Annual Center/Lab Review

2023-2024

NAME OF CENTER/LAB: **Computational Social Science Lab (CSSLab)**
CONTACT IN CASE OF QUESTIONS: **CSSLab@seas.upenn.edu**

RESEARCH UPDATES:
The CSSLab is excited to announce new projects in its four research areas: PennMAP, Group Dynamics, Human Mobility (COVID-Philly), and Common Sense. PennMAP was launched to increase media transparency and accountability by creating content that engages those interested in media bias. It also hosts the "Living Journal," a set of dashboards based on PennMAP research displaying data that updates in real-time, providing users with live insights into the information ecosystem. The first installment of the Living Journal was the YouTube Politics Dashboard, a set of interactive visualizations highlighting the extent of political content consumption on YouTube and the increase in user engagement with far-right channels on the platform. Currently there are two dashboards in progress, which will be discussed in further detail under the "**NEW PROJECTS**" section of this report, which are "Mapping the (Political) Information Ecosystem" and the "Media Bias Detector." A blog page will also be added to the PennMAP section of the CSSLab website, where you can find posts summarizing key takeaways of each dashboard and short posts relevant to media bias. Group Dynamics involves designing high-throughput experiments or large-scale studies that produce replicable and generalizable data to better integrate with existing experiments. There are two main projects involved: Group Deliberation and Nudge Cartography. Group Deliberation refers to small groups that get together to reach a certain decision or achieve a specific goal and Nudge Cartography is the mapping of experiments on "nudges," or subtle ways of influencing people's behavior and decisions without restricting their freedom of choice. A Nudge Cartography map would resemble a living database of commensurable experiments, helping researchers understand the factors influencing

decision-making in specific contexts. In COVID-Philly, the team created a set of dashboards based on public data to visualize Philly's COVID data and patterns of movement in the city. So far, there are three dashboards: "Find vaccine locations for each CT (local neighborhood)," "Explore the places people visit," and "Understand how communities move." Now the focus has expanded from COVID research to improving epidemiological modeling. Lastly, Common Sense revolves around one question: "How "common" is common sense?" A term usually seen as more of an abstraction, the team sought to develop a framework to measure common sense by creating a repository of statements considered common sense and having participants rate them to yield more insights on the nature of common sense.

NEW PROJECTS

Media Bias Detector Dashboards- This is our newest dashboard project, also funded by Richard J Mack (W'89). The bias detector will feature a series of interactive data visualizations that will quantify and surface media bias at the level of the information ecosystem. Our lab has partnered with *Polygraph*, a data visualization studio specializing in visual storytelling; their team will be developing the visualizations and narrative for users using research and data that we provide to them. On the CSSLab side, we have engineers working on the project's back end, and several PhD students and research assistants are making contributions. The final product will be a website hosting dashboards that are centered on what is happening in the news (facts), who is quoted in the media, a glimpse into authors and their opinions, and a publisher database that displays information about each publisher including their fact/opinion ratio and political lean.

News Consumption Dashboard (PennMAP)- The CSSLab is putting the final touches to its second dashboard project: "Mapping the (Political) Information Ecosystems." It consists of four data visualizations highlighting who is watching the news, what they are watching, and how many Americans are living in echo chambers? These data visualizations are based on the findings of two of the CSSLab's previous publications: "Evaluating the fake news problem at the scale of the information Ecosystem" and "Quantifying partisan news diets in Web and TV audiences," authored by Baird Howland, who is a Ph.D. student at the Annenberg School for Communication. Each visualization contains interactive tools to help users look at US news consumption trends over time and by state. The first visualization allows users to compare news consumption among different mediums: TV, desktop, and mobile (phones and tablets). The second and third ones show national and state echo chamber numbers, including the most viewed television programs. The last chart shows changes in news audiences over the last six years, and the main takeaways from this were that the news audience is shrinking but becoming more polarized as those still watching TV regularly tend to watch programs with a left or right-partisan leaning. This suggests that some Americans find themselves in echo chambers because the media is biased. TV news makes up the majority of news consumption, insinuating that the pervasiveness of fake news is minuscule to that of cable news, as some people consume fake news but these people are rare and watch very little online news. The team behind this project includes Baird Howland, Yuxuan Zhang, Delphine Gardiner, and Felippe Rodrigues (front-end developer), and the anticipated launch date for this dashboard is April 19, 2024.

Quantifying Partisanship in TV News (PennMAP)- Second-year Ph.D. student, Upasana Dutta, has started this new project aimed at analyzing news shows of various political leanings and determining whether large language models can accurately identify biases within these programs. The dataset is comprised of TV news transcripts from 24 different shows, with 12 left-leaning programs and 12 right-leaning programs. Before using Language Learning Models (LMMs) to classify these shows, all shows were first annotated by humans who manually went through transcripts and labeled them as "Liberal," "Neutral," or "Conservative." So far, it has been found that GPT-4 surpasses models in closeness to human judgment when it comes to identifying shows compared to the other models employed, LLaMa-7B and LLaMa-13B. In addition, episodes from five different shows were taken, and Dutta had GPT -4 assign these episodes partisan scores (from 0 to 100). GPT-4 was able to discern liberal, conservative, and neutral shows effectively. These analyses are an ongoing project that will eventually result in a growing dataset and will go through additional testing in order to corroborate results.

Investigating the Impact of Media Bias on News Readers (PennMAP)- Amir Tohidi examines bias in media and how that impacts readers, as the selection and framing of journalists on both sides of the political spectrum creates different "realities" for their respective readers as they come out with different perspectives of the same news topic. Selection refers to which topics and facts journalists choose to include or omit from their work, and framing is the way in which journalists present these topics and facts. GPT-4 was used to extract facts from news articles and then write new articles based on these facts for specific events (such as a particular debate) with controlled amounts of bias, such as positive and negative biases. Then, human participants were asked questions about these generated articles so that Tohidi could better understand whether articles with a subtle bias could shape public opinion and whether generated articles with a negative bias tended to have a more significant impact on a reader's opinion on news topics.

Do Viewers Want Nationalized, Partisan Commentary in their Local Television Newscasts? (PennMAP)- Local television news has garnered more attention in recent years, as Sinclair Broadcast News (which owns close to 200 local news stations) was found to have included "must-run" segments, or pre-written scripts that contained right-wing discourse. This was a cause of concern as local news is regarded as less biased compared to cable news, but there is little research on how local news impacts consumers, which is what led to Sam Wolken's interest in the infiltration of partisan cable news content into local news stations. Wolken has two questions in mind for his new research: 1) whether or not such partisan content resulted in changes in local news' audience sizes, and 2) does this content affect the demographic makeup of those who watch local news? Wolken combines two datasets: local TV news transcripts from 2012 to 2022 and television consumption data. Both sets of information allow Wolken to determine if news consumption habits changed after a viewer watched the "must-run segments" that were broadcasted by Sinclair. Using regression models, it was found that exposure to these Sinclair segments led to a decrease in news consumption associated with Sinclair local news. Additional research will also look into the demographics of these viewers and their individual media consumption habits.

A research agenda for encouraging prosocial behavior on social media (PennMAP)- Although antisocial behavior (such as bullying or the spread of misinformation) is widely studied, little research exists on its counterpart, prosocial behavior, and what prompts prosocial behavior on social media. Prosocial behavior on social media can be increased by facilitating connections between online communities, allowing for collective problem-solving, and expanding opportunities for philanthropy. But this proves to be a challenging task so more research is necessary to expand prosocial behavior, and Timothy Dorr has laid out different aspects of social media to further research: 1) the sizes of the social networks, 2) platform features and capabilities, 3) platforms that have clear goals and rewards for users that have positive interactions with other users, and 4) the integration of prosocial behavior into business models of social media platforms which would encourage the development of interventions that promote prosocial behavior. The influence of social media is of concern due to its adverse impacts, but now there are more efforts to change this narrative through using social media as a force of positive change. The collaboration of academia and industry and an increase in data sharing would advance knowledge on prosocial behavior and bring people closer together. (Status: Submitted for review)

Uncovering Anti-democratic Norm Justifications (PennMAP)- despite the general U.S. population supporting democracy and its principles, some political elites have demonstrated antidemocratic tendencies; so even when policies are viewed as anti-democratic, people can still consider them democratic if they align with their political views. Unfortunately, previous studies on this phenomena neglect to take political ideology into account, as in real-life scenarios, people will have different opinions toward someone's actions if they know the political affiliations of that person. Timothy Dorr tested this by creating a survey analysis where pro-democracy Americans who supported some of these antidemocratic justifications completed an open-ended questionnaire about their conflicting beliefs. 13,000 responses were collected over 13 months and 300 of these responses were coded manually to create a codebook that could be used to train an LLM to categorize all the responses. The most common responses from this survey were that antidemocratic actions were a necessary evil, denying that the actions were antidemocratic, and blaming the other party (or staying loyal to their political party). In addition, from the methodology side, this research also gives insight into the effectiveness of utilizing LLMs to categorize open-ended responses.

Narrative License in the Social Sciences: An Exploration with GPT-4 (PennMAP)- Honest scientific communication is necessary for advancing knowledge and disseminating findings to the broader community. But sometimes researchers, to gain more recognition, may employ Narrative Licensing, the conscious decision of authors to make claims that are not supported by the evidence found in their papers. This is usually done to make the results seem more intriguing but it leads to exaggerated and less credible scientific findings that deviate from the actual results. Timothy Dorr, Calvin Isch, and Neil Fasching are leading a new project to further examine Narrative Licensing and how it impacts research in academia. Based on 100 academic papers (in economics, politics, sociology, and psychology) that the researchers sampled, Narrative Licensing is a spectrum; some papers stick strictly to the paper's findings, while others will deviate and make claims not mentioned in the paper. Moreover, through the use of LLMs, it is possible to measure the prevalence of Narrative Licensing found in these research papers.

Hypothetical Nudges (Group Dynamics)- In recent years, nudges have gained attention from policymakers, managers, and marketers. However, behavioral experiments are often costly to run and are complex. In these cases, researchers have relied on online surveys for results by asking behavioral questions under hypothetical scenarios, such as asking what participants "would choose" or how they "would behave." Unfortunately, the results obtained from these hypothetical scenarios have appeared less credible than the actual experiment, as evidence shows that people tend to exaggerate their answers. Different techniques for constructing design scenarios should be explored to reflect the real treatment effects of the experiments accurately, thus Linnea Gandhi constructed four styles of hypothetical scenarios: simple-generic, complex-generic, simple-specific, and complex-specific, to find the methods that consistently work under different behavioral domains. The results show that no single design accurately aligns with it, indicating that hypothetical experiments are only helpful in finding the general "direction" of research.

Deliberation Lab (Group Dynamics)- Open science project led by James Houghton, Duncan Watts, and Xinlan (Emily) Hu in collaboration with Netta Weinstein (University of Reading), Dean Knox (Wharton), and Will Shultz (Princeton). Small group deliberation involves carefully discussing a topic to make a collective decision or reach a mutual understanding and three pilot studies on group deliberation are contributing to the project's first fully documented, commensurable dataset. These test cases include "Can good listening be taught (quickly)?", "Is the leader a bad influence?" and "Am I talking too much?" which will incorporate various modes of communication- video, voice, and chat- to facilitate participant interactions in a user-friendly interface. The insights from these studies will play a crucial role in using research to encourage meaningful dialogue among multiple viewpoints; the CSSLab has moved on to the second phase of the grant period, focusing on transitioning the Deliberation Lab to a publicly available platform. The Deliberation Lab was established to understand the underlying processes behind reducing polarization in small-group deliberation. To bring this to fruition, CSSLab is designing a collaborative Researcher Portal, providing the tools for researchers to build high-throughput experiments and carry them out within our open-science platform. The Researcher Portal also allows the social sciences community to build off other's designs while contributing to a growing set of shared experimental data.

Trajectory Data Mining and Debiasing sparse human mobility data (COVID-Philly)- Jorge (Paco) Barreras'recent research has been focused on using GPS data to gain better insights into movement patterns. An important aspect to studying movement is being able to identify when people stop at a given point in order to distinguish mobility behaviors and identify which locations people visited. There are algorithms that detect these stops but there is sparsity or gaps in these datasets that could negatively affect algorithm performance. To test this, Barreras stimulated GPS trajectories and used two stop-detection algorithms, finding that sparsity and parameters can affect stop detection, for example if the parameters are too fine and someone stops at a large location the algorithm would fail to detect that stop. Another aspect that Barreras is studying is bias in GPS data, specifically data-missingness bias, or missing data points that can skew analyses. He introduced deepMrP, a technique which uses both

survey learning and deep learning to fill in the missing data. In comparison to other models used to reduce variance in datasets, deepMrP provided more accurate estimates of mobility data.

Quantifying Bias and Feedback Interventions- Bethany Hsiao is in the early stages of a new project with Duncan Watts and Hamsa Bastani (Wharton OID professor) about fairness and feedback interventions to inform people of their biases and mitigate those biases. One of the challenges that Hsiao wants to address in this project is that in some lab experiments, participants may be overly aware of biases related to race and gender. A proposed alternative is to focus on attractiveness and weight, which participants are usually less aware of in terms of bias, so this study is focused on whether faces affect decisions or not. This will be tested by examining individual biases among MTurk workers to see if they are potentially biased in similar ways and can also be tested using LLMs for the same experiments to see if these models display similar biases. Pilot studies are underway and full-scale experiments will be carried out subsequently.

Do Language Models Have Common Sense?- Spearheaded by Josh Nguyen, this new study highlights that common sense is not well-defined, and even early AI researchers had an interest in common sense. With the shift to more data-driven research methods, there is a new focus on Learning Language Models (LLMs) but we are still unable to quantify the extent in which AI possesses common sense. This led to the question: "Do language models have common sense?" Different LLMS were tested by being given prompts (common sense statements) and asked if they agreed with the prompt and if most people would agree with it. The LLMs' responses were compared to human responses yielded from Whiting and Watt's new study by directly comparing human commonsensicality scores to those of three well-known LLM models: LLaMA-2, Gemini 1.5 Pro, and GPT-4. Human participants scored very high, with an average score of 79.43%. In comparison, the models scored slightly lower (78.59% and 74.98%), displaying lower levels of commonsensicality compared to humans, with larger models demonstrating more robust linguistic capabilities and better at predicting human responses than the other models. Going forward, Josh is seeking to further explore prompt engineering and how it can be used to improve LLM prediction of human thinking as well as studying additional LLMs to determine if common knowledge is shared between these models.

COLLABORATIONS:

- **Polygraph (industry)**- For our newest Media Bias Detector Dashboard Project, we are collaborating with Polygraph, a data journalism team which specializes in visual storytelling and using data to create narratives that users can interact with. Polygraph will create two dashboards based off of CSSLab research as well as a companion editorial which provides supplementary material about the Bias Detector. They are also creating a media kit composed of static and animated assets that we can use for press promotion.
- **Pollfish (industry)**- The CSSLab is also collaborating with pollfish as part of the bias detector project; we are using pollfish to create a tracking survey which will yield news topics which

will have the highest impact; this is to streamline topic selection, as one of the dashboards will include a section on topic coverage and whether some topics appear to be more covered than others.

- **Amazon Web Services (platform)**- Hosts three data repositories: Nielsen, TVEyes, and PeakMetrics. The CSSLab's data is stored and processed on this platform.
- **Nielsen (industry)**- Nielsen is a data analytics company that provides U.S. media consumption data of multiple mediums, including TV and mobile.
- **TVEyes (industry)**- TV data provider that monitors the production of 2,500 TV and radio stations globally in real time and provides closed-caption transcripts.
- **PeakMetrics (industry)**- Provides web publisher data including a continuously updating collection of news websites.
- **Harmony Labs**- Media research lab that provides preprocessed web data and TV data.
- **Safegraph (industry)**- Provides location and mobility data for regions based in the US.
- **Center for Innovation and Sustainability in Local Media (community)**- Has been helping with the "mapping the media landscape" project by providing data on local and regional newspapers across the country, which has helped the news publishers database grow.
- **NewsGuard (industry) / AllSides (industry) / Media Bias / Fact Check (industry)**- All three of these partnerships provide us with political bias and credibility ratings for several thousand news outlets.
- **Empirica (platform)**- Where high-throughput experiments take place, experiment designers can focus on the user interface and how participants interact during the experiment while the platform takes care of data storage and other background tasks.
- **Wharton Behavioral Lab (platform within Penn)**- This lab has provided other resources for carrying out high-throughput experiments.
- **Open Science Foundation (platform)**- Serves as a hub for researchers to manage and share their data with others.
- **The COVID-19 High Performance Computing Consortium (platform)**- Led by the White House Office of Science and Technology, the U.S. Department of Energy, and IBM, this platform is providing IT resources to our lab for epidemiological research.
- **City of Philadelphia Office of Innovation and Technology (community)**- The City of Philadelphia is an application partner that oversees technology projects in the City and we have partnered with them in building a set of dashboards that will track human mobility patterns at the community scale.
- **Spectus (industry)**- Provided location and mobility that made up the initial datasets for these human mobility dashboards.

OUTPUT:

This output section includes CSSLab website posts, three new publications from the lab and coverage we have received on lab research, both internally within Penn and externally, as we aim to reach out to a larger community beyond academia.

CSSLab WEBSITE POSTS

- **"Joe Biden's (but not Donald Trump's) age: A case study in the New York Times' inconsistent narrative selection and framing"** (March 8, 2024)- This piece was written by David Rothschild, Jenny Wang, and Duncan Watts, looking into the *New York Times* and their choice of narrative when talking about Joe Biden's age in comparison to other news topics of interest to Americans.

- **"Hyperpartisan consumption on YouTube is shaped more by user preferences than the algorithm"** (March 5, 2024)
    - Summary of new publication: "Casually estimating the effect of YouTube's recommender system using counterfactual bots"

- **"New Insights on Common Sense Take the Spotlight on Canadian Radio"** (February 7, 2024)- Press release talking about a podcast episode where Mark Whiting was interviewed to talk about his latest research on quantifying common sense.

- **"Commonsensicality: A Novel Approach to Thinking about Common Sense and Measuring it"** (January 17, 2024)- Summary of new publication: "A framework for quantifying individual and collective common sense"

- **Repost: Warped Front Pages** (November 29, 2024)- originally published in the Columbia Journalism Review, David Rothschild, Elliot Pickens, Gideon Heltzer, Jenny Wang, and Duncan J. Watts discuss the impact of what editors choose to put on the front pages of their newspapers and emphasize the importance of transparency in what topics to emphasize in order to mitigate misinformation.

- **Repost: Mapping the Murky Waters: The Promise of Integrative Experiment Design** (November 16, 2023)- originally published in the Mack Institute Collective Impact, this piece features Abdullah Almatouq, one of the CSSLab's affiliate researchers at the Massachusetts Institute of Technology who talks about his PhD journey and how exploring literature review and the multitudes of contradictory research led him to propose the "integrative experimental design," a framework that creates a more comprehensive understanding of research topics.

- **"Are experimental designs one-size-fits-all? Or should they be modified to encapsulate the complexity of human behavior?"** (November 10, 2023)- Summary of journal article: Beyond

Playing 20 Questions with Nature: Integrative Experiment Design in the Social and Behavioral Sciences (December, 2022)

- **"The Unintended Consequence of Deplatforming on the Spread of Harmful Content"** (November 1, 2023)- Publication post covering journal article: Deplatforming did not decrease Parler users' activity on fringe social media (March, 2023).

NEW JOURNAL ARTICLES

Causally estimating the effect of YouTube's recommender system using counterfactual bots (*PNAS*, February 13, 2024)- As one of the largest social media platforms, YouTube has drawn attention and concern from observers due to the large amount of partisan content found on it. But is it the algorithm driving consumption of extreme content or is there something else at play? To test this, Homa Hosseinmardi and coauthors introduced a novel experimental method called "counterfactual bots." These bots are programmed to simulate what a real user would see if they were to only follow algorithmic recommendations. First, all the bots mimicked "real" user behavior by watching the same video sequence so that the YouTube algorithm could learn the user's preferences. Then, one "control" bot continued on the real user's path, and the other three "counterfactual" bots switched to relying exclusively on recommendations. The bots programmed to follow rule-based recommendations consumed less partisan content than the bots following the real users; those who consumed very high levels of hyperpartisan content experienced a substantial moderating effect. When the authors found that the algorithm has a moderating effect, they consequently asked: how long does it take for the algorithm to exert this effect? They measured this by quantifying the number of videos the bots watched until extreme content disappeared from their recommendation page, which was done by having bots initially watch video sequences from heavy consumers of far-right content and switch to moderate content. After about 30 videos, the algorithm recommended moderate content, but this moderation took longer for content characterized by higher degrees of partisanship, regardless of political leaning. While YouTube is a medium through which hyperpartisan content is readily available, there is a responsibility that lies in users and what they choose to consume. This study is unique from previous literature in that it implements bots that mimic real users and then are instructed to follow parallel paths of both a real user and algorithmic viewership, leading to the discovery that the role of the algorithm isn't as influential as previously thought.

A framework for quantifying individual and collective common sense (*PNAS*, January 16, 2024)- Mark Whiting and Duncan Watts had 2,046 participants from diverse backgrounds rate 4,407 different claims, or common sense statements. They also introduced two measures: Commonsensicality, or the state of being in alignment with common sense, at the level of individual people or claims, and *pq* common sense, a measure of how large the cliques of shared beliefs are within a given population and corpus of claims. In the individual setting, a Commonsensicality score of 1 would indicate that raters agree on most claims and assume that others agree with them as well, whereas Commonsensicality closer to zero would appear in a setting where common sense is not so common. To measure individual

commonsensicality, each participant was given a set of claims and had to answer two questions for every claim: 1) whether they agreed or disagreed with the claim and 2) have them predict if the majority of participants would agree or disagree with the claim. These responses were compiled into a belief graph in order to measure collective common sense by looking at the relationship between raters and claims. Findings from the graph show that common sense is in fact, rare, as lots of people can agree that a specific claim is common sense, but as the number of claims increase, the number of people who agree on all of those claims decreases. Mark Whiting has also launched a website for those interested in learning more about commonsensicality.

The Effects of Group Composition and Dynamics on Collective Performance (*Topics in Cognitive Science,* November 5, 2023*)* - Duncan Watts and coauthors explore the relationships between different combinations of group composition attributes (skill level, skill diversity, social perceptiveness, and cognitive style diversity) by utilizing the Empirica virtual lab. Participants were randomly assigned to different "rooms" and communicated with their teams (groups of three) via chat services. In these rooms, participants completed tasks of varying levels of complexity and were subjected to a social perceptiveness test to assess whether participants could correctly identify the corresponding emotion based on looking at different pairs of eyes. Afterwards, every participant was evaluated based on skill, social perceptiveness, and cognitive style. Then these participants moved on to phase 2 of the study, where they performed the same room assignment tasks as in phase 1 but were grouped so that participants of varying skill levels worked together. In general, groups that exhibited more "turn-taking" between team members were more efficient, as increased communication delayed the amount of time it took to complete tasks. Surprisingly, skill level and social perceptiveness did not have a direct effect on communication within the team, as the authors anticipated that there would be a clear association between communication and these attributes. These findings raise the possibility of extraneous factors influencing group performance, such as diversity and demographics.

PENN COVERAGE

Causally estimating the effect of YouTube's recommender system using counterfactual bots

- **Analytics at Wharton Researcher Spotlight: Estimating the Effect of YouTube Recommendations with Homa Hosseinmardi** (March 25, 2024)- This March, the Women in Wharton Analytics campaign was launched in honor of Women's History Month by highlighting women who have made contributions to the field of analytics. They spoke to Homa to learn more about her newest publication, "Causally estimating the effect of YouTube's recommender system using counterfactual bots."
- **The YouTube Algorithm isn't Radicalizing People** (February 20, 2024)- originally posted in Annenberg News and reposted in Penn Today, Penn Engineering Today and Knowledge at Wharton (title rewritten as The YouTube Algorithm isn't Radicalizing People: Why User Choice Matters on Social Media).

A framework for quantifying individual and collective common sense

- **The commonalities of common sense** (January 23, 2024), published in Penn Today and reposted in Penn Engineering Today and Knowledge at Wharton.

Nudge Cartography Project

- **Nudge Cartography: Building a map to navigate behavioral research** (August 2, 2023)- Linnea Gandhi (Wharton OID PhD Student) and Anoushka Kiyawat (research assistant) were featured in Penn Today.

EXTERNAL MEDIA COVERAGE

Causally estimating the effect of YouTube's recommender system using counterfactual bots

- **YouTube Algorithm Steers People Away from Radical Content** (March 13, 2024)- Reason Magazine
- **The YouTube algorithm isn't radicalizing people, says bots study** (February 20, 2024) - Tech Xplore
- **Does YouTube's Algorithm Radicalize Young Americans? New Study has Answers** (February 1, 2024)- Tech Times

A framework for quantifying individual and collective common sense

- **New Research Demonstrates That Common Sense Is Not So Common After All** (February 25, 2024)- SciTechDaily
- **Appeals to Common Sense Don't Convince Anyone** (February 23, 2024)- Bloomberg
- **UPenn researcher on why 'we need to have a way to quantify common sense'- and what that means for AI** (February 19, 2024)- Fortune
- **Commentary: Appeals to common sense don't convince anyone** (February 5, 2024)- Union Bulletin
- **Flam: Appeals to common sense unconvincing** (February 1, 2024)- The Columbian
- **Commonsense appeals do not convince anyone- Something vital gets lost when a pundit or politician insists that the answer to a polarizing question is 'common sense'** (January 27, 2024)- Taipei Times
- **How Common is Common Sense?** (January 24, 2024)- Technology Networks
- **Common sense is not actually very common** (January 17, 2024) - the Economist

- **Common Sense? There May Not Be Anything Common About It** (January 17, 2024)- IFLScience
- **Common sense may depend on one's point of view** (January 16, 2024))- Phys.org
- **There's no such thing as common sense** (January 16, 2024)- El Pais
- **The common sense shared by the whole of society does not exist** (January 16, 2024)- The Limited Times
- **Common Sense May Not Be Commonly Shared, Sociologists Find** (January 15, 2024)- Newsweek
- **Study Challenges Universality of 'Common Sense,' Offers New Perspectives** (January 15, 2024)- Breaking News Network (BNN)

PODCASTS

- **Common sense is not that common, but is quite widely distributed** (January 19, 2024)- Mark Whiting was featured on Quirks and Quarks, an award-winning science and technology podcast based in Canada. Host Bob McDonald interviews Mark about his research and the main takeaways of the publication in this episode.

PRESENTATIONS:

CONFERENCE PRESENTATIONS

- **Association for Interdisciplinary Meta-Research and Open Science (AIMOS) 2023** (November 21-23, 2023)
  - "Narrative License in the Social Sciences: An Exploration with GPT-4.0"- presented by Calvin Isch
  - "Research Cartography: Building a Map to Navigate and Generalize Behavioral Science"- presented by Linnea Gandhi

- **Society for Judgement and Decision making conference** (November 17-20, 2023)
  - "Research Cartography: Building a Map to Navigate and Generalize Behavioral Science"- presented by Linnea Gandhi

- **Conference on Harmful Online Communication (CHOC2023)** (November 16-17, 2023)
  - "Perspectives of harmful online content: Hate and dehumanization"- presented by Homa Hosseinmardi

- **Conference on Digital Experimentation @MIT (Code@MIT)** (November 10-11, 2023)
    - "High-throughput experiments in small-group deliberation"- presented by James Houghton and Duncan Watts

- **INFORMS Annual Meeting** (October 15-18, 2023)-
    - "Habits in Social Media Use: Entropy as a Habit Measure"- presented by Amir Tohidi

- **International Conference on Computational Social Science (IC2S2)** (July 17-20, 2023)
    - "From shared to separate: Tracking polarization in the production and consumption of American TV news"- presented by Homa Hosseinmardi and Sam Wolken
    - "Disentangling algorithms and user preferences with counterfactual bots"- presented by Homa Hosseinmardi
    - "Population-level common sense"- presented by Mark Whiting and Duncan Watts
    - "Matter Orders: Structural cues affect readers' interpretations independently of word choice"- presented by Mark Whiting
    - "Tasks Beyond Taxonomies: A Multidimensional Embedding Space of Team Tasks"- presented by Xinlan (Emily) Hu
    - "High-throughput experiments in small-group deliberation"- presented by James P. Houghton
    - "Comparing Methodologies for Constructing Epidemic Networks from GPS Mobility Data"- presented by Jorge Barreras Cortes and Bethany Hsiao
    - "Normic Webs of Belief"- presented by Mark Whiting

- **NETSCI 2023: INTERNATIONAL SCHOOL AND CONFERENCE ON NETWORK SCIENCE** (July 10-14, 2023)
    - "Counterfactual bots (abstract)"- presented by Homa Hosseinmardi, Amir Ghasemian, Miguel Rivera-Lanas, and Duncan Watts
    - "Enmity paradox (abstract)"- presented by Amir Ghasemian

- **International Conference on Science of Science and Innovation** (June 26-29, 2023)
    - "Research Cartography: Building a Map to Navigate and Generalize Behavioral Science"- presented by Linnea Gandhi

- **International Association for Conflict Management** (June 23-26, 2023)
    - "High-throughput experiments in small-group deliberation"- presented by James Houghton

- **Society for the Improvement of Psychological Science** (June 22-24, 2023)
    - "Collaborative high-throughput experiments in small-group discussion"- presented by James Houghton

- **2023 MIT IDE Annual Conference** (May 18, 2023)
    - "Quantifying the Impact of Anti-vaccine Information on Facebook"- presented by Jennifer (Jenny) Allen

- **Midwestern Political Science Association** (April 13-16, 2023)
    - "High-throughput experiments in small-group deliberation"- presented James Houghton, Dean Knox (Wharton Professor and affiliate researcher with CSSLab)

POSTER SESSIONS

- **New Directions in Analyzing Text as Data (TADA)** (November 9-10, 2023)
    - "Measuring Diversity in Online News"- presented by Samar Haider and Duncan Watts
- **International Conference on Computational Social Science (IC2S2)** (July 17-20, 2023)
    - "Saying the Right Thing at the Right Time: Task-Sensitive Theories of Team Communication Processes"- presented by Xinlan (Emily) Hu
    - "Exogenity and Endogenity: the amplification of ideas in the News Cycle"- presented by Baird Howland
    - "Mapping the U.S. News Landscape"- presented by Samar Haider
    - "Measuring Diversity in Online News"- presented by Samar Haider
    - "Is luck random?"- presented by Mark Whiting

INVITED TALKS

- **Fairness and Safety Speaker Series, YouTube, NYC** (September 2023)- Homa Hossenmardi (Associate Research Scientist) was invited to speak about her paper: "Causally estimating the effect of YouTube's recommender system using counterfactual bots."

UPCOMING TALKS

- **Political and Information Networks** (April 25-26, 2024)- Homa Hosseinmardi has been invited to a table talk at "Political and Information Networks," hosted by the Center for Information Networks and Democracy which will take place at the Annenberg School of Communication.

EVENTS & GATHERINGS:

The International Conference on Computational Social Sciences (IC2S2) is an annual world-wide conference where researchers from the fields of "economics, sociology, political science, psychology, cognitive science, management, computer science, and statistics" come together to share their latest findings and addressing current challenges in the field of computational social science.

International Conference on Computational Social Science (IC2S2) (July 18-20, 2024)- This year's IC2S2 conference will be held at the University of Pennsylvania. Conference organizers affiliated with the CSSLab include Jeanne Ruane, Homa Hosseinmardi, Amir Ghasemian, Sam Wolken, Samar Haider, and Chris Callison-Burch.

International Conference on Computational Social Science (IC2S2) (July 17-20, 2023)- held at the University of Copenhagen.

AWARDS:

Duncan Watts (CSSLab Director) was one out of four Penn faculty to be elected into the United States National Academy of Sciences.

Samar Haider (Computer Science PhD student) recently received the Outstanding Reviewer Award from EMNLP (Empirical Methods in Natural Language Processing), one of the most prestigious conferences in this field.

Xinlan (Emily) Hu (Wharton OID PhD student) was awarded the Winkelman Fellowship. This is given to Wharton doctoral candidates working in the broad area of applied economics.

Baird Howland (Annenberg PhD student) was chosen as a doctoral fellow by the Annenberg Center on Digital Culture and Society. He was recognized for his work on "Measuring the emotions of political discourse" which uses Natural Language Processing to identify narratives and measure emotions within news articles and social media posts.

STUDENTS:

Josh Nguyen established a collaboration with the Mathematics and Computer Science Division of Argonne National Laboratory. His visiting student appointment allows him to make use of Argonne's supercomputing facility to study how language models perform commonsense reasoning.