Research Statement - July 2025

Aron Culotta | Professor of Computer Science | Tulane University
Director, Center for Community-Engaged Artificial Intelligence

Highlights:

- $5.03M total funding ($4.19M external): 8 NSF (7 as PI, 1 as co-PI); 1 NEH (co-PI); 1 LA Board of Regents (PI). [NSF Profile]
- 62 rigorously refereed publications (18 journal articles and 44 conference proceedings) at top venues in artificial intelligence, natural language processing, and web and social media analysis (AAAI, HLT/NAACL, ACL, EMNLP, ICWSM, WWW, CHI, CSCW, ICML, KDD).
- 3 outstanding paper awards (AAAIx2, CSCW)
- Google Scholar statistics (07/2025): 8,221 citations h-index=39 i10-index=60

My primary research interests span **natural language processing**, **large language models**, **machine learning**, **causal inference**, and **social network analysis**. As these areas of artificial intelligence have become critical throughout society, my research has similarly evolved to have strong cross-disciplinary, funded collaborations with experts in **public health, emergency management, marketing, communications, law,** and **political science**. This **computational social science** research has not only provided me the opportunity to advance AI in a diverse range of applications with societal importance, but it has also helped illuminate key methodological shortcomings in AI that have motivated more foundational research. This inspired me to lead the creation of the Center for Community-Engaged Artificial Intelligence, a cross-disciplinary research initiative focusing on new ways of designing AI systems in partnership with diverse stakeholders. Below, I summarize my past, present, and future research directions in these areas, organized by four overarching questions.

Q1: What can we learn about society by studying social media?

The field of **social media analysis** combines natural language processing, language models, machine learning, statistics, and databases to explore what we can infer from the behavior of social media users. Such analysis can provide insights into public health, finance, politics, social unrest, and natural disasters. In addition to these societal benefits, the immediacy, openness, and volume of social media present great opportunities and challenges to natural language processing research.

The goals of this research are: (1) to leverage this unprecedented source of natural language data to advance research in automated processing of informal human communication; and (2) to apply these techniques to analyze trends in social media and produce socially beneficial technology. My work analyzes millions of users and their messages from sites such as Twitter, Instagram, Reddit, and Facebook for a number of socially-relevant tasks, summarized below.

Public health Our work was among the first to demonstrate that Twitter can be used to track temporal trends in influenza-like illness. This requires developing NLP algorithms to detect self-reported symptoms online and has contributed to a new branch of science using social media to understand public health. Subsequently, we have estimated county health statistics from Twitter, investigated how attitudes towards electronic cigarettes have evolved, and developed matching algorithms to better adjust for confounding variables in online observational studies. Publications: [1, 2, 3, 4]

Emergency management A key challenge during natural disasters like hurricanes is measuring and forecasting evacuation behavior. While traditional approaches rely on small post-disaster surveys to understand evacuation decision making, we instead combine social media, traffic sensors, and government surveys to enable real-time monitoring and forecasting of evacuations. Additionally, by building language models to analyze the content of evacuation-related messages, we can discover self-reported reasons for either evacuating or not. Publications: [5, 6, 7, 8]

Consumer perception Measuring public perceptions and how they change over time is a central problem in marketing, public health, and politics. While traditional methods rely on costly surveys, online social networks offer an attractive alternative: real-time perceptions can be estimated from public, online activity and compared with an entity's communications to quantify how public messaging affects perception. While prior algorithmic approaches rely purely on text-based sentiment analysis, we develop novel methods based on the insight that an entity's online social connections are indicative of how they are perceived (e.g., "birds of a feather flock together"). Thus, rather than typical one-dimensional measures of sentiment, we instead investigate public perception with respect to multiple characteristics of an entity (e.g., is it seen as pro-environment, pro-health, etc.). We applied our approach to analyze the phenomenon of "greenwashing," a deceptive marketing practice in which firms market their products or policies as more environmentally friendly than they truly are. For example, we have found that we can accurately estimate how environmentally friendly a brand is perceived based on its Twitter followers. Publications: [9, 10, 11, 12].

Relevant funding:

- **III: Small: Quantifying Multifaceted Perception Dynamics in Online Social Networks**. role: PI, 2016–2019, NSF-IIS #1618244, $471,992 co-PI: Jennifer Cutler (Marketing), Northwestern-Kellogg
- **Collaborative Research: Predicting Real-time Population Behavior during Hurricanes Synthesizing Data from Transportation Systems and Social Media**. role: PI, 2019–2022, NSF-HDBE #1917112, $299,995 co-PI: Samiul Hasan (Civil Engineering), Claire Knox (Emergency Management), Naveen Eluru (Civil Engineering), University of Central Florida

Q2: How do recommendation systems and social media influence news engagement?

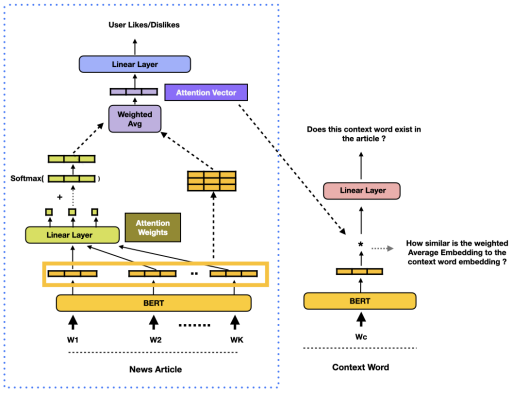

Figure 1: A multi-task attention network for more diverse news recommendation [13].

Computer algorithms are widely used by online sites to determine the content users see, for example, by curating news articles or recommending social media posts. These algorithms are primarily designed to improve user experience by showing to users the content that they are likely to be interested in. However, there is growing evidence that these algorithms may have unintended side effects. For example, by showing users only content that conforms with their preexisting perceptions and beliefs, users may receive a biased subset of all content, possibly increasing intellectual isolation, a phenomenon known as a "filter bubble." In the first branch of this research, we investigate how and why filter bubbles form and develop new algorithms to prevent them. The technical approach of the project focuses on three enhancements to content recommendation algorithms: 1) improving transparency by informing the users of their reading habits, what the recommendation model thinks of them, and why particular items are recommended; 2) supporting rich user interactions by enabling the user to provide feedback on model predictions and explanations; and 3) designing new neural network-based recommendation systems that reduce political homogeneity of recommendations (Fig. 1).

In a subsequent project, we focus on how people interact with news online and how their behaviors evolve over time. First, we develop neural network language models to identify various types of news engagement behaviors and their progression stages, innovating advanced language and user modeling techniques to predict future behavior patterns. Second, we establish a technical framework for estimating causal relationships between different news engagement behaviors, combining natural language processing with causal inference methods to estimate treatment effects from observational

data. Third, the project tests socio-technical hypotheses regarding strong positions on issues, trust, and information reliability using this framework. Publications: [14, 15, 16, 13, 17, 18]

Relevant funding:

- **Collaborative Research: HCC: Small: Socio-linguistic modeling to understand the long-term dynamics of news engagement in online media**. role: PI, 2024–2027, NSF-IIS #2333537, $277,178 ($600,000 total) co-PI: Mustafa Bilgic (CS), Matthew Shapiro (Political Science), Illinois Institute of Technology
- **EAGER:AI-DCL: Understanding the Relationship between Algorithmic Transparency and Filter Bubbles in Online Media**. role: co-PI, 2019–2021, NSF-IIS #1927407, $299,871 co-PI: Mustafa Bilgic (CS), Matthew Shapiro (Political Science), Illinois Institute of Technology

Q3: How can we design robust NLP systems for social-scientific studies?

The applied research described above has in turn motivated a number of methodological contributions that focus on how to improve NLP algorithms to generally be more robust, reliable, and appropriate for answering questions about society.

Robust text classification Conducting health studies using online data requires accurately classifying user data. Furthermore, the errors should not be correlated with important variables like race or age, as this will invalidate the conclusions of the health study. In this line of research, we develop new text classification algorithms that are robust to omitted variable bias, which occurs when important demographic variables confound the text classification algorithm. By combining ideas from causal inference with machine learning, we introduce methods to remove the impact of confounding variables on classification, even if the confounding variables are measured with noise. Publications: [19, 20, 21].

Learning from label proportions As we worked more closely with social scientists, it became apparent that to conduct valid health studies using online data, it is necessary to identify socio-demographic information for users in the sample. As these are typically not provided by the user, we have developed new machine learning algorithms to infer the location and demographics of social media users based on textual content and social connections. We have subsequently used this information to reduce bias in observational studies that use social media. Two publications [23, 24] were given best paper awards; one was extended from an AAAference publication into a JAIR journal article. The innovation of this work is that it does not rely on user-annotated data; instead, models are built from population statistics. This has led to several contributions to a branch of machine learning called "Learning from Label Proportions." This is a significant deviation from common practice, which requires first manually labeling thousands of training instances with the desired class labels. Publications: [23, 25, 26, 24, 27, 28]

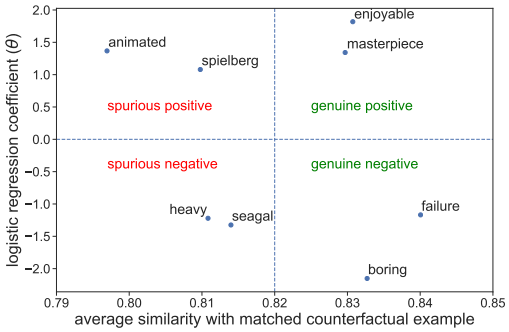

Figure 2: Using counterfactual examples to reduce spurious correlations in text classification [22].

From correlation to causation Nearly all machine learning methods learn correlations between inputs and outputs. But, we know that causal knowledge can be more reliable, efficient, and robust to data drift. In this line of research, we investigate machine learning methods that borrow ideas from the causal inference literature to improve robustness. For example, we find that many text classifiers rely on spurious correlations between words and class labels (e.g., "Steven Spielberg" correlates with positive movie reviews, but does not semantically indicate positive sentiment). We introduce several approaches to address this problem, including generating "counterfactual" training examples (e.g.,

negative reviews mentioning Spielberg) to reduce the impact of such words (Fig. 2). In another line of work, we develop individual treatment effect estimates to quantify the impact each word has on the final label assigned to a document. By making NLP methods more robust, we can increase their utility in computational social science studies. Publications: [29, 22, 30, 31]

Relevant funding:

- **Collaborative Research: III: Small: Reducing Classifier Bias in Social Media Studies of Public Health**. role: **PI**: 2015–2018, NSF-IIS #1526674, $304,725 ($499,251 total). co-PI: Sherry Emery, UIC, Institute for Health Research and Policy

Q4: How can we collaboratively develop responsible AI with diverse stakeholders?

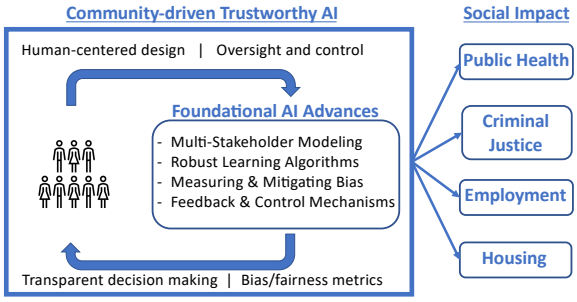

Figure 3: Overview of Center research projects.

My work with social scientists has shown me the potential of artificial intelligence to benefit society in numerous areas. However, ensuring that the benefits of AI are shared equitably among society and reducing any potential harmful effects is of critical importance given (i) the concentration of technical capacities into the hands of a few, often highly educated practitioners; (ii) long-standing socio-technical challenges of algorithmic discrimination, transparency, and accountability that underpin all AI systems; leading to (iii) a pervasive mistrust of AI technology by the public. These trends indicate that AI will succeed only with the trust and support of the communities it affects, especially those from historically underserved groups.

Motivated by these factors, I led the creation of the Tulane Center of Excellence for Community-Engaged Artificial Intelligence (CEAI). Established in 2022 with seed funds from Tulane's Office of Research, the Center leads cross-disciplinary research, facilitates project-based learning experiences, conducts extensive community outreach, hosts workshops and symposiums, and awards seed grants to support AI research across the university. A recent online profile summarizes this work.

The Center pursues AI approaches (1) that are demonstrably fair, transparent, and accountable, (2) that establish meaningful relationships with diverse communities in all stages of the AI process, from design through deployment, and (3) that augment, rather than replace, human interaction, i.e., are human centered. The Center's overarching goal is to conduct cutting edge research into impactful AI systems that can be responsibly transitioned into local communities. Achieving this vision requires a transformative approach to AI design. The Center is a multi-disciplinary team of technologists, designers, social scientists, and community partners for examining and affording new evidence-based socio-technical frameworks to create and deploy AI systems.

In order to scaffold these community-driven projects, I have partnered with Tulane's Center for Public Service to develop a Senior Capstone Service Learning course, in which I lead undergraduates in year-long AI and data science projects, working closely with non-profits to support their mission. Not only has this course provided students with enriching educational experiences, learning both about the work of non-profits and how to deploy real-world AI solutions, but it has also allowed the Center to build long-term relationships with non-profits, establishing trust and technical infrastructure that supports more long-term research projects.

Two ongoing Center projects that I lead exemplify this approach. In the first, we have partnered with Court Watch NOLA, the country's oldest court-watching organization, which trains and deploys hundreds of volunteers annually to objectively monitor court proceedings and to report key variables

that are only accessible by being present in the courtroom. Since 2020, my students and I have worked with Court Watch NOLA to create new data infrastructure that enables real-time analysis of the criminal court system. With this infrastructure in place, CWN has been able to discover and report on emerging trends in the court system. Building on this foundation, I am now leading a team of computer scientists, sociologists, and legal scholars to expand on this work, made possible through a grant from NSF's Smart and Connected Communities program. The goal of the proposed work is to expand the utility, accessibility, and analytical rigor of this type of system, so that it may serve as a transferable template for other jurisdictions. In particular, we are (i) developing new methods and best practices for integrating multiple community stakeholders across the data science pipeline; (ii) developing foundational neural network models to better represent case trajectories in criminal court; (iii) proposing new methods to perform causal inference in legal domains; (iv) applying the resulting framework to identify determinants of outcomes in the criminal legal system.

In a second project, I am working with Eye on Surveillance, a civic transparency group in New Orleans, to build a system that uses large language models to support rich querying of the transcripts of City Council meetings. Doing so would allow citizens to better monitor civic activities, supporting greater transparency and accountability. Given the importance of the reliability of generated answers in this context, we have been developing semi-automated tools that use specially trained LLMs to validate the output of other LLMs. We are pursuing additional funding to broaden the scope of this project to design general-purpose LLMs that enhance civic transparency across multiple dimensions.

Through these and other projects, CEAI is leading cutting edge research into impactful AI systems that can be responsibly transitioned into local communities. Our goal is to become an international leader in community-engaged artificial intelligence

Relevant funding:

- **SCC-IRG Track 2: Supporting Transparency and Equity in the Criminal Legal System through a Community-Driven Digital Platform**. role: PI: 2024–2027, NSF-SCC #2427237, $1,499,743. co-PI: Andrea Boyles (Tulane Sociology), Nicholas Mattei (Tulane CS), Andrea Armstrong (Loyola Law), Darrin Browder (Court Watch NOLA)

Additional Ongoing Research

- **Industry-University Research Partnerships:** Building on an NSF Planning Grant, I am leading efforts to build an Industry-University Cooperative Research Center focused on AI and Health. With partners as ULL, Georgia Tech, and University of Florida, we are partnering with health organizations like Ochsner and Humana to establish a research center focused on human-centered AI technologies to make quality healthcare ubiquitous, equitable and accessible.
- **Causal inference with LLMs:** I am investigating ways to adapt LLMs to estimate treatment effects when treatment, outcome, or confounding variables are represented with text.
- **Digital Humanities:** Through an NEH Digital Humanities Enhancement Grant, I am working with Alexis Culotta (Art History) to develop social network analysis algorithms that can track the spread of influence through artistic production networks.
- **Online Perceptions of Violent crime:** With Katherine Theall of the Violence Prevention Institute, I am developing NLP methods to mine online data to understand how circumstances of violent crime influences the perception and reactions of online communities.
- **Integrating Algorithmic Fairness with Audit Studies:** With Patrick Button (Economics) and Nicholas Mattei (CS), I am working to develop new ways of using audit studies (e.g., resume correspondence studies) to better train and measure the fairness of classification models.

References

[1] **Aron Culotta**. Lightweight methods to estimate influenza rates and alcohol sales volume from Twitter messages. *Language Resources and Evaluation, Special Issue on Analysis of Short Texts on the Web*, 2013. (1.7 impact factor; 118 citations in Google Scholar).

[2] **Aron Culotta**. Estimating county health statistics with Twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2014. (23% accepted; 195 citations in Google Scholar).

[3] Virgile Landeiro Dos Reis and **Aron Culotta**. Using matched samples to estimate the effects of exercise on mental health from Twitter. In *Twenty-ninth National Conference on Artificial Intelligence (AAAI)*, 2015. (531/1991=27% accepted; 65 citations in Google Scholar).

[4] Elaine Cristina Resende and **Aron Culotta**. A demographic and sentiment analysis of e-cigarette messages on Twitter. In *Workshop on Computational Health Science at the 6th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*, 2015.

[5] Tanmoy Bhowmik, Naveen Eluru, Samiul Hasan, **Aron Culotta**, and Kamol Chandra Roy. Predicting hurricane evacuation behavior synthesizing data from travel surveys and social media. *Transportation Research Part C: Emerging Technologies*, 165:104753, 2024. (7.6 impact factor).

[6] Xintian Li, Samiul Hasan, and **Aron Culotta**. Identifying hurricane evacuation intent on twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, pages 618–627, 2022.

[7] Kamol Chandra Roy, Samiul Hasan, **Aron Culotta**, and Naveen Eluru. Predicting traffic demand during hurricane evacuation using real-time data from transportation systems and social media. *Transportation Research Part C: Emerging Technologies*, 131:103339, 2021. (7.6 impact factor; 68 citations in Google Scholar).

[8] Xuan Song, Haoran Zhang, Rajendra A. Akerkar, Huawei Huang, Song Guo, Lei Zhong, Yusheng Ji, Andreas Lothe Opdahl, Hemant Purohit, Andre Supkin, Akshay Pottathil, and **Aron Culotta**. Big data and emergency management: Concepts, methodologies, and applications. *IEEE Transactions on Big Data*, 2020. (7.5 impact factor; 44 citations in Google Scholar).

[9] **Aron Culotta**, Jennifer Cutler, and Junzhe Zheng. Finding truth in cause-related advertising: A lexical analysis of brands' health, environment, and social justice communications on Twitter. *The Journal of Values-Based Leadership*, 8(2), 2015.

[10] **Aron Culotta** and Jennifer Cutler. Mining brand perceptions from Twitter social networks. *Marketing Science*, 2016. (4.0 impact factor; 352 citations in Google Scholar).

[11] Zhao Wang, Jennifer Cutler, and **Aron Culotta**. Are words commensurate with actions? Quantifying commitment to a cause from online public messaging. In *Proceedings of the ACUMEN Workshop at the 17th IEEE International Conference on Data Mining (ICDM)*, 2017.

[12] Jennifer Cutler and **Aron Culotta**. Using online social networks to measure consumers' brand perception. *Applied Marketing Analytics*, 2(4):312–321, 2017.

[13] Karthik Shivaram, Ping Liu, Matthew Shapiro, Mustafa Bilgic, and **Aron Culotta**. Reducing cross-topic political homogenization in content-based news recommendation. In *Proceedings of the 16th ACM Conference on Recommender Systems*, 2022. (39/231=26.9% accepted; 11 citations in Google Scholar).

[14] Shivaram, Karthik, Mustafa Bilgic, Matthew A Shapiro, and **Culotta, Aron**. Characterizing online criticism of partisan news media using weakly supervised learning. In *Proceedings of the International Conference on Web and Social Media*, 2024.

[15] Shivaram, Karthik, Mustafa Bilgic, Matthew A Shapiro, and **Culotta, Aron**. Forecasting political news engagement on social media. In *Proceedings of the International Conference on Web and Social Media*, 2024.

[16] Liu, Ping, Shivaram, Karthik, **Culotta, Aron**, Matthew A Shapiro, and Mustafa Bilgic. How does empowering users with greater system control affect news filter bubbles? In *Proceedings of the International Conference on Web and Social Media*, 2024.

[17] Siva K Balasubramanian, Mustafa Bilgic, **Aron Culotta**, Libby Hemphill, Libby, Anita Nikolich, and Matthew A Shapiro. Leaders or followers? a temporal analysis of tweets from ira trolls. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, pages 2–11, 2022.

[18] Ping Liu, Karthik Shivaram, Matthew Shapiro, **Aron Culotta**, and Mustafa Bilgic. The interaction between political typology and filter bubbles in news recommendation algorithms. In *Proceedings of the Web Conference 2021*, 2021. (357/1736=20.6% accepted; 59 citations in Google Scholar).

[19] **Aron Culotta**. Reducing sampling bias in social media data for county health inference. In *JSM Proceedings*, 2014. (52 citations in Google Scholar).

[20] Virgile Landeiro and **Aron Culotta**. Robust text classification in the presence of confounding bias. In *Thirtieth National Conference on Artificial Intelligence (AAAI)*, 2016. (549/2132=26% accepted; 52 citations in Google Scholar).

[21] Virgile Landeiro and **Aron Culotta**. Controlling for unobserved confounds in classification using correlational constraints. In *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017)*, 2017.

[22] Zhao Wang and **Aron Culotta**. Identifying spurious correlations for robust text classification. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, 2020. (66 citations in Google Scholar).

[23] Reid Priedhorsky, **Aron Culotta**, and Sara Y. Del Valle. Inferring the origin location of tweets with quantitative confidence. In *17th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW)*, 2014. (134/497=27% accepted; **Best Paper Honorable Mention** (given to 15 of 134 accepted papers); 141 citations in Google Scholar).

[24] **Aron Culotta**, Nirmal Kumar Ravi, and Jennifer Cutler. Predicting Twitter user demographics using distant supervision from website traffic data. *Journal of Artificial Intelligence Research*, 55:389–408, 2016. (5 impact factor; 74 citations in Google Scholar).

[25] Ehsan Mohammady Ardehaly and **Aron Culotta**. Inferring latent attributes of Twitter users with label regularization. In *Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics (HLT/NAACL)*, 2015. (117/402=29% accepted; 32 citations in Google Scholar).

[26] **Aron Culotta**, Nirmal Ravi Kumar, and Jennifer Cutler. Predicting the demographics of Twitter users from website traffic data. In *Twenty-ninth National Conference on Artificial Intelligence (AAAI)*, 2015. (531/1991=27% accepted, **Outstanding Paper Honorable Mention** (given to 1 of 531 accepted papers); 216 citations in Google Scholar).

[27] **Aron Culotta**. Training a text classifier with a single word using Twitter lists and domain adaptation. *Social Network Analysis and Mining*, 6(1):1–15, 2016. (2.3 impact factor).

[28] Ehsan Mohammady Ardehaly and **Aron Culotta**. Learning from noisy label proportions for classifying online social data. *Social Network Analysis and Mining*, 8(1):2–22, 2018. (2.3 impact factor; 12 citations in Google Scholar).

[29] Zhao Wang and **Aron Culotta**. When do words matter? Understanding the impact of lexical choice on audience perception using individual treatment effect estimation. In *Proceedings of the Thirty Third National Conference on Artificial Intelligence (AAAI 2019)*, 2019. (1150/7095=16% accepted); 14 citations in Google Scholar.

[30] Wang, Zhao, Kai Shu, and Aron Culotta. Enhancing model robustness and fairness with causality: A regularization approach. In *EMNLP First Workshop on Causal Inference & NLP*, 2021. (27 citations in Google Scholar).

[31] Zhao Wang and **Aron Culotta**. Robustness to spurious correlations in text classification via automatically generated counterfactuals. In *Proceedings of the Thirty Fifth National Conference on Artificial Intelligence (AAAI 2021)*, 2021. (1,692/7,911=21% accepted; 92 citations in Google Scholar).