Article:

Bates, J. orcid.org/0000-0001-7266-8470, Clough, P., Jäschke, R. et al. (2 more authors) (2020) Guest editorial : special section on “social and cultural biases in information, algorithms, and systems”. Online Information Review, 44 (2). pp. 321-323. ISSN 1468-4527

https://doi.org/10.1108/OIR-04-2020-400

Online Information Review

Special Section on "Social and Cultural Biases in Information, Algorithms, and Systems"

Journal:	*Online Information Review*
Manuscript ID	OIR-03-2020-0106
Manuscript Type:	Editorial

Editorial

Special Section on "Social and Cultural Biases in Information, Algorithms, and Systems"

Computer algorithms and analytics play an increasing role in citizens' lives, as they underlie the popular information services and "smart" technologies, which are rapidly being adopted across sectors of society, from transportation to education and healthcare. Algorithms allow the exploitation of rich and varied data sources, in order to support human decision-making and/or take direct actions; however, there are increasing concerns surrounding their transparency and accountability. There is growing recognition that even when designers and engineers have the best of intentions, systems relying on algorithmic processes can inadvertently result in serious consequences in the social world, such as biases in their outputs that can result in discrimination against individuals and/or groups of people. Recent cases in the news and media have highlighted the wider societal effects of data and algorithms, and have exposed examples of gender, race and class biases in popular information access services.

It is important to note the complexity of social and cultural biases in algorithmic processes. For instance, recent research shows that word embeddings, a class of natural language processing techniques that enable machines to use human language in sensible ways, are quite effective at absorbing the accepted meaning of words (Caliskan et al., 2017). These algorithms also pick up on the human biases, such as gender stereotypes (e.g., associating male names with concepts related to career, and female names with home/family) and racial stereotypes (e.g., associating European-/African-American names with pleasant/unpleasant concepts) embedded in our language use. These biases are "accurate" in that they are comparable to those discovered when humans take the Implicit Association Test, a widely used measure in social psychology that reveals the subconscious associations between the mental representations of concepts in our memory (Greenwald et al., 1998).

The biases inherent in word embeddings provide a good illustration for the need to promote algorithmic transparency in information systems. Word embeddings are extensively used in services, such as Web search engines and machine translation systems (e.g., Google Translate), which rely on the technique to interpret human language in real time. It may be infeasible to eradicate social biases from algorithms while preserving their power to interpret the world, particularly when this interpretation is based on historical and human-produced training data. In fact, another way of viewing such unconscious biases is as sources of 'knowledge diversity'; what one thinks are the true facts of the world, and how one uses language to describe them, is very much dependent on local context, culture and intentions. An alternative approach would be to systematically trace and represent sources of 'knowledge diversity' in data sources and analytic procedures, rather than eliminate them (Giunchiglia et al., 2012). Such approaches would support accountability in algorithmic systems (e.g., a right to explanation of automated decisions, which to date has proven very challenging to implement). In addition, these approaches could facilitate the development of more "fair" algorithmic processes, which take into account a particular user's context and the extent of "informedness" (Koene et al., 2017).

This special section brings together researchers from different disciplines who are investigating and tackling bias within their discipline, arising from the data, algorithms and the methods they use. From the ten submissions we received, after peer review, the following four articles were selected:

- Situated Algorithms: A Sociotechnical Systemic Approach to Bias
- Antagonistic Bias: Developing a Typology of Agonistic Talk on Twitter Using Gun Control Networks
- An Investigation of Biases in Web Search Engine Query Suggestions
- Algorithmic Equity in the Hiring of Underrepresented IT Job Candidates

In their conceptual paper, *Situated Algorithms: A Sociotechnical Systemic Approach to Bias,* Claude Draude, Goda Klumbyte, and Phillip Lücking argue that efforts to address bias in algorithmic systems require a sociotechnical approach that necessitates translational work between disciplines. In this paper, they draw specifically upon the fields of gender and diversity studies, to explore the ways in which conceptual frameworks such as situated knowledges, standpoint theory and strong objectivity are a useful lense for understanding and reframing the discourse around bias, and addressing the ways in which inequalities in algorithmic systems ought to be accounted for. In examining these issues, the authors provide the reader with a novel and concrete approach that can be used to develop situated understandings of algorithmic bias in different contexts.

Negotiating disagreement and conflict is a part of effective communication and integral to democratic discourse. Jose Marichal and Richard Neve propose a method for increasing civil dialogue by encouraging agonistic online talk. *Antagonistic Bias: Developing a Typology of Agonistic Talk on Twitter Using Gun Control Networks* presents a typology of modes that identify antagonistic / agonistic discourse on Twitter. The authors offer a number of ways that this typology could be used to create a more democratic environment. Using AI and NLP it would be possible to counter the proliferation of antagonistic tweets and “fake news” by identifying opposite poles of a discourse space and introducing gradations of the argument or introducing bias scores. Ultimately, developing such interventions can reset public dialogue to civil discourse.

In *An Investigation of Biases in Web Search Engine Query Suggestions* Bonart et al. focus on web search engines – one of the most widely used entry points to the vast amounts of information available on the World Wide Web. Specifically, they propose an approach to identify and analyse potential biases within query suggestions provided by search engines related to person names. Such suggestions reach a wide audience and can potentially bias people via exposure or by leading them to Web pages that contain further biased information. As the suggestions are typically computed from people’s query histories, this would constitute an example of second-order bias – a phenomenon that can lead to self-reinforcing feedback loops. The article of Bonart et al. lays the groundwork for better understanding, and ultimately avoiding, such bias originating from a subtle, but powerful mechanism that feature as a key part of today’s search engines.

Much has been said recently about social bias in algorithmic decision support in the context of hiring and recruitment, with some high-profile companies such as Amazon abandoning the practice.[1] Nonetheless, some still hold out hope that such tools could promote equality, by reducing human biases in the hiring process. In *Algorithmic Equity in the Hiring of Underrepresented IT Job Candidates* the authors Lynette Yarger, Fay Cobb Payton, and Bikalpa Neupane consider the case of talent acquisition software used in the case of IT hires. Using feminist design thinking as a theoretical lens, they uncover several sources of bias in the software. Their analysis demonstrates that data – and data-driven analyses – are only one source of information that should be used in hiring decisions; human experience and expertise must always complement the use of these tools.

Dr. Jo Bates, Information School, University of Sheffield, UK
Prof. Paul Clough, Information School, University of Sheffield, UK
Prof. Robert Jäschke, Humboldt-Universität zu Berlin, Germany
Prof. Jahna Otterbacher, Open University of Cyprus
Prof. Kristene Unsworth, Drexel University, Philadelphia, USA

References

Aylin Caliskan, Joanna J. Bryson, and Arvind Narayanan. 2017. Semantics derived automatically from language corpora contain human-like biases. *Science* 356.6334: 183-186.

Anthony G. Greenwald, Debbie E. McGhee, and Jordan LK Schwartz. 1998. Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology* 74.6: 1464-1480.

Fausto Giunchiglia, Vincenzo Maltese, Anthony Baldry, and Biswanath Dutta. 2012. Living Knowledge. Università degli Studi di Trento. *Technical Report* # DISI-12-015.

Ansgar Koene, Elvira Perez, Sofia Ceppi, Michael Rovatsos, Helena Webb, Menisha Patel, Marina Jirotka, and Giles Lane. 2017. Algorithmic Fairness in Online Information Mediating Systems. In *Proceedings of the 2017 ACM on Web Science Conference*, pp. 391-392. ACM Press.

[1] https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G