HDSI | Harvard Data Science Initiative

# TRUST AND MISTRUST IN SCIENCE
# MAY 2020 HDSI WORKSHOP SUMMARY

The mission of the Harvard Data Science Initiative is to engage and activate data science pioneers, at Harvard and around the world, to address major challenges facing humanity. We are committed to transformational research that will have widespread impact, bringing to bear the full spectrum of Harvard's diverse data science methodology and domain expertise to the thorniest issues of our time.

One such issue is understanding public trust and mistrust in science. This is not a new topic, but takes on a new urgency in a time of ever more complex circumstances. While studies and polls show that a significant share of, for example, the US population does trust in science[1], many people may lack confidence in the scientific method and the integrity of researchers. In the data science era, we have an unprecedented opportunity to understand whether and why this is the case, and unlock new understandings of the root causes of trust and mistrust in science and possible mitigations.

With our Corporate Members, in 2020 the HDSI will launch the HDSI Project on Trust in Science, a multi-year, multi-disciplinary, and multi-perspective research program to make measurable progress on this topic. Our goal: To increase public trust in science, by shedding light on the causes of public mistrust in science and developing interventions that support enduring cultural change.

In May 2020, the HDSI hosted a faculty workshop as a first organizing step towards the larger, multi-year research program. Our purpose was to energize our community around the topic of Trust in Science and surface project ideas that could be knit together as a cohesive research plan. Here, we offer a summary of the broad and creative discussion about how we can use the new science of data to unlock insights around a topic of timely and of critical importance to us all.

## OVERVIEW OF WORKSHOP DISCUSSION

Participants were clustered into discussion groups of eight to ten with facilitators from the Harvard Data Science Initiative. Discussion groups were organized to represent a breadth of expertise, applications, and perspectives from academia or industry. From this diverse participant set, certain themes, challenges, and opportunities emerged that we intend to use to guide future work on the topic.

### What guiding themes should ground all work on data-driven approaches to understanding trust and mistrust in science?

Trust is relational and context dependent. It is a complex, non-linear response built on, for example, audiences' perception of the credibility of scientists, the relevance of findings to audiences' concerns, and the integrity of the work.

Audiences are not monoliths. Information confers different meanings to different audiences with different worldviews. In seeking to engender greater "public" trust in science, we must be cognizant of different "publics" each with their own priorities and frames of reference.

[1] Pew Research Center, August 2019, "Trust and Mistrust in Americans' Views of Scientific Experts"

**Science is not built on certainty, but on consensus.** Built-in uncertainty in science may conflict with audiences' need for authoritative information from communicators, for example policymakers and the media.

## What challenges exist to improving trust in science?

**Disconnects exist between scientific findings and behavior**, such as public policy decisions or personal actions. The nuance of scientific uncertainty may be lost as information becomes repackaged for different audiences' use.

**Communication outpaces scientific consensus.** For example, during the Covid-19 pandemic, the high frequency of sometimes conflicting messages and visualizations has led to fatigue and lessening trust in some audiences. The current trend toward open science, while positive, can unintentionally bypass the process of scientific consensus, for example the distribution of pre-prints before peer review.

**Media, whether traditional or social media, can amplify misinformation,** reducing trust in science. News media, in an attempt to offer a balanced view, may promote outlier views that contradict scientific consensus. Social media can create echo chambers for both good and bad scientific information. As a result, audiences may not know who to trust or understand motives for communication.

**The infrastructure and methodology of science is often opaque.** Increased transparency including, for example, funding sources, data sources and the publication process, can help dispel myths around researchers' motivations and address bias in scientific output.

**Cherry-picking information undermines public trust in science.** The selection of experts and results that reinforce specific agendas erodes credibility, contributing to skepticism about long-held scientific consensus, for example around the topic of climate change.

## What are some examples of opportunities to address the above challenges?

**Designing new tools to indicate trustworthiness.** Participants discussed the potential for new indicators of whether a study could be trusted along dimensions of reliability and reproducibility of data-driven research. Furthermore, participants perceived an opportunity for a new system of checks and balances to research, beyond the current peer review standard.

**Improving communication about the process of science.** Participants suggested promoting transparency around "how science gets done," including acknowledging the limitations of research studies. Participants also emphasized the need for visible corrections when new information replaces previous accepted consensus.

**Creating deeper understandings of audiences.** Participants cited the need to develop a more nuanced view of what influences audiences to trust or distrust science. Dimensions could include, for example, media diets, science literacy, and existing value systems. Participants were interested in further defining the insularity of echo chambers, particularly those that exist in online networks.

**Tracing the dissemination of information.** Participants contemplated novel opportunities to track how information spreads, for example unlocking unstructured text and other untapped sources for a historical record on science, or filtering conflicting information to reconstruct scientific consensus. Relatedly, data science could offer new ways to examine the language of belief and trust in science, such as tracing how calls to action propagate, in some cases from a common source.

## SUGGESTED READING LIST

This list was collected from workshop participants, and reflects a breadth of the expertise and concerns shared with the HDSI.

Trust and Mistrust in Americans' Views of Scientific Experts
The Pew Research Center (2019)

An update on our research into trust, facts and democracy
Michael Dimock, Pew Research Center (2019)

A Closer Look at Public Trust in Scientists
ScienceCounts (2019)

Trust in Research
Elsevier Global Communications (2019)

Unite behind science!
Matthias Berninger (2020)

### Communication and Information

Why False Claims About COVID-19 Refuse to Die
Cailin O'Connor & James Owen Weatherall (2020)

Visualization Design Principles for the Pandemic
Michael Correll (2020)

Why outbreaks like coronavirus spread exponentially, and how to "flatten the curve"
Harry Stevens, The Washington Post (2020)

7 Ways to Explore the Math of the Coronavirus Using The New York Times
Patrick Honner, The New York Times (2020)

A Kernel of Truth: Determining Rumor Veracity on Twitter by Diffusion Pattern Alone
Nir Rosenfeld, Aron Szanto, David C. Parkes (2020)

Monetizing disinformation in the attention economy: The case of genetically modified organisms (GMOs)
Camille D.Ryan, Andrew J.Schaul, RyanButner, John T.Swarthouta (2020)

The Stability of Twitter Metrics: A Study on Unavailable Twitter Mentions of Scientific Publications
Zhichao Fang, Jonathan Dudek, Rodrigo Costas (2020)

Data Science: A guide for society and other projects from Ask For Evidence/Sense About Science
Sense About Science (2019)

The Communications Revolution and Health Inequalities in the 21st Century: Implications for Cancer Control
K. Viswanath, Rebekah Nagler, Cabral Bigman-Galimore, Michael McCauley, Minsoo Jung, and Shoba Ramanadhan (2012)

## Science and Policymaking

Embrace Experimentation in Biosecurity Governance
Sam Weiss Evans et al. (2020)

Big Data in Context: Addressing the Twin Perils of Data Absenteeism and Chauvinism in the Context of Health Disparities Research
Edmund W J Lee and Kasisomayajula Viswanath (2020)

Virtual, visible, and actionable: Data assemblages and the sightlines of justice
Sheila Jasanoff (2017)

Boundary work for sustainable development: Natural resource management at the Consultative Group on International Agricultural Research (CGIAR)
William C. Clark, Thomas P. Tomich, Meine van Noordwijk, David Guston, Delia Catacutan, Nancy M. Dickson, and Elizabeth McNie (2016)

Crafting usable knowledge for sustainable development
William C. Clark, Lorrae van Kerkhoff, Louis Lebel, and Gilberto C. Gallopin (2016)

## Scientific Process and Transparency

Nongenic cancer-risk SNPs affect oncogenes, tumour-suppressor genes, and immune function
Maud Fagny, John Platig, Marieke Lydia Kuijjer, Xihong Lin & John Quackenbush (2020)

Reproducibility and Replicability in Science
The National Academies (2019)

Estimating Sample-Specific Regulatory Networks
Marieke Lydia Kuijjer, Matthew George Tung, GuoCheng Yuan, John Quackenbush, Kimberly Glass (2019)

Training replicable predictors in multiple studies
Prasad Patil and Giovanni Parmigiani (2018)

Statistical paradises and paradoxes in big data (I): Law of large populations, big data paradox, and the 2016 US presidential election
Xiao-Li Meng (2018)

Gene Regulatory Network Analysis Identifies Sex-Linked Differences in Colon Cancer Drug Metabolism
Camila M. Lopes-Ramos, Marieke L. Kuijjer, Shuji Ogino, Charles S. Fuchs, Dawn L. DeMeo, Kimberly Glass and John Quackenbush (2018)

## WORKSHOP PARTICIPANTS

* Denotes HDSI Corporate Member

Shankar Ananthakrishnan Amazon*

Matthew Baum Harvard Kennedy School of Government

Francine Berman Rensselaer Polytechnic Institute; Harvard Radcliffe Institute for Advanced Studies

Matthias Berninger Bayer*

Elizabeth Bruce Microsoft

Jennifer Chow Harvard Data Science Initiative

William Clark Harvard Kennedy School of Government

Ethan Cowan Harvard Data Science Initiative

Amir Danak McKinsey & Company*

Anita de Waard Elsevier*

Melissa Dell Harvard Faculty of Arts and Sciences

Raymond Dobert Bayer*

Francesca Dominici Harvard T.H. Chan School of Public Health; Harvard Data Science Initiative

Sam Evans Harvard Kennedy School of Government

Rob Faris Harvard Kennedy School of Government

Sarah Figge Hussain Harvard Weatherhead Center for International Affairs

Nick Fowler Elsevier*

Archon Fung Harvard Kennedy School of Government

Jon Giebel Bayer*

Elena Glassman Harvard John A. Paulson School of Engineering and Applied Sciences

Dave Harrington Harvard T.H. Chan School of Public Health

David Homa Harvard Business School Digital Initiative

Rafael Irizarry Harvard T.H. Chan School of Public Health, Dana-Farber Cancer Institute

Sheila Jasanoff Harvard Kennedy School of Government

Gary King Harvard Faculty of Arts and Sciences; Harvard Institute for Quantitative Social Sciences

Elizabeth Langdon-Gray Harvard Data Science Initiative

Monika Lessl Bayer*

Bekah Mahan Bayer*

Rebecca McLeod Harvard Data Science Review

Xiao-Li Meng Harvard Faculty of Arts and Sciences; Harvard Data Science Review

Carl Nising Bayer*

Sharon-Lise Normand Harvard Medical School

Bo Yun Park Harvard Faculty of Arts and Sciences

David Parkes Harvard John A. Paulson School of Engineering and Applied Sciences; Harvard Data Science Initiative

Giovanni Parmigiani Harvard T.H. Chan School of Public Health; Dana-Farber Cancer Institute

Chirag Patel Harvard Medical School

Hanspeter Pfister Harvard John A. Paulson School of Engineering and Applied Sciences

John Pham Cell Press/Elsevier*

Andrew Plume Elsevier*

Jeffrey Pobst Bayer*

John Quackenbush Harvard T.H. Chan School of Public Health; Dana-Farber Cancer Institute

Meredith Roman Bayer*

Paige Sammartino Harvard Data Science Review

Carmel Shachar The Petrie-Flom Center at Harvard Law School

Laurie Shaw Harvard Data Science Initiative

Paul Viera Harmony Analytics and EARNEST Partners*

K. "Vish" Viswanath Harvard T.H. Chan School of Public Health; Dana-Farber Cancer Institute

Christopher Winship Harvard Faculty of Arts and Sciences

## ABOUT THE HDSI TRUST IN SCIENCE PROJECT

The Harvard Data Science Initiative is currently engaged in planning for the launch of a research program on the topic of trust and mistrust in science. Our overarching goal: To increase public trust in science by leveraging data science to shed light on issues related to public mistrust of science, and developing interventions that support enduring cultural change.

Work supported by this program and the related HDSI Research Fund for Trust in Science will focus on advancing data-driven understandings of trust and mistrust in science.

### Ethics and Governance

The HDSI views relationships with industry as critical to our mission of transformation through data science. The problems we tackle can be informed by the most difficult challenges facing industry - challenges that can be solved in partnership with academia. At the same time, we recognize that industry-academia collaboration gives rise to its own ethical and governance questions. When partnering around a complex and impactful topic such as trust in science, these issues are more pronounced. Therefore, the HDSI's governance of this research program will be advised by an independent advisory board whose members are external to Harvard.

### The HDSI Research Fund for Trust in Science

The Harvard Data Science Initiative Research Fund for Trust in Science will provide philanthropic support for research that advances understanding of trust and mistrust in science by leveraging data science, toward the goal of creating actionable insights. Reflecting the breadth and potential impact of this work, the Fund is open to gifts from multiple donors who seek to catalyze this progress. Initial seed funding for the Fund has been provided by Bayer, an HDSI Corporate Member.

For more information, please visit https://datascience.harvard.edu/trust-science.