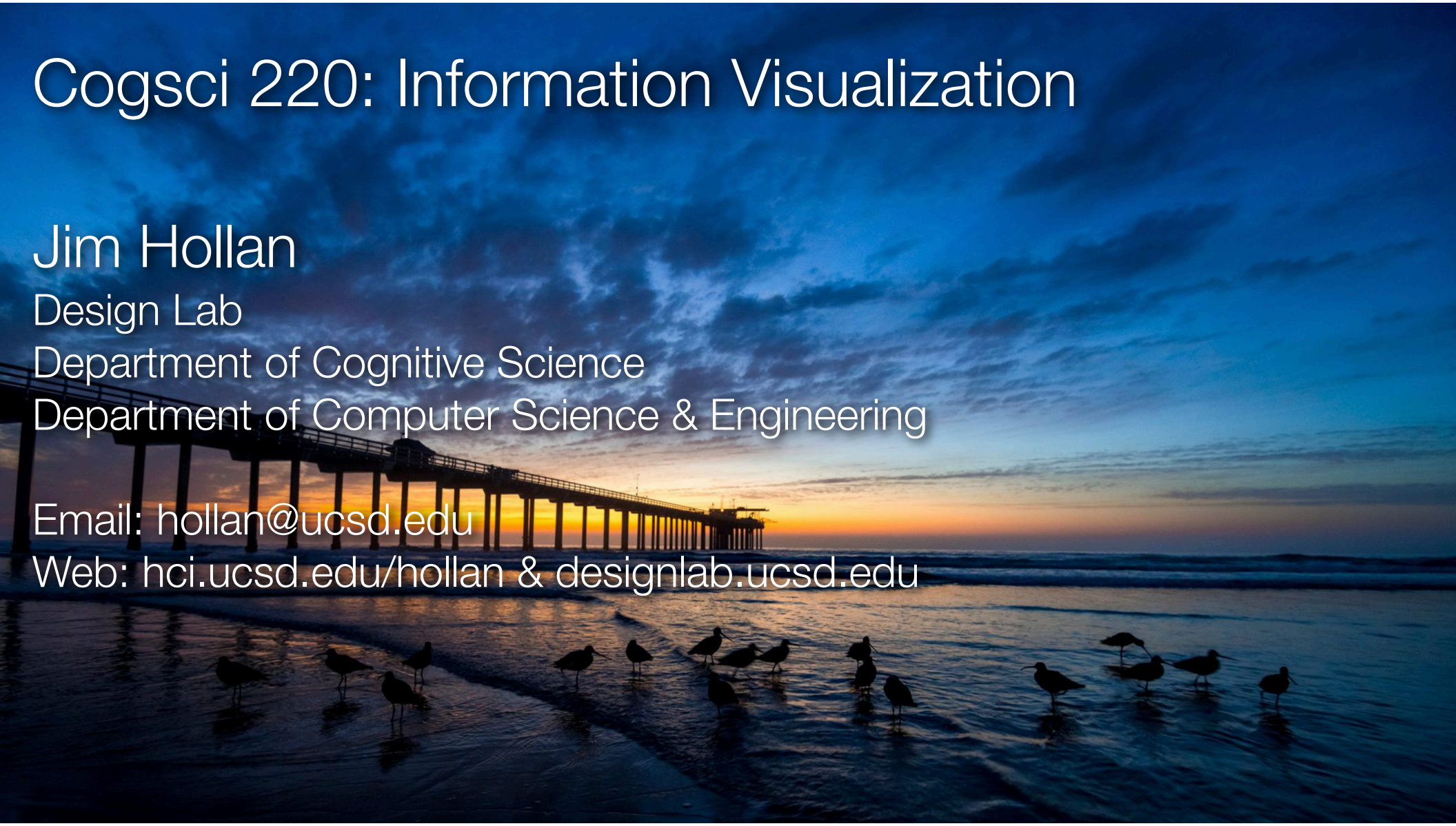


Cogsci 220: Information Visualization

Jim Hollan
Design Lab
Department of Cognitive Science
Department of Computer Science & Engineering

Email: hollan@ucsd.edu
Web: hci.ucsd.edu/hollan & designlab.ucsd.edu

Cogsci 220: Information Visualization

Welcome to week five.

Hope everyone continues to be safe in these challenging times.

Today:

I sent a link to the interview with Arvind Satyanarayan.

I need to end early. Main goal today is to briefly review each group's project.

Starting next week we will focus on one group and they will lead the discussion of their selected readings. We will also have very brief status updates/issues from other groups.

Group Schedule

Feb 9: Garrett, Sloan, and Xuhan

Feb 16: Chen, Martin, Naba, and Tommy

Feb 23: Kendall, Martha, and Michael

March 2: Casey, Lu, and Mary Anne

March 9: Brinley, Isabel, Kristin, and Talia

Provide overview of background literature related to your project.

Summarize other especially relevant papers you came across.

Lead discussion of papers we all read.

Good to be critical but goal is interesting and useful takeaways.

Discuss issues you are facing in your project.

Quick updates/issues from other projects.

Feel free to exchange weeks if both groups agree. Must be set by Thursday this week.

Project Journal

Beginning this week you should create a personal project journal (a Google Doc). I will send a Google Form for you to provide a link.

Make at least a weekly entry by each Sunday evening. I encourage you to add more regularly. It should be informal and cover what you did during the current week, plans for next week, and your thinking about your project. Make it a useful activity rather than just a requirement.

One suggestion is to try voice input.

Another is to try free form sketching and then just take a pic.

Or some combination of techniques. Explore.

At the end of the quarter, you will add a summary assessment of your and your teammates contributions.

Garrett, Sloan, and Xuhan

We interested in exploring ways of representing activity histories particularly focusing on a more intuitive organization of tabs. The core idea being to cluster tabs into related ideas or categories. And perhaps automating that feature when examining a browser history.



Project Goals

As more of our work shifts online, we have more and more tabs open with no good way to organize them.We feel that it's essential to create a more natural and information dense structure to provide this organization. The best way to accomplish that is to convert a 1D list of tabs into a 2D undirected graph with nodes representing webpages and clustered around tasks or related ideas. This will enable a user to rapidly scan they're existing tabs to find what they want.

Questions

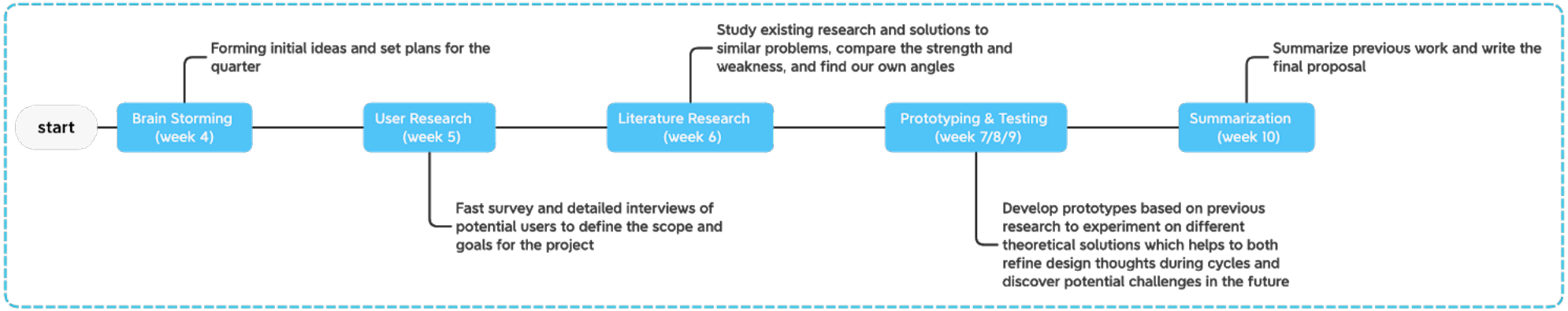
The questions we will be exploring are how to best create this graph. More explicitly how to tune certain parameters like how large nodes should be, how far they should be from other nodes, what color they should be, what opacity, etc. Each of these parameters encode information about the webpage, like how frequently that page is visited, how long the user spends on the page, how long it's been since they last visited etc.

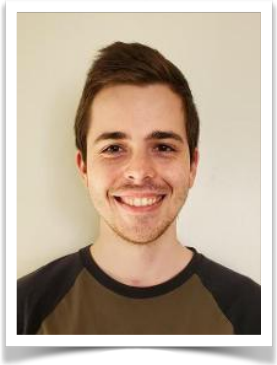
We also want to explore larger questions, like how much of the organization should be manual vs automated? Should there be a clear distinction between "open" tabs and the users entire web browsing history. When should a page be part of a cluster and when should it branch off into its own cluster.

We plan to find answers to these questions through user surveys and interviews.

Garrett, Sloan, and Xuhan

Timeline





Garrett, Sloan, and Xuhan

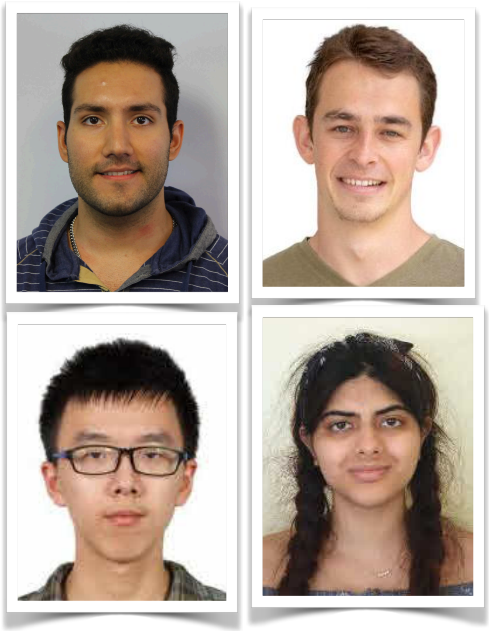
Tues 2/9 **Organizing Tabs (Garrett, Sloan, and Xhun)** Slides

- REQUIRED Brehmer, M., Carpendale, S., Lee, B., & Tory, M. Pre-design empiricism for information visualization: Scenarios, methods, and challenges, 2014. In Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization (pp. 147-151).
  This paper advocates for pre-design empirical work so that visualization researchers can design with a deep understanding of the user, their needs, and their environment. It then provides four examples of pre-design work that illustrate different pre-design methods that provided necessary insights for making sure you have the right design, a design that is appropriate for the given situation.
- REQUIRED Koshman, S. (2006). Visualization-based information retrieval on the web. Library & Information Science Research, 28(2), 192-207.
  This paper reviews current existing visualization tools that were designed to graphically represent related documents found during web searches. While this isn't exactly what we are trying to do, it still shows what design features are possible and provides recommendations and critiques.
- REQUIRED Lam, H., Bertini, E., Isenberg, P., Plaisant, C., & Carpendale, S. (2011). Seven guiding scenarios for information visualization evaluation. University of Calgary Techreport, 2011-992-04.
  This paper should help everyone with writing their proposals. It reviews seven types of evaluations in information visualization: evaluating visual data analysis and reasoning, evaluating user performance, evaluating user experience, evaluating environments and work practices, evaluating communication through visualization, automated evaluation of visualizations, and evaluating collaborative data analysis . They review 800 visualization publications and describe current best practices for each of these purposes, provide advice on what to evaluate for a given project, and provide examples from the literature.

Chen, Martin, Naba, and Tommy

Describe and motivate your project focus.
Why is it interesting and important?

Institutional education podcasts (e.g., lecture recordings) are a popular information resource in higher education. Instructors can easily record and broadcast audio content of their lectures (without the need for intensive technology infrastructure) and students can later listen at their time and place of convenience. Students reviewing these audio resources can encounter difficulties in finding specific segments, due to the number, length, and nature of these recording files. The following scenario showcases this problem:



Alex -- an Undergraduate student taking an introductory calculus class -- prepares for the final exam. As he reviews the content, he struggles with the Taylor series and the related trigonometric functions. Alex opens an online folder with all of the audio lecture recordings which are organized solely by date and don't present any information about the content. He opens one lecture and jumps around to see if he gets lucky and finds the part where his professor discusses representing sine and cosine functions.

After a minute, Alex gives up and tries another lecture. Then another and another until Alex knows he must have already passed it. Alex then goes back to the first lecture and skips the beginning where people are getting settled, and listens to the rest of the lecture at 2x speed, but doesn't find the content he is looking for. Moving onto the next lecture, Alex follows the same procedure and finally finds what he was looking for! He then slows the lecture down to 1x speed and listens to the content.

As demonstrated above, reviewing lecture recordings can be challenging: (1) each class contains many recordings in a semester, (2) these recordings are typically long and monotone, and (3) navigating audio files is time-based (i.e., the scrub bar) while students search for specific content.

Chen, Martin, Naba, and Tommy

Detail the central question(s) you are addressing. Not only what the questions are but how you plan to address them.

In this project, we aim to explore effective and efficient meta-data visualization techniques for helping students navigate, explore and search among the lengthy audio podcasts.

Our key research questions include:

What are the barriers for students to navigate audio podcasts?

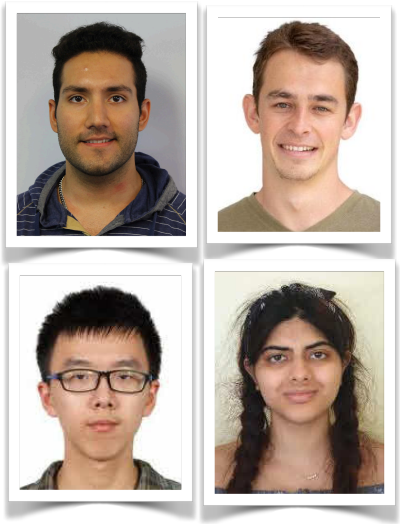
To address this, we will conduct a need-finding study. We will have a small population of conveniently-sampled subjects watch two information-dense lectures on human hormone signaling pathways. After watching the videos, participants will be given a series of questions to answer. As they search for the answers to the questions, they will be asked to think out loud. Finally, they will be directly asked to imagine a tool that would have helped them with the process.

This will help us understand users' navigational cues in educational audio content.

What kind of meta-data visualization techniques would be most efficient? In which situation and why?

Inspired by the OSI model (the network stack), we will explore the meta-data visualization techniques for audio podcast in terms of 3 layers:

1. The signal layer: the spectrogram, pitch, rate of speech etc.;
2. The semantic layer: the transcribed text and highlights of aggregated topics;
3. The user layer: the annotations and browsing activities of a specific user;



With a proof-of-concept prototype, we will observe and evaluate how students navigate audio podcasts with the three types of meta-data visualization techniques.

Although our work focuses on the audio-only podcast, we believe our contribution can be integrated into online class platforms (e.g., Coursera) that blend audio and video, LMS (e.g., Panapto), as well as general podcasts (e.g., https://podcasts.google.com).

Chen, Martin, Naba, and Tommy

Provide a timeline for the rest of the quarter.
Convince us (and yourselves) your plan is doable.
Be as specific as you can. Of course your plan will change.

W5- Quick Need Finding around navigating audio data

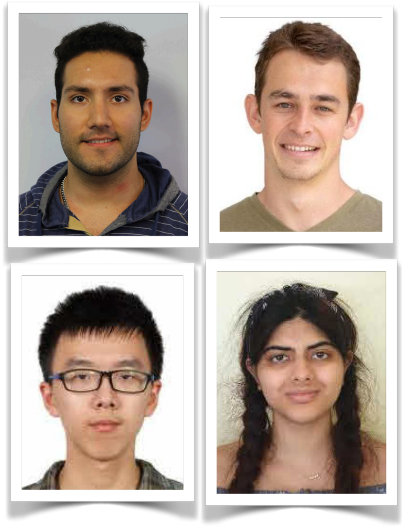
W6- Analysis of need-finding results proposal of three visualizations. We will make sketches from the collected data after the need-finding results. The sketches will inform the preliminary design of our visualization interface.

W7- Implement / WOz / high fidelity prototype of behavior. The goal is to get a better understanding of the benefits and limitations of our proposed solution to help guide the proposal writing process, such as how users interact with the visualizations.

W8- Write Related Work, Introduction, Intellectual Merit

W9- Write Proposed Evaluation, Broader Impact, and any other discussion

W10- Wrap up



Chen, Martin, Naba, and Tommy

Week 7

Tues 2/16 **Podcast Visualization (Chen, Matin, Naba, and Tommy)** Slides

Visualizing Multi-Participant Meetings

- REQUIRED Yang Shi, Chris Bryan,Sridatt Bhamidipati, Ying Zhao, Yaoxue Zhang, and Kwan-Liu Ma. MeetingVis: Visual Narratives to Assist in Recalling Meeting Context and Content. IEEE Transaction on Visualization AND Computer Graphics, 1918-1929.
- REQUIRED Senthil Chandrasegaran, Chris Bryan, Hidekazu Shidara, Tungyen Chuang, Kwan-Liu Ma. TalkTraces: Real-Time Capture and Visualization of Verbal Content in Meetings. CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.

Visualization Techniques for Videos

- REQUIRED Amy Pavel, Colorado Reed, Björn Hartmann, and Maneesh Agrawala. Video digests: a browsable, skimmable format for informational lecture videos. UIST '14: Proceedings of the 27th annual ACM symposium on User interface software and technology, 573-582, 2014.
- REQUIRED Justin Matejka, Tovi Grossman, and George Fitzmaurice. Video lens: rapid playback and exploration of large video collections and associated metadata. UIST14: Proceedings of the 27th annual ACM symposium on User interface software and technology.

Visualization Techniques for Storytelling Podcasts

- REQUIRED Haijun Xia, Jennifer Jacobs, and Maneesh Agrawala. Crosscast: Adding Visuals to Audio Travel Podcasts. UIST20: Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology.

Finding Video Scenes via Closed Captions

- REQUIRED Amy Pavel, Dan B. Goldman, Björn Hartmann, and Maneesh Agrawala. SceneSkim: Searching and Browsing Movies Using Synchronized Captions, Scripts and Plot Summaries. UIST15: Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, 181-190, 2015.

Kendall, Martha, and Michael

Our project focus is the Representation and Behavior of Information Entities topic. We decided to specialize to make it customized (based on a few short questions to understand a user's data-viewing background, motivation for viewing data representations, etc) representation of data (whether it be summarized or cleaned data or raw CSV files).

By using these to change representation of the Ordata, we hope to make the data more accessible and therefore more widely usable. More accessible data allows individuals in the community to be more informed citizens with independent access to data.

Kendall, Martha, and Michael

How to make public data more easily interpretable and accessible?

We plan to address this question by:

Viewing the use case of San Diego Association of Governments (SANDAG) collected or provided datasets and which local residents view these and how they are used. Datasets are currently presented in this form: https://rdw.sandag.org/Account/gisdtview?dir=Health

Adding visualizations where none exist.

Visualizing publicly available data tailored to individual users. We will focus on users' existing skill set, interests, background, and need for data in order to determine the best representation of the data.

Kendall, Martha, and Michael

Timeline

Week 5: Literature review of data narration and visualizations, competitive analysis of current data breakdown online tools, and peeling through SANDAG website to see if we can make a specific use-case scenario

Week 6: Writing background/motivation/broader impacts sections of proposal.

Week 7: Prototyping and making clickable-interface on Figma and do short pilot testing

Week 8: Develop demo of project on Qualtrics (survey platform)

Week 9: Writing research approach/intellectual merit

Week 10: Writing research plan/evaluation

Final: Presentation! Woo!

Kendall, Martha, and Michael

Papers (these are all the papers that we thought are relevant, but we will narrow the list down before we present in class)**:**

1. **Datafication and data fiction: Narrating data and narrating with data:** https://journals.sagepub.com/doi/full/10.1177/2053951718784083

2. **Narrative hermeneutics: in search of narrative data**

    1. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1471-6712.2010.00846.x?casa_token=q13pGxpKh20AAAAA%3Ar-quU1Phno11-AE_Njgp_lCsA1vOf_nN86itmAusiR7fPsadXy-QMSyoZrP-rkfRxSbH5NC0z1om97c

3. **Narrative Visualization: Telling Stories with Data**

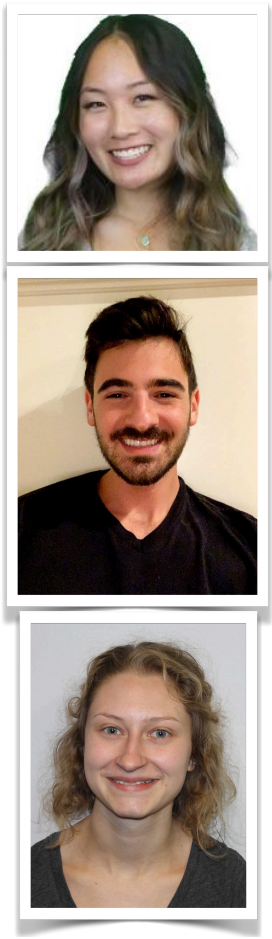
    1. https://ieeexplore.ieee.org/abstract/document/5613452?casa_token=cL7p5NsXUvEAAAAA:QXwUP7kKYzV9LJyWyn4qMD6D7bPmQo7h7LKrSs61bqrK3rBdBaBMIh9PDZZY_wGOgUAolJ-yrQ

4. **A survey on automatic infographics and visualization** recommendations:https://www.sciencedirect.com/science/article/pii/S2468502X20300292

5. **LinkDaViz – Automatic Binding of Linked Data to Visualizations:** https://files.ifi.uzh.ch/ddis/iswc_archive/iswc/ab/iswc2015.semanticweb.org/sites/iswc2015.semanticweb.org/files/93660129.pdf (more technical than we need but worth looking at in terms of visualization recommendations since they implement it as a webapp)

6. https://onlinelibrary.wiley.com/doi/full/10.1111/cgf.13195?casa_token=BHwMbyMhfnoAAAAA%3AUUlaaXk0dB1Duuss4azZGaVgKkBsoUzbGn8L74CMQwgE-P-2qGwqCIJIpfGOxdPtNa2uyv0s6PY_jO8

7. https://ieeexplore.ieee.org/abstract/document/5613452?casa_token=xxlj3kMPclEAAAAA:FJxSY6_aWHxYgf5cmfGBwmjfjRwskSoqYY6tFQZGAFcG1tq81SP6zp3BQemLzgS0U7rweDGYeQ



Casey, Lu, and Mary Anne

We are looking to explore personal information spaces from the standpoint of user privacy. It is now common knowledge that our online and offline behavior is being monitored and processed for the purposes of ad targeting and even investment banking. So much of our data is being monitored in so many mystifying ways, it is easy to feel helpless and apathetic: “either be surveilled or be a hermit”. In this project, we want to apply information visualization fundamentals towards making our personal surveilled information manageable. We want to make it intuitive and easy to ask questions like “Who monitors me? How often? What kinds of data are recorded? How is my data privacy connected to that of my friends and family?”. By making our personal surveilled information more accessible, we hope to impart a newfound sense of efficacy in both tracking the trackers and in demanding new privacy policy.

Casey, Lu, and Mary Anne

Detail the central question(s) you are addressing. Not only what the questions are but how you plan to address them.

How can interactive data visualizations help shape users' understanding of and attitudes toward privacy-related issues? We can address this question by developing visualizations and conducting interviews with users. We won't have time to do user studies, but we can discuss potential user studies in our proposal.

Which personal data types are both prevalent in data surveillance and easy to visualize (e.g. location data, browsing data)? We can download data logs from Google, weather apps or browser and then categorize their types. We can visualize different types of data and then invite users to make decisions on which part of data is sensitive to data privacy.

How do we demonstrate the intersections of who collects what data and when? Ideally, we can additionally demonstrate why that data might be private (e.g. browsing history).

How can we visually connect individual experiences of invasions of privacy with the larger collective landscape of mass data collection? As a starting point, we can look at the animated visualizations used by the New York Times to show the extent of smartphone-based location tracking. They use animation to start with a high-level view of all the data collected from multiple people throughout Central Park and then zoom in to show the datapoints collected from a single individual.

Casey, Lu, and Mary Anne

Week 5: Brianstorm the research questions. Conduct literature review. Find potential data sources related to the usage of App, location data or social data. Write background, research question of the proposal.

Week 6: Find or Scrape data sources. Sketch and wireframe 3 versions of data visualization.

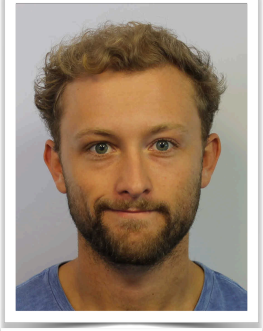
Week 7: Prototype the visualization using visualization tools, such as D3.js or VegaLite.

Week 8: Iterate on the prototype using visualization tools, such as D3.js or VegaLite.

Week 9: Iterate on the prototype. Conduct a small pilot study. Write system design session and the plan for user study session of the proposal.

Week 10: Finalize research proposal.

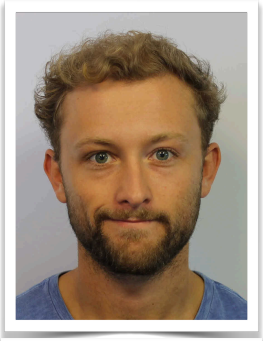
Final: Present our proposal.

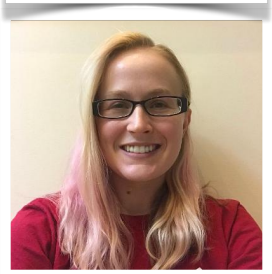


Casey, Lu, and Mary Anne

- Max G. Van Kleek, Reuben Daniel Binns, Jun Zhao, Adam Slack, Sauyon Lee, Dean Ottewell, Nigel R Shadbolt. X-Ray Refine: Supporting the Exploration and Refinement of Information Exposure Resulting from Smartphone Apps.

- Ben Weinshel, Miranda Wei, Mainack Mondal, Euirim Choi, Shawn Shan, Claire Dolin, Michelle L. Mazurek, Blase Ur. Oh, the Places You've Been! User Reactions to Longitudinal Transparency About Third-Party Web Tracking and Inferencing





Brinley, Isabel, Kristin, and Talia

Our group plans to implement a tagging system and potentially a hover-over feature that will help visualize and organize activities within teacher.desmos.

Research questions include:

1) Address the challenges of searching for and representing research based activities in teacher.desmos.

2) Propose design changes that work to correct these challenges.

3) Illustrate some features in a vega-lite prototype.

Brinley, Isabel, Kristin, and Talia

Week 4: Lit review; find relevant papers & reading

mathy/visualization

Week 5: Write overview/intellectual merit/broader impact (aka abstract)

Week 6: Write introduction scenario, challenges of developing

Week 7: Write research agenda, approach, similar work

Week 8: Write research plan, why worth funding; implementation in Vega Lite

Week 9: Edit and wrap up; implementation in Vega Lite

Week 10: submit

Brinley, Isabel, Kristin, and Talia

We have begun a literature review, here are some the papers we think will be most interesting to the class:

- The Challenge of Information Visualization Evaluation https://dl.acm.org/doi/pdf/10.1145/989863.989880?casa_token=bLgEKfRXoJEAAAAA:QE7v6OKeh0QGpY3oF3qAvACD2iyKUaJz0AFpbMgArzknNclVRXPc8v5w5Kg_INQptHwePe1HBVGb

- Increased Participation and Conversation Using Networked Devices https://pdfs.semanticscholar.org/c3c8/f6a578fe8e64972792060af2c72233e7875d.pdf

- The Role of Visual Representation in the Learning of Mathematics http://cimm.ucr.ac.cr/ojs/index.php/eudoxus/article/viewFile/154/152

- Decision making and informing of information visualization for teachers and students: https://www-learntechlib-org.libproxy.sdsu.edu/p/32722/