

LOGOS - A Multidisciplinary UCR initiative for modeling and improving online dialogue

PIs: Emiliano De Cristofaro, CSE emilianodc@cs.ucr.edu
Kevin M. Esterling, SPP and Political Science kevin.esterling@ucr.edu
Richard M. Carpiano, SPP, richard.carpiano@ucr.edu
Vassilis Tsotras, CSE, vtsotras@cs.ucr.edu
Benjamin Stewart, ECE, bstewart@ucr.edu
Shaun Bowler, Political Science, shaun.bowler@ucr.edu
Michalis Faloutsos, CSE, michalis@cs.ucr.edu
Vagelis Papalexakis, CSE, epapalex@cs.ucr.edu
Diogo Ferrari, Political Science, diogof@ucr.edu

Introduction

How can we model, quantify, and improve online dialogue? Given the direct impact that online interactions have on polarization, misinformation, democracy, and radicalization, this is an existential societal mandate.

In its 30+ years of life, the Web has contributed to advancing modern societies in unimaginable ways. Thanks to the progress in Internet infrastructure and the increased affordability of user devices, information and knowledge have never been easier to spread and retrieve. Online social interactions have become an ever-increasing means of interpersonal engagement, as the supporting and enabling role of the Web in our everyday lives has never been clearer than during the recent COVID-19 pandemic.

Despite its positive impacts, the Web has also enabled the weaponization of information. Malevolent actors have used it to build toxic communities [1], spread hateful and extremist ideologies [2], perpetrate aggression [3], and wage information warfare [4]. Despite its increasingly relevant societal impact [4], research efforts around information weaponization have only scratched the surface. We anticipate that this problem will only be amplified in the future with the proliferation and accessibility of AI generative models (e.g., ChatGPT, Gemini, etc.). We lack a comprehensive, holistic understanding of the socio-technical issues at the heart of the problem, as well as multi-faceted techniques that can be leveraged to build holistic mitigation tools.

Information weaponization unfolds across multiple, entangled, and interdependent communities. Yet, available countermeasures look predominantly at platforms in isolation (at a message and account granularity) and almost exclusively from the point of view of a single discipline. This applies to methodological approaches as well. For instance, social science-focused work has heavily relied on small sample qualitative analysis and randomized experiments, which, though insightful, are limited in nature, size, and generalizability; whereas, computer science research mainly uses quantitative approaches, which might fail to capture the complex societal and behavioral nuances of the very issues it sets out to study. Furthermore, recent reviews have

identified the need for a) applying a broader range of theoretical approaches to create interventions to address mis-/disinformation [5] and b) focusing greater attention on how to best evaluate the effectiveness of online interventions [6,7].

With this motivation in mind, several UCR faculty have recently begun joining forces in establishing a cross-discipline interest group around information weaponization, aimed at modeling and improving online dialogue. We have called this initiative, LOGOS, from the Greek noun “logos” (derived from the verb legein, ‘to say’ something significant). Over the years, the term logos developed a wide variety of senses, including ‘reason,’ ‘reasoning power,’ ‘principle,’ and ‘speech.’¹

The LOGOS initiative is in its early stages of development, with “founding” members including Emiliano De Cristofaro and Michalis Faloutsos (Computer Science and Engineering) and Kevin Esterling (Public Policy and Political Science), with close and active involvement of Prof. Yue Dong, Evangelos Papalexakis, Mariam Salloum, Vassilis Tsotras (Computer Science) as well as Shaun Bowler and Diogo Ferrari (Political Science), Benjamin Stewart (Electrical and Computer Engineering), and Richard Carpiano (Public Policy).

Since early January, we have held a bi-weekly meeting, led by one rotating member, to discuss academic papers, relevant news, and initiatives; we foster cross-disciplinary collaborations and will move to explore funding opportunities. This trajectory is fully in line with the development path of efforts in other inter-disciplinary efforts at this stage of development. Our next step is to build on these promising foundations and take the LOGOS initiatives to the next level, aiming to 1) build a solid, well-established, and impactful center within UCR and 2) support dedicated funding efforts, as discussed below.

Objectives

As discussed above, our main objective is to foster and grow the LOGOS initiative. We will leverage our unique cross-disciplinary collaborations and expertise in ways that synthesize the insights of quantitative and qualitative methods. The collaborative protocols we establish will offer a sustainable approach to high-quality, high-impact research.

More specifically, we aim to:

1. Create a center or institute within UCR and increase its visibility both internally and externally;
2. Seek extramural grant funding to support PhD students and post-doctoral researchers and help bootstrap, lead, and collaborate on ambitious, long-term, cross-disciplinary research initiatives.

¹ <https://www.rep.routledge.com/articles/thematic/logos/v-1>

3. Seek internal (by consolidating existing grants) and extramural grant funding to support dissemination and outreach efforts.
4. Seek collaborations with relevant industry partners (e.g., in the form of PhD student internships, academic collaborations, contracts, etc.)

Approach

We are applying for the CoR grant and seeking \$10,000 in seed-funding support to help achieve the three objectives above.

More specifically, we intend to leverage CoR support for a combination of equipment, event planning, and travel expenditures.

Equipment (\$1,500). We believe that the ability to conduct hybrid meetings and public events will tremendously help us with our objectives. To this end, we wish to purchase an Owl Camera, which provides a 360-degree camera, high-quality ambient microphones and speakers, featuring automatic speaker focus and smart zooming. At the moment, we either all have to be in the same room (which is hard given the size and nature of the interest group) or all remote on Zoom.

The hybrid setup will also enable us to invite external researchers who are not at UCR to participate and present at the meetings, thus allowing us to expand in scope, relevance, and visibility.

Public Launching Event (\$4,500). We would like to officially launch the LOGOS initiative with a 1-day event hosted at UCR and live-streamed on Zoom/YouTube, featuring a mix of internal presentations and invited external speakers. This will allow us to establish and further our visibility and guide our next steps via open brainstorming sessions and the participation (and guidance) of key external figures from industry, academia, policy, and nonprofits. To this end, we need support to provide coffee breaks and refreshments, as well as cover expenses for keynote speakers.

Travel Expenditures (\$4,000). Finally, we plan to support travel to conferences and events in this initial phase of our initiative to support networking and scoping. More precisely, researchers are typically at a disadvantage with respect to attending events in other disciplines as they do not publish in their venues, or those events are not considered core to their research; thus, these types of travel expenditures are much harder to cover from research grants. Examples of events we plan to attend include the International Conference on Computational Social Science (IC2S2), the International Conference on Web and Social Media (ICWSM), the annual meeting of the American Association for the Advancement of Science and the annual research meeting of the Association for Public Policy and Management. As LOGOS members become more tightly embedded with the different interdisciplinary research communities, our intent is to pursue attracting and organizing one of such scientific meetings in Riverside, to increase the visibility of LOGOS and UCR.

Relevance to extramural grant application

As mentioned, we plan to identify and target cross-disciplinary funding opportunities. There are already examples of very relevant program solicitations, e.g., the Growing Convergence Research (GCR) from the National Science Foundation ([NSF 24-527](#)). We will further focus on other related CFPs coming from NSF's CISE (Computer and Information Science and Engineering) and SBE (Social, Behavioral, and Economic Sciences) directorates, as well as from DARPA, etc.

While we are not yet at a point where we have a solid structure to pursue such ambitious, multidisciplinary grants, we do believe that, with the support of a CoR grant, we will soon be able to set forward a coherent, ambitious, and impactful research agenda. This, in turn, will enable us to pursue these kinds of grants. In fact, note that the NSF 24-527 solicitation will have another deadline in February 2025.

References

- [1] S. Mohan, A. Guha, M. Harris, F. Popowich, A. Schuster, and C. Priebe. The impact of toxic language on the health of reddit communities. In Canadian Conference on Artificial Intelligence, 2017.
- [2] N. Chetty and S. Alathur. Hate speech review in the context of online social networks. *Aggression and violent behavior*, 40, 2018.
- [3] K. R. Allison. Social norms in online communities: formation, evolution and relation to cyber-aggression. In ACM CHI, 2018.
- [4] J. Prier. Commanding the Trend: Social Media as Information Warfare. *Strategic Studies Quarterly*, 11(4), 2017.
- [5] Ziemer CT, Rothmund T. Psychological underpinnings of misinformation countermeasures: A systematic scoping review. *Journal of Media Psychology: Theories, Methods, and Applications*. 2024 Jan 23.
- [6] Tay LQ, Lewandowsky S, Hurlstone MJ, Kurz T, Ecker UK. A focus shift in the evaluation of misinformation interventions. *Harvard Kennedy School Misinformation Review*. 2023 Oct 5.
- [7] Blair RA, Gottlieb J, Nyhan B, Paler L, Argote P, Stainfield CJ. Interventions to counter misinformation: lessons from the global north and applications to the global south. *Current Opinion in Psychology*. 2023 Nov 9:101732.