

September 7-9, 2021

Weblink: <http://cdltr.uohyd.ac.in/iceie>

Abstract



**Dr. Galit Wellner**

Tel Aviv University, Israel  
[galit\\_wellner@yahoo.com](mailto:galit_wellner@yahoo.com)

### Gender and Racial Biases in AI

The rise of AI algorithms has been accompanied by the belief that these systems are impartial and do not suffer from the biases that humans and previous technologies express. It becomes evident, however, that gender and racial biases can (and practically do) exist in AI algorithms. The question is where the bias is rooted – in the training dataset or in the algorithm? Is it a linguistic issue (and hence related to the dataset) or a broader sociological current (and hence can be also in the algorithm itself)? Several solutions have been proposed to the problem of biases in AI. The paper arranges them into five categories: (a) "no gender or race" - ignoring and omitting any reference to gender and race from the dataset; (b) transparency - revealing the considerations that led the algorithm to reach a certain conclusion; (c) designing algorithms that are not biased; (d) involving humans in the process; or (e) "machine education" that complements "machine learning" by adding value sensitivity to the algorithm.

The paper will assess the feasibility of the five types of solutions and consequently will provide policy recommendations according to two of them: Although type (b) solutions (transparency) are complicated to implement due to the current regime of intellectual property as well as cultural and business considerations, such solutions can lead to policy guidelines similar to those of the pharmaceutical industry in the process of drug approval. Type (d) solutions ("human in the loop") can be implemented in several points in the lifecycle of an AI system starting from the development stage, to the usage stage in which an "ombudsman" mechanism can be applied to ensure the biases detected by the users are dealt with and overcome by the companies who develop and run the algorithms.

#### Brief Bio :

Dr. Galit Wellner, PhD., is an adjunct professor at Tel Aviv University and at Bezalel Academy of Art and Design. Galit studies digital technologies and their inter-relations with humans. She is an active member of the Postphenomenology Community that studies philosophy of technology. She published several peer-reviewed articles and book chapters. Her book *A Postphenomenological Inquiry of Cellphones: Genealogies, Meanings and Becoming* was published in 2015 in Lexington Books. She translated to Hebrew Don Ihde's book *Postphenomenology and Technoscience* (Resling 2016). She also co-edited *Postphenomenology and Media: Essays on Human–Media–World Relations* (Lexington Books, 2017). In the last years, she published several articles and book chapters on gender bias in AI and algorithmic imagination. She currently works on two books on attention and on imagination.