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
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 GAMES 110  Lee Hammock  Instructor  AU14
UW Bothell GAMES 210  Jay Schneider  Instructor  WI15
UW Bothell GAMES 220 A  Daniel Smith  Instructor  SP14
Executive Master Bus Admin Group GEMBA 514 G  Patrick Bettin  Lecturer  SP14
Genome Sciences GENOME 351 A  Leo Pellanck  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  Maitreyya 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**

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

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
<table>
<thead>
<tr>
<th>Instructor</th>
<th>Overall</th>
<th>Content</th>
<th>Amount Learned</th>
<th>Teaching</th>
<th>Grading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allison Obourn</td>
<td>3.97</td>
<td>4.16</td>
<td>3.21</td>
<td>3.73</td>
<td>4</td>
</tr>
<tr>
<td>Allison Obourn</td>
<td>4.3</td>
<td>4.4</td>
<td>4.49</td>
<td>4.26</td>
<td>4.06</td>
</tr>
<tr>
<td>Benson Limketkai</td>
<td>3.95</td>
<td>3.99</td>
<td>4.18</td>
<td>4.16</td>
<td>3.75</td>
</tr>
<tr>
<td>Helene Martin</td>
<td>4.66</td>
<td>4.62</td>
<td>4.87</td>
<td>4.72</td>
<td>4.28</td>
</tr>
<tr>
<td>Helene Martin</td>
<td>4.05</td>
<td>4.95</td>
<td>4.81</td>
<td>4.94</td>
<td>4.68</td>
</tr>
<tr>
<td>Helene Martin</td>
<td>4.84</td>
<td>4.84</td>
<td>4.85</td>
<td>4.85</td>
<td>4.67</td>
</tr>
<tr>
<td>Scott Petersen</td>
<td>3</td>
<td>3.33</td>
<td>3.25</td>
<td>3</td>
<td>2.13</td>
</tr>
<tr>
<td>Stuart Reges</td>
<td>4.63</td>
<td>4.56</td>
<td>4.64</td>
<td>4.74</td>
<td>4.17</td>
</tr>
<tr>
<td>Stuart Reges</td>
<td>4.7</td>
<td>4.71</td>
<td>4.78</td>
<td>4.84</td>
<td>4.45</td>
</tr>
<tr>
<td>Tyler Rigsby</td>
<td>4.3</td>
<td>4.37</td>
<td>4.64</td>
<td>4.66</td>
<td>4.61</td>
</tr>
<tr>
<td>Whitaker Brand</td>
<td>4.16</td>
<td>4.27</td>
<td>4.35</td>
<td>4.32</td>
<td>3.84</td>
</tr>
<tr>
<td>William Paul Osborne</td>
<td>4.28</td>
<td>4.21</td>
<td>4.59</td>
<td>4.25</td>
<td>4.2</td>
</tr>
<tr>
<td>Zorah Fung</td>
<td>4.29</td>
<td>4.31</td>
<td>4.61</td>
<td>4.59</td>
<td>4.06</td>
</tr>
</tbody>
</table>
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

CSE 142: Computer Programming I

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

<table>
<thead>
<tr>
<th>Instructor</th>
<th>Overall</th>
<th>Content</th>
<th>Amount Learned</th>
<th>Grading</th>
<th>Teaching</th>
<th>% Filled Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allison Obourn</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
</tr>
<tr>
<td>Benson Limbodi</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
</tr>
<tr>
<td>Helene McMorris</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
</tr>
<tr>
<td>Scott Peterson</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
</tr>
<tr>
<td>Stuart Rogers</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
</tr>
<tr>
<td>Tyler Rigby</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
</tr>
<tr>
<td>Whitaker Brand</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
</tr>
<tr>
<td>William Osbourne</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>~</td>
</tr>
</tbody>
</table>
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

• showed size of vocabulary for various artists and groups

• incorporated regional data for artists, but displayed with no geographical context
RELATED WORK:

COLLABORATION NETWORK

- 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

• 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
Storyboard: regional view

1996
NYC

Storyboard: collaboration view

- Beastie Boys
- Kool Herc

- Track Name, 06-15-1975
- Track Name, 06-15-1975
- Track Name, 06-15-1975
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

Current Progress

Name: ODESSA
Zip Code: 99159
Utility: Avista
Carbon Impact: 0.83
Current Progress

The Overall Picture

The View by Income

KWH per year

0.0

5.0k

10k

20k

25k

30k

≥30k

Electricity Usage

Sample Mean

- Sampling weight: 0.20%
- Average KWH: 17k
- Household income: $35,000 to $39,999
- Number of bedrooms: 5
- Own the house?: Rent
- Urban area?: Urban

Chart Key

Color shows electricity usage

Size shows sampling weight

- 0.50%
- 1.00%
- 2.00%
- 3.00%
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
CO2 Emission across Countries from 1960 to 2008

- 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
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
• Vaccines as a whole are ineffective for HIV.

• However, new techniques show they can be effective on certain strains
  o 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

SIEVE ANALYSIS GRAPHICS

Alignment Viewers
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
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

Mutation!
Challenge
Design

Subtitle of this Chart for Entrophy Analysis

Detail of Selected AA

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

Gene Composition

Sample 1

A  
B  
C  
D

Sample 2

2A 4B 5C 4D

4A 3B 4C 1D
Design plan: Start with traditional visualizations of these data

Lactobacillus 97%

Transporters 10%
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.

<table>
<thead>
<tr>
<th>Types of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial orientation and diameter of cells</td>
</tr>
<tr>
<td>Cell types and lineages</td>
</tr>
<tr>
<td>Gene expression for ~250 genes in each cell</td>
</tr>
</tbody>
</table>
General layout of visualization

- Playback
  - slow

- Highlighting
  - Cell lineage or cell type
  - Hide non-highlighted

- Expression
  - Gene name or function

Diagram:
- 3D coordinate system: x, y, z
- Gene expression visualization
- Lineage tree nodes
- Gradient color scale: 0% to 100%
- Gene names: Gene 1, Gene 2, Gene 3, Gene 4, Gene 5, Gene 6
- Hide non-highlighted option
- Gene name or function input
- Slow playback option

Helpful features:
- About
- Show me how
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
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.

Browse By Diagnosis
Explore counts of patient lives by different diagnosis categories, driven by ICD-9 Codes.

Data Dictionary
Explore our data dictionary.

Patient Demographics

<table>
<thead>
<tr>
<th>Gender</th>
<th>Ages 0-12</th>
<th>Ages 13-18</th>
<th>Ages 19-25</th>
<th>Ages 26-45</th>
<th>Ages 46-64</th>
<th>Ages 65 plus</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>5,330</td>
<td>2,346</td>
<td>4,969</td>
<td>9,916</td>
<td>7,615</td>
<td>4,570</td>
<td>34,746</td>
</tr>
<tr>
<td>Male</td>
<td>5,693</td>
<td>2,161</td>
<td>2,230</td>
<td>6,412</td>
<td>5,875</td>
<td>3,119</td>
<td>25,280</td>
</tr>
<tr>
<td>Unknown</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>8</td>
</tr>
</tbody>
</table>

Contact Us for More Information
Comparison

Clinic F

Clinic G

- 42%
- 8%
- 29%
- 21%
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

Find the anomalies

Cordeiro et. al, 2015
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, business 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 center, career curriculum, and a post secondary education resource booklet.
- To support community care facilities for OVCs.

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

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

This organisation’s info updated on: Thu, Jun 18, 2009. Please disregard the date below.
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?
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
<table>
<thead>
<tr>
<th>Leaderboard</th>
<th>Affirm</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Name</td>
<td>Told Him</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Trust Move

Deny
Idea #2: Spaghetti

AFPmedia: Reports that the AFP is conducting search warrants in the Sydney suburb of Lakemba are incorrect.
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

The circle represents an individual tweet. It coalesces from the creation date/thumbnail for a second at its

# of followers

it then implodes on itself

The re-tweet firework emerges. Each line represents a re-tweet, the length of the line represents the number of followers of the re-tweeted.

As time passes and more fireworks explode, past fireworks fade.

followers

re-tweet

color = code
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
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
Data Resource

- Instagram API:
  - [https://instagram.com/developer/endpoints/](https://instagram.com/developer/endpoints/)

- Scripting Data:
  - [https://github.com/rldw/instastats/](https://github.com/rldw/instastats/)
Filters - Locations

Popular Filter on each region.

- Hudson
- 1977
- Walden
- Kelvin

Different Color
Different Filters.

Click each region. Filters rank details show up.
Filters – Tags and Likes

- Search #
- Hash tag Search
- Search Results
- Different Filters
- Kelvin
- Hudson
- Toaster
- Hefe
- Sierra
- Inkwell
- Sutro
- Darker Color
- More Likes
- Larger Region
- More Uses
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
Related Work: Paper-Level Citation 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