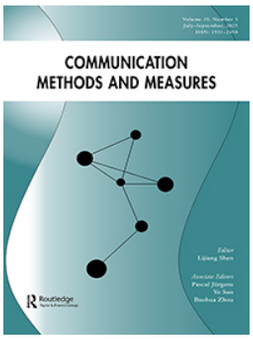

Routledge
Taylor & Francis Group

Communication Methods and Measures

ISSN: 1931-2458 (Print) 1931-2466 (Online) Journal homepage: www.tandfonline.com/journals/hcms20

Critical, but constructive: defining, detecting, and addressing bias in Computational Social Science

Valerie Hase, Marko Bachl & Nathan TeBlunthuis

To cite this article: Valerie Hase, Marko Bachl & Nathan TeBlunthuis (29 Oct 2025): Critical, but constructive: defining, detecting, and addressing bias in Computational Social Science, Communication Methods and Measures, DOI: 10.1080/19312458.2025.2575468

To link to this article: https://doi.org/10.1080/19312458.2025.2575468

Published online: 29 Oct 2025.

COMMUNICATION METHODS AND MEASURES
https://doi.org/10.1080/19312458.2025.2575468

Routledge
Taylor & Francis Group

Critical, but constructive: defining, detecting, and addressing bias in Computational Social Science

Valerie Hase [a], Marko Bachl [b], and Nathan TeBlunthuis [c]

[a]Department of Media and Communication, LMU Munich, Munich, Germany; [b]Institute for Media and Communication Studies, Freie Universität Berlin, Berlin, Germany; [c]School of Information, University of Texas at Austin, Texas, USA

ABSTRACT

Computational Social Science (CSS) increasingly engages in critical discussions about bias in and through computational methods. Two developments drive this shift: first, the recognition of bias as a societal problem, as flawed CSS methods in socio-technical systems can perpetuate structural inequalities; and second, the field's growing methodological resources, which create not only the opportunity but also the responsibility to confront bias. In this editorial to our Special Issue on CSS and bias, we introduce the contributions and outline a research agenda. In defining bias, we emphasize the importance of embracing epistemological pluralism while balancing the need for standardization with methodological diversity. Detecting bias requires stronger integration of bias detection into validation procedures and the establishment of shared metrics and thresholds across studies. Finally, addressing bias involves adapting established and emerging error-correction strategies from social science traditions to CSS, as well as leveraging bias as an analytical resource for revealing structural inequalities in society. Moving forward, progress in defining, detecting, and addressing bias will require both bottom-up engagement by researchers and top-down institutional support. This Special Issue positions bias as a central theme in CSS – one that the field now has both the tools and the obligation to address.

In its early stages, Computational Social Science (CSS) was accompanied by bold claims about its potential, with some scholars predicting "an entirely new scientific approach for social analysis" (Conte et al., 2012, p. 327). Margolin noted that many social scientists regarded this as an "opening salvo" (2019, p. 232), which positioned CSS as a potential rival to established methodologies. Critics were quick to challenge such sweeping promises of "computational means to know everything (I), of anything (II), free from bias (III), with a high degree of certainty (IV)" (Rieder & Simon, 2017, p. 91). For example, boyd and Crawford argued that just because CSS can quantify social phenomena, it "does not necessarily have a closer claim on objective truth" (2012, p. 667), with critical voices having only grown louder in recent years (Shugars, 2024). In parallel, scholars empirically challenged claims that CSS data and methods would necessarily serve as a new, unbiased gold standard (Jürgens et al., 2020).

More than a decade later, CSS has transitioned from an emerging to an established approach, enriching rather than displacing existing methodologies (Van Atteveldt & Peng, 2018). This maturing process has brought critical voices more into the mainstream and underscored an important recognition: *CSS is by no means free of bias*. In line with this, previous Special Issues on CSS have touched on bias (Van Atteveldt & Peng, 2018), for example, in relation to digital traces (Conrad et al., 2021; Loecherbach & Otto, 2024), their integration with surveys (Stier et al., 2020), or LLMs (Davidson & Karell, 2025). More recently, Weiß et al. (2025) have more explicitly discussed the issue of often

CONTACT Valerie Hase valerie.hase@ifkw.lmu.de Department of Media and Communication, LMU Munich, Akademiestr. 7, Munich 80799, Germany

insufficient data quality in digital traces. Against this backdrop of numerous Special Issues devoted to computational methods, one might reasonably ask: Why devote another to bias in CSS?

A critical view: bias in and through computational methods

We argue that attending to bias as a central concern rather than as a secondary theme is essential, and for multiple reasons. Like other disciplines, CSS is embedded in social contexts marked by structural inequalities, to which it is not immune. These include the predominance of Western voices in shaping knowledge production (Yi & Zhang, 2023), persistent gender disparities in access to computational methods (Van Der Velden & Dolinsky, 2024), and the often criticized narrowness of research agendas (Waldherr et al., 2025) – all of which risk perpetuating bias within the field. Second, CSS draws upon a diverse set of disciplinary traditions. Each community brings with it its own epistemological orientations and methodological standards (Theocharis & Jungherr, 2021), shaping competing understandings of what constitutes bias and how it ought to be handled. This heterogeneity is magnified by the field's reliance on new data (e.g., digital traces) and methods (e.g., text-as-data), which partly demand new methods for detecting bias to ensure methodological rigor. Found data and black-box methods developed by large technology firms, in particular, further limit the extent to which researchers control samples and measures. Finally, because CSS increasingly underpins socio-technical systems that shape everyday life, its bias can produce harmful effects beyond academia (Wagner et al., 2021). Examples include the silencing of marginalized voices through algorithmic content moderation (Sap et al., 2019) or racially biased health outcomes produced by medical algorithms (Obermeyer et al., 2019). In these contexts, CSS too often projects an illusion of objectivity while obscuring its tangible real-world consequences. Recognizing and confronting bias, then, is not only a methodological imperative but also an ethical one. In short, we *should* be concerned about bias in CSS.

Being constructive: why now is the best possible time to confront bias in CSS

At the same time, this is an excellent moment to address bias in CSS. As the field matures, it has accumulated more resources than ever to tackle this challenge. Among these is a growing embracement of epistemological pluralism (Kitchin, 2014; Shugars, 2024), which provides a foundation for defining bias in CSS from different vantage points. Simultaneously, the "palpable need for developing proper guidelines and standards" (Waldherr et al., 2025, p. 14) has been met with methodological guidelines for CSS (Carrière et al., 2025), different validation methods (Bernhard-Harrer et al., 2025; Birkenmaier, Lechner, et al., 2024), and research infrastructures (Balluff et al., 2023). At the very least, this illustrates the field's growing capacity to confront bias based on a set of different, yet equally valued methods: whether inherited from non-computational traditions (e.g., surveys, content analysis) or those developed further within CSS (e.g., digital traces, text-as-data, machine learning).

Growing concerns about bias in CSS, combined with the resources now available to confront it, place a new responsibility on the field. Having recognized the problem of bias and developed the means to tackle it, we can no longer rest on criticizing where CSS falls short of its initial promises. The task ahead is to *transform critique into constructive contribution*. For communication research, this entails not only defining and detecting bias but actively shaping the methods by which it can be mitigated. We see this Special Issue as a forum for such *critical yet constructive engagement with CSS and bias*. Alongside introducing the contributions featured in this Special Issue, we provide readers with an overview of scholarship on bias and sketch directions for a future research agenda.

Defining bias in CSS

A central question when discussing bias in CSS is whether we are all, in fact, referring to the same phenomenon. Definitions not only diverge across disciplines – such as the social sciences

(Hammersley & Gomm, 1997) and technical fields (Blodgett et al., 2020) – but also within them. From a measurement perspective, bias refers to a systematic difference between the true value of a quantity for a population and how a study observes it. From a social perspective, bias is framed in terms of outcomes, namely as the systematic, unjust treatment of individuals or groups due to flawed CSS methods (Ntoutsi et al., 2020). Both perspectives are interrelated (Boelaert et al., 2025): Statistical bias can feed directly into unfair social outcomes. At the same time, they invoke different quality criteria: bias as a threat to validity in scientific research or as a threat to fairness in society. This becomes evident when discussing examples of bias in and through CSS: Machine-learning classifiers, for example, may misclassify and flag minority-group speech as hate speech (TeBlunthuis et al., 2024), particularly when training data fails to adequately represent specific groups (Sap et al., 2019). As a result, minority-group content may more likely be removed on digital platforms and, thus, statistical bias may cascade into social bias. Similarly, text-as-data approaches and computer vision methods often misclassify or ignore non-binary genders (Siemon, 2025). This undermines both scientific accuracy and social equity, with flawed algorithms disadvantaging individuals in contexts such as online job searches based on their gender (Shekhawat et al., 2019). In many ways, such differences – for example, with respect to validity or fairness as relevant quality criteria for evaluating concerns about bias – parallel established debates outside CSS, such as those between quantitative and qualitative traditions. We argue that grappling with ambiguity in definitions requires two things: (1) embracing epistemological pluralism and (2) balancing the need for standardization and methodological pluralism.

Embracing epistemological pluralism in CSS

Difficulty in defining bias arises from epistemological pluralism within CSS (Kitchin, 2014). Although epistemological debates on bias are hardly new (Hammersley & Gomm, 1997), they are amplified in CSS (Shugars, 2024). Methodologies inevitably rest on epistemological assumptions about what counts as valid knowledge, and those assumptions differ across traditions. As a field where the social sciences and technical fields intersect, CSS embodies the tension between divergent traditions. Communication scholars, for example, often operate within (post-)positivist epistemologies (Walter et al., 2018), relying on empirical data while acknowledging its imperfections. Technical sciences, by contrast, have partly favored empiricist approaches that assume data speaks for itself (Orlikowski & Baroudi, 1991). When considering how CSS can accurately represent reality, this pluralism also forces CSS scholars to grapple with ontological questions. Consider the use of CSS to study latent constructs like racism (Kathirgamalingam et al., 2025): perspectives on racism – and whether CSS can measure this – may differ profoundly depending on participants' lived experiences. As such, researchers must ask themselves: Is there a single truth that researchers can uncover, with bias as deviation that is to be corrected? Or are there multiple truths, shaped by the perspectives of those who produce them, with bias reflecting structural inequalities that ought to be understood rather than eliminated (Shugars, 2024)?

Fortunately, CSS scholars have not only scrutinized epistemological tensions but also advanced related conceptual work, in line with what we consider a constructive approach. This includes efforts to anchor CSS in critical traditions (Shugars, 2024; Siemon, 2025; Törnberg & Uitermark, 2021), synthesizing interpretivist approaches with large-scale approaches (Nelson, 2020), and re-conceptualizing gold standards that integrate multiple perspectives (Baden et al., 2023; Cabitza et al., 2023). Such conceptual work extends to more agnostic approaches: favoring the relative, goal-dependent validity of methods instead of assuming a single true underlying data-generating process a single method can accurately capture (Bernhard-Harrer et al., 2025; Grimmer et al., 2021), something similarly described as "fitness for purpose" (Weiß et al., 2025, p. 934). The challenge ahead will be to integrate such debates more strongly across disciplinary silos. Epistemological diversity has long coexisted alongside dominant paradigms in both the social sciences (Walter et al., 2018) and technical disciplines (Eden, 2007), with critical voices becoming especially prominent in the latter (Bender et al.,

2021). We therefore believe that CSS is well-positioned to embrace such pluralism even more explicitly in the future.

Balancing standardization and methodological pluralism

Another challenge in defining bias stems from the tension between demands for standardization and the reality of methodological pluralism. On the one hand, establishing common standards, including terminologies for quality criteria and bias, is seen as essential to consolidating CSS (Waldherr et al., 2025) and avoiding a "fragmentation" (Birkenmaier, Daikeler, et al., 2024, p. 5) of the field. On the other hand, CSS thrives precisely because it develops multiple, sometimes competing approaches to similar problems. For instance, the collection of digital traces can be pursued through APIs, tracking, or data donation. Similarly, text-as-data methods include dictionaries, supervised machine learning, and few-shot classification. However, this methodological pluralism within and across methods challenges shared standards. Because different CSS methods are, for example, linked to distinct forms of bias, it is difficult – and sometimes unhelpful – to impose overarching terminologies.

By taking shared concepts as a point of departure, researchers have constructively charted ways to adapt definitions for method-specific understandings of bias. For example, the field often draws on the Total Survey Error Framework, which distinguishes between types of bias, and adapts it to define method-specific bias in data donation (Boeschoten et al., 2022) or platform data (Sen et al., 2021). But even when building from shared frameworks, risks of fragmentation persist: Daikeler et al. (2024) identified more than 20 error frameworks for CSS data, often without sufficient exchanges across methods and disciplines. As CSS continues to expand, the challenge ahead will be to strike a balance: shared definitions of quality criteria and bias that are broad enough to unify, yet flexible enough to capture methodological pluralism.

Detecting bias in CSS

A second question in CSS concerns methods, metrics, and thresholds for detecting bias: Given the study design, which methods allow us to identify bias, what metrics can quantify it, and where should we set thresholds beyond which bias undermines substantive findings? Despite the urgency of these questions, the lack of guidelines for bias detection in CSS is striking, with only a few notable exceptions (Moons et al., 2025; Tay et al., 2022). Most discussions remain folded into broader debates about data quality – encompassing different criteria such as reliability and validity – or validity itself, which is often equated with the absence of bias. Related reviews map approaches for detecting data quality and validity in text-as-data (Bernhard-Harrer et al., 2025; Birkenmaier, Lechner, et al., 2024) or agent-based modeling (Collins et al., 2024), but often without an explicit focus on bias. Strengthening bias detection in CSS involves two key steps: (1) enhancing validation procedures with targeted bias detection and (2) developing shared metrics and thresholds.

Validation needs more explicit bias detection: choosing among biased methods

In the absence of dedicated methods for bias detection, researchers often borrow validation procedures. Given the close relationship between validity and bias, this strategy is useful – but not sufficient (TeBlunthuis et al., 2024). Often, though not always, validation is conducted through comparisons to presumably less biased gold standards, against which researchers approximate the extent to which CSS results deviate from "true" population values. For example, text-as-data measures are validated against manual coding (Bernhard-Harrer et al., 2025), samples in data donation studies to survey samples (Hase & Haim, 2024), and LLM-generated synthetic data against survey data (Boelaert et al., 2025). In such cases, non-CSS data sources, such as decisions by human coders, are privileged as gold standards, based on the belief that researcher control reduces bias. In other cases, however, CSS data itself serves as the benchmark – for instance, when

digital traces are used to detect bias in self-reports or when CSS is used to create synthetic, "debiased" datasets (for a critical debate, see Toh & Park, 2025). Crucially, the idea of a gold standard rests on epistemological judgments about what counts as valid knowledge. Again, such evaluations can and sometimes should be purpose-specific (Weiß et al., 2025). In line with broader epistemological and ontological debates, scholars increasingly even question the very notion of gold standards – whether derived from CSS (Jürgens et al., 2020) or non-CSS methods (Bachl & Scharkow, 2017). As such, focusing only on deviations from such standards may be insufficient to fully capture bias in CSS data and methods.

In response to critical perspectives, scholars are adopting constructive approaches that more explicitly integrate bias detection in existing validation methods. They argue that methods vary in their strengths and limitations depending on the task (Bernhard-Harrer et al., 2025; Grimmer et al., 2021), a point not new, but one sometimes overlooked with the rise of CSS as a supposedly novel approach. Digital trace data, for instance, may provide more accurate measures of app- or device-specific usage, whereas surveys may often better capture activity across platforms and devices (Wenz et al., 2024). Cernat et al. (2024) demonstrate this with the Multitrait Multimethod approach, revealing that both survey and digital trace data introduce errors that vary with the latent concept under investigation. The key is to identify which biases attach to which methods – and selecting the method whose biases least undermine one's research question or even integrating different methods (Wenz et al., 2024). Generally, such mixed methods approaches promise to build stronger evidence through "triangulation" when results from data analyses with different sorts of biases converge (Creamer, 2018). Validation remains a cornerstone, but it must be paired with greater attention to when, how, and why CSS may misrepresent the phenomena it seeks to capture. Taken together, these developments point toward a more pragmatic vision of bias detection: one that abandons the pursuit of perfect benchmarks in favor of comparative assessments of biases across CSS and non-CSS methods.

Metrics and thresholds for bias detection: how good is good enough?

Another point of contention is whether metrics and thresholds borrowed from validation methods can also function as safeguards against bias in CSS. To date, such measures often emphasize whether differences between methods exist, rather than whether and how those differences are systematic. Take recall and precision, for example: while these metrics can show how well text-as-data classifiers align with manual annotations overall, they provide little insight into systematic shortcomings, such as consistent misclassification of certain languages or specific sub-indicators of a latent concept, unless explicitly examined. For this reason, critics caution that transparency in the form of general validation metrics, while important, is ultimately not enough (TeBlunthuis et al., 2024).

Fortunately, progress has been made in integrating more bias-sensitive metrics into CSS. Researchers have adapted metrics from survey research to quantify nonparticipation bias in tracking (Keusch et al., 2020) and advanced approaches for differential error in text-as-data (Egami et al., 2023; TeBlunthuis et al., 2024). Similar efforts extend to synthetic data generated via LLMs, such as comparing the distance of probability distributions between LLM-generated and survey data (Boelaert et al., 2025). To understand thresholds at which bias may substantially threaten validity, studies largely illustrate impacts on results. Some show how bias may change descriptive quantities of interest, such as the prevalence of substantial concepts like political polarization (Chen et al., 2025) or platform use (Bosch et al., 2025). Others examine how bias cascades into inferential relationships, for instance, when explaining toxicity on social media (TeBlunthuis et al., 2024) or patterns of digital platform use (Bosch et al., 2025). Still, procedures and guidelines specifically tailored to bias detection are more established outside CSS – in survey methodology (The American Association for Public Opinion Research, 2023) or in machine learning research, where subgroup and intersectional fairness metrics are more common (Caton & Haas, 2024). The challenge ahead for CSS is therefore to develop standard metrics for bias detection and thresholds for when bias substantially undermines validity.

Addressing bias in CSS

Finally, how should researchers handle bias once they detect it? Here, the field has to address two pressing issues: (1) testing whether non-CSS error correction can be transferred to CSS and (2) thinking about bias as an analytical resource.

Turning found data into designed data: adapting traditional bias correction for CSS?

With growing recognition of bias in CSS, scholars have developed error-correction strategies to address it. These fall into two categories: ex-ante strategies implemented prior to data collection and a-posteriori strategies applied afterward (Caton & Haas, 2024; Hase & Haim, 2024). These approaches seek to reintroduce control – effectively aiming to transform "found" data into more "designed" data. As such strategies borrow heavily from established social science practices, they meet calls for social scientists to take a more active role in CSS. Ex-ante strategies include survey design techniques, such as incentivization or framing, to mitigate nonparticipation bias in studies collecting digital trace data (Hase & Haim, 2024). They also encompass instructions for LLMs via promptbooks grounded on traditions of manual content analysis (Stuhler et al., 2025) or changing the conditions under which coders annotate manual gold standards (Van Der Velden et al., 2025). A-posteriori strategies rely on statistical corrections, including weighting (Pak et al., 2022) or adjustments for missing data (TeBlunthuis et al., 2024).

Efforts to mitigate bias in CSS are promising and represent constructive advances, especially when explicitly building bridges across disciplines. This includes interdisciplinary adaptations of methods, such as machine learning (Lavelle-Hill et al., 2025), or collaborative policy initiatives pushing for improved data access (Hase et al., 2024). However, the effectiveness of such mitigation strategies is far from settled. Error-correction techniques depend on high-quality validation data – which, as debates about gold standards show, are hard to find – and knowledge of the processes through which bias emerges. Researchers cannot model "unknown unknowns": nonparticipation in tracking or data donation, for instance, is unlikely to be corrected via sociodemographic variables typically used in statistical weighting. Since reliable predictors of selective drop-out are not yet well established – especially for actual participation rather than hypothetical behavior measured via vignette experiments – error correction risks being ineffective or even generating additional bias. Survey design interventions face similar hurdles: most techniques appear to have little to no effect on bias in user-centric data collection (Hase & Haim, 2024; Struminskaya et al., 2021), likely because attrition in CSS settings follows dynamics distinct from traditional surveys. We believe that both ex-ante and a-posteriori strategies remain worth pursuing, but their successful application will require stronger evidence on types and predictors of bias in CSS. Still, abandoning bias mitigation simply because it is difficult is not an option.

Bias as a feature of reality: from transparency to analytical resource

Moreover, scholars influenced by critical theory argue that bias should be used rather than treated exclusively as a problem to erase (Shugars, 2024). As Toh and Park (2025) warn, CSS has developed a tendency to frame bias as a purely technical issue, seeking technical fixes while neglecting the structural forces that produce it. Since CSS data and methods rely on real-world information, they inevitably reflect existing social inequalities. From this standpoint, bias can be reframed more productively as a lens for scientific analysis (Hammersley & Gomm, 1997), one that highlights structural inequalities in society.

Scholars have addressed this critique along two lines. The first is a push toward a more transparent error culture through documentation tools. Technical fields have pioneered, for example, model cards for machine learning classifiers (Mitchell et al., 2019) and data statements in NLP (Bender & Friedman, 2018), though these also gain traction in the social sciences

(Kostygina et al., 2023). In essence, a transparent error culture requires researchers to explicitly document the shortcomings of their CSS data and methods and outline the conditions under which their methods should – or should not – be applied. Only through transparent documentation can the field accumulate the evidence needed for systematic reviews and meta-analyses for identifying common predictors and mechanisms of bias in CSS. These technical reports can be paired with positionality statements that disclose researchers' own characteristics, such as their epistemological orientation or relationship to the subject. The second line of response is to turn bias into an analytic resource by using biased CSS methods to reveal structural inequalities, for instance, detecting bias in text-as-data approaches to uncover gender stereotypes in society represented in their underlying training data (Garg et al., 2018).

Contributions in this Special Issue

In 2024, we launched a Call for Papers in *Communication Methods and Measures* on the theme of CSS and bias. From 57 abstract submissions, 8 author teams were invited to submit full manuscripts, and 4 were ultimately accepted for publication in this Special Issue. The accepted contributions largely converge on the topic of bias detection. They include studies focusing on gender as a site of bias in text-as-data approaches (Siemon, 2025) and evidence that such classifiers perform unevenly in detecting incivility directed toward different genders (Stoll et al., 2025). Other contributions examine how bias affects the study of political polarization when comparing APIs and data donations as CSS data sources (Chen et al., 2025) or how coder characteristics, such as political orientation, influence the creation of gold standards (Van Der Velden et al., 2025). By contrast, fewer contributions also advance definitions of bias or strategies for mitigation (Siemon, 2025; Van Der Velden et al., 2025). In what follows, we briefly introduce each of the four articles (for an overview, see Table 1).

Chen et al. (2025) demonstrate that different CSS data sources may carry distinct errors, with important consequences for answering substantive questions in the social sciences. Drawing on Hungarian data donations combined with survey responses and the YouTube API, they contribute to this Special Issue by detecting bias in CSS measures of political polarization. Their findings show that visible digital trace data (e.g., comments) exhibits more political polarization than invisible data

Table 1. Contributions per paper.

Authors and title	Defining bias	Detecting bias	Addressing bias	Example quote
Chen et al. (2025): The Public that Engages Invisibly: What Visible Engagement Fails to Capture in Online Political Communication		✓		*"Visible engagement on YouTube portrays an incomplete picture of political communication, potentially leading to inaccurate and often exaggerated estimates of polarization."*
Siemon (2025): Beyond the Binary? Automated Gender Classification of Social Media Profiles	✓	✓	✓	*"To critically evaluate gender classification methods, we need to integrate theoretical knowledge of feminist and queer theory with methodological expertise in standardized and computational research."*
Van Der Velden et al. (2025): Whose Truth is it Anyway? An Experiment on Annotation Bias in Times of Factual Opinion Polarization	✓	✓	✓	*"To improve annotated data for automated text analyses, and for stance detection models in particular, we need to critically evaluate how we create our gold standards."*
Stoll et al. (2025): Classification Bias of LLMs in Detecting Incivility Towards Female and Male Politicians In German Social Media Discourse		✓		*"If comments containing vulgarity and direct insults are classified more reliably than those involving stereotyping or subtle discrimination, LLMs risk reinforcing a narrow view of incivility [. . .]. This bias in classification can lead to the marginalization of social groups."*

Note. The format of this table was inspired by Van Rooij et al. (2024).

(e.g., video consumption). As a result, the choice of data sources and measurement strategies can lead researchers to over- or underestimate phenomena such as political polarization and filter bubbles. Their study contributes to comparative assessments of biases, here across different CSS data sources, and to understanding metrics for bias detection. Moreover, their extensive robustness tests exemplify the type of transparency culture that CSS increasingly requires.

Siemon (2025) advances text-as-data methods for gender classification. Drawing on feminist and queer theory, she urges researchers to critically reflect on their epistemological assumptions when defining gender as a research category, particularly because such classifications can affect marginalized communities. Through a case study of German Twitter data, she offers a set of guidelines for text-as-data research, addressing steps like setting up the research design, method selection, data preparation, analysis, and validation. She further illustrates how bias detection can be implemented, for example, by identifying systematic misclassification across groups and by noting which populations remain unclassifiable and thus invisible. As such, her study underlines the importance of epistemological pluralism for defining bias as well as approaches to empirically detecting it.

Contributing to definitional work on bias in CSS, Van Der Velden et al. (2025) propose that coder disagreement – often dismissed as mere measurement error – should be treated as an informative signal in text-as-data methods. This is particularly relevant when dealing with latent concepts that are prone to political polarization. Through experiments in the Netherlands and the United States, the authors examine how annotator characteristics, such as coders' ideological distance, as well as the latentness of the concept under study, can affect bias. Van der Velden et al. also evaluate error correction strategies, such as masking text elements that may trigger ideologically biased annotations, though this mitigation's effect was limited in the current study. The study contributes to defining bias while also advancing approaches for detecting and mitigating it, thereby addressing several key themes of this Special Issue.

Finally, Stoll et al. (2025) analyze the unequal performance of LLMs in detecting incivility toward politicians of different genders. Using user comments from YouTube and X/Twitter, they investigate how both the nature of the concept under study (e.g., the latentness of indicators of incivility) and its object (e.g., the gender of the politician) shape bias. Their findings show that classifiers often overlook more implicit forms of incivility, which are disproportionately directed toward women. As a result, classifier performance is worse for these groups and may reinforce gender bias. Beyond contributing to debates on validation, their study also speaks to fairness in and through CSS, particularly by advancing methodological discussions on metrics for bias detection through an in-depth analysis of false-negative classifications and their underlying causes.

Charting the road ahead: bias in and through computational methods

Based on our review of existing scholarship on defining, detecting, and addressing bias in CSS, together with the contributions to this Special Issue, we argue that CSS researchers must adopt a dual orientation: on the one hand, a critical stance committed to defining and detecting bias; on the other, a constructive stance that seeks to address or even leverage it. If CSS is to maintain its credibility, confronting bias must become a central, not peripheral, endeavor – and this Special Issue represents a step in that direction. Meeting this challenge requires joint efforts of bottom-up commitment by individual researchers, complemented by top-down institutional engagement.

How to move forward: bottom-up commitment by CSS researchers

CSS researchers can, and already do, engage with bias in ways that strengthen the field. For defining bias, this entails advancing conceptual work from multiple epistemological standpoints (Siemon, 2025). This could extend to critically reassessing CSS studies – for instance, research on Twitter/X, which remains used as evidence of human online behavior despite biases in the populations it represents. Finally, we urge researchers to more explicitly define bias in their CSS studies, both

conceptually (e.g., related to valid statistical estimation or societal fairness) and formally (related to bias calculation, see, for example, Keusch et al., 2020). Ultimately, this also means that CSS will need to more explicitly integrate broader quality criteria beyond validity, such as its impact on societal fairness (Wagner et al., 2021).

Relatedly, extending validation practices to incorporate bias detection broadens the field's scope from bias *in* CSS to bias *through* CSS. Researchers could extend theory-informed work with studies that also, from an applied perspective, aim at solving societal problems (Lazer et al., 2020). As socio-technical systems governed by big tech expand their influence, CSS has both the opportunity and, arguably, the obligation to inform policymakers and practitioners (Wagner et al., 2021). In addition, more agnostic approaches (Bernhard-Harrer et al., 2025; Grimmer et al., 2021) increasingly underline that no method, CSS or otherwise, is perfect. Different approaches carry distinct strengths and biases. Moving forward, this requires more methodological work on comparative bias detection across methods – both within different CSS approaches but also across CSS and non-CS work.

Finally, and in relation to addressing bias, it has long been emphasized that CSS should complement rather than replace traditional methods (Van Atteveldt & Peng, 2018). At present, however, this potential remains only partially realized: complementarity has mainly taken the form of comparison (e.g., juxtaposing survey and digital trace data to assess bias, see Wenz et al., 2024) or sequential use (e.g., applying manual coding alongside text-as-data for different tasks, see Nelson, 2020). A next step would be integration: using CSS and non-CSS methods simultaneously for the same task to account for different biases across them. As Törnberg and Uitermark (2021) put it, such methodological pluralism could serve to "bring into view different dimensions of social reality" (p. 8), just as diverse coder perspectives can enrich annotation (Cabitza et al., 2023). To fully realize this potential, CSS has to advance methods for integrating different CSS methods, as well as CSS and non-CSS approaches.

How to move forward: top-down support by institutions

We recognize that this requires additional work on top of other calls, such as those for greater engagement with open science and reproducibility practices in CSS (Waldherr et al., 2025). The burden of confronting bias cannot rest solely on the shoulders of individual CSS researchers; it also requires systemic, top-down action from institutions. Such support could take the form of institutions providing standardized bias metrics and reporting frameworks, modeled on longstanding practices in survey research (The American Association for Public Opinion Research, 2023). Journals, which "are in a prime position to set standards" (Waldherr et al., 2025, p. 22), could incorporate these into review processes. This might include providing templates for positionality statements and bias reports, as well as the introduction of quality badges that reward researchers who proactively engage with bias detection. It could also involve developing shared data infrastructures that supply either high-quality or explicitly biased datasets, both of which could function as valuable benchmarks for bias detection. Finally, infrastructures such as the KODAQS initiative by GESIS (Dehne et al., 2024) illustrate how training and guidelines on data quality and bias can empower researchers and provide lasting support for the field.

What this Special Issue could not do

Although this Special Issue seeks to advance debates about bias in CSS, several limitations remain. First, structural inequalities within the field are mirrored in the contributions. Of the published pieces, a majority is led by female researchers, yet none by scholars outside Europe. This outcome partly reflects submission patterns but nonetheless reproduces the Western-centrism of CSS (Yi & Zhang, 2023). Second, despite our call for both critical and constructive work, future research could place even greater emphasis on advancing epistemologically diverse definitions of bias *and* strategies for addressing it. Third, most contributions prioritize implications for scientific validity over questions of

societal fairness, indicating that this issue engages primarily with bias *in* CSS rather than bias *through* CSS. Even so, we hope that the range of perspectives assembled – spanning definitions, detection methods, and strategies for addressing bias – will provide a foundation for future debates on CSS and bias.

Acknowledgments

We would like to thank everyone who submitted to this Special Issue. We are also grateful to the reviewers for their valuable feedback and to the authors for thoughtfully engaging with it. Finally, we thank the journal, and especially the editor-in-chief, Lijiang Shen, for trusting us with this Special Issue. The authors used GPT-5 (OpenAI, accessed October 2025) to assist with language refinement during manuscript preparation. All content was initially written by the authors and, after language refinement by GPT-5, reviewed and approved by them.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Notes on contributors

Dr. Valerie Hase is a postdoctoral researcher at the Department of Media and Communication, LMU Munich. Her research focuses on automated content analysis, digital trace data, bias in computational methods, and digital journalism.

Marko Bachl is an Assistant Professor at the Institute for Media and Communication Studies, Freie Universität Berlin. His research focuses on evaluating and improving methods for communication research and their application in political and health communication.

Nathan TeBlunthuis is an Assistant Professor at the School of Information, University of Texas at Austin. His research focuses on online communities, digitally mediated organization, collaborative knowledge work, and computational methods for communication research.

ORCID

Valerie Hase http://orcid.org/0000-0001-6656-4894
Marko Bachl http://orcid.org/0000-0001-5852-4948
Nathan TeBlunthuis http://orcid.org/0000-0002-3333-5013

References

Bachl, M., & Scharkow, M. (2017). Correcting measurement error in content analysis. *Communication Methods and Measures*, *11*(2), 87–104. https://doi.org/10.1080/19312458.2017.1305103

Baden, C., Boxman-Shabtai, L., Tenenboim-Weinblatt, K., Overbeck, M., & Aharoni, T. (2023). Meaning multiplicity and valid disagreement in textual measurement: A plea for a revised notion of reliability. *Studies in Communication & Media*, *12*(4), 305–326. https://doi.org/10.5771/2192-4007-2023-4-305

Balluff, P., Lind, F., Boomgaarden, H. G., & Waldherr, A. (2023). Mapping the European media landscape - Meteor, a curated and community-coded inventory of news sources. *European Journal of Communication*, *38*(2), 181–194. https://doi.org/10/grmt54

Bender, E. M., & Friedman, B. (2018). Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Transactions of the Association for Computational Linguistics*, *6*, 587–604. https://doi.org/10.1162/tacl_a_00041

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? 🦜 In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Bernhard-Harrer, J., Ashour, R., Eberl, J.-M., Tolochko, P., & Boomgaarden, H. (2025). Beyond standardization: A comprehensive review of topic modeling validation methods for computational social science research. *Political Science Research and Methods*, 1–19. https://doi.org/10.1017/psrm.2025.10008

Birkenmaier, L., Daikeler, J., Fröhling, L., Gummer, T., Lechner, C., Lux, V., Schwalbach, J., Silber, H., Weiß, B., Weller, K., Wolf, C., Abel, D., Breuer, J., Dietze, S., Dimitrov, D., Döring, H., Hebel, A., Hochman, O., Jünger, S. Ziaja, S. . . . Ziaja, S. (2024). *Defining and evaluating data quality for the social sciences: Position paper*. GESIS papers. https://doi.org/10.21241/SSOAR.96764

Birkenmaier, L., Lechner, C. M., & Wagner, C. (2024). The search for solid ground in text as data: A systematic review of validation practices and practical recommendations for validation. *Communication Methods and Measures*, *18*(3), 249–277. https://doi.org/10.1080/19312458.2023.2285765

Blodgett, S. L., Barocas, S., Daumé, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of "bias" in NLP (No. arXiv: 2005.14050). *arXiv*. https://doi.org/10.48550/arXiv.2005.14050

Boelaert, J., Coavoux, S., Ollion, É., Petev, I., & Präg, P. (2025). Machine bias. How do generative language models answer opinion polls? *Sociological Methods & Research*, 00491241251330582. https://doi.org/10.1177/00491241251330582

Boeschoten, L., Ausloos, J., Möller, J. E., Araujo, T., & Oberski, D. L. (2022). A framework for privacy preserving digital trace data collection through data donation. *Computational Communication Research*, *4*(2), 388–423. https://doi.org/10.5117/CCR2022.2.002.BOES

Bosch, O. J., Sturgis, P., Kuha, J., & Revilla, M. (2025). Uncovering digital trace data biases: Tracking undercoverage in web tracking data. *Communication Methods and Measures*, *19*(2), 157–177. https://doi.org/10.1080/19312458.2024.2393165

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. https://doi.org/10.1080/1369118X.2012.678878

Cabitza, F., Campagner, A., & Basile, V. (2023). Toward a perspectivist turn in ground truthing for predictive computing. *Proceedings of the AAAI Conference on Artificial Intelligence*, *37*(6), 6860–6868. https://doi.org/10.1609/aaai.v37i6.25840

Carrière, T. C., Boeschoten, L., Struminskaya, B., Janssen, H. L., De Schipper, N. C., & Araujo, T. (2025). Best practices for studies using data donation *Quality & Quantity*, 59(S1), 389–412. https://doi.org/10.1007/s11135-024-01983-x

Caton, S., & Haas, C. (2024). Fairness in machine learning: A survey. *ACM Computing Surveys*, *56*(7), 1–38. https://doi.org/10.1145/3616865

Cernat, A., Keusch, F., Bach, R. L., & Pankowska, P. K. (2024). Estimating measurement quality in digital trace data and surveys using the MultiTrait MultiMethod model. *Social Science Computer Review*, *43*(5), 08944393241254464. https://doi.org/10.1177/08944393241254464

Chen, Y., Kmetty, Z., Iñiguez, G., & Omodei, E. (2025). The public that engages invisibly: What visible engagement fails to capture in online political communication. *Communication Methods and Measures*, 1–19. https://doi.org/10.1080/19312458.2025.2542725

Collins, A., Koehler, M., & Lynch, C. (2024). Methods that support the validation of agent-based models: An overview and discussion. *Journal of Artificial Societies and Social Simulation*, *27*(1), 11. https://doi.org/10.18564/jasss.5258

Conrad, F. G., Keusch, F., & Schober, M. F. (2021). New data in social and behavioral research. *Public Opinion Quarterly*, *85*(S1), 253–263. https://doi.org/10.1093/poq/nfab027

Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., Loreto, V., Moat, S., Nadal, J.-P., Sanchez, A., Nowak, A., Flache, A., San Miguel, M., & Helbing, D. (2012). Manifesto of computational social science. *The European Physical Journal Special Topics*, *214*(1), 325–346. https://doi.org/10.1140/epjst/e2012-01697-8

Creamer, E. G. (2018). *An introduction to fully integrated mixed methods research*. SAGE Publications, Inc. https://doi.org/10.4135/9781071802823

Daikeler, J., Fröhling, L., Sen, I., Birkenmaier, L., Gummer, T., Schwalbach, J., Silber, H., Weiß, B., Weller, K., & Lechner, C. (2024). Assessing data quality in the age of digital social research: A systematic review. *Social Science Computer Review*, *8944393241245395*. https://doi.org/10/g9rnqk

Davidson, T., & Karell, D. (2025). Integrating generative artificial intelligence into social science research: Measurement, prompting, and simulation. *Sociological Methods & Research*, 00491241251339184. https://doi.org/10/g9hn9m

Dehne, J., Daikeler, J., Silber, H., Rammstedt, B., Keusch, F., Kreuter, F., Blumenberg, J., Lechner, C., Römer, L., Schoch, D., Siegers, P., Jünger, S., & Weller, K. (2024). *Die Vermittlung von Best Practices zur Messung von Datenqualität in den quantitativen Sozialwissenschaften*. https://doi.org/10/g9rnqm

Eden, A. H. (2007). Three paradigms of computer science. *Minds and Machines*, *17*(2), 135–167. https://doi.org/10.1007/s11023-007-9060-8

Egami, N., Hinck, M., Stewart, B., & Wei, H. (2023). Using imperfect surrogates for downstream inference: Design-based supervised learning for social science applications of large language models. *Advances in Neural Information Processing Systems*, *36*, 68589–68601.

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, *115*(16). https://doi.org/10.1073/pnas.1720347115

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2021). Machine learning for social science: An agnostic approach. *Annual Review of Political Science*, *24*(1), 395–419. https://doi.org/10.1146/annurev-polisci-053119-015921

Hammersley, M., & Gomm, R. (1997). Bias in social research. *Sociological Research Online*, *2*(1), 7–19. https://doi.org/10.5153/sro.55

Hase, V., Ausloos, J., Boeschoten, L., Pfiffner, N., Janssen, H., Araujo, T., Carrière, T., De Vreese, C., Haßler, J., Loecherbach, F., Kmetty, Z., Möller, J., Ohme, J., Schmidbauer, E., Struminskaya, B., Trilling, D., Welbers, K., & Haim, M. (2024). Fulfilling data access obligations: How could (and should) platforms facilitate data donation studies? *Internet Policy Review*, *13*(3). https://doi.org/10.14763/2024.3.1793

Hase, V., & Haim, M. (2024). Can we get rid of bias? Mitigating systematic error in data donation studies through survey design strategies. *Computational Communication Research*, *6*(2), 1–29. https://doi.org/10.5117/CCR2024.2.2.HASE

Jürgens, P., Stark, B., & Magin, M. (2020). Two half-truths make a whole? On bias in self-reports and tracking data. *Social Science Computer Review*, *38*(5), 600–615. https://doi.org/10.1177/0894439319831643

Kathirgamalingam, A., Lind, F., & Boomgaarden, H. G. (2025). Measuring racism and related concepts using computational text-as-data approaches: A systematic literature review. *Annals of the International Communication Association*, *49*(3), 241–256. https://doi.org/10.1093/anncom/wlaf013

Keusch, F., Bähr, S., Haas, G.-C., Kreuter, F., & Trappmann, M. (2020). Coverage error in data collection combining mobile surveys with passive measurement using apps: Data from a German National survey. *Sociological Methods & Research*, *52*(2), 004912412091492. https://doi.org/10.1177/0049124120914924

Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 2053951714528481. https://doi.org/10.1177/2053951714528481

Kostygina, G., Kim, Y., Seeskin, Z., LeClere, F., & Emery, S. (2023). Disclosure standards for social media and generative artificial intelligence research: Toward transparency and replicability. *Social Media + Society*, *9*(4), 20563051231216947. https://doi.org/10.1177/20563051231216947

Lavelle-Hill, R., Smith, G., & Murayama, K. (2025). Bridging traditional-statistics and machine-learning approaches in psychology: Navigating small samples, measurement error, nonindependent observations, and missing data. *Advances in Methods and Practices in Psychological Science*, *8*(3), 25152459251345696. https://doi.org/10.1177/25152459251345696

Lazer, D., Pentland, A., Aral, S., Athey, S., Contractor, N., Freelon, D., Gonzalez-Bailon, S., King, G., Margetts, H., Nelson, A., Salganik, M. J., Strohmaier, M., Vespignani, A., & Wagner, C. (2020). Computational social science: Obstacles and opportunities. *Science*, *369*(6507), 1060–1062. https://doi.org/10.1126/science.aaz8170

Loecherbach, F., & Otto, L. (2024). Introduction to the special issue on participant-centered behavioral traces. *Computational Communication Research*, *6*(2), 1. https://doi.org/10.5117/CCR2024.2.1.LOEC

Margolin, D. B. (2019). Computational contributions: A symbiotic approach to integrating big, observational data studies into the communication field. *Communication Methods and Measures*, *13*(4), 229–247. https://doi.org/10.1080/19312458.2019.1639144

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220–229). Association for Computing Machinery. https://doi.org/10/gftgjg

Moons, K. G. M., Damen, J. A. A., Kaul, T., Hooft, L., Andaur Navarro, C., Dhiman, P., Beam, A. L., Van Calster, B., Celi, L. A., Denaxas, S., Denniston, A. K., Ghassemi, M., Heinze, G., Kengne, A. P., Maier-Hein, L., Liu, X., Logullo, P., McCradden, M. D.Liu, N. . . . Van Smeden, M. (2025). Probast+AI: An updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*, e082505. https://doi.org/10.1136/bmj-2024-082505

Nelson, L. K. (2020). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, *49*(1), 3–42. https://doi.org/10.1177/0049124117729703

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T.Heinze, C. . . . Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *WIREs Data Mining and Knowledge Discovery*, *10*(3), e1356. https://doi.org/10.1002/widm.1356

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, *2*(1), 1–28. https://doi.org/10.1287/isre.2.1.1

Pak, C., Cotter, K., & Thorson, K. (2022). Correcting sample selection bias of historical digital trace data: Inverse probability weighting (IPW) and type II tobit model. *Communication Methods and Measures*, *16*(2), 134–155. https://doi.org/10.1080/19312458.2022.2037537

Rieder, G., & Simon, J. (2017). Big data: A new empiricism and its epistemic and socio-political consequences. In W. Pietsch, J. Wernecke, & M. Ott (Eds.), *Berechenbarkeit der Welt? Philosophie und Wissenschaft im Zeitalter von Big Data* (pp. 85–105). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-12153-2_4

Sap, M., Card, D., Gabriel, S., Choi, Y., & Smith, N. A. (2019). The risk of racial bias in hate speech detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 1668–1678). Association for Computational Linguistics. https://doi.org/10/gf9wr8

Sen, I., Flöck, F., Weller, K., Weiß, B., & Wagner, C. (2021). A total error framework for digital traces of human behavior on online platforms. *Public Opinion Quarterly*, *85*(S1), 399–422. https://doi.org/10.1093/poq/nfab018

Shekhawat, N., Chauhan, A., & Muthiah, S. B. (2019). Algorithmic privacy and gender bias issues in Google ad settings. In *Proceedings of the 10th ACM Conference on Web Science* (pp. 281–285). Association for Computing Machinery. https://doi.org/10.1145/3292522.3326033

Shugars, S. (2024). Critical computational social science. *EPJ Data Science*, *13*(1), 13. https://doi.org/10.1140/epjds/s13688-023-00433-2

Siemon, M. (2025). Beyond the binary? Automated gender classification of social media profiles. *Communication Methods and Measures*, 1–19. https://doi.org/10.1080/19312458.2025.2554864

Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2020). Integrating survey data and digital trace data: Key issues in developing an emerging field. *Social Science Computer Review*, *38*(5), 503–516. https://doi.org/10.1177/0894439319843669

Stoll, A., Yu, J., Andrich, A., & Domahidi, E. (2025). Classification bias of LLMs in detecting incivility towards female and male politicians in German social media discourse. *Communication Methods and Measures*, 1–19. https://doi.org/10.1080/19312458.2025.2551693

Struminskaya, B., Lugtig, P., Toepoel, V., Schouten, B., Giesen, D., & Dolmans, R. (2021). Sharing data collected with smartphone sensors. *Public Opinion Quarterly*, *85*(S1), 423–462. https://doi.org/10.1093/poq/nfab025

Stuhler, O., Ton, C. D., & Ollion, E. (2025). From codebooks to promptbooks: Extracting information from text with generative large language models. *Sociological Methods & Research*, *54*(3), 794–848. https://doi.org/10.1177/00491241251336794

Tay, L., Woo, S. E., Hickman, L., Booth, B. M., & D'Mello, S. (2022). A conceptual Framework for investigating and mitigating machine-learning measurement bias (MLMB) in psychological assessment. *Advances in Methods and Practices in Psychological Science*, *5*(1), 25152459211061337. https://doi.org/10.1177/25152459211061337

TeBlunthuis, N., Hase, V., & Chan, C.-H. (2024). Misclassification in automated content analysis causes bias in regression. Can we fix it? Yes we can! *Communication Methods and Measures*, *18*(3), 278–299. https://doi.org/10.1080/19312458.2023.2293713

The American Association for Public Opinion Research. (2023). *Standard definitions: Final dispositions of case codes and outcome rates for surveys* (10th ed.). AAPOR. https://aapor.org/wp-content/uploads/2024/03/Standards-Definitions-10th-edition.pdf

Theocharis, Y., & Jungherr, A. (2021). Computational social science and the study of political communication. *Political Communication*, *38*(1–2), 1–22. https://doi.org/10.1080/10584609.2020.1833121

Toh, S.-L., & Park, J. (2025). Fake it till you make it: Synthetic data and algorithmic bias. *International Journal of Communication*, vol. *19*, 1852–1858.

Törnberg, P., & Uitermark, J. (2021). For a heterodox computational social science. *Big Data & Society*, *8*(2), 20539517211047725. https://doi.org/10.1177/20539517211047725

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. https://doi.org/10.1080/19312458.2018.1458084

Van Der Velden, M. A. C. G., & Dolinsky, A. O. (2024). Is Matilda playing it safe?: Gender in computational text analysis methods. *Computational Communication Research*, *6*(1), 1–26. https://doi.org/10.5117/CCR2024.1.4.VAND

Van Der Velden, M. A. C. G., Loecherbach, F., Van Atteveldt, W., Fokkens, A., Reuver, M., & Welbers, K. (2025). Whose truth is it anyway? An experiment on annotation bias in times of factual opinion polarization. *Communication Methods and Measures*, 1–18. https://doi.org/10.1080/19312458.2025.2562034

Van Rooij, I., Devezer, B., Skewes, J., Varma, S., & Wareham, T. (2024). What makes a good theory? Interdisciplinary perspectives. *Computational Brain & Behavior*, *7*(4), 503–507. https://doi.org/10.1007/s42113-024-00225-5

Wagner, C., Strohmaier, M., Olteanu, A., Kıcıman, E., Contractor, N., & Eliassi-Rad, T. (2021). Measuring algorithmically infused societies. *Nature*, *595*(7866), 197–204. https://doi.org/10.1038/s41586-021-03666-1

Waldherr, A., Weber, M., Wu, S., Haim, M., Van Der Velden, M. A. C. G., & Chen, K. (2025). Between innovation and standardization: Best practices and inclusive guidelines in computational communication science. *Journalism & Mass Communication Quarterly*, *102*(1), 13–36. https://doi.org/10.1177/10776990241301676

Walter, N., Cody, M. J., & Ball-Rokeach, S. J. (2018). The ebb and flow of communication research: Seven decades of publication trends and research priorities. *Journal of Communication*, *68*(2), 424–440. https://doi.org/10.1093/joc/jqx015

Weiß, B., Leitgöb, H., & Wagner, C. (2025). Conceptualizing, assessing, and improving the quality of digital behavioral data. *Social Science Computer Review*, *43*(5), 927–942. https://doi.org/10.1177/08944393251367041

Wenz, A., Keusch, F., & Bach, R. L. (2024). Measuring smartphone use: Survey versus digital behavioral data. *Social Science Computer Review*, *43*(5), 08944393231224540. https://doi.org/10.1177/08944393231224540

Yi, J., & Zhang, W. J. (2023). Mapping the global flow of computational communication science scholars. *Journal of International Communication*, *29*(1), 144–171. https://doi.org/10.1080/13216597.2022.2160780