

4th Kavli Symposium on Science Journalism

Austin, Texas
19-21 February 2018



THE PRIMERS OF THE 4TH KAVLI SYMPOSIUM

on

Technology & Truths

How to ethically harness the potential of open source data and avoid fake news

Organized by the
World Federation of Science Journalists
&
The Kavli Foundation

Symposium Primers

This document contains four primers for each of the themed sessions at the upcoming 4th Kavli Symposium. As an introduction, it is not intended to encompass all available information, viewpoints and opinions. Instead, we expect you to bring your expertise and views on missing but vital topics to the event to build new knowledge and initiatives on the topic.

Introduction

The science journalism landscape is continually shifting as information on science becomes more accessible. Open source data has given rise to data journalism and an era where journalists need to vet the authenticity of data often before scientific experts can provide their consensus. The situation is further complicated by media hackers and propagandists seeking to advance their particular agenda at the expense of the truth. Technological advances are available to help detect fake news and protect sources, but they are not all created equal. It's vital to be aware of the possible pitfalls and how to circumvent them.

Goal of the Symposium

To develop collaborative projects that make the best use of current technology to elevate the communication of science so that it reflects truth and prioritizes privacy for data sources.

4th Kavli Symposium on Science Journalism

Austin, Texas
19-21 February 2018



WFSJ
World Federation of Science Journalists

THE  **KAVLI FOUNDATION**

Session 1: DATA JOURNALISM

Moderator: Richard Stone, International News Editor, Science

Speakers: Steven Doig, Journalism Professor, Arizona State University

Jennifer LaFleur, Senior Editor of Data Journalism, Reveal

As society becomes ever-more digitized, journalists are presented with new challenges and opportunities to access data and present it to the public. Investigative journalists have tended to dominate the data-rich story scene as they have acquired the tools and experience needed to distill stories from datasets. In essence, they have learned how to harness and analyze data.

Now, the time has come for science journalists to get in on the action. A few frontrunners already have. At the KS4, two experts on working with data – Jennifer LaFleur and Stephen Doig – will guide you on how to tap the potential of open source data and produce more incisive stories without sacrificing accuracy.

The session will commence by encapsulating the challenges and opportunities of data journalism – with an explanation of what our experts set out to do and how they accomplished it. Then our session leaders will plunge into the nitty gritty of data journalism. Doing data-based stories is less about what buttons to click and more about knowing how to find and wrangle data and then what to do with it. You should investigate data like you would a person: Do a background check and see what it says, and then back it up with real examples. The steps for getting there will be outlined in two segments of this session – “Getting Data” and “Working with Data.”

Here’s a teaser on what each segment will cover:

Getting Data

- Government data/FOIA (Do’s and Don’ts when negotiating for data)
- Data accessible online through dashboards and other tools
- Scraping websites
- Making your own data

Working with Data

- The tools (spreadsheets, databases, stats, maps)
- The process (including: the importance of the literature review and understanding the documentation)
- Vetting data - This is particularly important at a time when journalists are being accused of reporting false information.

4th Kavli Symposium on Science Journalism

Austin, Texas
19-21 February 2018



Session 2: PROTECTING DATA SOURCES AND PERSONAL DATA

Moderator: Ivan Oransky, New York University, Retraction Watch

Speaker: Simson Garfinkel, Senior Computer Scientist for Confidentiality and Data Access at the US Census Bureau

“It’s frequently easier to reveal secrets than to keep them,” says [Simson Garfinkel](#). He has published research articles on computer security, digital forensics and privacy, is a fellow of the Association for Computing Machinery, and holds a PhD in Computer Science from MIT.

The goal of this session is to help journalists think about the best ways to protect confidential sources and the data that they volunteer. Critical topics to be discussed include:

- **Communications security:** How do you obtain confidential information with your sources?
- **Storage security:** How do you maintain your secrets?
- **Publication security:** How do you control the information released by your publication to prevent the inadvertent release of confidential information?

Although there is a wide variety of privacy and anonymity tools available for use on the Internet today, including programs like Tor, Signal and SecureDrop, these systems can be difficult to use securely and can be by bypassed by some adversaries. Garfinkel will discuss usable approaches for achieving operational security that is sufficient for many journalists.

Storage security is an aspect of security that is frequently overlooked. In 2003, Garfinkel showed that many computer users failed to take basic precautions when disposing of their computer equipment: he purchased hundreds of used hard drives and other storage devices and showed that many still contained the data of their original users. Follow-up research has shown that this problem persists even today – 15 years later. Other kinds of security problems exist with cloud-based email and storage systems. You can’t avoid these systems, so it’s important to know how to use them securely.

Finally, journalists need to understand that if they aren’t careful, they can inadvertently betray their confidential sources with innocent-seeming publications. That’s because publications can reveal all sorts of information inadvertently, the way a digital photo reveals not just what was photographed but potentially where the photographer was standing and when the picture was taken. Garfinkel will discuss how data journalism and some government open data programs have likewise revealed private data about individuals. To help understand these issues, Garfinkel will discuss a few high-profile incidents involving personal data and the media, then show how the science of differential privacy can help individuals and organizations that want to publish data without compromising privacy.

4th Kavli Symposium on Science Journalism

Austin, Texas
19-21 February 2018



Session 3: FACT-CHECKING AND FIGHTING MISINFORMATION

Moderator: Deborah Blum, Director Knight Science Journalism Program, MIT

Rapporteur: Kathleen Hall Jamieson, Elizabeth Ware Packard Professor of Communication, Director, Annenberg Public Policy Center, University of Pennsylvania

We live in a time when there are increasing signs of an anti-science shift in modern culture, with indicators ranging from the [publicly hostile](#) stance of the current federal administration to [alarmed discussions](#) by worried scientists around the world. Now more than ever, it is imperative that the journalists who tell the stories of science get those stories meticulously right. If we hold to that standard, we better serve both the reputations of science and of journalism.

Yet, at the moment, it appears that professional fact-checking in science journalism is thin at best. Much of the evidence for this is, of course, anecdotal. As a recent example, at a fall Kavli Symposium/Workshop on Science Editing, a roomful of twenty-some science editors was asked if their publication employed a staff fact-checker. Only three raised their hands. The New York based science writer Brooke Borel, author of [The Chicago Guide to Fact-Checking](#) recently surveyed more than 200 magazines that produce science copy on the subject of staff fact-checkers. She counted the positive responses at 10. And this [brief survey story](#) in the magazine Undark found that that even basic science fact checking was minimally available to most publications.

So, how do we best address and improve on this situation? This session will explore both the problem and possible solutions across a range of possibilities. Keynote speaker, Kathleen Hall Jamieson, will raise some of the key issues in her talk “Pre-emptive Reporting: Increasing Public Understanding of What and How Science Knows in a Polarized Climate.” Jamieson, Director of the Annenberg Public Policy Center at the University of Pennsylvania, is also the Elizabeth Ware Packard Professor of Communication at the Annenberg School for Communication, and co-founder of FactCheck.org. She is an expert both in understanding miscommunication techniques and in finding ways to combat them.

The session will be coordinated by Deborah Blum, a Pulitzer-Prize winning journalist, Director of the Knight Science Journalism Program at MIT, and publisher of its award-winning magazine, Undark. Blum will introduce Kathleen Hall Jamieson and the session by providing an overview of the critical issues in integrity of story-telling in science reporting, some additional context on professional fact-checking as well as the promise and limits of technological solutions. As Jamieson has noted in earlier presentations, the timing is critical to address such problems, not only in science reporting but across the political spectrum. In a talk last year, she pointed out that the risks of allowing fake news to continue unchecked are high: “Deception is problematic because it can mobilize national action which you might not take in the absence of deception, it can mislead the electorate, it can invite non-responsive policy, ... it can impugn character, and it can even endanger lives.”

4th Kavli Symposium on Science Journalism

Austin, Texas
19-21 February 2018



WFSJ
World Federation of Science Journalists

THE  **KAVLI FOUNDATION**

Session 4: A.I. IMPLICATIONS IN SCIENCE JOURNALISM

Moderator: Laura Helmuth, Health, Science and Environment Editor, The Washington Post

Speaker: Sam Han, Director of Big Data in the engineering department of The Washington Post

As technological advances are implicated more and more in online news, what is the next step? Artificial intelligence (AI) is certainly poised to be the next game changer. For instance, local news journalism has certainly suffered because by nature it is of interest only to a limited number of people and smaller markets spell smaller profit margins. Local news operations often don't have the staff to write local stories and large centralized operations are not likely to take on writing localized stories. Could AI be the answer?

A Press Association project known as RADAR (Reporters and Data and Robots) uses open data sources from government, local authorities and public services to generate local stories. Initially, the idea was to use AI to generate stories for national media using open data sources, but it unexpectedly proved to be very well suited to generating local stories. What other potential new roles are there for AI in the science journalism area?

This session will discuss the following tools and their utility in news making:

1. Virality prediction engine.
This service predicts the popularity of an article.
2. ModBot.
Automatic moderation of user comments.
3. Heliograf.
Automatic story generation tool.