International Journal of Communication 13(2019), 3876–3884 1932–8036/20190005

Outlining the Way Ahead in Computational Communication Science: An Introduction to the IJoC Special Section on "Computational Methods for Communication Science: Toward a Strategic Roadmap"

EMESE DOMAHIDI[1]
Technische Universität Ilmenau, Germany

JUNGHWAN YANG
University of Illinois at Urbana-Champaign, USA

JULIA NIEMANN-LENZ
Hanover University of Music, Drama and Media, Germany

LEONARD REINECKE
Johannes Gutenberg University Mainz, Germany

Over the past two decades, processes of digitalization and mediatization have shaped the communication landscape and have had a strong impact on various facets of communication. The digitalization of communication results in completely new forms of digital traces that make communication processes observable in new and unprecedented ways. Although many scholars in the social sciences acknowledge the chances and requirements of the digital revolution in communication, they are also facing fundamental challenges in implementing successful research programs, strategies, and designs that

Emese Domahidi: emese.domahidi@tu-ilmenau.de
JungHwan Yang: junghwan@illinois.edu
Julia Niemann-Lenz: julia.niemann-lenz@ijk.hmtm-hannover.de
Leonard Reinecke: leonard.reinecke@uni-mainz.de
Date submitted: 2019–08–09

[1] Our sincere thanks go out to the reviewers of this Special Section for their commitment as well as their invaluable feedback. We would also like to thank all submitting authors for sharing their work and their perspective on the field of CCS with us. Special thanks also go to IJoC Editor in Chief, Larry Gross, for making this Special Section possible, as well as to IJoC Managing Editor, Arlene Luck, for her invaluable support during all stages of the production process. This Special Section was initiated as part of the CCS 2018 Conference in Hanover, Germany, and the call for papers was distributed among the conference attendees. Both the CCS Conference as well as this Special Section were funded by a grant from Volkswagen Foundation in the funding line "International Research in Computational Social Sciences." We would like to thank Volkswagen Foundation for its generous financial support as well as the conference organizers and all conference attendees for their support and their continuing dedication to this fast-developing field.

are based on computational methods and "big data." This Special Section aims at bringing together seminal perspectives on challenges and chances of computational communication science (CCS). In this introduction, we highlight the impulses provided by the research presented in the Special Section, discuss the most pressing challenges in the context of CCS, and sketch a potential roadmap for future research in this field.

Keywords: computational communication science, CCS, data science, computational methods, big data

The enhanced capacity to collect data coupled with computational methods have created a new impetus in communication science. However, computational methods are no longer "the next big thing" that they once were. In fact, they are reaching an ever-new level of maturation and are by now ubiquitous in contemporary research discussions in various scientific fields (Lazer et al., 2009).

The current stakes and expectations are high. Researchers believe that large and complex data sets coupled with new methods and algorithms can provide answers to fundamental questions (Grimmer, 2015), and this "computational turn" not only complements traditional social science methods but also offers various advantages and opens up new possibilities for scientific investigation (Lazer et al., 2009). More recently, popular packages of the R and Python programming languages and various other programs with graphical user interfaces provide numerous out-of-the-box analytic tools that make computational methods even more accessible for social scientists. With this prospect, computational methods have been discussed and implemented in different neighboring research fields, such as political science (Grimmer & Stewart, 2013; Monroe, 2013), psychology (Yarkoni & Westfall, 2017), and information science (Diesner, 2013), and the lines between social sciences and "hard" sciences have become blurred as they now share similar research questions and methods (Grimmer, 2015). Research funding sources, such as NSF, Russell Sage Foundation, and Volkswagen Foundation, have been a catalyst of recent interdisciplinary collaborations, as they understand that different areas of expertise are key for successful computational projects (Wallach, 2018).

Communication science is especially sensitive to this methodological change (van Atteveldt & Peng, 2018), as communication and media data—the main interest of our field—have become increasingly available in a digital format and have tremendous economic, societal, and political value (Lazer et al., 2009). Thus, the adoption of computational methods has reinvigorated communication science and made computational communication science (CCS) a flourishing field of activity in computational social science (CSS). From the perspective of our discipline, this is both a challenge and a chance. Communication scholars have to deal not only with adapting a whole new spectrum of methods but also adapting to new influences and practices of other scientific disciplines.

Definitions of CSS and CCS emphasize the aforementioned changes in scientific research, such as increased availability of digital data (Anderson, 2008), new tools, and advanced computing infrastructure, as well as the capability for processing "big data" (Lazer et al., 2009). Most recently, van Atteveldt and Peng (2018) described CCS as involving "(1) large and complex data sets; (2) consisting of digital traces and

other 'naturally occurring' data; (3) requiring algorithmic solutions to analyze; and (4) allowing the study of human communication by applying and testing communication theory" (p. 82). Of course, it is self-evident that a CCS must include computational methods, though it remains vague how large and complex a data set must be to really require advanced analytic methods. Further, the data used in CCS might not always "occur naturally," as it is possible for researchers to create digital traces by incorporating their own cookies or apps (e.g., Montag et al., 2015). However, the last and fourth point of the definition by van Atteveldt and Peng is essential. It emphasizes not only the object of CCS—human communication—but also that theory is a key aspect of CCS and that it is by no means merely data-driven research. On contrary to the early days of "big data," when some scholars believed that social theory would be obsolete (Anderson, 2008), many computational social scientists believe that either "big data" or computational methods alone cannot be solely capable of explaining human behavior independent of existing theories (boyd & Crawford, 2012; Lewis, Zamith, & Hermida, 2013). Instead, CCS researchers are cognizant of the architecture and politics of data and believe that computational methods can inform new theory when they are understood in a larger societal, political, economic, and cultural context.

Consequently, more scholars are introducing and discussing digital data and computational methods to the communication science context (e.g., Hopp, Schaffer, Fisher, & Weber, 2019; Maier et al., 2018; Rudkowsky et al., 2018; Scharkow, 2013; Trilling & Jonkman, 2018) and applying these techniques to a wide range of communication science research questions (e.g., González-Bailón, Banchs, & Kaltenbrunner, 2012; Günther, & Domahidi, 2017; Kwon, Chadha, & Wang, 2019; Shah et al., 2016). This engagement leads to the emergence of new divisions within scientific societies (e.g., ICA) as well as specialized conferences (IC2S2, ICWSM) and results in various publications in major academic journals of our research field. Several Special Sections on different aspects of the topic have been announced or have already been released in our core communication science journals (e.g., *International Journal of Communication, Communication Methods and Measures, Political Communication, Public Opinion Quarterly, Social Science Computer Review*). Even a new peer-reviewed open-access journal, *Computational Communication Research* (van Atteveldt, Margolin, Shen, Trilling, & Weber, 2019), has been established to cover applications of computational approaches in addressing communication science issues. New positions with a specific focus on CCS have been created in the academic job market as a growing number of universities have started to offer formal education for undergraduate and graduate students. It is time to acknowledge that communication science has already adapted computational methods and to take stock of how the discipline has dealt with the chances and challenges arising from this computational reorientation so far.

Current Challenges in the Maturing Field of Computational Communication Science and the Focus of This Special Section

CCS as a field of study has entered a maturing stage. Gathering and analyzing digital data and applying various computational methods in communication science projects have become part of the regular research procedures. The challenges that were common in the early stages of the CCS research, such as a general lack of resources, skills, tools, and talents (Lazer et al., 2009), are less urgent today. However, in the current stage of its development, CCS faces new challenges. For instance, it is still not clear how CCS can be incorporated in the wider field of communication science and be distinguished from other subfields

of CSS. To address these issues, we need to investigate how the community of communication scientists defines and applies computational methods as well as how it aims to contribute to further development of CCS. Another issue that the CCS field faces today is uncertainty in how the computational turn is changing the field of communication and how it contributes to finding meaningful answers to our core research questions.

The main goal of this Special Section is to contribute to this discussion by reflecting on the opportunities and challenges in methodological and theoretical developments and to extend the discussion on how the community of CCS researchers can design the field for the future. The nine articles published in this Special Section provide new perspectives and valuable insight in two different contexts of CCS research: (1) community building and the status quo of the CCS discipline, and (2) advances in computational methods and their theoretical contributions. In the following, we elaborate on these aspects, provide a short introduction to the individual articles of the Special Section, and clarify how they relate to these challenges and in which ways they offer advancements for CCS.

Community Building for Computational Communication Science

Key elements of successful evolvement of the new scientific field include building a strong community of CCS researchers, institutionalizing CCS methods at universities and academic societies, and defining standards and procedures for research and teaching. These challenges are urgent for the community because they influence research and teaching in important ways. In "Crafting a Strategic Roadmap for Computational Methods in Communication Science: Learnings From the CCS 2018 Conference in Hanover," Niemann-Lenz and colleagues provide a check of the status quo of the CCS community. With the help of the Volkswagen Foundation funding line "International Research in Computational Social Sciences," the authors invited international CCS researchers to the CCS 2018 Conference in Hanover, Germany, to discuss the state of CCS and enable strategic networking and community building. The key findings and learnings from the conference are documented in the article that intends to stimulate organizational and collaborative efforts for CCS.

Because access to valid and meaningful data is a critical point in the conduct of any empirical study, Possler, Bruns, and Niemann-Lenz, who co-organized the CCS 2018 conference, provide insights into how communication scientists are changing their workflows in this regard. In their article "Data Is the New Oil—But How Do We Drill It? Pathways to Access and Acquire Large Data Sets in Communication Science," they interviewed 22 computational communication scientists to ask how those researchers define CCS as a field and what challenges they experienced in their research processes. They particularly discussed various possibilities and challenges regarding data acquisition that many CCS researchers are facing.

In "Computational Communication Science: A Methodological Catalyzer for a Maturing Discipline," Hilbert and 15 coauthors reviewed how computational methods can inform three fundamental pillars of scientific research methods—observational approaches (i.e., digital trace data), theoretical approaches (i.e., computer simulations), and experimental research (i.e., virtual labs and field experiments)—and discussed practical and theoretical implications of each method for advancing communication research.

Informed by the recent replication crisis in psychology (Munafò et al., 2017; Nosek et al., 2015), growing numbers of communication researchers advocate transparency and reproducibility in the research process (Bowman, & Keene, 2018). Although open science is a problem of communication science as a whole and all neighboring disciplines in general, computational methods offer unique challenges and possibilities in this context. In "Toward Open Computational Communication Science: A Practical Road Map for Reusable Data and Code," van Atteveldt, Strycharz, Trilling, and Welbers raise awareness for conducting open and reproducible science, as CCS research relies on complex models and algorithms that often seem indecipherable. They discuss goals and challenges of open science and provide practical solutions for code and data sharing.

Advancements in Computational Communication Science Methods and Their Theoretical Contributions

In the maturing stage of CCS, the question is not whether we are capable of conducting research using computational approaches, but how we can apply the right methods and develop new approaches to answer fundamental research questions of communication science. More rigorous applications of a wide range of advanced computational methods are necessary to address theoretical and practical concerns of communication research and also to meet increasingly higher standards of top-tier social science journals. Thus, CCS scholars need to implement and extend various computational methods and to actively adapt and develop computational techniques that are suited for our research questions (van Atteveldt et al., 2019). Advancing CCS in this sense requires not only knowledge of computational methods but also strong domain-specific expertise to be able to define relevant problems and tailor specific solutions for the question.

Therefore, communication scholars need to broaden their scope of theories because CCS operates in an interdisciplinary environment. Thus, in this Special Section, Waldherr, Geise, and Katzenbach discuss actor–network theory (ANT) as a useful conceptual framework for theorizing interdependencies in CCS projects in their piece "Because Technology Matters: Theorizing Interdependencies in Computational Communication Science with Actor–Network Theory." Using an extensive literature review, the authors identify three key contributions of the ANT and reflect on potential merits of the theory in CCS.

In "Bridging the Gaps: Using Agent-Based Modeling to Reconcile Data and Theory in Computational Communication Science," Waldherr and Wettstein introduce the foundations and applications of agent-based modeling (ABM). Because ABM is still rarely used in communication science, this article highlights its potential to advance theoretical work in CCS and link micro and macro perspectives. The authors provide an overview of ABM, discuss potentials and challenges, and put a special emphasis on the reliability and validity of the method.

In "When the Journey Is as Important as the Goal: A Roadmap to Multilingual Dictionary Construction," Lind, Eberl, Heidenreich, and Boomgarden discuss dictionary construction as an increasingly relevant challenge for comparative CCS research that involves multiple languages. Using the example of migration frames in the news coverage in seven languages, the authors offer insights into the art of multilingual dictionary construction and discuss different strategies on the way to a reliable and valid dictionary in that context.

In "The Temporal Turn in Communication Research: Time Series Analyses Using Computational Approaches" Wells et al. emphasize the importance of temporal dynamics for our field's research interests thanks to the availability of temporal data from the Web and social media, and discuss a number of issues to consider when generating sequential data for time-series analysis. The authors not only specify appropriate techniques for the analysis of temporal data but also discuss practical and methodological challenges in these processes and provide resources for interested scholars.

Joo, Bucy, and Seidel discuss the importance of visual images in understanding communicative behavior in politics. In "Automated Coding of Televised Leader Displays: Detecting Nonverbal Political Behavior With Computer Vision and Deep Learning," they develop a pipeline of using computational methods to automatically classify fine-grained facial expressions and physical gestures. This case study shows that computational methods can replicate human coding with a high degree of accuracy and also suggests that their methods can dramatically scale up existing manual coding procedures. This has an important implication for rapid progress in quantitative visual communication research.

Quo Vadis Computational Communication Science? Putting the Insights in a Context

As outlined above, computational methods are incorporated in various ways in our field and are expected to become even more important for communication science in the future. Here, we outline and discuss central insights from the Special Section and provide an outlook into future challenges and directions for CCS researchers.

First, we argue that community building is a key element in advancing CCS. Only a strong community can collectively solve challenges that are related to computational methods in research, like developing strategies to overcome limitations related to the collection of digital data that in most cases belong to private companies and come with inherent privacy and ethical concerns (see Niemann-Lenz et al. and Possler et al. in this Special Section). In addition, the institutionalization of CCS is related to challenges in teaching, developing (interdisciplinary) publication strategies, and adjusting diverging expertise of researchers in computational methods. Building a strong and active community of CCS researchers enables the definition of challenges and implementation of solutions that are suited for our field of communication science. Some of the challenges are not specific for the field of CCS, but are critical to social sciences in general—like open science practices—and can be driven by the community of CCS researchers (see van Atteveldt et al. in this Special Section).

Second, CCS has to go beyond simply applying computational methods; instead, researchers need to adapt and develop these methods according to the central research questions for our field. In our Special Section, we present the introduction, adaptation, and evaluation of various computational methods for a maturing field (see Hilbert et al.; Joo et al.; Lindt et al.; and Wells et al.). It becomes evident from these articles that communication science needs to widen the focus of analysis to comparative, multilingual settings, and to temporal and pictorial data. Furthermore, theoretical work and methods like ABM, which can help to test central theoretical assumptions of communication science, need to be incorporated in our standard CCS repertoires (see Waldherr et al. and Waldherr & Wettstein in this Special Section). Future

questions might refer to how to enrich these and other computational methods, like supervised machine learning for automatic text analysis, by our communication science domain expertise (e.g., by specific feature engineering; Kuhn & Johnson, 2019).

Third, open and reusable science practices are central for CCS to overcome inherent challenges of digital data acquisition and computational methods (see van Atteveldt et al. in this Special Section). However, the emphasis on open and reusable science practices is needed not only for reproducing CCS research results but also to collectively advance computational methods for communication science. This means sharing scripts to spread knowledge on the application of certain methods (e.g., Wells et al. in this Special Section) and also sharing the development of methods that are important for the advancement of the field, but that might (yet) not be perfect (a noticeable example in this Special Section is the article by Lind et al.). By doing so, CCS researchers can foster the collective development and evaluation of methods that might be too complex or time-consuming to be developed or validated by individual researchers.

We sincerely believe that the diverse perspectives provided in the articles published in this Special Section can be a useful starting point and provide numerous invaluable impulses and ideas for solving the problems the field of CCS is currently facing. We are optimistic that the CCS community will continue to strive and take on these challenges, and we hope that this Special Section will provide orientation and inspiration for the future journey of our discipline.

References

Anderson, C. (2008). The end of theory, will the data deluge makes the scientific method obsolete? *Wired*. Retrieved from https://www.wired.com/2008/06/pb-theory/

Bowman, N. D., & Keene, J. R. (2018). A layered framework for considering open science practices. *Communication Research Reports, 35*(4), 363–372. doi:10.1080/08824096.2018.1513273

boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15*(5), 662–679. doi:10.1080/1369118X.2012.678878

Diesner, J. (2013). From texts to networks: Detecting and managing the impact of methodological choices for extracting network data from text data. *KI - Künstliche Intelligenz, 27*(1), 75–78. doi:10.1007/s13218-012-0225-0

González-Bailón, S., Banchs, R. E., & Kaltenbrunner, A. (2012). Emotions, public opinion, and U.S. presidential approval rates: A 5-year analysis of online political discussions. *Human Communication Research, 38*(2), 121–143. doi:10.1111/j.1468-2958.2011.01423.x

Grimmer, J. (2015). We are all social scientists now: How big data, machine learning, and causal inference work together. *PS: Political Science & Politics, 48*(01), 80–83. doi:10.1017/S1049096514001784

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis, 21*(3), 267–297. doi:10.1093/pan/mps028

Günther, E., & Domahidi, E. (2017). What communication scholars write about: An analysis of 80 years of research in high-impact journals. *International Journal of Communication*, *11*, 3051–3071.

Hopp, F. R., Schaffer, J., Fisher, J. T., & Weber, R. (2019). *iCoRe: The GDELT Interface for the Advancement of Communication Research*. doi:10.31235/osf.io/smjwb

Kuhn, M., & Johnson, K. (2019). *Feature engineering and selection: A practical approach for predictive models.* Boca Raton, FL: CRC.

Kwon, K. H., Chadha, M., & Wang, F. (2019). Proximity and networked news public: Structural topic modeling of global Twitter conversations about the 2017 Quebec mosque shooting. *International Journal of Communication*, *13*, 2652–2675.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., . . . Alstyne, M. V. (2009). Computational Social Science. *Science, 323*(5915), 721–723. doi:10.1126/science.1167742

Lewis, S. C., Zamith, R., & Hermida, A. (2013). Content analysis in an era of big data: A hybrid approach to computational and manual methods. *Journal of Broadcasting & Electronic Media*, *57*(1), 34–52. doi:10.1080/08838151.2012.761702

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., . . . Häussler, T. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures*, *12*(2/3), 93–118. doi:10.1080/19312458.2018.1430754

Monroe, B. L. (2013). The five Vs of big data political science: Introduction to the virtual issue on big data in political science political analysis. *Political Analysis*, *21*(V5), 1–9. doi:10.1017/S1047198700014315

Montag, C., Błaszkiewicz, K., Lachmann, B., Sariyska, R., Andone, I., Trendafilov, B., & Markowetz, A. (2015). Recorded behavior as a valuable resource for diagnostics in mobile phone addiction: Evidence from psychoinformatics. *Behavioral Sciences, 5*(4), 434–442. doi:10.3390/bs5040434

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Du Sert, N. P., . . . Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*(1), 0021. doi:10.1038/s41562-016-0021

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., . . . Yarkoni, T. (2015). Promoting an open research culture. *Science, 348*(6242), 1422–1425. doi:10.1126/science.aab2374

Rudkowsky, E., Haselmayer, M., Wastian, M., Jenny, M., Emrich, Š., & Sedlmair, M. (2018). More than bags of words: Sentiment analysis with word embeddings. *Communication Methods and Measures*, *12*(2/3), 140–157. doi:10.1080/19312458.2018.1455817

Scharkow, M. (2013). Thematic content analysis using supervised machine learning: An empirical evaluation using German online news. *Quality & Quantity*, *47*(2), 761–773.

Shah, D. V., Hanna, A., Bucy, E. P., Lassen, D. S., Van Thomme, J., Bialik, K., . . . Pevehouse, J. (2016). Dual screening during presidential debates. *American Behavioral Scientist, 60*(14), 1816–1843. doi:10.1177/0002764216676245

Trilling, D., & Jonkman, J. G. (2018). Scaling up content analysis. *Communication Methods and Measures*, *12*(2/3), 158–174. doi:10.1080/19312458.2018.1447655

van Atteveldt, W., Margolin, D., Shen, C., Trilling, D., & Weber, R. (2019). *A roadmap for computational communication research.* doi:10.31235/osf.io/4dhfk

van Atteveldt, W., & Peng, T. Q. (2018). When communication meets computation: opportunities, challenges, and pitfalls in computational communication science. *Communication Methods and Measures, 12*(2/3), 81–92. doi:10.1080/19312458.2018.1458084

Wallach, H. (2018). Computational social science ≠ computer science + social data. *Communications of the ACM, 61*(3), 42–44. doi:10.1145/3132698

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122. doi:10.1177/1745691617693393