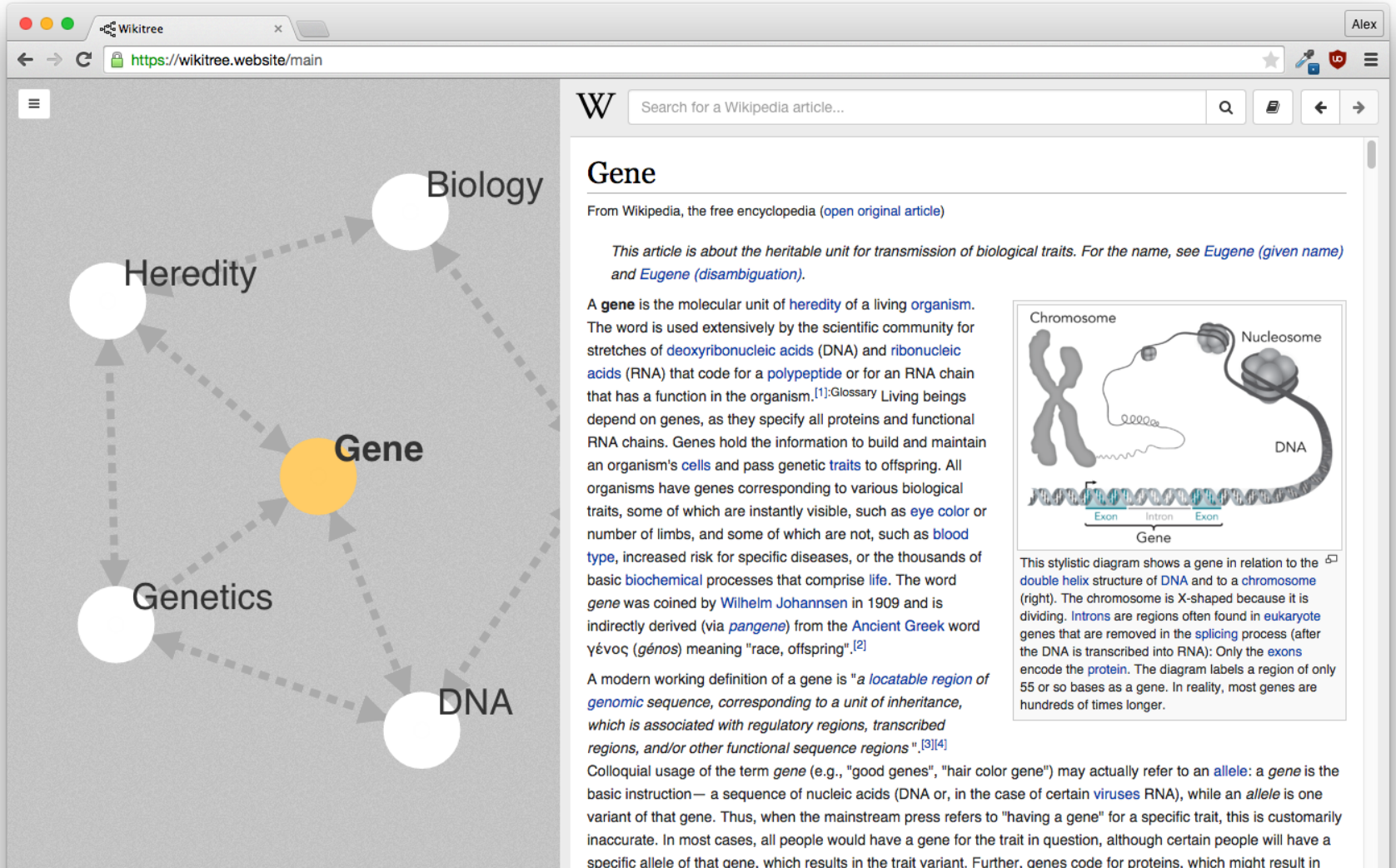
Wikitree represents a user's navigation paths using the D3 force-directed graph layout



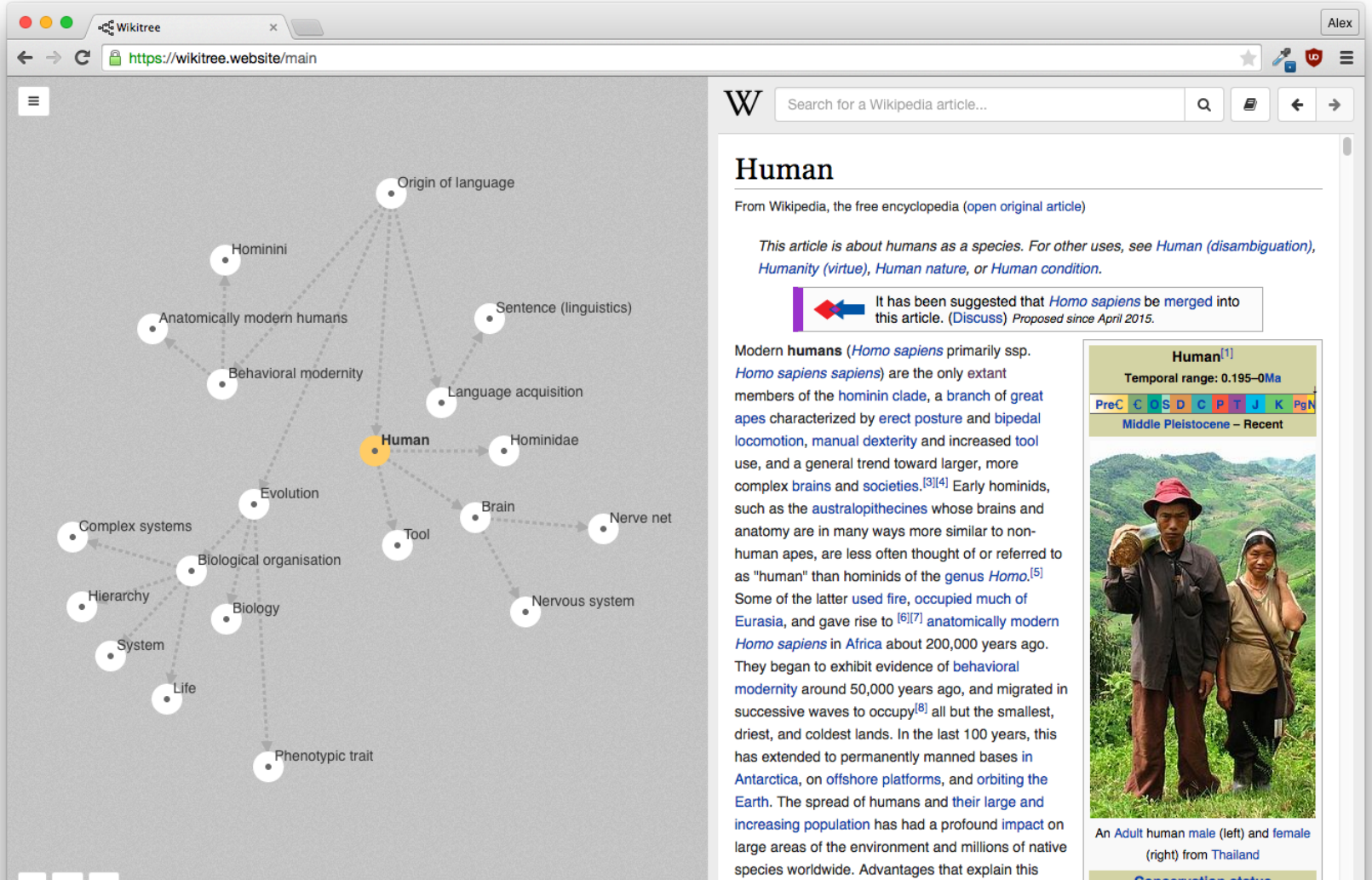
The directed graph can allow a user's navigation trail to loop back on itself



Articles may also share links to each other



Users can rearrange nodes and “pin” them, fixing nodes to a (x, y) location



Next steps

- Use **text analysis** to determine the relevancy of a current article's links against all articles in the graph, and offer the user **"suggested" nodes**
- Add more "appy" features such as server-side **user accounts** and **full Wikipedia search**
- With server-side user data we could **gather usage patterns** and use it to build **suggestion paths**

Questions

- Would you use Wikitree as part of your normal routine?
  - Do you already make average (or extensive) use of Wikipedia?
- What would help bring Wikitree into your life?
  - Other data sets, missing features?
- What features would you like to see in Wikitree?
  - Ex: Article text highlights & annotations
  - Ex: Drag-to-select multiple nodes (move, un/pin, delete)
- What information encodings could be useful?
  - Ex: Size of node proportional to length of article
  - Ex: Cluster nodes by topic/category/keywords
  - Ex: Set node value/opacity by relevancy