

Disrupt, Ally, Resist, Embrace (DARE): Action Items for Computational Social Scientists in a Changing World

Kokil Jaidka,¹ Eni Mustafaraj,² David Schoch,³ Kenneth Joseph⁴

¹ Department of Communications and New Media, National University of Singapore

² Department of Computer Science, Wellesley College, Wellesley, MA, USA

³ Department of Computational Social Science, GESIS, Germany

⁴ Department of Computer Science and Engineering, University of Buffalo, NY, USA

jaidka@nus.edu.sg, emustafaraj@wellesley.edu, david.schoch@gesis.org, kjoseph@buffalo.edu

Abstract

In the past decade, a number of sophisticated AI-powered systems and tools have been developed and released to the scientific community and the public. These technical developments have occurred against a backdrop of political and social upheaval that is both magnifying and magnified by public health and macroeconomic crises. These technical and socio-political changes offer multiple lenses to contextualize (or distort) scientific reflexivity. Further, to computational social scientists who study computer-mediated human behavior, they have implications on what we study and how we study it. How should the ICWSM community engage with this changing world? Which disruptions should we embrace, and which ones should we resist? Whom do we ally with, and for what purpose? In this workshop co-located with ICWSM, we invited experience-based perspectives on these questions with the intent of drafting a collective research agenda for the computational social science community. We did so via the facilitation of collaborative position papers and the discussion of imminent challenges we face in the context of, for example, proprietary large language models, an increasingly unwieldy peer review process, and growing issues in data collection and access. This document presents a summary of the contributions and discussions in the workshop.

Motivation

The Disrupt, Ally, Resist, and Embrace (DARE) Workshop at AAAI-ICWSM¹ was organized to address some of the issues affecting computational social science researchers and their research agendas. Broadly, the workshop engaged with the *processes of* and the *ethical principles underlying* computational social science research. These are often, and repeatedly, disrupted by platform politics, new technologies, their implications, and their unknowables. For instance, the increasingly turbulent techno-political online environment has seen a few key developments that have affected the scope and characteristics of computational social science research centered on social media. The global pandemic, a looming climate change crisis, violent populist events such as those that occurred on January 6, 2021 (the first ever attack on the US Capitol), and January 8, 2023, in Brazil (a copy-cat attack on Brazil's Praça dos Três Poderes), the enduring

war between Russia and Ukraine, and ubiquitous conspiracy theories surrounding everything have spurred more discussions about access, inclusivity, privilege, and propaganda than ever before.

Furthermore, these real-world events have been accompanied (and often closely entangled) with changes in the online world: the rise of TikTok and the fall of Facebook, Twitter's takeover by Elon Musk, and new AI technologies (DALL-E, Stable Diffusion, Open AI's ChatGPT, GitHub CoPilot, and others). In particular, Large Language Models (LLMs) and their applications are being widely discussed in academia and media as potential "disruptors" of scientific integrity, especially the jobs of knowledge workers.

As computational social scientists, we must study and discuss these events in light of our roles as creators and stewards of relevant bases of knowledge. But, how should we do so? Which disruptions should we embrace, and which ones should we resist? Whom do we ally with, and for what purpose? These are not philosophical questions any longer. They are real, and they need to be addressed.

Workshop Structure

The workshop consisted of five sessions, listed below:

1. The disruption of academic publishing: how will Large Language Models affect knowledge production and peer reviewing - a group discussion.
2. Contributions from the open call for submission: four accepted papers and one invited presentation.
3. Building alliances - the IPIE initiative, presented by Dr. Saiph Savage (Northeastern).
4. Embracing the post-API world, new practices in data sharing, a panel with invited speakers Dr. David Lazer (Northeastern, Harvard) and Dr. Kiran Garimella (Rutgers).
5. The Opportunities and Dangers of Large Language Models (LLMs), a panel with invited speakers Dr. João Sedoc (New York University), Dr. Laura Nelson (University of British Columbia), and Dr. Lisa P. Argyle (Brigham Young University)

Discussion on Peer Review

In the first session of the day, as a community building activity, participants first discussed the challenges of the peer

review process in small groups, before sharing them with the whole group. The groups identified various current and future issues of the peer review process for authors, reviewers, and editors. One challenge is the current lack of guidance regarding what type of AI-assisted tools are acceptable to use and what not. For instance, participants suggested that tools such as Grammarly are tolerable, but generating an entire paper with ChatGPT is not. For reviewers, a consensus was that generating a review for a paper with LLMs should not be permitted, but synthesizing a set of reviews into one could be acceptable. Editors could use LLMs to check whether a new paper is a good fit for the journal. However, there is a danger that novel research will be rejected because similar research does not appear in the training data. Another helpful tool for reviewers and editors would check if relevant references exist and are cited in a meaningful way.

What needs to be considered is that reviewing is typically “free labor” by researchers. Often, this labor goes to the benefit of for-profit publishing companies, a recurring controversial issue for the peer reviewing system. The appearance of LLMs will only intensify the ethical aspects of what it means to provide human peer reviewing of high quality.

Another point that was raised pertained to the capabilities of LLMs to make contributions to the knowledge production endeavour. If they can do this, should we accept their contributions? Will this turn us in the role of the interpreter for knowledge generated by LLMs? The interplay of humans and LLMs in knowledge production is a new ethical dilemma to ponder.

Overall, participants concurred that while LLMs – in their current state – should not replace any part of the publication pipeline, they may act as helpful assistants for the peer review process for authors, reviewers, and editors.

Submitted Contributions

Four peer-reviewed papers and one extended abstract were accepted for presentation at the workshop. The first contribution was *Remote Sensing and the Lack of Local Context* by Ingmar Weber (Saarland University). He talked about advances in remote sensing, using images from earth observation satellites, as well as in social sensing, using signals from social media to make it easier to understand what is happening in a particular location while being located thousands of miles away. This is increasingly creating situations where the headquarters of a prominent NGO or UN organization might feel they know the best course of action without requiring direct consultation of those concerned. Weber articulated the dangers of relying simply on satellite data without local contextual knowledge, and shared interesting case studies illustrating these points. During the discussion, Weber mentioned Bellingcat,² as an example of an intelligence gathering organization that strikes the right balance in building collaborations with people on the ground.

The second contribution was *If the data do not speak for themselves, how ought we to speak for the data?* by Ian Van Buskirk, Brian Zaharatos, Aaron Clauset, and Daniel Larremore (University of Colorado Boulder). The authors dis-

cuss the growing challenge of the convergence between scientific scholarship and politics. The authors wonder if they should heed calls from some researchers that it is necessary to abandon the pretense of objectivity and neutrality in research and instead “embrace a political orientation.” To deal with their dilemma, the authors propose an adaptation of Winner’s framework, which asks three questions about a technology: which practices does it enable, encourage, and make inevitable? By evaluating the potential impacts of a technology from an impartial standpoint, it becomes possible to move beyond a simplistic good-versus-bad evaluation and consider the possible, probable, and inevitable outcomes.

As a specific case, the authors examine name-based gender classification (NBGC), as an example of a technology with complex implications. NBGC enables the study of gender inequality in various domains, but also reinforces a gender binary and leads to misgendering or exclusion. The authors argue that despite the drawbacks, the benefits of NBGC outweigh the harms in some context. They conclude by emphasizing the need for researchers to be accountable for their evaluations and decisions, and highlight the difficulty of reconciling divergent ethical frameworks, suggesting open and good-faith discussions to navigate differences of opinion and ensure accountability.

The third contribution was *Why Don’t We Ever Talk About Education When We Talk About Computational Social Science?* by Diliara Valeeva (University of Amsterdam). The author addresses the challenges and opportunities of teaching CSS at the university level while emphasizing the growing demand for CSS expertise in various sectors, which necessitates the preparation of a new generation of CSS practitioners. However, there appears to be a lack of discussion regarding effectively teaching CSS. The author argues that CSS researchers should view teaching and learning CSS as educators, not just researchers. The paper also raises questions about what constitutes a CSS question and how to incorporate open science principles and ethical considerations into CSS education.

The fourth contribution *Hey, ChatGPT, how does CSS research need to change? Research questions in CSS in the wake of the perceived LLM singularity* by Olga Zagovora and Ralf Lämmel (GESIS). The authors discuss the impact of large language models and artificial general intelligence (AGI) on CSS research. The authors propose new research questions and challenges for CSS in light of these developments. While some researchers have expressed concerns and suggested coping strategies, there are also potential benefits of LLMs in CSS research, such as their use in labeling tasks that previously relied on human annotators (see also the fifth contribution). The authors propose three new types of CSS research questions.

First, the study of the agency of humans and AIs aims to understand how humans and AIs complement each other in CSS research and explore questions related to the benefits of human-AI collaboration and conditions under which it is beneficial. Second, the study of how humans are challenged by AIs, which focuses on the challenges faced by humans due to the increasing strength of AI. It examines the effects

²<https://en.wikipedia.org/wiki/Bellingcat>

on individuals regarding uncertainty, confusion, and other disruptions. The authors emphasize the need to distinguish between actual human behavior and AI behavior and measure the impact of AI on human behavior. Third, the study of the “imminent singularity” addresses the potentially existential threat of AIs or a singularity event. The authors argue that CSS research should help inform society about this threat.

The fifth contribution was *From Humans to Machines: Can ChatGPT-like LLMs Effectively Replace Human Annotators in NLP Tasks?* by Surendrabikram Thapa, Usman Naseem, and Mehwish Nasim (Virginia Tech). The paper discusses the increasing demand for NLP applications and the need for large amounts of labeled data to train machine learning models. Human annotators are currently used for text classification, sentiment analysis, and named entity recognition tasks. However, this process is costly and time-consuming, and the quality of annotations can vary.

The emergence of LLMs has led to the exploration of whether these models can effectively replace human annotators in NLP tasks. While LLMs can potentially reduce annotation costs and time, they may only partially replace human annotators in some NLP tasks. LLMs have shown exemplary performance in annotation for various tasks, but they can be biased and struggle with nuances and context-specific meanings in language. Additionally, LLMs require large amounts of high-quality training data, and biases present in the training data can result in biased annotations. LLMs may also need domain-specific knowledge for specific annotation tasks and help with complex linguistic constructions.

LLMs have the potential to automate the annotation process. However, some challenges and limitations must be addressed to ensure the reliability and accuracy of the annotations. Human-in-the-loop validation, domain-specific LLMs, and leveraging LLMs to improve existing annotations are suggested as possible approaches to overcome these challenges.

The papers presented at the workshop can be accessed at the workshop proceedings webpage.³

Presentation and Discussion about IPIE

Dr. Saiph Savage, Assistant Professor in the Khoury College of Computer Sciences at Northeastern University, presented the *International Panel for the Information Environment (IPIE)*⁴ organization, with which she is affiliated as a co-chair of the membership panel. The IPIE aims to establish a global scientific effort to independently analyze systems of information manipulation and bias and evaluate the best policy solutions for addressing threats to that environment. Within this broader aim, one of the main goals of the IPIE is to engage researchers from the Global South. In her talk, Dr. Savage discussed strategies for building alliances, and participants brainstormed about methods for ensuring the credibility and reliability of information contributed by crowd experts, including validation processes and peer verification. Discussion at the workshop centered around the

challenges to ensuring that the IPIE could simultaneously engage with existing organizations that organize and encourage knowledge sharing amongst researchers (e.g., AAAI or ACM) while at the same time ensuring that scholars left out of these “traditional” sites of knowledge sharing are still heard and their contributions included in the quest for improving the global information environment.

Panel on Data Sharing

Dr. David Lazer, Professor of Political Science and Computer Sciences at Northeastern University, and Dr. Kiran Garimella, Assistant Professor in the School of Communication and Information at Rutgers University, discussed emerging challenges for data gathering in the “Post-API” age. Dr. Lazer is currently the PI in a five-year, \$15.7 million research grant funded by the NSF that aims to establish a National Internet Observatory,⁵ to observe online human behavior across many digital platforms. The Observatory will provide access to privacy-preserving human data to third-party researchers to study questions of broad interest. Dr. Garimella’s current research is focused on building innovative, opt-in, privacy-preserving data-collection tools for platforms like WhatsApp to develop interventions to stop or at least slow the spread of misinformation. Both Dr. Lazer and Dr. Garimella discussed the difficulties associated with constructing methods for data collection, with a particular focus on 1) the difficulties of collection in both mobile and desktop settings, 2) challenges with creating collections that are globally representative, 3) difficulties with the recruitment of participants, 4) security and ethical challenges that may inhibit deep research, and 5) the increased difficulty of such work as social media platforms become more hostile to external data collection efforts. Also discussed by Dr. Garimella, and echoed by Dr. Lazer, was the high-risk nature of conducting major data collection efforts for junior CSS scholars and what could be done to address these concerns both in the career trajectory of individual researchers, such as evaluating their contribution in the tenure review process, as well as in the trajectory of research projects, such as the potential to recycle (or reuse) data (Hemphill et al. 2022) and infrastructure in subsequent efforts.

Panel on Large Language Models

The panel on large language models included Dr. João Sedoc, Assistant Professor of Information Systems in the Department of Technology, Operations and Statistics at New York University Stern School of Business, Dr. Laura Nelson, Assistant Professor of sociology at the University of British Columbia, and Dr. Lisa P. Argyle, Assistant Professor of Political Science at Brigham Young University. Panelists had a robust discussion about the dangers and opportunities of LLMs and their ability to capture a rich and complex web of social biases in the data they are trained on. Panelists agreed that whether this ability should be considered a danger or an opportunity centers mainly on how the LLM is used. For example, Dr. Sedoc emphasized that gender and racial biases in LLMs could significantly increase existing patterns

³<https://workshop-proceedings.icwsm.org/index.php?year=2023>

⁴<https://www.ipie.info>

⁵<https://nationalinternetobservatory.org>

of social inequality if LLMs were used in a decision-making context or to provide knowledge to end users uncritically. In contrast, Dr. Nelson and Dr. Argyle emphasized, with pointers to their existing work, that when interrogated using a sociological and political scientific lens, the biases embedded in LLMs can help us to understand better, simulate, and measure society in ways that may ultimately challenge existing structures of power. Referenced in the discussion was Nagel’s *The View from Nowhere* (Nagel 1989), where the panelists emphasized that the construction of a single, “un-biased” AI is philosophically unreasonable. This also leads to the problematic trade-off between computational precision and societal insights when models are debiased, which some of the panelists perceived as exchanging one kind of bias for another.



Figure 1: Many of the invited panelists were unable to attend the workshop physically, but the setup made it easy for them to engage with the audience. The photo shows Dr. Savage’s zoom feed projected in the workshop room, as well as the room video feed displaying the audience to her.

The Hybrid Format

DARE was co-located with AAAI ICWSM 2023, which took place in Limassol, Cyprus.⁶ While the majority of workshop participants were physically on-site, most of the invited speakers and panelists joined remotely via Zoom. When considering the risks of climate change and our responsibility for mitigating it, the decision to discourage long flights for short engagement in an one-day workshop seems appropriate. In-person workshops and conferences are however very valuable to young researchers, as sites for sharing their work, getting feedback from peers, engaging with senior researchers in the field, and establishing a network of future collaborators. To facilitate all of the above, we were prepared with the right technical setup to support virtual speakers and in-person collaboration, as shown in Figure 1

⁶<https://www.icwsm.org/2023/index.html/>



Figure 2: The workshop had interactive sessions in which the participants discussed in small groups. Here they are photographed during the session with Dr. Savage. Interjecting in-person interaction in a virtual session proved a good strategy for engagement.

and Figure 2. We believe that this hybrid format made it easy to invite senior researchers from outside of our community, for whom ICWSM is not their primary conference.

Conclusion

Many academic venues are now exploring the problems around ethics in AI with respect to its impacts in the public, but we are interested in expanding this discussion to the ethical principles of our own work as CSS researchers and practitioners. We (the organizers of this workshop, listed as authors in this paper) are invested in the success of ICWSM as a premier venue for CSS, therefore, we believe that these conversations should be an integral part of ICWSM gatherings. We hope to contribute toward diversifying participation in ICWSM (Nurse et al. 2021) and provide mentoring and networking opportunities for young researchers. Given the success of the first iteration of this workshop, we are energized to continue offering it in the coming years, while finding more creative ways to broaden the participation either in-site or virtually and deepen the reach of our discussions.

References

Hemphill, L.; Pienta, A.; Lafia, S.; Akmon, D.; and Bleckley, D. A. 2022. How do properties of data, their curation, and their funding relate to reuse? *Journal of the Association for Information Science and Technology*, 73(10): 1432–1444.

Nagel, T. 1989. *The view from nowhere*. OUP.

Nurse, J. R.; Vargas, F.; Hassan, N.; Wang, D.; Andriotis, P.; Ghenai, A.; Jaidka, K.; Mustafaraj, E.; Joseph, K.; and Foucault Welles, B. 2021. Towards A Diverse, Inclusive, Accessible and Equitable AAAI International Conference on Web and Social Media (ICWSM). *AAAI ICWSM Workshop Proc.*