# Theory-Grounded Computational Text Analysis

**Arya D. McCarthy**[*♦] and **Giovanna Maria Dora Dore**[*♠]
[♦]Center for Language and Speech Processing, Johns Hopkins University
[♠]Krieger School of Arts and Sciences, Johns Hopkins University

## Abstract

In this position paper, we argue that computational text analysis lacks and requires organizing principles. A broad space separates its two constituent disciplines—natural language processing and social science—which has to date been sidestepped rather than filled by applying increasingly complex computational models to problems in social science research. We contrast descriptive and integrative findings, and our review of approximately 60 papers on computational text analysis reveals that those from *ACL venues are typically descriptive. The lack of theory began at the area's inception and has, over the decades, grown more important and challenging. A return to theoretically grounded research questions will propel the area from both theoretical and methodological points of view.

## 1 Introduction

Computational text analysis methods—an umbrella combining natural language processing with social science—are in a honeymoon period (Lazer and Radford, 2017; van Atteveldt and Peng, 2018). Today's social scientist might reach for the tools of computer science for their speed, scale, granularity, and consistency; for instance, natural language processing offers "to analyze signals ranging from simple lexical cues to word clusters to choices of syntactic structure" (Boydstun et al., 2014). The numerical outputs tell a story that is simple, easy to make sense of, and in that regard comforting. Conversely, today's computer scientist may see the problems of social science as answerable by objectivity and reductionism, eschewing interpretation for quantitative analysis.

The conclusion of this reasoning, and the dominant stance in computational social science, is a reliance on machines alone to answer questions in the field, surrendering to their supposed objectivity or impartiality. Can a machine's output go beyond descriptive catalogs of evidence, accelerating understanding of processes and motivations? From our experience, computers are nowhere near supplanting humans in interpreting social science results.[1]

An interdisciplinary inquiry must go farther than matching computational techniques to social science questions (O'Connor et al., 2011; Nguyen et al., 2020). It embraces synergistic methodology and connects the norms and standards of evidence from both. This means partnering computer science's preference for the structured, generalizable, and objective with the unstructured, critical, and contextual which the social sciences champion. This level of interdisciplinarity addresses the question raised by descriptive findings: *So what?*

We see theory as the solution, empowering rather than shackling investigations. What this paper advocates is not one particular theory—certainly these are myriad, and "even subject matter which has been under intensive and prolonged study remains at the unsettled periphery of research" (Nagel, 1963). Instead, we expand on our prior work (Dore and McCarthy, 2022) to clarify calls echoed for decades by computational and social science (McDermott, 1976; Jelinek, 2005; Hajič and Hajičová, 2007; Hofman et al., 2018; Lipton and Steinhardt, 2019; Baden et al., 2021). Underlying each, we find, is the urge to return to theory, which we espouse herein.

## 2 Description vs. Integration

We contrast descriptive findings and theoretical analysis. An example of a descriptive finding is that an apple falls, or that it falls faster when pushed than dropped, or even that it falls at a particular rate estimated with some standard error by a complex

[*] Equal contribution.

[1]See, e.g., Noam Chomsky's remark on GPT-3: "You can't go to a physics conference and say: I've got a great theory. It accounts for everything and is so simple it can be captured in two words: 'Anything goes.' All known and unknown laws of nature are accommodated. . . Of course, everything impossible is accommodated also. That's GPT-3." [link]



*Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*
*Volume 2: Short Papers*, pages 1586–1594

interpolation. A theoretical analysis of the same phenomenon, credited to Newton, is that a fundamental force acts upon the apple, and that this same force governs the motion of the heavens. The theoretical analysis links the finding about the world critically to a broader body of knowledge and context.

Despite advances in causal inference in NLP, the descriptive is all that a machine can provide to the social sciences (Feder et al., 2021). Certainly the methods of computational text analysis have advanced since the General Inquirer (Stone and Hunt, 1963) and Mosteller and Wallace's statistical inference of text authorship (1963). But methods are means, not ends. They uncover more descriptive findings in data: the rate of an apple's fall, the topics of refugees' tweets (Walk et al., 2022), the space given to marginalized groups in textbooks (Lucy et al., 2020), or patterns of state censorship (Bamman et al., 2012; King et al., 2013).

The foils to descriptive findings are *integrative findings* (Hofman et al., 2021), which offer causal explanations that enable future predictions—a theory, or as a 'model' in the sense of the Standard Model, rather than of a statistical model. Integrative findings can either offer new theories or couch their explanations in existing theories—but the theory is essential either way.

## 3 We Don't Integrate

To contrast descriptive and integrative findings, we reviewed approximately 60 papers in computational text analysis published in *ACL venues. In Table 1, we describe several of these in terms of their descriptive or theory-grounded contributions.[2] Descriptive papers may refer to social science theories or make generalizable claims, as when Demszky et al. (2019) write, "The shooter's race appears to play a role in topic preference: if the shooter is white, Democrats become more likely to focus on shooter's identity," but they do not link to the two to each other.

An excellent theory-grounded quantitative work is Nelson (2021); she confirms some of the most compelling features of identity theory, specifically that identities based on race were most distinguished by cultural discourse, whereas those based on gender by the domestic and the economic discourse. Similarly, we conducted theory-grounded quantitative work to investigate the application of the protest paradigm and thematic framing in how western- and Hong Kong based newspapers portray protests in Hong Kong (McCarthy et al., 2021; McCarthy and Dore, 2022). Generally, it remains challenging to find computational social science papers in *ACL venues that go beyond description and prediction, advancing theory. Why is this? We believe it stemmed from the field's "empirical turn".[3]

Few remember when the meetings of ACL offered a few dozen papers, all entrenched in formalisms and linguistic theories. Arguably, 1996 was a turning point when the founders of SIGDAT held the first EMNLP at Penn under the auspices of the ACL.[4] This gave a spotlight to the few but growing empiricists in the field and drew in more.

EMNLP began a half-decade of measurable reorganization the field (Anderson et al., 2012). That EMNLP remains affiliated with ACL keeps the language-focused machine learning practitioners in our tent. The slow blurring of boundaries between each *ACL conference's expectations (Church, 2020) increases this unity. Both groups belong under this tent. But without a doubt, one group's voice is becoming less heard.

Publication venues within the ACL focus on methods over theory.[5] Techniques are taken off the shelf without critical examination because these are "the best" (often "state of the art") for their purposes (Ethayarajh and Jurafsky, 2020). This widens the gap between theoretical and empirical work.[6] Hopkins and King (2010) claim, "computer scientists may be interested in finding the needle in the haystack. . . social scientists are more commonly interested in characterizing the haystack"—evincing the value of broader context.[7] Wallach (2018), quoting Hopkins and King, explains that the two groups

[2]Following Lipton and Steinhardt (2019), we only describe papers by established researchers to "avoid singling out junior students. . . who lack the opportunity to reply symmetrically".

[3]A lesser reason is the challenge of serving two masters: adequately covering both the theoretical and methodological components within 8 pages. We recently received two reviews for an *ACL submission: one advocating for more of the social science context in the main text by eschewing methods to the appendix, and the other instructing us to do the opposite.

[4]And its predecessor the Workshop on Very Large Corpora.

[5]This is due to the outsized influence of computer science, often seen as the science of method (Hoare and Jones, 1989; Shapiro, 2001), when not instead seen as an engineering discipline (Rapaport, 2005).

[6]A related criticism is that empirical research has narrowed to focus on 'easy' questions that its tools can address (Coleman, 1986; Baden et al., 2021), especially when research questions are baked into the design of the task.

[7]As evidence, see Siegel (2018): "We usually don't know about causation, and we often don't necessarily care. . . the objective is more to predict than it is to understand the world. . . It just needs to work; prediction trumps explanation."

| | |
|---|---|
| **Descriptive** | |
| Chang et al. (2009) | The article presents new quantitative methods to measure semantic meaning in inferred topics. The authors emphasize the qualitative relevance of their findings as it validates the use of topics for corpus exploration and information retrieval. However, their working hypothesis and **empirical findings are not connected to the extremely relevant field of communication theory**. |
| Bamman et al. (2012) | The article presents the first large–scale analysis of political content censorship in social media. The authors **miss the opportunity to relate their hypothesis and findings to censorship theory, a natural theoretical context for the research**, which would strengthen the relevance and generalizability of the findings. |
| Field et al. (2018) | The article discusses media manipulation in Russia in the context of agenda-setting and framing, the tools that Russian state-owned (or heavily influenced) media outlets use to distract public attention from domestic economic politics. The authors **implicitly refer to propaganda theory and autocratic theory throughout the article even though their findings are not discussed in relation to these theories**. |
| Demszky et al. (2019) | The article applies "a more comprhensive NLP framework to study linguistic aspects of polarization in social media". While the article implicitly refer to theories of social conformity and social conflict, **the findings are not linked or discussed (either explicitly or implicitly) to the theoretical frameworks** that the authors touch on in their §1. |
| **Integrative** | |
| DiMaggio et al. (2013) | The article describes how topic models of newspaper articles help to study the politicization of government support for arts organizations and artists in the late 1980s in the US. The authors clearly define the theoretical context of their investigation and **emphasize the relationship between theory and method** throughout the paper. |
| Bamman et al. (2014) | The article validates an empirical model that "employs multiple effects to account for the influence of extra-linguistic information (such as author)" by **testing specific parameters against a variety of theory-based hypotheses** derived from writing styles theories of England between 1700 and 1899. |
| Nelson (2021) | The article argues that the full potential of machine learning can be better realized by "leveraging the epistemological alignment between machine learning and inductive research." The author empirically demonstrates this by **anchoring in identity theory** a word embedding model of first-person narratives of the nineteenth-century U.S. South. |

Table 1: Contrast between work in computational text analysis with descriptive findings versus integrative findings.

are interested in very different research questions, and that computational social science must be more than computer science with social data; it must strive for valid explanatory models. In the same vein, at ACL 2022, ACL fellow Eduard Hovy remarked that NLP must be more than "just machine learning on corpora".

Social scientists are also coming to terms with the meaning of computational techniques applied more often in social science (Bail, 2014; Biernacki, 2015; Lee and Martin, 2015; Spillman, 2015). The focus of the debates, however, is on which methods are best suited to extract meaning from text, without addressing any theoretical considerations related to the methods or whether a theoretical framework for those methods even exists. The discussions on whether computational methods make social science research more efficient, reliable, and reproducible overtake attempts at theory-building.

## 4 Moving Forward

We are not denying the value of computational approaches to analyzing text. Certainly, computing can be an instrumental approach for modeling and understanding social complexity. This does not mean that other approaches, such as historical, ethnographic, or mathematical, become irrelevant. On the contrary, computational methods necessarily (whether awarely or not) rely on these earlier approaches to add value, in terms of improving our explanations and understanding (Radford and Joseph, 2020).

As we are a field that prioritizes methods, consider the seminal book on methods in science: Abbott (2004) taxonomizes scientific ways of knowing. Its five broad categories are ethnography, historical narration, standard causal analysis, small-$N$ comparison, and formal modeling. We in NLP myopically choose the third and fifth of these, ignoring the value of the others. But the broader point of *Methods of Discovery* is not methods. It is the research question. Any methodology should be grounded in the question, not incremental tweaks and reviewers' comfort (Church, 2020). This admits even qualitative or mixed-method approaches to text analysis.

The role of humans in scientific inquiry is nothing new. Using qualitative analysis to complement quantitative techniques has its roots in Achen and Snidal (1989)'s recommendation to use historical case studies as a complement to statistical research.[8] Their plea was strengthened by Verba's work in the early 1990s (Verba et al., 1993, 1995; Verba, 1996) and Tarrow (1995), who openly called for bridging qualitative and quantitative modes of research in social science. In doing so, they have enriched the field with critical methodological innovations (Gerring, 2004), benefiting from the recognition that "quantitative methods must augment humans, not replace them" (Grimmer and Stewart, 2013, 4).

The field can draw more from social science's rich tradition of inductive theory-building and interpretation to develop its theoretical approach—to prize either induction or deduction alone is a myth of scientific procedure (Thagard, 1988), but the melding of the two opens new doors. Rather than eschewing the complexity (a criticism leveled by Baden et al., 2021), it should put complexity at the center of its ontology on the basis that there are no immutable laws in social life or optimal solutions to social problems.

Skepticism can linger toward findings not drawn from the standard practices of one's own field; indeed, social science was long skeptical of computational contributions (Armstrong, 1967). We believe that this drives the hyperfocus on improving a few accepted methods instead of exploring more broadly. If the doorway between disciplines is only narrowly open, this reflects a lack of appreciation for each field's ways of knowing. The disciplinary divide keeps computational researchers from embracing methods beyond *standard causal analysis* or *formal modeling*, so the interpreter-centric richness allowed by histories, ethnographies, and small-$N$ exploration are precluded.

## 5 Conclusion

We have explained the distinction between descriptive and theoretical findings as it pertains to computational text analysis. The bulk of work we found provided vast descriptive findings, often of high quality, but not giving back to questions of theory. We offer several suggestions on how to 'push the pendulum back' by prioritizing theory-building or theory-affirming research questions and accepting whichever methods are best suited toward answering it—not only the familiar and entrenched ones.

We are not the first to advocate for a shift in the patterns of applying computational techniques to real-world problems. There is a steady drumbeat from voices in the field advocating careful approaches (Nagel, 1963; McDermott, 1976; Jelinek, 2005; Hajič and Hajičová, 2007; Hofman et al., 2018; Lipton and Steinhardt, 2019; Baden et al., 2021). What we see underlying all of these—those writing against 'mathiness' and speculation, advocating for clear evaluation over anecdotes, criticizing textual researchers' dilution of conceptual standards, highlighting work that ties linguistic information into complex models—is an unspoken, perhaps unrealized, call for a return to theory.

Not only do we aver that incorporating theory is essential; but also, other fields have strengthened themselves when espousing organizing principles beyond those of their progenitors. Behavioral economics is a success story here. It transcended the neat (but psychosocially stripped) mathematics it draws from to acknowledge deviations from rationality and blend economics with cognitive science (Kahneman and Tversky, 1979; Thaler, 1980; Thaler and Sunstein, 2009).

For *scientific*—not simply *engineering*—advances to arise from the *ACL community, authors and reviewers alike must resist the temptation toward incremental, 'safe' research and follow Church (2005): "Controversial papers are great; boring unobjectionable incremental papers are not." In reviewing new research, we should privilege not only work that presents new and unusual computational methods, but also interactions between computational and humanistic approaches to answering research questions. EMNLP was founded because of reviewing biases at ACL against groundbreaking methodological advances, and since then the two have homogenized; "EMNLP reviewing is no longer much of a differentiator" (Church, 2020). We found that theoretically grounded findings in text analysis are often published in non-*ACL venues (Table 1), but ACL sets the standard for work involving computational text analysis and NLP. Is there no home for groundbreaking integrative or interdisciplinary work in *ACL, such that a new venue is required? Or can we adapt our standards to invite deeper connections to theory and new ways of knowing?

[8] Expertise plays a role as well (Shing et al., 2018), which is why Mechanical Turk doesn't fill the need for qualitative analysis. This is exemplified by Radford and Joseph (2020)'s observation of "non-expert annotators provid[ing] unreliable annotations, even after a discussion period".

## Acknowledgments

This publication was made possible in part by a grant from the American Political Science Association to A.D.M. and G.M.D.D. The statements made and views expressed are solely the responsibility of the authors. A.D.M. is supported by an Amazon Fellowship and a Frederick Jelinek Fellowship.

## Limitations

The key limitation of our work is that, when conducting the review of approximately 60 papers (by searching through the ACL Anthology for works in computational social science since 2010), we encountered a skewed distribution of descriptive versus integrative works. In fact, it was relatively simple to find descriptive works, and that section of Table 1 could have been much longer. We also recognize that, due to the mixed nature of our field, scientific and integrative findings are not the only goal—our 'big tent' includes engineers as well, who value gains in performance indicators. Finally, the fact that we have few examples of papers showing a return to theory renders the possibility that our central claim is misinterpreted in a normative way as a mandate.

## References

Andrew Delano Abbott. 2004. *Methods of discovery: Heuristics for the social sciences (contemporary societies)*. WW Norton & Company.

Christopher H. Achen and Duncan Snidal. 1989. Rational deterrence theory and comparative case studies. *World Politics*, 41(2):143–169.

Ashton Anderson, Dan Jurafsky, and Daniel A. McFarland. 2012. Towards a computational history of the ACL: 1980-2008. In *Proceedings of the ACL-2012 Special Workshop on Rediscovering 50 Years of Discoveries*, pages 13–21, Jeju Island, Korea. Association for Computational Linguistics.

J. Scott Armstrong. 1967. Derivation of theory by means of factor analysis or Tom Swift and his electric factor analysis machine. *The American Statistician*, 21(5):17–21.

Christian Baden, Christian Pipal, Martijn Schoonvelde, and Mariken A. C. G van der Velden. 2021. Three gaps in computational text analysis methods for social sciences: A research agenda. *Communication Methods and Measures*, 0(0):1–18.

Christopher A. Bail. 2014. The cultural environment: measuring culture with big data. *Theory and Society*, 43(3):465–482.

David Bamman, Brendan O'Connor, and Noah Smith. 2012. Censorship and deletion practices in chinese social media. *First Monday*, 17(3).

David Bamman, Ted Underwood, and Noah A. Smith. 2014. A Bayesian mixed effects model of literary character. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 370–379, Baltimore, Maryland. Association for Computational Linguistics.

Richard Biernacki. 2015. How to do things with historical texts. *American Journal of Cultural Sociology*, 3(3):311–352. Copyright - © Palgrave Macmillan, a division of Macmillan Publishers Ltd 2015; Last updated - 2018-09-25.

Amber E Boydstun, Dallas Card, Justin Gross, Paul Resnick, and Noah A Smith. 2014. Tracking the development of media frames within and across policy issues. Unpublished.

Jonathan Chang, Jordan Boyd-Graber, Sean Gerrish, Chong Wang, and David M. Blei. 2009. Reading tea leaves: How humans interpret topic models. In *Proceedings of the 22nd International Conference on Neural Information Processing Systems*, NIPS'09, page 288–296, Red Hook, NY, USA. Curran Associates Inc.

Kenneth Church. 2005. Last words: Reviewing the reviewers. *Computational Linguistics*, 31(4):575–578.

Kenneth Ward Church. 2020. Emerging trends: Reviewing the reviewers (again). *Natural Language Engineering*, 26(2):245–257.

James S. Coleman. 1986. Social theory, social research, and a theory of action. *American Journal of Sociology*, 91(6):1309–1335.

Dorottya Demszky, Nikhil Garg, Rob Voigt, James Zou, Jesse Shapiro, Matthew Gentzkow, and Dan Jurafsky. 2019. Analyzing polarization in social media: Method and application to tweets on 21 mass shootings. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 2970–3005, Minneapolis, Minnesota. Association for Computational Linguistics.

Paul DiMaggio, Manish Nag, and David Blei. 2013. Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics*, 41(6):570–606. Topic Models and the Cultural Sciences.

Giovanna Maria Dora Dore and Arya D. McCarthy. 2022. Learning to play with the machines in social science research: Bringing the theory back in. In *ICML 2022 Workshop on Human-Machine Collaboration and Teaming*, Baltimore, Maryland.

Kawin Ethayarajh and Dan Jurafsky. 2020. Utility is in the eye of the user: A critique of NLP leaderboards. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4846–4853, Online. Association for Computational Linguistics.

Amir Feder, Katherine A. Keith, Emaad Manzoor, Reid Pryzant, Dhanya Sridhar, Zach Wood-Doughty, Jacob Eisenstein, Justin Grimmer, Roi Reichart, Margaret E. Roberts, Brandon M. Stewart, Victor Veitch, and Diyi Yang. 2021. Causal inference in natural language processing: Estimation, prediction, interpretation and beyond. *CoRR*, abs/2109.00725.

Anjalie Field, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia Tsvetkov. 2018. Framing and agenda-setting in Russian news: a computational analysis of intricate political strategies. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3570–3580, Brussels, Belgium. Association for Computational Linguistics.

John Gerring. 2004. What is a case study and what is it good for? *American Political Science Review*, 98(2):341–354.

Justin Grimmer and Brandon M. Stewart. 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3):267–297.

Jan Hajič and Eva Hajičová. 2007. Some of our best friends are statisticians. In *Text, Speech and Dialogue*, pages 2–10, Berlin, Heidelberg. Springer Berlin Heidelberg.

C. A. R. Hoare and C. B. Jones. 1989. *Essays in Computing Science*. Prentice-Hall, Inc., USA.

Jake Hofman, Miro Dudík, and Daniel G. Goldstein. 2018. Perspective annotation for numerical representations. United States Patent Application.

Jake M Hofman, Duncan J Watts, Susan Athey, Filiz Garip, Thomas L Griffiths, Jon Kleinberg, Helen Margetts, Sendhil Mullainathan, Matthew J Salganik, Simine Vazire, et al. 2021. Integrating explanation and prediction in computational social science. *Nature*, 595(7866):181–188.

Daniel J. Hopkins and Gary King. 2010. A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, 54(1):229–247.

Frederick Jelinek. 2005. Some of my best friends are linguists. *Language Resources and Evaluation*, 39(1):25–34.

Daniel Kahneman and Amos Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Gary King, Jennifer Pan, and Margaret E. Roberts. 2013. How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, 107(2 (May)):1–18.

David Lazer and Jason Radford. 2017. Data ex machina: Introduction to big data. *Annual Review of Sociology*, 43(1):19–39.

Monica Lee and John Levi Martin. 2015. Coding, counting and cultural cartography. *American Journal of Cultural Sociology*, 3(1):1–33.

Zachary C. Lipton and Jacob Steinhardt. 2019. Troubling trends in machine learning scholarship: Some ML papers suffer from flaws that could mislead the public and stymie future research. *Queue*, 17(1):45–77.

Li Lucy, Dorottya Demszky, Patricia Bromley, and Dan Jurafsky. 2020. Content analysis of textbooks via natural language processing: Findings on gender, race, and ethnicity in Texas U.S. history textbooks. *AERA Open*, 6(3):2332858420940312.

Arya D. McCarthy and Giovanna Maria Dora Dore. 2022. Hong Kong: Longitudinal and synchronic characterisations of protest news between 1998 and 2020. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 2891–2900, Marseille, France. European Language Resources Association.

Arya D. McCarthy, James Scharf, and Giovanna Maria Dora Dore. 2021. A mixed-methods analysis of western and Hong Kong–based reporting on the 2019–2020 protests. In *Proceedings of the 5th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pages 178–188, Punta Cana, Dominican Republic (online). Association for Computational Linguistics.

Drew McDermott. 1976. Artificial intelligence meets natural stupidity. *SIGART Bull.*, (57):4–9.

Frederick Mosteller and David L. Wallace. 1963. Inference in an authorship problem. *Journal of the American Statistical Association*, 58(302):275–309.

Ernest Nagel. 1963. The structure of science: Problems in the logic of scientific explanation. *Mind*, 72(287).

Laura K. Nelson. 2021. Leveraging the alignment between machine learning and intersectionality: Using word embeddings to measure intersectional experiences of the nineteenth century U.S. South. *Poetics*, 88:101539. Measure Mohr Culture.

Dong Nguyen, Maria Liakata, Simon DeDeo, Jacob Eisenstein, David Mimno, Rebekah Tromble, and Jane Winters. 2020. How we do things with words: Analyzing text as social and cultural data. *Frontiers in Artificial Intelligence*, 3.

Brendan O'Connor, David Bamman, and Noah A Smith. 2011. Computational text analysis for social science: Model complexity and assumptions. In *Proc. of the NIPS Workshop on Comptuational Social Science and the Wisdom of Crowds*.

Jason Radford and Kenneth Joseph. 2020. Theory in, theory out: The uses of social theory in machine learning for social science. *Frontiers in Big Data*, 3.

William J Rapaport. 2005. Philosophy of computer science: An introductory course. *Teaching philosophy*, 28(4):319–341.

Stuart C. Shapiro. 2001. Computer science: The study of procedures. Technical report, Department of Computer Science and Engineering, University of Buffalo.

Han-Chin Shing, Suraj Nair, Ayah Zirikly, Meir Friedenberg, Hal Daumé III, and Philip Resnik. 2018. Expert, crowdsourced, and machine assessment of suicide risk via online postings. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, pages 25–36, New Orleans, LA. Association for Computational Linguistics.

David A. Siegel. 2018. Analyzing computational models. *American Journal of Political Science*, 62(3):745–759.

Lyn Spillman. 2015. Ghosts of straw men: A reply to Lee and Martin. *American Journal of Cultural Sociology*, 3(3):365–379.

Philip J. Stone and Earl B. Hunt. 1963. A computer approach to content analysis: Studies using the general inquirer system. In *Proceedings of the May 21-23, 1963, Spring Joint Computer Conference*, AFIPS '63 (Spring), page 241–256, New York, NY, USA. Association for Computing Machinery.

Sidney Tarrow. 1995. Bridging the quantitative-qualitative divide in political science. *American Political Science Review*, 89(2):471–474.

Paul Thagard. 1988. *Computational Philosophy of Science*. MIT Press.

Richard Thaler. 1980. Judgement And Decision Making Under Uncertainty: What Economists Can Learn From Psychology. Risk Analysis in Agriculture: Research and Educational Developments, January 16-18, 1980, Tucson, Arizona 271572, Regional Research Projects > W-149: An Economic Evaluation of Managing Market Risks in Agriculture.

Richard H. Thaler and Cass R. Sunstein. 2009. *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Wouter van Atteveldt and Tai-Quan Peng. 2018. When communication meets computation: Opportunities, challenges, and pitfalls in computational communication science. *Communication Methods and Measures*, 12(2-3):81–92.

Sidney Verba. 1996. The citizen as respondent: Sample surveys and american democracy. *American Political Science Review*, 90(1):1–7. Presidential Address, American Political Science Association, 1995.

Sidney Verba, Kay Lehman Schlozman, Henry Brady, and Norman H. Nie. 1993. Citizen activity: Who participates? what do they say? *The American Political Science Review*, 87(2):303–318.

Sidney Verba, Kay Lehman Schlozman, and Henry E Brady. 1995. *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.

Erin Walk, Elizabeth Parker-Magyar, Kiran Garimella, Ahmet Akbiyik, and Fotini Christia. 2022. Social media narratives across platforms in conflict: Evidence from Syria. MIT Political Science Department Research Paper No. 2022-2, available at SSRN.

Hanna Wallach. 2018. Computational social science ≠ computer science + social data. *Commun. ACM*, 61(3):42–44.

## ACL 2023 Responsible NLP Checklist

### A For every submission:

☑ A1. Did you describe the limitations of your work?
*Unnumbered; appears on page 5.*

☒ A2. Did you discuss any potential risks of your work?
*This is a position paper.*

☑ A3. Do the abstract and introduction summarize the paper's main claims?
*Left blank.*

☒ A4. Have you used AI writing assistants when working on this paper?
*Left blank.*

### B ☒ Did you use or create scientific artifacts?

*Left blank.*

☐ B1. Did you cite the creators of artifacts you used?
*Not applicable. Left blank.*

☐ B2. Did you discuss the license or terms for use and / or distribution of any artifacts?
*Not applicable. Left blank.*

☐ B3. Did you discuss if your use of existing artifact(s) was consistent with their intended use, provided that it was specified? For the artifacts you create, do you specify intended use and whether that is compatible with the original access conditions (in particular, derivatives of data accessed for research purposes should not be used outside of research contexts)?
*Not applicable. Left blank.*

☐ B4. Did you discuss the steps taken to check whether the data that was collected / used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect / anonymize it?
*Not applicable. Left blank.*

☐ B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?
*Not applicable. Left blank.*

☐ B6. Did you report relevant statistics like the number of examples, details of train / test / dev splits, etc. for the data that you used / created? Even for commonly-used benchmark datasets, include the number of examples in train / validation / test splits, as these provide necessary context for a reader to understand experimental results. For example, small differences in accuracy on large test sets may be significant, while on small test sets they may not be.
*Not applicable. Left blank.*

### C ☒ Did you run computational experiments?

*Left blank.*

☐ C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?
*Not applicable. Left blank.*

*The Responsible NLP Checklist used at ACL 2023 is adopted from NAACL 2022, with the addition of a question on AI writing assistance.*

☐ C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
*Not applicable. Left blank.*

☐ C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?
*Not applicable. Left blank.*

☐ C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation), did you report the implementation, model, and parameter settings used (e.g., NLTK, Spacy, ROUGE, etc.)?
*Not applicable. Left blank.*

**D ☒ Did you use human annotators (e.g., crowdworkers) or research with human participants?**

*Left blank.*

☐ D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?
*Not applicable. Left blank.*

☐ D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?
*Not applicable. Left blank.*

☐ D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating? For example, if you collected data via crowdsourcing, did your instructions to crowdworkers explain how the data would be used?
*Not applicable. Left blank.*

☐ D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
*Not applicable. Left blank.*

☐ D5. Did you report the basic demographic and geographic characteristics of the annotator population that is the source of the data?
*Not applicable. Left blank.*