This is a repository copy of *Intersectional approaches to data: the importance of an articulation mindset for intersectional data science*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/203863/

Version: Published Version

Article:

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Intersectional approaches to data: The importance of an articulation mindset for intersectional data science

Caitlin Bentley[1] iD , **Chisenga Muyoya**[2] iD , **Sara Vannini**[2] iD ,
Susan Oman[2] iD **and Andrea Jimenez**[2] iD

Abstract
Data's increasing role in society and high profile reproduction of inequalities is in tension with traditional methods of using social data for social justice. Alongside this, 'intersectionality' has increased in prominence as a critical social theory and praxis to address inequalities. Yet, there is not a comprehensive review of how intersectionality is operationalized in research data practice. In this study, we examined how intersectionality researchers across a range of disciplines conduct intersectional analysis as a means of unpacking how intersectional praxis may advance an intersectional data science agenda. To explore how intersectionality researchers collect and analyze data, we conducted a critical discourse analysis approach in a review of 172 articles that stated using an intersectional approach in some way. We contemplated whether and how Collins' three frames of relationality were evident in their approach. We found an over-reliance on the additive thinking frame in quantitative research, which poses limits on the potential for this research to address structural inequality. We suggest ways in which intersectional data science could adopt an articulation mindset to improve on this tendency.

Keywords
Intersectionality, underrepresented groups, data collection, data analysis, data science methods, relationality

Introduction

The 2030 Agenda for Sustainable Development is predicated on the principle of 'leaving no one behind'. Data are seen as having a key role in this aspiration. Yet, the deportation of 30 Kurdish asylum seekers from the UK to Iraq demonstrates how, when data renders them visible, certain populations' vulnerability can become a death sentence, with one man claiming, 'in my country some people wouldn't think twice about shooting us in the head' (Taylor, 2022: n.p.). Visibility risks for those already vulnerable are not new. Techniques that progressed HIV detection in the 1980s were the same used by the Amsterdam City Office of Statistics that identified the Jewish population to Nazi occupation in 1941 (Oman, 2021). This is not only an issue of visibility in data but also the right to choose to be 'left behind'.

Progress in data science about inequalities is promising for those who want to further knowledge on disparities and discriminations for social change, but prohibitive for those concerned with the populations most at risk. More research on how data science, as a field of research and practice, should engage with these issues is needed.

Intersectionality is currently promoted by governments and international policy-makers, following recommendations to address multiple inequality-related issues with data: the Inclusive Data Charter's (2022) five recommendations[1] centre on promoting equity across a data value chain; the Inclusive Data Taskforce (2022: n.p) develop a set of eight principles[2] for inclusive data infrastructures, foregrounding sufficient data for 'intersectional analysis', to 'broaden the range of methods used' and 'create new approaches to understand everyone's experience'. Yet, recommending intersectionality as a solution in data-related policy and practice often lacks clarity on how to apply it to specific datasets or data contexts and does not engage intersectionality's origins as praxis.

[1]Department of Informatics, King's College London, London, UK
[2]Information School, The University of Sheffield, Sheffield, UK

Corresponding author:
Caitlin Bentley, Department of Informatics, King's College London, London, UK.
Email: caitlin.bentley@kcl.ac.uk

Intersectionality has a long history as a theory and a praxis that is political in its focus on dismantling interlocking systems of oppression (Collins and Bilge, 2020). Whilst theories differ, most agree it enables comprehension of multiple aspects of identity, which overlap in context, influencing experienced (dis)advantage (Crenshaw, 1990). Intersectional data practices therefore promise possibilities for data to better understand the systemic roots of privilege, prejudice and inequality, as well as requisite steps to alleviate them, whilst also acknowledging potential harms to underrepresented groups.

Intersectionality research is growing across disciplines and sectors but without guidance on its implementation. This article's concern is this lack in data science and in their ethics in practice and theory. Calls for improved diversity in the workforce have seen little improvement (HM Government, 2022; Young et al., 2021). The *who* of data science, as presented by D'Ignazio and Klein (2020), thus, still remains an important question.

There is also little evidence on how intersectional analysis works, or should be done, how it is being done, differently, by whom. Thus, data science's potential to adopt intersectionality in its broadening of methods and advocacy for the potential of reparatory work is stymied. The authors share experience of these concerns and of our ambitions being sidelined in teaching and research. These shared experiences emerged despite our individual, distinctive progressions as early career women in academia. We come from four continents, having worked with communities across geographical contexts and with various international policy-makers. Our observed sidelining of feminist science and technology studies (STS) and feminist data science (e.g. D'Ignazio and Klein, 2020; Haraway, 1985; Harding, 1986) led to a shared urgency to change the status quo.

Through key roles developing a new data science programme, on its core steering group and promoting decolonization and education for sustainable development, we encountered protests that learning ethical or responsible data science was a distraction from the requisite technical skills. Contrastingly, our experiences through research, and wider collaborations with organizations, found them keen to embrace intersectional data science in ways that were resisted in teaching and learning. Instead of ethics, equity, decolonization and sustainability being embedded into data science education, it is often appended, such that key texts incorporating feminist approaches (D'Ignazio and Klein, 2020), context and well-being (Oman, 2021) and ethics (Martens, 2022) are provided in 'professional skills' or 'data and society' courses rather than in 'data modelling' or 'data analysis' courses, which favour technical or mathematical texts (e.g. Christen, 2012; Hastie et al., 2009). The lack of key texts describing or demonstrating responsible data science and the scarce guidance on

intersectional data science led us to review these as connected issues.

To grasp the potential for building intersectional data science theory/praxis, we wanted to understand the prominence, strengths and weaknesses of intersectional approaches in current research. We did this through a literature review, applying critical discourse analysis (CDA) to 172 journal articles that self-identified as intersectional. We coded these using Collins' (2019) three notions of relationality (additive thinking, articulation, and co-formation) as central to intersectionality theory/praxis. We also coded elements of their data practice, to evaluate these practices in relation to their intersectional approach.

We found that additive thinking, in which separate lines of inquiry (around gender, race, class, etc.) are added together, dominated the intersectionality research examined and disproportionately so in quantitative examples. We argue that articulation has the potential to improve on the limitations of additive thinking in data science because it engages with the mechanisms of structural inequality. We contribute three traits of an articulation mindset to suggest how articulation can be enacted in data science: by making the historical antecedents of inequality, discrimination or disadvantage more explicit; by examining links between individuals and social and institutional contexts in a systemic or integrative fashion; and by attending to the role and positionality of researchers in research. The next section outlines the intersectionality concepts that were used to frame this analysis.

Intersectionality and intersectional data science

Intersectionality has many origin stories, all linking to the research and activism of critical black feminists, who first denounced the inadequacy of frameworks that aimed to address oppression at the time. They demonstrated that practices and knowledge systems that did not conform to white perspectives have been silenced and discredited (Erete et al., 2021). Our understanding is informed by core texts published by Crenshaw (1990) and Collins (2019), acknowledging their preceding feminist activist authors, histories and genealogies. We explain how and why these works have framed our exploration of intersectional praxis in data science and present the reasoning behind our focus on relationality as a key way into investigating praxis.

The intersectionality field expanded beyond feminism and critical race studies and is now cited as interpreting identity within a broader social structural framework (e.g. Harris and Patton, 2019; Meier, 2020). Thus, intersectionality extends ideas of identity from personality characteristics or unique experiences and also focuses on how they are shaped by overlapping institutional, political and social

systems of privilege and oppression. The vocabulary and concepts of intersectional thought have been adopted by disciplines across the humanities, social sciences and medicine enabling the analysis of lived experience, social relations and inequality.

Data science, we argue, must also address these issues, through attending to the voices and perspectives of underrepresented groups, defining 'underrepresented groups' as those at risk of marginalization or discrimination in or through data.[3] We aim to understand how research that calls itself 'intersectional' defines and investigates experiences of underrepresented groups and to draw lessons for data science because 'lack of data typically correlates with issues affecting those who are most vulnerable in that context' (Onuoha, 2018: n.p.).

Much can be learned from feminist STS scholars' engagements with intersectional praxis, suggesting lessons for how intersectional data science should function. Feminist STS scholars have long advocated for embodied research and the importance of standpoint epistemologies that recognize different perspectives as valid sources of knowledge (Haraway, 1985; Harding, 1986; Mirza, 2009). Feminist STS also commonly challenges dominant discourses and knowledge systems in research (Alcoff, 2013; Hartsock, 1998; Wynter, 2003). The experiences of black feminist STS scholars, especially in the US and UK, furthered Haraway (1985), Harding (1986) and Hartsock (1998) by emphasizing the deconstruction of knowledge-building processes; advocating for listening to multiple standpoints, especially non-white ones; and putting non-white perspectives at the centre of knowledge production spaces, to shape the field in more just directions (Erete et al., 2021; Mirza, 2009; Reynolds, 2002).

Moreover, building relationships with communities affected by research is prioritized (Costanza-Chock, 2018; Llewellyn and Llewellyn, 2015). Researchers and practitioners need to be cognizant of how relationships between individuals and within institutional environments may be sites of oppression and/or resistance and how these relationships are shaped by larger socio-historical forces (Llewellyn and Llewellyn, 2015). Similarly, Costanza-Chock (2018) highlighted the importance of building relationships with research(ed) communities and engaging participants as co-researchers and co-producers of knowledge. She argued that when the aim is social justice, building trust with communities, accountability, transparency and accessibility to data and outcomes need to be part of the research design. However, these authors were focused on higher education and design contexts broadly, which may need consideration within data science specifically.

D'Ignazio and Klein's (2020) framework for 'data feminism' is one significant example of how feminist scholars engage with intersectionality in data science. They offered seven general principles to inform data science praxis, including, for instance, examining power, rethinking binaries and hierarchies and considering context (D'Ignazio and Klein, 2020). They also emphasized how reflexivity should be incorporated by asking questions about who is doing data work, represented in the data, who is excluded, and who benefits or is harmed by data collection and analysis. Likewise, the Inclusive Data Taskforce (2022: n.p.) found that 'greater representation across relevant groups and populations within the research community would ensure better understanding of different cultures, address barriers to participation and reduce the risk of burdening participants with duplication of research'. Questions surrounding the growing popularity of intersectionality, as it moves 'beyond' black women, arguably erasing or displacing 'their centrality within the discourse' (Cooper, 2016) are crucial as well. Thus, it is not just the representation of data subjects that requires attention, but of data science itself, which we address in this paper.

To build on the work of D'Ignazio and Klein (2020), whose principles were drawn from intersectional feminist theory, we critically review intersectionality articles to investigate, empirically, whether and how researchers are enacting similar principles in their praxis. In the next section, we introduce Collins' (2019) three notions of relationality as the way in which we frame this investigation.

Towards intersectional data science: the importance of relationality

A main challenge in data science for Stevens et al. (2021) is that it is established on core tenets antagonistic to intersectionality. They showed how relational database systems were theorized to interface with computer data storage mechanisms in ways that delinked data from their context, simplifying relationships between data, systems and people. Crucially, they showed how this decontextualized thinking reinforced dominant 'colourblind' perspectives, thus how data modelling theory hides the ways that race structures society.[4]

Stevens et al.'s (2021) revision of relational data modelling coincides with increasing calls from scholars to confront underlying assumptions around 'objectivity' and 'impartiality' in data science (D'Ignazio and Klein, 2020; Oman, 2021). For Keith (2021), there are four main problems with the prevailing paradigm of data analytics; data are: (a) not impartial; (b) often insufficient in and of themselves; (c) not the only starting point for analysis and (d) impossible to interpret fully without context. Keith's solution is to put data scientists in a dialogue with data.

Intersectionality could provide a theory/praxis for conceptualizing such dialogic engagement around the relationships between data, systems and people. In this paper, we focus on relationality because, as Collins (2019: 226) states, '[relationality is] so central to intersectionality,

[that] developing a more comprehensive analysis of this core theme is an essential task for intersectionality's theory-building project'. We therefore introduce Collins' (2019) three notions of relationality – additive thinking, articulation and co-formation. Notably, Collins (2019) does not present these relational concepts as pristine, static or exclusive, nor does she prioritize one over the other. Instead, she emphasizes their intricacies and how they help construct diverse, but not mutually exclusive, relationality approaches, and thus, we expand on the concepts immediately below.

Additive thinking. Issues resulting from a person's identity such as class, race, gender or sexuality have historically been treated as separate lines of inquiry by scholars and practitioners (Collins, 2019; Crenshaw 1990). Additive thinking considers these issues – and their intersections – together as crucial to understanding social position, questioning assumptions and dominant frameworks. It also shows what is lost when these intersectional categories are excluded from analyses. To illustrate, Buolamwini and Gebru (2018) showed that facial recognition technology failed to detect the faces of women of colour disproportionately to other groups. This failure was only detected by combining gender and race categories together.

Additive thinking can be problematic for Collins (2019: 227) when it relies on a 'logic of segregation' in that 'everything has one place, a place has meaning only in relation to other places, and every place has its rank'. Collins further contends that due to the difficulty in establishing these categories and the dangers of giving one category precedence over another, additive thinking can be challenging. For us, these problems are also symptomatic of the decontextualized approaches to data collection and analysis that dominate data science. Although additive thinking is tied to categorical thinking and classification, our interpretation is that it should also rely on situating meaning (including how categories are constructed) in a power hierarchy.

Articulation. Collins' (2019) second framing of relational thinking is based on Hall's (1980) dual-meaning concept of articulation which analyzes social formations by identifying links and by the language we use to explain these links. This framing outlines a more holistic and reflexive perspective on relationship dynamics than additive thinking. Rather than aiming to combine different categories that were formerly seen as separate, articulation acknowledges that categories coupled together result in new configurations that are always changing and incomplete. Hall's first meaning of articulation is a unique connection that links, or articulates, one or many other pieces together in a whole. Hall's (1980) second notion of articulation, Collins (2019: 233) argued as being important for unpacking 'how language "articulates" or brings new ideas by combining existing ideas into new patterns, by attaching

new connotations to them, or both'. This is important for exploring how data practices communicate meaning and occur in specific social contexts, whilst promoting socio-political agendas. For example, Oman (2021) demonstrated how data discourses operate and shape society, culture and policy, often justifying political decisions by presenting them as objective and neutral, whilst they are, in fact, rather ideological or habitual. Concepts and the language that articulate social contexts can often create pretense for data practice.

Co-formation. Co-formation is the final framing of relationality. 'It's meaningless to argue that race and gender co-form one another without assuming they are separate entities', argued Collins (2019: 241). Co-formation eliminates categorical thinking altogether and 'posits holistic analysis of a seamless process of mutual construction of race, class and gender as phenomena. It seems to describe a reality that it can neither observe nor study empirically' (Collins, 2019: 241). In other words, we intuitively know that intersecting systems of power influence and shape the privilege or discrimination individuals experience, and vice-versa. Masquelier (2022) argued that co-formation resonates with decolonial approaches and indigenous ways of thinking, which have always embraced the interconnectedness and interdependence of reality. They suggested that dialogue is crucial for emancipation and empowerment, enabling diverse people to organize collectively and to claim equity in cooperative terms. Yet, the implication is that certain topics of discussion as well as modes of political representation and democratic process underpin such a dialogue. Whilst we agree that decolonial approaches to data science are important, and that a co-formative frame of relationality seems capable of supporting them theoretically, translating such a dialogic process to data practice does not seem straightforward. Just as Collins (2019) explained, co-formation serves as a starting point for thinking beyond categories, yet it struggles with standard social science research tools because they are not suited to this ontological approach. Nonetheless, we consider these issues in the context of data science.

We summarize Collins' (2019) three notions of relationality (additive thinking, articulation and co-formation) in Table 1. In the next section, we explain how we reviewed intersectionality literature to examine whether and how these relational frames are enacted as a means of examining potential benefits and drawbacks of adopting such relational frames within data science.

Methodology

Our search of the Scopus database in November 2021 found that more than 2000 articles had been published since 2017 with 'intersectional' used within its title/abstract. A full

Table 1. Collins' three notions of relationality.

	Additive	Articulation	Co-formation
Aim	To combine what was formerly seen as distinctive. It helps signal what is missing from excluding a category.	To understand complex interrelationships in the context of power dynamics, and the importance of context and language in shaping those interrelationships.	To provide a holistic examination of intersecting power relations. It suggests engaging with decolonial, indigenous ways of thinking.
Nature of relationality	Suggests analytical and political integrity of categories (i.e. race and gender cannot substitute each other)	Society is dynamic, with moving parts and shifting power dynamics among multiple power structures. Parts can be distinct. Not just adding categories, but also acknowledging that categories combined result in a new 'articulation' that is constantly changing. Naturally adaptable, unfinished, resisting closure.	Assumes categories cannot be separated from one another because they never were in the first place. Focuses on the interconnectedness and interdependence of things.
Issues	Can be considered as adding with no transformation to each. There can be disagreement on the meaning of each category (i.e. what is gender?) Can lead to privileging ones' favourite category for analysis.	May be more visible during times of crisis. Aims to provide provisional analysis to explain social events that can always be changed.	Intellectually, it is far easier to envisage than it is to achieve methodologically. Incompatible with Western epistemological and methodological tools, which were designed to destroy holistic entities thus making sense of them.

explanation of our inclusion and exclusion criteria, article search methods, abstract screening and discipline coding methods are provided in Appendix 1 of the supplementary material. An overview of this process is shown in Figure 1.

The articles were coded based on a tailored content analysis procedure. The code book had previously been tested within Bentley's work on the Inclusive Data Charter's (2022) white paper, and some adjustments were made in this iteration. For example, categories were added to many of the codes to simplify the coding process. The complete code book is given in Appendix 2 of the supplementary material. We also iterated the coding process and refined our coding procedure through reflexive practice in which we discussed our biases and assumptions, our interests and understandings of the materials.

Given our focus on examining data science as discipline and discourse, we used a CDA review methodology (Wall et al., 2015) foregrounding Collin's (2019) approach to intersectionality over Foucauldian and Habermasian CDA frames. This approach critically examines scientific discourse, prompting researchers to question underpinning assumptions and to challenge ideological premises.

Using Collins' (2019) relationality frames of additive thinking, articulation and co-formation, we also examined the article sample set for evidence of how relationality was enacted in the research holistically, coding the articles for the relationality frame used along a spectrum. Whereas coding data practices required attention to detail and specifics, examining the adopted relationality frames required a general reading of author intentions, conceptual frameworks, analysis or discussion taken together. Often, there were articles in which we were unsure, so we recorded our reasoning and measured the level of agreement between two coders.

Overall, we adopted reflexive, relational and dialogic practice throughout the research, recording how data were deleted, combined or reorganized in some way. The authors analyzed the substantive codes using a qualitative thematic analysis procedure (Saldaña, 2014), moving iteratively between examining an individual code, and then the abstract of a particular article, or the article in full for more 'data context'. We held collaborative writing sessions in which we discussed thematic analysis and how concepts inspired or influenced our analysis, and how to handle limitations. We discussed further layers of themes and re-coded articles again accordingly.

Findings

Article sample characteristics

The final sample examined consisted of 172 articles: the majority had produced and used data in the context of empirical research (123/172 (72%)). The remainder included literature or discourse as data (37/172 (22%)) or practitioner or researcher reflections as data (12/172 (7%)). In terms of methodological orientation, 66% were qualitative (114/172), 14% were mixed methods (24/172), and 19% were quantitative (33/172). Figures 2 to 4 present three visualizations: geographical location of the
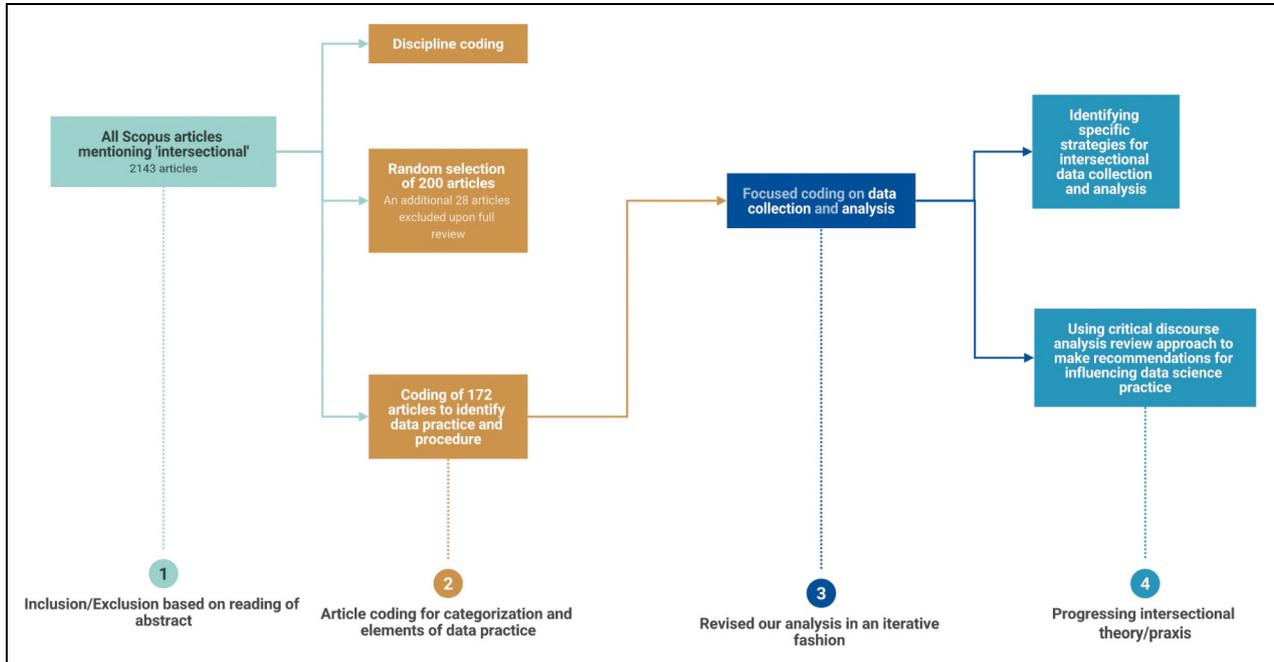
Figure 1. Flow diagram of article search, coding and analysis process.
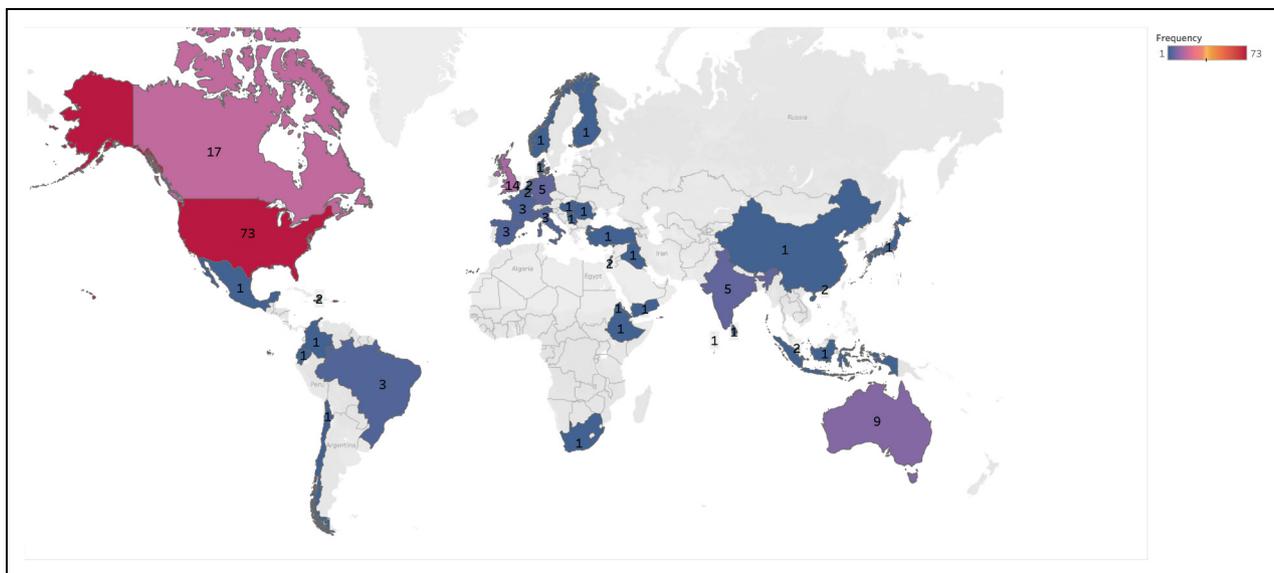

Figure 2. Frequency of articles taking place across countries.

article (coded by where the research took place), discipline of the article (coded by journal) and a network graph of identity markers included in the article. These are provided to give context to the reader and to highlight limitations within this body of work.

Figure 2 shows that intersectionality research takes place mostly within the US (73 articles). Canada, UK and Australia accounted for 9–17 articles each. India, Brazil and some mainland European countries have between

three and five publications each, whilst most of 'the rest' have one or zero. This likely reflects sample biases which focused on academic publishing, or perhaps indicates intersectionality is not a concept of universal interest.

Figure 3 shows the journal disciplines represented in the set (purple) and sample (green) of articles. The most common journal disciplines represented in the sample were social science (49), humanities (27), psychology (26), medicine (21), education (15), and philosophy (12), totaling 150 of 247
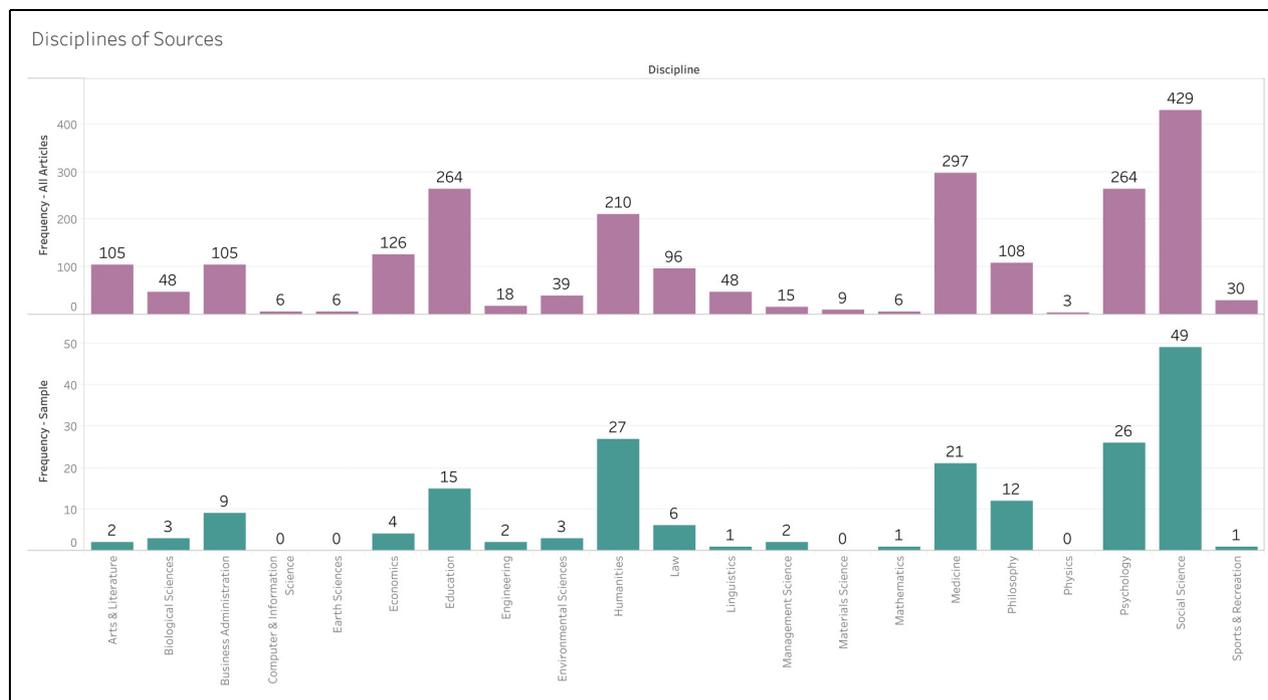
Figure 3. Frequency of selected journal article sources categorized by discipline using JournalSeek.

(61%) unique sources (Figure 3). These charts also show limited engagement with intersectionality in data science, computer and information science, or engineering disciplines.

In terms of the identity markers explored, a network graph (Figure 4) shows that most of the articles included ethnicity, gender, class, migratory status, women/girls and disability centrally or together in their research. In contrast, intersectionality research focusing on transgender, indigenous, youth, elderly or climate change–affected people were less likely. There was also a tendency to focus on one identity factor, such as 'indigenous' participants, without further differentiating these cases.

Table 2 summarizes a range of articles from the 172 reviewed; these included autoethnographic (Ashlee et al., 2017), ethnographic (Sharp, 2021) and narrative accounts (de Regt, 2017), qualitative interview-based studies (Meier, 2020; Peretz, 2017; Thorjussen and Sisjord, 2020), quantitative survey designs (Lord et al., 2019; Rahman et al., 2020) and analysis of secondary data (Earnshaw et al., 2021; Ugidos et al., 2020). Table 2 also identifies the main intersectionality-related findings in the selection of studies, highlighting how each study could be grouped together in different ways around engagement with intersectional methods (Ashlee et al., 2017; de Regt, 2017; Wiens et al., 2020), the use of intersectionality to contribute new insights in a disciplinary field or context of study (Lord et al., 2019; Thorjussen and Sisjord, 2020), to better understand social injustices or inequality (Meier, 2020; Rahman et al., 2020; Ugidos et al., 2020), or to engage in acts of resistance or liberation (Ashlee et al., 2017; Peretz, 2017). We explore in greater

detail how and why some of these articles adopt additive thinking and articulation relational frames in the next section.

Examining relationality frames adopted: what is an 'articulation mindset'?

In this section, we give an overview of the relationality frames adopted by authors in the articles examined. In about 43% of the papers, additive thinking was used as the dominant relational frame. Most of the remaining papers showed evidence of using the articulation frame, which we shall examine and comment on. There were some articles that were not easily categorized, falling in between additive thinking and articulation (15%) or articulation and co-formation (4%). Yet, there was strong inter-rater agreement, as the percent agreement between two coders was 90% (Table 3). Due to the limited scope of co-formation articles identified, we focus our analysis on examining the qualities and characteristics of motivations, practices and techniques used by authors adopting the additive thinking and articulation frames of relationality. However, Ashlee et al. (2017) and Wiens et al. (2020) (Table 2) are two examples of co-formation that demonstrate original methods and insights into intersectional theory/praxis.

The dominance of the additive thinking frame across the 75 articles coded thus is reflected in the researchers' use of multiple traits or characteristics (i.e. gender, class, ethnicity together) to select participants by defining boundaries around groups of individuals who are similar or different
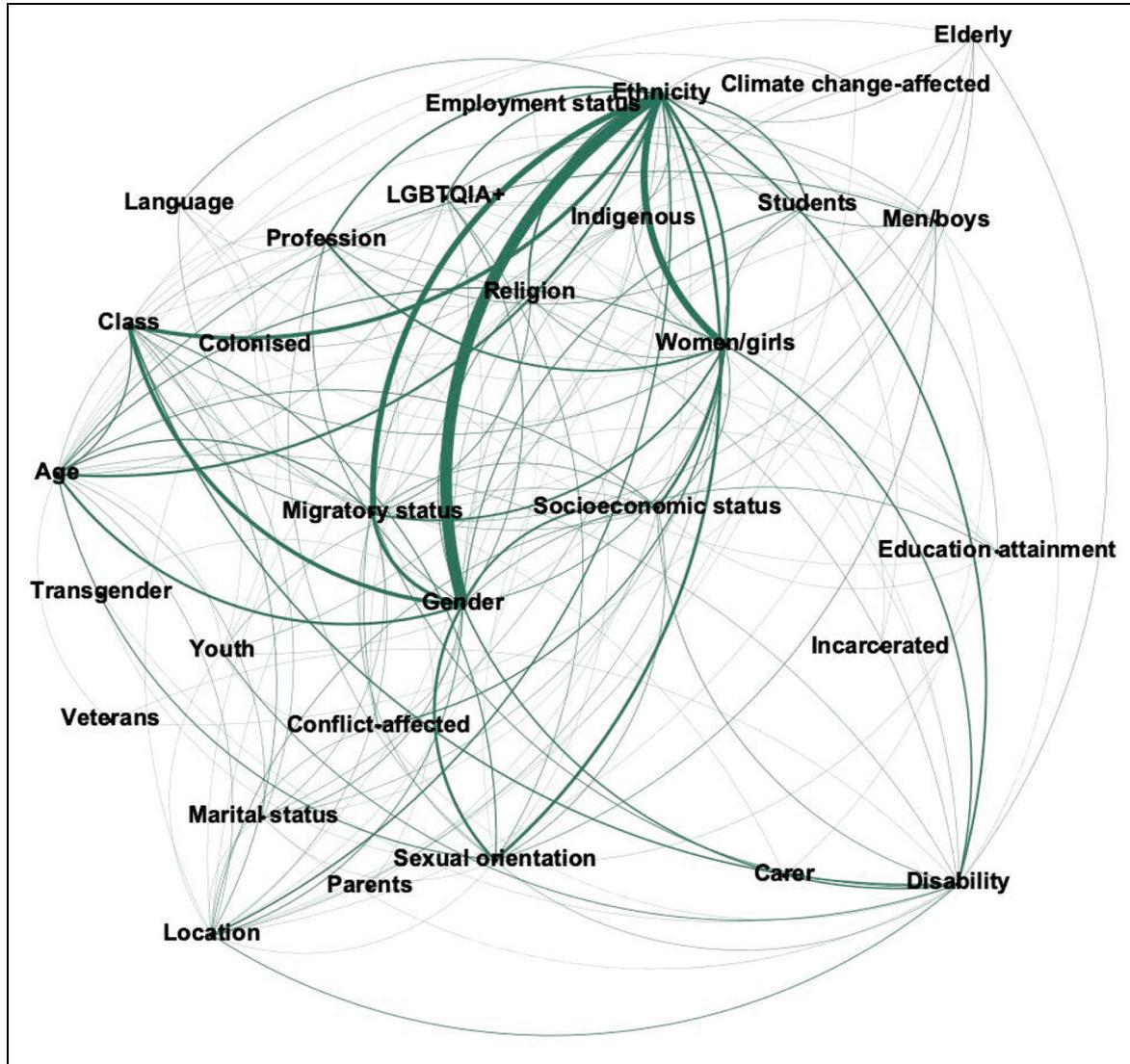
Figure 4. Weighted network graph of relations between personal characteristics included in intersectionality research.

in some way. In line with Collins' (2019) view, analyzing experience through the additive thinking frame can indeed enable new insights that are beneficial for a range of purposes such as progressing an academic field (e.g. Thorjussen and Sisjord, 2020), understanding inequalities in social outcomes (e.g. Earnshaw et al., 2021) or understanding the needs of particular underrepresented groups that have otherwise been ignored (e.g. Meier, 2020). Such attention to neglected perspectives can be a vital first step researchers can take towards understanding oppressions or reducing inequalities. However, there is a clear tendency towards additive thinking within quantitative articles (Figure 5), which we investigate further.

Within articulation articles, there was more variation and fewer standard practices methodologically speaking. Most of the articulation papers were discursive, using CDA,

literature reviews or practitioner reflections to expose linkages and their articulations (Figure 5). Whilst Collins (2019) and Hall (1980) give a blueprint for what articulation accomplishes, Collins (2019) is not prescriptive about the elements of articulation or their linking characteristics. The results of our analysis therefore identify three emergent traits of an articulation mindset to progress this exploration of intersectional research.

The first trait is the importance that the historical context plays in shaping the framing and analysis of the problem. Many authors examine historical antecedents that have shaped places, practices and systems of oppression (de Regt, 2017; Gurusami, 2017; Quinn and Ferree, 2019). In an expansive family history, de Regt (2017) situates these stories as a means of investigating how people's lives connect with global discourses in general

Table 2. Summary of selected intersectionality research.

	Study context	Main intersectionality-related findings
de Regt (2017)	Family history of a Yemeni woman of Ethiopian–Yemeni descent, exploring the link between micro-histories and macro-historical developments.	Highlights how family histories enable analysis of individual experiences and how they are linked to changing political economies and cultural dynamics.
Sharp (2021)	How local punk art and music events in Australia and the UK become queer(ed).	Develops a reflexive and intersectional 'insighter' ethnography methodology to contemplate researcher positionality in youth research.
Meier (2020)	Disabled ex-combatant populations and their experiences of reintegration versus ex-combatants without disability. .	How disability intersects with gender and location to create multiple and varied forms of marginalization.
Thorjussen and Sisjord (2020)	Students' experiences in a multi-ethnic co-educational physical education (PE) context.	Uses intersectionality to select students to include in analysis, such that their narratives showed how their multiple identities intersect and shape their PE experience.
Peretz (2017)	Why gender justice groups are disproportionately present in Atlanta, GA.	Examines how and why geographies shape activism from the perspectives of socially marginalized men who participate in men's gender justice groups.
Ugidos et al. (2020)	Perceived discrimination and internalized stigma during COVID-19 lockdown among the general population in Spain.	Uses an intersectional discrimination measure to evaluate discrimination on the basis of gender, ethnicity, mental health diagnosis or COVID-19.
Lord et al. (2019)	Student patterns within and through engineering disciplines and majors, examines engineering student outcomes based on discipline, race/ethnicity and sex.	Compares and contrasts models of student experience using an intersectional approach, focusing on persistence.
Rahman et al. (2020)	Bangladesh Violence against Women Survey.	Examines the probability that women that hold differing intersectional 'locations' may be more or less likely to experience violence. Locations are used over identities as the data they have represents their position rather than their self-held beliefs of their identity.
Earnshaw et al. (2021)	Understand and address discrimination and health inequities in diverse US communities.	Uses latent class analysis to explore discrimination across urban community spaces (rather than only grouping individuals together).
Ashlee et al. (2017)	Autoethnography of three 'womxn' of colour graduate students.	They use critical collaborative autoethnography as an act of resistance, and to foster 'wokeness', or a critical consciousness to intersecting systems of oppression.
Wiens et al. (2020)	Introduces materializing data as a useful methodology in intersectional feminist digital humanities.	Motivated to change the way digital humanities is done, to be more focused on social transformation, and foregrounds intersectional feminist principles.

and political and historical events in particular. Likewise, Gurusami (2017) gave a historical sketch illustrating precisely how labour and racial capitalism shaped the link between slavery, the emergence of women's prisons, welfare reform, and the imprisonment of black women in the US today. Whilst this mindset trait is not unique to articulation articles, it seems a more prominent feature because grasping the holistic or inner-workings of systems of oppression needs to examine the full story to appreciate the intricacies and complexities of lived experience. Furthermore, the significance of historical context allows us to comprehend how articulations emerge differently depending on context. In other words, we cannot predict structural outcomes because societies will be articulated differently in each context,

Table 3. Summary coding of articles by relationality frame adopted.

Relationality frame	Coder 1	%	Coder 2	%	Average
Additive thinking	75	44%	72	42%	43%
Between articulation and additive thinking	22	13%	27	16%	15%
Articulation	65	38%	64	37%	38%
Between articulation and co-formation	7	4%	6	4%	4%
Co-formation	3	2%	3	2%	4%
Totals	172		172		

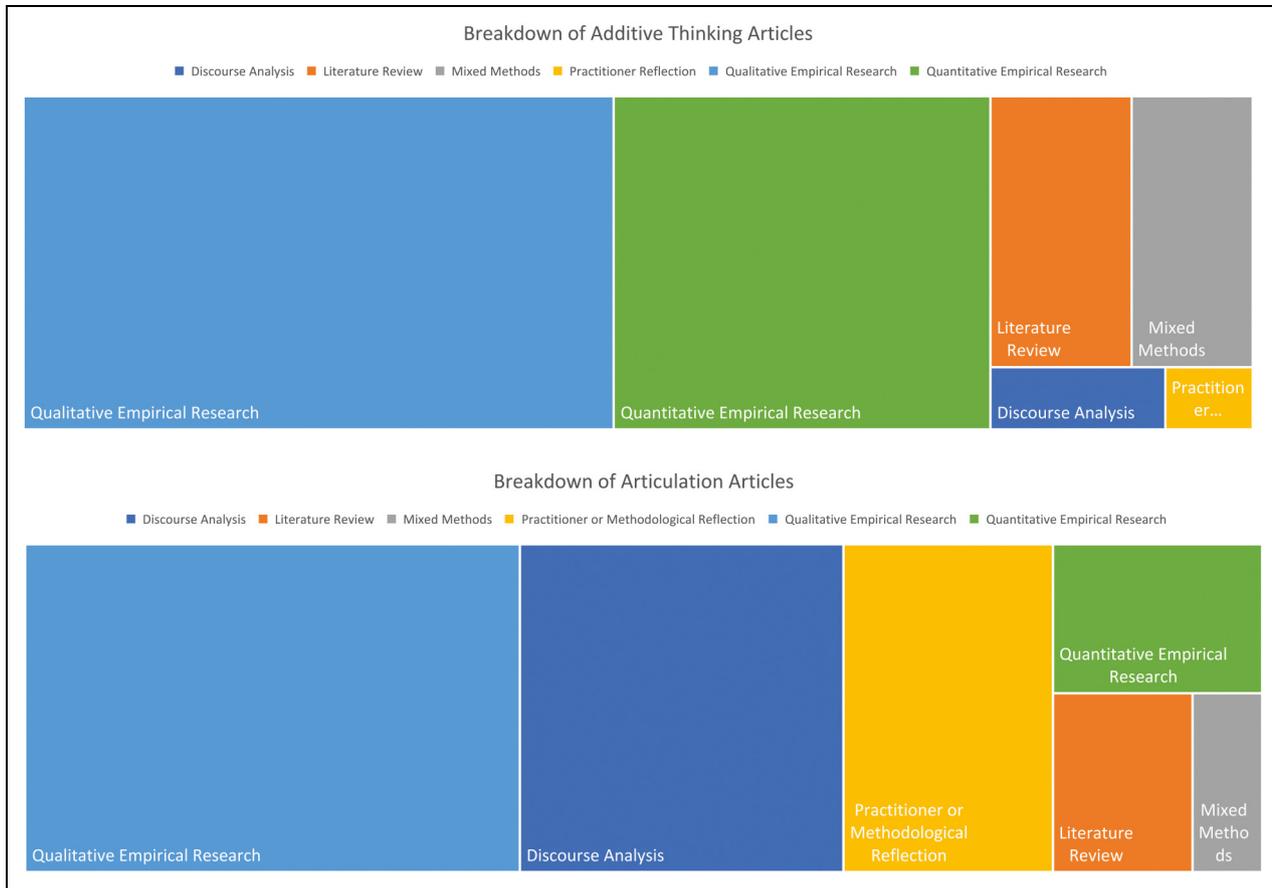**Level of agreement using percent agreement between two coders: 90%**

Figure 5. Treemaps of the breakdown of additive thinking and articulation articles by methodological approach.

and with this, their intersecting power relations will differ too.

The second mindset trait is about systems or integrative thinking. Typically, systems thinking involves focusing on dynamics between different levels of analysis (e.g. individual vs institutional (de Regt, 2017; Tomlinson et al., 2019)). Additionally, authors used a range of ideas to visualize or conceptualize linkages other than comparing levels of analysis, such as, inequality regimes (Quinn and Ferree, 2019), ecosystems (Harvey et al., 2018; Lord et al., 2019), or events, interventions or mutually constituted processes that reproduce or explain conditions (Maharani et al., 2019; Migliarini and Annamma, 2020; Stock and Birkenholtz, 2020). In contrast, Peretz (2017: 1447) described intersectionality as involving 'multiple categories and the intersections between categories, diversity within groups, time dynamics, relationships between individuals and institutions…[Functioning] at multiple levels of analysis, from structural analyses of power relationships, through institutions and groups, down to the level of individual identity'. For Peretz (2017), the articulation mindset may not only involve systems thinking[5] but also consider how boundaries are drawn or conceptualized, how various aspects interact, bleed together, or integrate in other words. The authors who used this trait tended to privilege the perspectives of underrepresented groups, focusing on challenging power relations and how they shaped discrimination or disadvantage.

The third mindset trait is to reflect on the association between positionality of the researcher and their membership to an intersectional community of inquiry. Within the scope of articulation, it is generally recognized that the relationships between the researcher, underrepresented groups and methodology must be acknowledged and reflected upon (D'Agostino, 2018; Sharp, 2021; Stock and Birkenholtz, 2020). As part of a broader attempt to confront inequalities and discrimination within intersectional communities of inquiry, Rankin and Thomas (2019) argued to cite and collaborate with black women and collaborate with the people written about. By adopting an articulation mindset, one's positionality within an intersectional community of inquiry is a vital interface serving this type of critical reflection. Moreover, it is often combined with taking a social-justice orientation (e.g. Gurusami, 2017; Peretz, 2017; Quinn and Ferree, 2019).

Of note for consideration within data science is the extent and depth at which authors collect and combine multiple data sources. For instance, Crenshaw (1990) drew on many angles and data sources (narratives, law, social science methods, and practice) to illustrate that constellations of

actors and institutions were operating on a 'single-axis framework' in her theorization of intersectionality. Similarly, de Regt (2017) and Peretz (2017) amongst others used a similar practice to supplement, triangulate and explore multiple angles and perspectives by combining data sources. This holistic scope for data practice, whilst potentially offering richer insights into mechanisms of structural inequality, requires further debate in data science, as it introduces new risks for underrepresented groups when the scope and scale of data collection and use are potentially increased.

Overall, we find that authors demonstrating an articulation mindset show greater engagement with structural mechanisms of inequality, which improves on the limitations of additive thinking. We now turn to examining how this mindset could be operationalized in data science, given most authors adopting the articulation mindset used qualitative methodologies.

Beyond additive thinking

In this section, we further examine the limitations of additive thinking for data science, arguing that the articulation mindset could strengthen data science methods to have a greater focus on mechanisms of structural inequality from a holistic social-justice perspective. We focus on the subset of 42 quantitative or mixed methods articles for this purpose. These articles align with common analysis techniques in data science. Most quantitative intersectionality research loses some of its power to combat oppression and systemic injustice when it adopts an additive thinking approach – favouring data practices that are necessary for appropriate statistical inference or analytics. Having an articulation mindset helps data scientists to consider not only what is necessary for data analysis but also the complete breadth of data practice. We outline how additive thinking is limited in the majority of quantitative publications, and we outline techniques that have the potential to remedy some of these issues.

As stated in Section 4.2, additive thinking offers a useful first step in research. However, it should not be the end goal of the research because it serves to identify the problem and does not explain why it exists. Intersectionality seeks to explain why people are underprivileged or discriminated against, and those answers are nuanced and complex. However, in our analysis, most articles were identifying issues through various analysis approaches rather than unpacking built-in structural inequality. Table 4 shows that most articles used regression analysis (60% (25/42)). Additive statistical methods entailed using regression models to analyze relationships between an independent identity variable and dependent social variables. A main effects regression model might have a main axis of inequality, such as race, correlating it with variables such as income or status that are used as proxies for social experience. For example, Nazareno et al.'s (2021) study predicted work experience differences between

Table 4. Summary of data analysis characteristics.

Aspects of quantitative data analysis	Prevalence
Analysis methods	
Descriptive statistics	14% (6/42)
Analysis of variance (*t*-tests and ANOVA)	10% (4/42)
Regression analysis	60% (25/42)
Linear regression (ordinary least squares, stepwise, multiple linear regression)	21% (9/42)
Logistic regression (multinomial, multivariable, binary, bivariate)	36% (15/42)
Latent class and latent profile analysis	7% (3/42)
Spatial methods: geographically weighted regression (GWR), spatial enumeration methodology, multivariable generalized estimating equations (GEEs)	7% (3/42)
Chi square	10% (4/42)
Prevalence ratio and calculated odds ratio	5% (2/42)
Multilevel approaches or nesting	33% (11/42)
Data context	
City/town/village	17% (7/42)
Global	2% (1/42)
N/A	2% (1/42)
National	60% (25/42)
Organization/university	7% (3/42)
Regional	12% (5/42)
Secondary/primary data	
Primary	31% (13/42)
Secondary	69% (29/42)

Filipino and white registered nurses in the US. Whilst models can consider multiple 'main' axes, this approach is additive and has been critiqued for reinforcing single axis thinking within intersectional studies (Evans, 2019).

Some articles used multiplicative regression analysis, which is said to improve on this limitation. These studies often explicitly attempted to move beyond main effects analyses by analyzing the points of intersection. For example, to investigate voter turnout intersectionally, Medenica and Fowler (2020) first built a main effects model with gender as the main axis. They then built a second main effects model with race as the main axis. Finally, they built a third model that considered the effects of gender and race combined – this is called an interaction model. Multiplicative methods have been criticized elsewhere for their reliance on main effects models because researchers often misspecify interpretations when interaction variables are not interpreted alongside main effects (Fehrenbacher and Patel, 2020). Misspecification occurs when there is a lack of engagement by the researcher with context. Unfortunately, it was difficult to assess whether articles reviewed made this error. Nevertheless, we still find that authors using multiplicative or additive regression analyses focused more on 'what' issues rather than 'why' issues.

Multilevel modelling is employed to engage more rigorously with how and why individuals are situated within social contexts and structures. For example, students can

be nested in schools and households can be nested in neighbourhoods. This makes a complex quantitative intersectional analysis that considers differences within and between social strata more feasible (e.g. Tomlinson et al., 2019). Multilevel models may also perform better than main effects models when more context-level units are included in the analysis (Evans and Erickson, 2019). Multilevel models may therefore address issues around additive thinking in which adding multiple identity axes can result in intersections with very small sample sizes, even in large datasets, which can lead to difficulty interpreting results as the number of interactions increase.

However, multilevel models require appropriate context-level data which may not be available, difficult to access, or could pose privacy risks. For example, Eco-Intersectional Multilevel (EIM)[6] modelling was developed to understand community-level environmental health risks in poor and minority communities (Alvarez and Evans, 2021). Alvarez and Evans (2021: 3) applied intersectionality and environmental justice theory to 'evaluate experiences of environmental injustice at community level'. As opposed to using individuals as the primary unit of analysis, EIM treats neighbourhoods[7] as the first level and intersectional strata of neighbourhoods[8] such as demographic and urbanization characteristics as the second level. Having access to large-scale context-level data sources enabled them to demonstrate that environmental injustice is 'socially patterned across numerous intersecting axes of marginalization, including axes rarely evaluated such as gendered family structure' (Alvarez and Evans, 2021: 1).

Whilst Alvarez and Evans' (2021) EIM approach has clear benefits for identifying structural mechanisms of inequality, the scope and scale of context-level data required are considerable. When compared to Nazareno et al.'s (2012) regression analysis, identified above as indicative of additive thinking, we argue that their approach improves on some limitations of additive thinking because they explained their results in the context of historical inequality, which is reflective of an articulation mindset. In some instances, choosing an approach like theirs may be safer because some aspects of structural inequality may still be revealed despite the quantitative data analysis approach. The articulation mindset could help both to guide these choices and enrich the analysis. It is not only a matter of considering what the best data analysis procedure should be.

Moreover, an articulation mindset supports consideration beyond data analysis techniques by prompting engagement with data collection processes and techniques as well. With 60% (25/42) of data analysis occurring at national levels, it was often unclear how findings or results will apply to underrepresented groups in specific locations. More than this, secondary data, used in 69% (29/42) of the articles, was not often reviewed or discussed in terms of the measures used to collect data in the first place, assuming that original data collection processes and associated categorizations were sufficient. Without greater reflection on the limitations and appropriateness of standard measures, additive thinking may aggravate inequalities through inappropriate logics of segregation.

Improvements and attention to data collection processes were highlighted in qualitative intersectional research. Qualitative researchers reflected on removing demographic information collection from the context of interviews or focus groups (Hagai et al., 2020), protecting the safety and privacy of data collection spaces (Alvi and Zaidi, 2017) adjusting the order or language of questions, considering reception of sensitive questions, or whether participants could express themselves in their own words (Quah, 2020). Seeing data collection as a holistic process in which collection methods are cognizant or linked to the emotional experience of participating in research may have additional benefits for both researchers and underrepresented groups. Worthen and Wallace's (2017) use of concurrent mixed methods in which quantitative measures were triangulated with a qualitative study shows how additional steps taken to confirm categorizations with participants could reflect the integrative thinking trait of the articulation mindset.

Moreover, how researchers develop relationships with other researchers and underrepresented groups is important from the articulation perspective. Doing preliminary research to develop an understanding of people, places and relationships beforehand was occasionally used (e.g. Peretz, 2017). This helped researchers understand data contexts prior to data collection and develop relationships with them before beginning. Many articles spoke of developing relationships with community organizations, often serving as recruitment partners, offering advice on reaching hard to reach people, advertising research and facilitating participation. Huot and Veronis' (2018) approach to holding 'town hall' meetings signalled a different approach to checking data with participants than emailing or calling participants and asking them to review recorded data. During meetings, they asked the community at large about the completeness, representativeness or equity of data. Although town hall meetings are a good example of how data scientists could engage in reflexive practice, we did not find that this example or others found in this review had progressed debates or practices around accountability towards communities most impacted by research surfaced by feminist scholars that were discussed in Section 2.

Ultimately, we have argued that when researchers select methods that give importance to statistical inference or analytics, most quantitative intersectionality research loses part of its capacity to address social injustices. This seems more likely to happen when additive thinking is used, which was found to be in most quantitative research articles. We discussed how taking an articulation mindset could enable reflection and evaluation on

the whole scope of data practice, not simply data analysis.

Discussion

Embarking on this research, we were initially interested in questions such as how is intersectionality understood? How should it be understood? And how could it be (more widely) practiced in data science? Yet, intersectionality functions as a multifaceted area of theory and praxis that is often contradictory. Intersectionality can be an expression of one's identity, which can be singular, multiple, and/or intersectional, rooted and stable, or changing constantly. It can constitute belonging and/or unpack marginalization and disadvantage. It can unite people in their endeavours and/or detonate struggles against systems of oppression, discrimination or persecution. It can be an abstract ideological project and/or rich and detailed experiences of existence. Crucially, it is a significant forum for investigating and transforming relationships between people, places and institutions, towards human rights, reduced inequality and social justice. Likewise, it can be and do none of those things – serving merely as a buzzword.

This paper considered interdisciplinary debates and studies of intersectionality to respond to the questions above. Table 1 illustrated how Collins' (2019) conceptualization around relationality captured adequately well the underlying logic of research across disciplines, even though her terms for the relationality frames appeared infrequently. Nevertheless, it was fairly evident which of the frames an article adopted since two independent coders agreed 90% of the time. The answer to the first question is therefore that additive thinking is used to understand intersectionality most of the time (Table 3). There was also a tendency towards additive thinking in quantitative intersectionality research (Figure 5). Clearly, by considering how understanding changes or grows by adding a dimension of analysis that has not previously been considered can be beneficial. For quantitative data analysis, numerous compelling methods for intersectional analysis have emerged, including multiple types of regression analysis and multilevel modelling. These advances should not be ignored and should be represented within intersectional data science.

However, in response to the second question, in line with feminist scholars that precede us (D'Ignazio and Klein, 2020; Erete et al., 2021; Mirza, 2009), we sought to consider how articulation could facilitate social-justice agendas explicitly in data science. For us, Collins' (2019) notion of articulation implies having a greater criticality not only about data science processes but views data scientists and their positionalities as vital links or actors in making praxis socially just. Section 4.2 contributed three traits of an articulation mindset, illustrating how authors engaged with mechanisms of structural inequality.

Mapping the historical origins of inequality, integrative or systems thinking and examining one's positionality in relation to the research participants or community of inquiry provide useful insights into how articulation should be enacted. These traits were often implicit in the approach, mirroring the feminist intersectional stance that 'there is no separation between our scholarly lives and systemic oppression' (Erete et al. 2021: 2).

These articulation mindset traits informed our evaluation of quantitative or mixed methods articles in Section 4.3. Bentley was motivated to learn about data analysis techniques that could go beyond additive thinking, in order to teach data science and artificial intelligence (AI) students such methods. If we are required within our institutions to teach from a data lifecycle standpoint, rather than our own, could we infuse the curriculum with methods that have transformational potential? In many articles reviewed, the general assumption is that differences in outcomes predicted can be viewed as disadvantages that are built into systems, institutions and/or societies. Although multilevel modelling can explain some of these disadvantages (Section 4.3), in teaching these methods without greater consciousness as to how inequalities become entrenched or what could be done to address them, serious concerns creep in. We suggest that the articulation mindset may help learners navigate these complexities.

Typically, authors adopting articulation often drew from several data sources, gathering robust sources of evidence (Section 4.2). This allowed for the triangulation of data, giving insights into complex webs of interrelated factors that influence inequity. Whilst our article sample was not appropriate for investigating the associated risks of increased data collection needed for multilevel or multivariable modelling, the articulation mindset could help to guide responsible data science choices. Vannini's research with Gomez and Newell (2019) on undocumented migrants showed that recording personal information may exacerbate the risks they face, including detention and deportation for themselves and their families, thus recommending data minimalism. Yet, it may be challenging to understand when data minimalism is needed, and not every circumstance requires it. Sometimes, reorienting data collection or analysis towards unpacking systemic injustice via other means could support undocumented migrants in their struggles for liberation. Whereas Costanza-Chock (2018) would argue that participatory approaches are vital for strengthening accountability, the articulation mindset traits – of unpacking the historical origins of inequality and/or using systems or integrative thinking to understand situated power dynamics – could establish accountability differently. Nevertheless, the articulation mindset also privileges building relationships with participants but recognizes that this may not always be possible or suitable.

As such, the articulation mindset highlights a holistic perspective on data practice, which creates space for

combining methods and insights in response to positionalities and relationships with participants. Oman's (2019) research with the UK cultural sector, which prides itself on being progressive and egalitarian, found some organizations were averse to including sexual orientation from its workforce data collection. Data workers misbelieved two aspects: firstly, that staff would see their sexuality and jobs as unrelated and reject these questions; secondly, that these data mattered less because their organization was inclusive. Qualitative research revealed informants were glad these data were being collected, as they had experienced discrimination that had been invisible. Thus, contextual specificities of data to address inequality should not be overlooked and the articulation mindset should improve this in data science.

Whilst articulation mindset traits proposed in this paper reflect preceding principles and theories (Collins, 2019; D'Ignazio and Klein, 2020), we showed through empirical review of intersectionality research how dominant approaches to intersectionality are enacted. We pointed to key examples that can be referenced, critically reviewing these articles' methodological techniques across disciplines and approaches. We build on D'Ignazio and Klein (2020) by suggesting how an articulation mindset could help data scientists make responsible choices at various stages of research and impact.

A limitation in this paper is that the co-formation frame of relationality was uncommon. Yet, in the case of Jimenez's research (see Jimenez et al., 2022 for an example), co-formation would help highlight issues from yet another vantage point. For example, whilst working with indigenous communities in Peru, the topic of gender was raised. The partners responded that their cosmovision does not differentiate gender from themselves or the environment. Such conflicts around communication, plurality or diversity in knowledges have not been considered in this review, but would still be useful to explore further. Another limitation is that innovative data science methods developed in related data justice, fairness, accountability or transparency fields may not have been considered because they did not describe their approach as intersectional. Whilst out of the scope of this review, we invite such researchers to the conversation and hope to see more of these researchers participate in intersectionality discourse.

Conclusion

Amidst calls for greater data-driven equity, societies generally remain unequal. Data gaps can obscure or misinterpret inequalities, whilst including marginalized individuals in data raises risks. Simplistic data categories often overlook the complex, intersectional nature of inequality, causing tension with traditional data science methods.

There is, therefore, a need to review intersectionality in data science, yet little is known of how it is used across disciplines. We undertook a literature review and coded intersectionality articles applying a tailored content analysis procedure. Using Collins' (2019) relationality frames of additive thinking, articulation and co-formation, we examined the articles for evidence of how intersectionality is being applied by researchers in order to understand how data science might further benefit from intersectional approaches for social justice.

We found that authors exhibiting an articulation mindset went beyond the confines of additive thinking by actively engaging with the mechanisms of structural inequality. However, most articles adopting the articulation frame were qualitative. We suggested that most quantitative intersectionality research loses some of its power to address social inequities because of the tendency towards additive thinking. We contributed three articulation mindset traits that may allow for a more responsible intersectional data science.

We need to encourage data scientists to adopt different practices motivated by more nuanced understandings of intersectionality. This paper investigated how intersectionality is being used across disciplines as a way to draw lessons for intersectional data science towards increasing social justice.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

ORCID iDs

Caitlin Bentley https://orcid.org/0000-0002-2602-601X
Chisenga Muyoya https://orcid.org/0000-0003-4048-8648
Sara Vannini https://orcid.org/0000-0003-1527-7494
Susan Oman https://orcid.org/0000-0002-5723-0830
Andrea Jimenez https://orcid.org/0000-0002-2166-8574

Supplemental material

Supplemental material for this article is available online.

Notes

1. The Inclusive Data Charter's (2022) five recommendations: (a) establish a commitment to centring the voices of individuals at greatest risk of marginalization or discrimination in all aspects of data systems and practice; (b) promote equity across the entire data value chain; (c) ensure that institutional data systems are inclusive and safe; (d) engage data to increase context awareness and reduce inequality and (e) build inclusive institutions.
2. The Inclusive Data Taskforce (2022) eight principles engage with environmental, institutional, statistical, infrastructural, methodological, privacy and regulatory issues.
3. The term 'underrepresented' is problematic because it can deny groups the opportunity to name themselves, erase differences, or imply unequal relationships between an overrepresented majoritized group and underrepresented minoritized populations (Walden et al., 2018). For this study which reviews literature enacting various methodologies and is concerned about how people at risk of marginalization or discrimination can be better represented by data science researchers and practice, we use this term per APA (2021) guidelines.
4. Hoffmann et al. is referring here to the US context, but this assertion applies to societies around the globe, albeit in different ways.
5. Systems thinking has seen significant paradigm shifts (Zexian and Xuhui, 2010), yet it is commonly used to break systems down into components and interrelationships between these and the environment, which may be viewed as reductive.
6. EIM is based on intersectional Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy (MAIHDA) (Merlo, 2018). MAIHDA applies multilevel models to study interactions at individual level (level 1) and within a social context (level 2) with the capability for analysis within and between social strata.
7. Census tracts as a geographical proxy for the neighborhood/community level.
8. Racial/ethnic composition, percent female-headed households, educational attainment, median household income level, and metro/non-metro locale.

References

Alcoff LM (2013) Philosophy and racial identity. In: Bulmer M and Solomos J (eds) *Ethnic and Racial Studies Today*. London: Routledge, 29–44.

Alvarez CH and Evans CR (2021) Intersectional environmental justice and population health inequalities: A novel approach. *Social Science & Medicine* 269: 113559.

Alvi S and Zaidi AU (2017) Invisible voices: An intersectional exploration of quality of life for elderly South Asian immigrant women in a Canadian sample. *Journal of Cross-Cultural Gerontology* 32(2): 147–170.

American Psychological Association (2021) Inclusive language guidelines. Available at: www.apa.org/about/apa/equity-diversity-inclusion/language-guidelines.pdf (accessed 26 August 2023).

Ashlee AA, Zamora B and Karikari SN (2017) We are woke: A collaborative critical autoethnography of three "womxn" of color graduate students in higher education. *International Journal of Multicultural Education* 19(1): 89–104.

Buolamwini J and Gebru T (2018) Gender shades: Intersectional accuracy disparities in commercial gender classification. In: Friedler SA and Wilson C (eds) *Conference on Fairness, Accountability and Transparency*. New York, USA, 23–24 February 2018. New York: PMLR, 77–91.

Christen P (2012) *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*. Berlin: Springer.

Collins PH (2019) *Intersectionality as Critical Social Theory*. Durham: Duke University Press.

Collins PH and Bilge S (2020) *Intersectionality*, 2nd Ed. Cambridge: Polity.

Cooper B (2016) Intersectionality. In: Disch L and Hawkesworth M (eds) *The Oxford Handbook of Feminist Theory*. Oxford: Oxford University Press, 385–486.

Costanza-Chock S (2018) Design justice: Towards an intersectional feminist framework for design theory and practice. In: Storni C, Leahy K, McMahon M, et al. (eds) *Design as a Catalyst for Change – Proceedings of DRS 2018 International Conference, Vol. 1*. Limerick, Ireland, 25–28 June 2018. London: Design Research Society, 529–540.

Crenshaw K (1990) Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review* 43: 1241–1299.

D'Agostino S (2018) Intersectional mobilization and the EU: Which political opportunities are there for Romani women's activism? *European Yearbook of Minority Issues Online* 15(01): 23–49.

De Regt M (2017) From Yemen to Eritrea and back: A twentieth century family history. *Northeast African Studies* 17(1): 25–50.

D'Ignazio C and Klein LF (2020) *Data Feminism*. Cambridge: MIT press.

Earnshaw VA, Reed NM, Watson RJ, et al. (2021) Intersectional internalized stigma among black gay and bisexual men: A longitudinal analysis spanning HIV/sexually transmitted infection diagnosis. *Journal of Health Psychology* 26(3): 465–476.

Erete S, Rankin YA and Thomas JO (2021) I can't breathe: Reflections from black women in CSCW and HCI. *Proceedings of the ACM on Human-Computer Interaction* 4(CSCW3): 1–23.

Evans CR (2019) Modeling the intersectionality of processes in the social production of health inequalities. *Social Science & Medicine* 226: 249–253.

Evans CR and Erickson N (2019) Intersectionality and depression in adolescence and early adulthood: A MAIHDA analysis of the national longitudinal study of adolescent to adult health, 1995–2008. *Social Science & Medicine* 220: 1–11.

Fehrenbacher AE and Patel D (2020) Translating the theory of intersectionality into quantitative and mixed methods for empirical gender transformative research on health. *Culture, Health & Sexuality* 22(sup1): 145–160.

Gurusami S (2017) Working for redemption: Formerly incarcerated black women and punishment in the labor market. *Gender & Society* 31(4): 433–456.

Hagai EB, Annechino R, Young N, et al. (2020) Intersecting sexual identities, oppressions, and social justice work: Comparing LGBTQ baby boomers to millennials who came

of age after the 1980s AIDS epidemic. *Journal of Social Issues* 76(4): 971.

Hall S (1980) Race, articulation and societies structured in dominance. In: Morley D (eds) *Essential Essays, Volume 1: Foundations of Cultural Studies*. Durham: Duke University Press, 305–345.

Haraway D (1985) A manifesto for cyborgs: Science, technology, and socialist feminism in the 1980s. *Socialist Review* 80: 65–108.

Harding SG (1986) *The Science Question in Feminism*. Ithaca: Cornell University Press.

Harris JC and Patton LD (2019) Un/doing intersectionality through higher education research. *Journal of Higher Education* 90(3): 347–372.

Hartsock NC (1998) Standpoint theories for the next century. *Women & Politics* 18(3): 93–101.

Harvey S, Voelker DK, Cope E, et al. (2018) Navigating the leadership labyrinth: Barriers and supports of a woman collegiate coach in a 20-year leadership role. *Sports Coaching Review* 7(1): 45–62.

Hastie T, Tibshirani R, Friedman JH, et al. (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.

HM Government (2022) *National AI strategy, Report, May*. London: HM Government.

Huot S and Veronis L (2018) Examining the role of minority community spaces for enabling migrants' performance of intersectional identities through occupation. *Journal of Occupational Science* 25(1): 37–50.

Inclusive Data Charter (2022) *Unpacking Intersectional Approaches to Data: A White Paper Produced by the Inclusive Data Charter, Champions and partners, Report, June*. London: Inclusive Data Charter.

Inclusive Data Taskforce (2022) *Leaving No One Behind – How Can We be More Inclusive in our Data? Report, September*. London: Office for National Statistics.

Jimenez A, Delgado D, Merino R, et al. (2022) A decolonial approach to innovation? Building paths towards buen vivir. *Journal of Development Studies* 58(9): 1633–1650.

Keith G (2021) The dark side of data science. Available at: https://towardsdatascience.com/the-dark-side-of-data-science-30b554bcf946 (accessed 26 August 2023).

Llewellyn J and Llewellyn KR (2015) A restorative approach to learning: Relational theory as feminist pedagogy in universities. In: Light TP, Nicholas J and Bondy R (eds) *Feminist Pedagogy in Higher Education: Critical Theory and Practice*. Waterloo: Wilfrid Laurier Press, 11–31.

Lord SM, Ohland MW, Layton RA, et al. (2019) Beyond pipeline and pathways: Ecosystem metrics. *Journal of Engineering Education* 108(1): 32–56.

Maharani CD, Moeliono M, Wong GY, et al. (2019) Development and equity: A gendered inquiry in a swidden landscape. *Forest Policy and Economics* 101: 120–128.

Martens D (2022) *Data Science Ethics: Concepts, Techniques, and Cautionary Tales*. Oxford: Oxford University Press.

Masquelier C (2022) Pluriversal intersectionality, critique and utopia. *The Sociological Review* 70(3): 616–631.

Medenica VE and Fowler M (2020) The intersectional effects of diverse elections on validated turnout in the 2018 midterm elections. *Political Research Quarterly* 73(4): 988–1003.

Meier LD (2020) An intersectional approach to the understanding of patterns of marginalization among ex-combatants with

disabilities in Sri Lanka. *Conflict, Security & Development* 20(4): 441–465.

Merlo J (2018) Multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA) within an intersectional framework. *Social Science & Medicine* 203: 74–80.

Migliarini V and Annamma SA (2020) Classroom and behavior management: (Re)conceptualization through disability critical race theory. In: Papa R (eds) *Handbook on Promoting Social Justice in Education*. Basel: Springer Cham, 1–22.

Mirza HS (2009) Plotting a history: Black and postcolonial feminisms in 'new times'. *Race Ethnicity and Education* 12(1): 1–10.

Nazareno J, Yoshioka E, Adia AC, et al. (2021) From imperialism to inpatient care: Work differences of Filipino and white registered nurses in the United States and implications for COVID-19 through an intersectional lens. *Gender, Work & Organization* 28(4): 1426–1446.

Oman S (2019) Improving data practices to monitor inequality and introduce social mobility measures – a working paper for the cultural sector. Report, The University of Sheffield, UK, July.

Oman S (2021) *Understanding Well-Being Data: Improving Social and Cultural Policy, Practice and Research*. Cham: Palgrave Macmillan.

Onuoha M (2018) On missing data sets. Available at: https://github.com/MimiOnuoha/missing-datasets (accessed 26 August 2023).

Peretz T (2017) Engaging diverse men: An intersectional analysis of men's pathways to antiviolence activism. *Gender & Society* 31(4): 526–548.

Quah SEL (2020) Transnational divorces in Singapore: Experiences of low-income divorced marriage migrant women. *Journal of Ethnic and Migration Studies* 46(14): 3040–3058.

Quinn JS and Ferree MM (2019) Schools as workplaces: Intersectional regimes of inequality. *Gender, Work & Organization* 26(12): 1806–1815.

Rahman L, Du Mont J, O'Campo P, et al. (2020) Intersectional community correlates of married women's experiences of male intimate partner physical violence in Bangladesh: A cross-sectional study. *Journal of Epidemiol Community Health* 74(2): 182–189.

Rankin YA and Thomas JO (2019) Straighten up and fly right: Rethinking intersectionality in HCI research. *Interactions* 26(6): 64–68.

Reynolds T (2002) Re-thinking a black feminist standpoint. *Ethnic and Racial Studies* 25(4): 591–606.

Saldaña J (2014) Coding and analysis strategies. In: Leavy P (eds) *The Oxford Handbook of Qualitative Research*. Oxford: Oxford University Press, 581–605.

Sharp M (2021) Insighters': The complexity of qualitative methods in youth music research. *Journal of Youth Studies* 24(6): 799–814.

Stevens N, Hoffmann AL and Florini S (2021) The unremarked optimum: Whiteness, optimization, and control in the database revolution. *Review of Communication* 21(2): 113–128.

Stock R and Birkenholtz T (2020) Photons vs. firewood: Female (dis) empowerment by solar power in India. *Gender, Place & Culture* 27(11): 1628–1651.

Taylor D (2022) UK about to deport up to 30 Kurdish asylum seekers to Iraq despite dangers. Available at: https://www.theguardian.com/uk-news/2022/may/30/uk-about-to-deport-up-

to-30-kurdish-asylum-seekers-to-iraq-despite-dangers (accessed 26 August 2023).

Thorjussen IM and Sisjord MK (2020) Inclusion and exclusion in multi-ethnic physical education: An intersectional perspective. *Curriculum Studies in Health and Physical Education* 11(1): 50–66.

Tomlinson J, Valizade D, Muzio D, et al. (2019) Privileges and penalties in the legal profession: An intersectional analysis of career progression. *British Journal of Sociology* 70(3): 1043–1066.

Ugidos C, López-Gómez A, Castellanos MÁ, et al. (2020) Evolution of intersectional perceived discrimination and internalized stigma during COVID-19 lockdown among the general population in Spain. *International Journal of Social Psychiatry* 68(1): 55–63.

Vannini S, Gomez R and Newell BC (2019) Mind the five: Guidelines for data privacy and security in humanitarian work with irregular migrants and vulnerable populations. *Journal of the Association for Information Science and Technology* 71(8): 927–938.

Walden SE, Trytten DA, Shehab RL, et al. (2018) Critiquing the "Underrepresented Minorities" Label. In: CoNECD-The Collaborative Network for Engineering and Computing Diversity Conference, Crystal City, USA, April 2018. Washington, DC: ASEE.

Wall JD, Stahl BC and Salam A (2015) Critical discourse analysis as a review methodology: An empirical example. *Communications of the Association for Information Systems* 37(1): 257–285.

Wiens B, Ruecker S, Roberts-Smith J, et al. (2020) Materializing data: New research methods for feminist digital humanities. *Digital Studies/Le champ numérique* 10(1): 1–22.

Worthen MG and Wallace SA (2017) Intersectionality and perceptions about sexual assault education and reporting on college campuses. *Family Relations* 66(1): 180–196.

Wynter S (2003) Unsettling the coloniality of being/power/truth/freedom: Towards the human, after man, its overrepresentation—an argument. *CR: The New Centennial Review* 3(3): 257–337.

Young E, Wajcman J and Sprejer L (2021) *Where are the Women? Mapping the Gender Job Gap in AI, Report, March.* London: The Alan Turing Institute.

Zexian Y and Xuhui Y (2010) A revolution in the field of systems thinking—a review of Checkland's system thinking. *Systems Research and Behavioral Science: The Official Journal of the International Federation for Systems Research* 27(2): 140–155.