

Computation+Journalism 2022

Workshop Descriptions

Presenters' names that are bolded and black will be attending in-person, while those that are bolded and blue will be remote.

Thursday, June 9 — 2:00 - 3:15 PM EDT

Numbers for audience understanding - Laura Santhanam (PBS NewsHour), **Jena Barchas-Lichtenstein** (Knology), John Voiklis (Knology), Erica Hendry (PBS NewsHour) and Travis Daub (PBS NewsHour)

There are tons of resources to help journalists get the numbers right...but right isn't enough if the audience doesn't get the point. In this workshop, participants will have hands-on opportunities to (a) interpret unfamiliar numbers to understand the audience perspective and (b) present numbers in ways that actually help audience members make sense of them. All of our recommendations were developed through a close collaboration between social scientists and journalists in various roles within a large public media organization. The workshop will be co-facilitated by at least two journalists (one data producer and one person in an editorial role) and two researchers; at least half of them will be women. Participants will have the chance to engage in small groups to give one another feedback as they consider how to present a wide range of numbers to a general audience. They will walk away with questions to ask themselves about statistics in news stories and strategies for writing and visualizing them that supports audience understanding. We will also share explainers we have developed that can be linked from other stories to support that understanding. This session will be valuable for anyone who writes, edits, or reviews stories that contain numbers.

AI for Everyone: Learnings from the Local News Challenge - Swapneel Mehta (New York University), **Christopher Brennan** (Overtone AI), **Zhouhan Chen** (New York University) and Matt Macvey (NYC Media Lab)

Journalists and media organizations have increasingly embraced AI to offer a special kind of value to their readers, listeners, and viewers, but what are the advances that can help smaller outlets with resource constraints, focused on their local communities? NYC Media Lab organized the AI Local News Challenge this spring to bring together five teams each approaching the issue in a different way, whether through the lens of localized stories for a network of newspapers, monitoring social network interactions, tracking misinformation, curating social conversation or creating qualitative data on news articles.

This session will see the presentation of results from the teams: Gannett-Localizer, Information Tracer, Overtone, and SimPPL, focusing on their work this spring. They will share their solutions as well as their learnings from months enmeshing themselves at the intersection of local news and technological advances. All used techniques centered around artificial intelligence, including natural language processing, text generation, cybersecurity, simulation intelligence, and large-scale data analytics.

Afterwards, the discussants will offer their own ideas to dig into the complexities of questions on topics such as “what does local really mean?” and “what technologies are best able to help newsrooms that are struggling?”. The participants will also welcome questions and comments from the audience to delve further into these topics, as well as discuss the practical implications of adopting these innovations at scale.

Thursday, June 9 — 3:30 - 4:45 PM EDT

Identifying social media manipulation with OSoMe tools - Kai-Cheng Yang and Christopher Torres-Lugo (Observatory on Social Media, Indiana University)

As social media become the major platforms for discussions of important topics such as national politics, public health, and environmental policy, there is a growing concern about the manipulation of these information ecosystems and their users. Malicious behaviors include astroturf, amplification of misinformation, and trolling. Such abuses can be carried out by humans and social bots --- inauthentic accounts controlled in part by software. The resulting biased reality can fool even professionals. While journalists and researchers are increasingly interested in detecting and studying these malicious activities, there are serious challenges. First, the collection and analysis of data from social media require significant storage and computing resources. Second, knowledge, experience, and advanced computational skills are necessary to find patterns and signals of suspicious behaviors in the collected data. In this workshop, we will present free tools that aim to help researchers, journalists, and the general public combat online manipulation from the Observatory on Social Media (OSoMe) at Indiana University (osome.iu.edu/tools). We will focus on Botometer that helps detect social bots on Twitter and Hoaxy that can track and visualize the diffusion of misinformation; other useful tools from the OSoMe family will also be covered. These tools utilize the massive social media data collected by OSoMe and they are equipped with state-of-the-art algorithms and user-friendly interfaces. They also provide public APIs to allow querying in bulk. They have helped thousands of users and served as the foundation for hundreds of research projects.

Temporal Topologies – Making journalistic authorship transparent through in-depth tagging of newsworthy events - Francesca Morini (Fachhochschule Potsdam)

Journalistic writing is a complex and delicate form of authorship encompassing various techniques to organize, make sense, and effectively craft stories to inform readers. Through their sense-making process, journalists decide which events are newsworthy, as well as frame them meaningfully. By choosing the news angle, journalists work to establish and render unique constellations of related events cohesively. This relational contextualization of events is what stays implicitly at the core of journalistic authorship and relies on journalists' commitment to truth and transparency. As readers struggle to orient themselves through increasingly nuanced, globally entangled phenomena, in which way can journalistic authorship become their guiding thread? We engage with this question by looking at journalistic authorship as an actionable concept for data visualization. Starting from Digital Humanities, this workshop combines analog sketching and collaging with computational thinking through the use of standardized ontologies for the Semantic Web – e.g. OWL Time (Cox & Little, 2020) – and temporal reasoning techniques like Allen's interval algebra (1983). Participants will be asked to structure a story from an archive of textual material working solely on temporal and relational instances. The workshop targets

journalists interested in approaching unstructured text with diagrammatic thinking. The goal of the exercise is to produce the temporal topology of one newsworthy event: its temporal layout, extensions, and relations among parts. Participants will gain knowledge on consistent tagging languages for deep-time relational entities and will explore the act of authoring, its inflections, and the implications of exposing their sense-making process to readers.

Using Networks Analysis and Visualization to Explain COVID-19 Spread through the Physical, Social, and Information Graphs - [Hong Qu](#) (Harvard Kennedy School)

Understanding network science and compartmental models in epidemiology are essential to understanding, explaining, intervening, and forecasting the pandemic. This workshop introduces core concepts and skills for network data analysis and visualization using Flourish and NetworkX. We begin with simple data sets that represent the flow of migration between countries using sankey, chord, correlation matrix heatmaps and, thereby, introduce the notion of nodes and edges in a graph. Next, we introduce network models such as small-world and scale-free networks to describe typical properties of complex networks, and demonstrate how to use Flourish to produce network visualizations, encoding the size of nodes as centrality and thickness of the links as edge weight. Then, we use NetworkX to analyze network statistics and run a simple agent-based model simulation of misinformation diffusion based on a simplified SIR epidemic model. By the end of the workshop, participants will fully appreciate the network science and design principles that inspired the the fam visual story created by Harry Stevens for the Washington Post: 'Why outbreaks like coronavirus spread exponentially, and how to "flatten the curve"'. We end with a brainstorming discussion of ways to combine interaction of three different layers in the multiplex network--physical mobility, social relationship, and infodemic beliefs--drive outcomes and trajectories for the ongoing pandemic.

Thursday, June 9 — 5:00 - 6:15 PM EDT

How can AI help you? Explore writing with a machine - [Lydia Chilton](#) (Columbia University, Computer Science)

Groundbreaking AI has enabled computers to write; it can finish your sentences in Gmail, but it can also finish paragraphs of text and write fiction on its own. Although there is a lot of hype around AI (both positive and negative), the truth is somewhere in between. We'll demonstrate how we use AI to summarize, find background, and discover story angles. Then we'll give you keys to the latest text-generating AI, and we can explore together the good, the bad, and the ugly of writing with a machine.

Share a Story: A Daily Practice for Teaching Healthy News Habits - [Blake Eskin](#) (Journalism + Design)

Share a Story is an application for journalism educators and their students to work on mindful news consumption and production while building community in their courses. Inspired by social reading apps, 100-day design projects, and Duolingo, Share a Story has been developed over the past five years with an undergraduate cohort that is passionate about storytelling but often lacking confidence in computational thinking. Share a Story aims to undergird an approach to journalism education focused on the individual's role as one node in a dynamic and unreliable information ecosystem rather than on the heroic reporter who goes out hunting for scoops.

Each day, Share a Story participants must share one — and only one — news story and say what happened, why it matters, and where they found it. At first students do in-class exercises based on the stories they share, such as researching the author of the story or the owner of the publication where it appeared, all while being nudged to share stories from more diverse sources. Over time, the accrued stories form a collaboratively created data set that students use to analyze their news habits, understand community information needs, look for story ideas, and prototype new news products on paper or using no code/low code tools.

This workshop session will demo a working prototype of Share a Story built in Airtable, discuss progress and challenges, and explore plans for further testing and development.