A Pathway towards Co-creating Responsible Standards for Digital Equity: A Case Study of Digitization of Women's Transit Safety in India

Swati Kumari[1], Raghvendra Singh Yadav[2]

[1]Institute for Advanced Studies on Science Technology and Society, Graz University of Technology, Austria

[2] Centre for Studies in Science Policy, Jawaharlal Nehru University, India

DOI 10.3217/978-3-85125-976-6-06

Abstract. It is well established that pre-existing biases of the socio-technical landscape get auto-embedded in algorithms leading to the persistent digital reproduction of social biases. Gender biases are one of the most prominent categories of existing biases in digital technologies. Gender biases are clearly reflected in the algorithms used in reproductive technologies, health technologies, employment and marketing platforms. This paper suggests the intervention before formulating logic for algorithms. In this context, the **primary objective** of this research paper is **to provide a methodological pathway for responsible co-creation of the standards for gender equity in digital technologies**. The research paper seeks to answer the question of **how to co-create inclusive responsible standards of gender equity for the digitization process.** To address the research question, the paper utilizes both primary and secondary data. Primary data is taken from the field survey done in Delhi, as a part of a larger project on women's safety in transportation. The survey was designed by using a modified **Responsible Research and Innovation Framework.** The survey has used a **mixed-method participatory approach**, where qualitative methods included, in-depth interviews and focus group discussions, and quantitative methods included statistical modeling using SPSS software. The universe of the field survey was women commuters of Delhi, belonging to different ages, classes, educational backgrounds, reproductive cycle stages, and locations. The paper provides a pathway for harmonious intersectional reconciliation of diverging interests for responsible standards of women's transit safety.

1 Introduction

Digitization is considered a key to unlocking the development potential among policymakers. Especially in developing countries, digitization is assumed to be the most efficient pathway for removing bottlenecks and boosting the services to the citizens.

Digitization in multiple sectors despite having efficiency advantages, is also riddled with fundamental questions such as ‘equity’ which manifests as a digital divide (Bakon et al., 2020). Socially disadvantaged groups including women are more affected by the digital divide in terms of access as well as opportunities (Acilar and Sæbø, 2023). To move towards an inclusive digitized future, it is critical to understand the type and nature of the relationship between social inequalities and the process of digitization. The two significant facets of the relationship between the digital divide and social inequalities include access to ICT services and the inherent design of the digital infrastructure. This paper focuses on the latter and attempts to develop an understanding of how the inherent design of technology development is influenced by preexisting social inequalities.

1.1 Digital Divide and Social Inequalities

New technologies come and operate in social contexts and digital technology is no different. Prevailing social and cultural norms have a defining impact on the shaping of these technologies. For instance, Bala and Singhal (2018) identified social norms, financial restrictions, and exclusion from fundamental technological skills as the causes of the gender gap in ICT usage. The family members' consent for women to use and access the internet is the most crucial social norm in India (Bala and Singhal, 2018). The influence of the prevailing social norms on embedded values of technology also means that the personal biases of the innovators or technology creators getting encapsulated in the digital products and services. Without any conscious efforts and awareness, there are high chances of auto-embedding social biases in digital products, processes, and services. Such auto-embedding effectively leads to the digital reproduction of social inequalities including gender biases, rural-urban divide, and racial and ethnic discrimination (Choi et al., 2022). Pre-existing social inequalities are reflected in digital platforms in multiple sectors including marketing, advertising, (Chang et al., 2021), financial sector, (Mogaji et al., 2020), and the criminal justice system (Ugwudike, 2022). Even WHO has come out with ethical guidelines for digital health platforms (Organization, 2021), and there is increasing demand for considering gender intersectionality in digital health services (Figueroa et al., 2021). Studies on the digital divide report the presence of algorithm biases as the major factor in perpetuating social inequalities, but, the big question is how it is happening.

1.2 Vicious Cycle of Algorithm Biases

Broadly speaking, a fundamental of the current digitization process is machine learning where, a large data set is used for finding patterns, developing logic and modeling that logic. Based on the logic, algorithms are developed which later translate into digital products (i.e. Android applications) or digital platform-based services. At every step of

the algorithm development process, there are very high chances of biases entering into the cycle (Marjanovic et al., 2022). For instance, how the data has been collected, has the data been sanitized from biases? If feed data is skewed and have biases then these biases get amplified at each step in developing algorithms. There is also the issue of the feedback loops of the data sets used across sectors for making predictive decisions which have implications for vulnerable groups in accessing services (Qureshi, 2020). For example, data from health apps are used by hospitals, and data from insurance companies are used by financial institutions for making financial credit decisions for individuals. Thus, trapping the vulnerable groups in a vicious cycle of deprivation.

The embedding of biases in algorithms intensifies the existing social inequalities because of the speed of diffusion of biases through digital platforms (Marjanovic et al., 2022). For instance, making a biased, discriminatory decision is as fast as downloading a digital App or a simple tap on an App. Another important concern regarding biases in algorithms is the 'universalization of social inequalities' (Choi and Cristol, 2021). Every day new communities are coming under the fold of digitization hence communities are also exposed to new biases. One of the biggest challenges for data-driven development and automation of services is to protect the privacy of highly vulnerable groups. Gender biases are one of the most prominent categories of existing biases in digital technologies. Gender biases are clearly reflected in the algorithms used in reproductive technologies, health technologies, employment, and marketing platforms. Thus, this paper focuses on gender biases in algorithms.

1.3 Objective

There are two potential approaches to counter gender biases in algorithms. One is a gender audit of the algorithms to identify and filter the biases. This approach relies on opening the black box of the algorithms from the perspective of, who designed the algorithms, why particular algorithms, and how it was designed. Another approach is to adopt a gender-responsive pathway for ideating and actualizing particular algorithms. This approach focuses on the intervention before formulating logic for algorithms. However, the meaningful application of both approaches requires a guiding framework setting standards for gender equity. In this context, the **primary objective** of this research paper is **to provide a methodological pathway for responsible co-creation of the standards for gender equity in digital technologies**.

2 Wicked Problem of Biased Algorithms

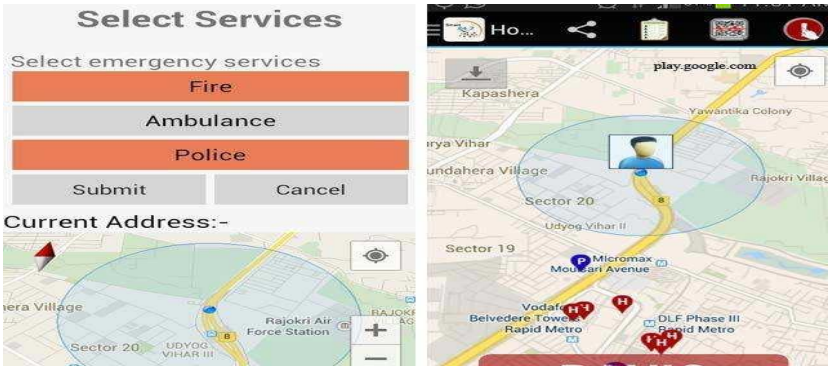
The problem of biases in algorithms is rooted in the complex web of interaction between the process of technology development and society. Before offering a solution, it is crucial to understand the reasons for the embedding of the biases in the algorithms. It is well understood in Innovation studies that the social sustainability of new technologies is critical for the long-term economic viability of the products (Missimer and Mesquita, 2022). Still, due to several reasons, firms do not pay adequate attention to filtering the biases. One of the most prominent causal factors is cut-throat competition among firms in digital sector innovations where 'time' is a critical determinant of profit (Akter et al., 2021). Though many companies have brought up the 'code of ethics for program developers, however, due to the persistent paucity of time such codes at large remain the passive document (Gal et al., 2020). In such situations, firms tend to use the earliest available data set without much attention to social biases in the data. Another issue is the short-term focus on profit in the firms based on maximum profit in minimum time, which may even be detrimental in the long term. Thus the 'ephemeral' nature of digital innovations has very limited scope for ethical concerns such as social biases (Gal et al., 2020). On the other hand, it has been observed that the higher the vulnerability of a particular group to biased algorithms lower the knowledge of the existence of the biases in that group. Understanding 'information asymmetry' from the perspective of vulnerable groups is crucial for understanding the circulation of gender inequality in the digital realm (Alvesalo-Kuusi et al., 2022). At the same time even if there is an intention to be responsible firms are so devoid of social realities and have cut-off work cultures that there is a lack of knowledge and methods to develop inclusive digital products. The biggest challenge is to address the question of 'intersectionality' (Bala and Singhal, 2018; Choi and Cristol, 2021; Figueroa et al., 2021) through participation. The research paper seeks to answer the question of **how to co-create inclusive responsible standards of gender equity for the digitization process.** The following sections are dealing with the empirical study conducted and subsequently suggested methodological pathway for developing the inclusive gender-sensitive algorithm for women's transit safety.

3 Research Design, Methods and Procedures

A Responsible Innovation Framework (Buhmann and Fieseler, 2021; Dreyer et al., 2017; Steen et al., 2021; Stilgoe et al., 2013) including the dimensions of anticipation, reflexivity, responsiveness, and participation has been used for research design. This research work has used both primary and secondary data. For collecting primary data, a multi-round

field survey has been conducted among women commuters of public transport in Delhi, India. The sample size of the survey is 326 women commuters belonging to diverse groups formed around different parameters (discussed in detail in section 4). The hybrid Purposive Sampling technique was used to give representation to diverse sections of women commuters. Structured written questionnaires and semi-structured interviews were used to ensure the participation of a diverse group of women commuters.

For secondary data literature survey is done. Also, the existing women transit safety apps[15] in Delhi, such as 24*7 Smart, 'Safety pin', and 'Himmat' were analyzed from the Responsible Innovation perspective and gaps have been identified. Table.1 gives the output images of these safety apps

App Name	Screen Picture
Himmat	
24*7 Smart	

[15] Transit safety apps in Delhi specifically aim to enhance women's safety in the city. However, the inherent logic of these apps is GPS-based surveillance. These apps usually also have a crisis button to inform the security agencies in case of any unsafe incidents. However, the functional utility of these apps has been very limited the logic and parameters used for constructing these apps along with daily usage do not reflect the issues and concerns of everyday women commuters. Currently, 112 India, My Safetypin, Sheroes, bSafe, Himmat, Smart 24x7, Shake 2 Safety are prominent women-oriented digital safety apps in Delhi.

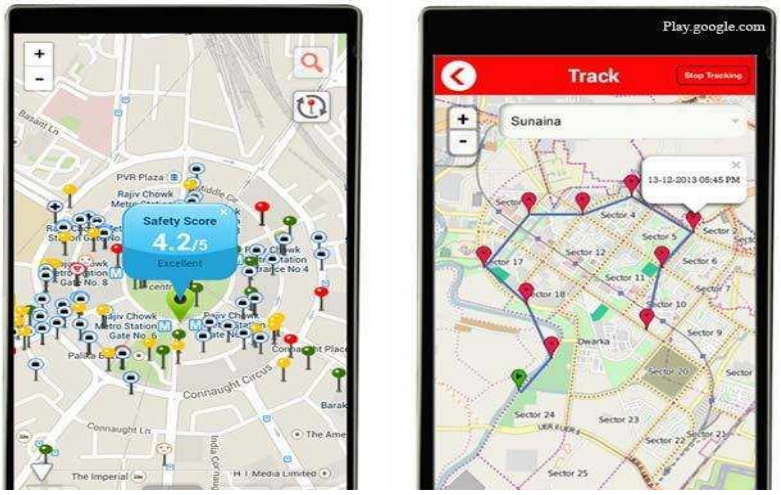
SaftyPin	

Table 1: Women safety aaps in Delhi screen shots

3.1 Gaps in the Digital Response to Women's Transit Safety

Existing apps and helplines promoted to address the crisis situation miss out on the intersectionality among women commuters and treat women commuters as a monolithic category. Hence diverse safety needs of women belonging to diverse groups cannot be addressed through these apps. Similarly, while developing the logic of these apps parameters like time and location were not factored in, while it is well established that there is a strong correlation between the women's perception of safety with time and location. The existing digital response focuses on the preservation of the physical being only and equates women's safety with physical safety only. This understanding is very limited as conceptual development in the field of gender transit safety shows that 'psychological safety' plays a significant role in perception formulation and making decisions on the mode of transportation among women commuters (Ceccato et al., 2022; KUMARI and SINGH, 2019). All these gaps necessitated a fresh understanding of women's safety from women's own perspective and translating that into an inclusive algorithm. The following section discusses the field study conducted to develop a logic of women's safety in transportation in an inclusive and participative way.

4 Field Survey Description

A multi-round field survey was conducted in Delhi's Public Transport System to collect cross-sectional data from women commuters. The city of Delhi has well well-developed Public Transport System including the rails, metro, buses, auto-rikshaws and e-rikshaws.

The city of Delhi has a notorious reputation for a worrisome situation of women's safety in public transportation especially in the bus transport system Mukherjee Mukherjee (KUMARI and SINGH, 2019; Mukerjee, 2019). Despite the government's efforts to improve the situation with multiple gender-sensitive schemes, the perception of women's safety has not improved. Thus, the Delhi bus transport system is an interesting case to study and find out the gap between the solutions offered for improving women's safety and the impact on the ground level. The survey was designed to capture the perceptions and expectations of the women using bus services in Delhi. Keeping the women commuters' socioeconomic and educational background, the survey was designed to collect qualitative and quantitative data. The qualitative data was further converted into quantitative data in consultation with the participants of the survey in multiple rounds.

To address the intersectionality among women users of public buses different categories were created based on age, income, education, and reproductive cycle stage. Within the category of age, data was collected from women belonging to the age group, up to 18 years, 18-40 years, 40-60 years, and above 60 years. Similarly, different subcategories were created to be inclusive of diverse educational backgrounds and different income levels among women commuters. The issues faced by women during various stages of the reproductive cycle have been completely absent from the transit safety discourses. Hence, under the reproductive cycle stage, data were collected from women experiencing or experienced, menstruation, pregnancy, lactation, and menopause-related issues while using public transportation.

It is well established that women's perception of safety is closely associated with time. Thus, data was collected to capture the women's expectation of safety during different times of transportation such as early morning, peak hours, evening, and late nights. However, the sense of time among women commuters depends on location. Like any other global megacity, Delhi also has dynamics of 'center & periphery' meaning the existence of the different transport cultures in the central and peripheral locations. For instance, the late-night time has a different meaning for women commuting in or to central and peripheral locations of the city. Thus, the survey was designed to include the voices of women commuters belonging to different parts of the city including the central part of Delhi and peripheral locations. The central and peripheral locations were determined by the geographical location and availability of the mode commutations.

The collected data has conversed into the 'values' women commuters aspired to be included in safety response. Thus, it is prudent to include the values related to women's safety in digital response towards enhancing women's safety in public transportation. Based on this data the statistical model is developed for analyzing various correlations between the safety expectations and issues to group characteristics. The following section discusses the major findings of the data analysis.

5 Findings

Analysis of collected data shows that different groups of women have different expectations in terms of transit safety. However, analysis of the logic of the existing digital apps shows that women's safety has been treated the same for all women without factoring in the time and location. The survey shows that since women are not a monolithic category, the safety needs of women commuters depend on a variety of factors. Such a gap indicates that understanding of women's safety among technology developers is assumptive and not reflective of the voices of women. In the absence of anticipation of parameters of safety from women's own experiences, existing apps remain generic and fail to enhance the sense of safety among women commuters. Therefore, the perception of the utility of these apps among women commuters is not favorable. Further, before constructing the logic these apps have not collected any primary data thus, they miss the community participation potential of responsible digital solutions. Apart from the heterogeneity and intersectionality present in the perception of safety, the study also finds out that for women commuters, psychological safety is equally important to physical safety. Which requires empathy and responsiveness for a meaningful solution as well as enhancing the perception of public transportation.

For the women commuters belonging to the lower economic income groups, economic safety in terms of affordability was found to be strongly correlated with overall safety needs. It was also found that pregnant women and young mothers are more concerned about the impact of pollution on the health of children. Thus, the study brought out four important components of women's overall transit safety as- Physical safety, Psychological safety, Economic safety, and Environmental safety. However, the data analysis shows that women in different parts of the city give variable weightage to these safety components. Even at the same location women belonging to the same group give different weightage to different safety components at different times. For instance, in some peripheral areas, women belonging to under 18 years of age in the low-income group give more weightage to physical safety during early morning and late night, but during the day economic safety concerns play a determining role in transit choices. Without including such nuances in constructing the logic for algorithms, the data collected by the safety apps remains skewed.

The study shows that there is inevitable friction of interest among different groups of women commuters. For example, the data shows that older women have a strong preference for enhancing surveillance through CCTV and security personnel whereas women with higher education (graduation and above) considered the increased surveillance ineffective and an infringement upon privacy. Similarly, women with high education understood the concept of autonomy and connected it with safety, while in the

rest of the groups, autonomy was one of the least preferred aspects of safety exhibiting a lack of utility of autonomy. Captive women commuters belonging to higher income groups prefer more privacy and comfort but women commuting from the peripheral areas for work to the city belonging to a low-income group (i.e. sanitation workers, domestic help) consider availability and reliability larger issues for safety. Thus, there exists a complex web of correlations embedded in the women's psychological process regarding feeling safe while using a public bus. It is challenging to balance the diverging safety requirements of women commuters, but without being responsive to the specific needs of a particular group the safety algorithms do not have any meaning to the women commuters. This is the reason most of the women commuting on buses do not use the existing safety app and many women are even found to be unaware of the existence of women's safety apps. This also indicates the lack of participation.

Most of the women respondents agreed that they would be able to trust the safety apps more if they had a way to participate in them. We propose a way for women commuters to participate by sharing their experiences in real-time which will act as feedback for other women commuters. In simple terms, women themselves decide the safety status of a particular location. The following sections discuss the proposed pathway for co-creating responsible digital responses for enhancing women commuters' sense of safety while using public transportation.

5.1 Proposed Pathway

The proposed pathway has interconnected multiple steps:

Step 1: Identify the different groups of the women commuters. For such groups, the possible parameters could be age, income, location (center or periphery), and reproductive cycle stage.

Step 2: Identify the safety issues and needs for each identified group of women commuters. We propose to identify the safety needs by ensuring the participation of each group and giving representation to each group's expectations. The group expectation needs to be categorized into parameters. Collaboration with social scientists and academic institutions at this stage can enhance the quality of the research. Define Safety for each group in terms of the underlying parameters.

Step 3: Define Safety for each group in terms of the underlying parameters. The parameters must reflect the women commuter's physical, psychological, economic, and biological safety needs.

Step 4: Statistical Modelling for establishing correlations. Each parameter is to be assigned a particular ranking point, for a particular safety parameter by each specific group. Weighted Averages would be applied to harmonize the diverse ranking of the parameters by different groups.

Step 5: Develop the algorithm based on correlations in the statistical model.

Step 6: Develop a digital product. Create an add-on based on the developed algorithm for the already existing transport Apps for specific cities. Here women commuters can share their experience of the commutations.

Step 7: Maintain the digital cloud from the data feed collected from the women commuters. The data cloud is the nodal point to provide feedback to women commuters based on the ranking of a particular transport area.

Figure 1 shows the entire suggested plan for developing an inclusive participative digital solution for transit safety apps.

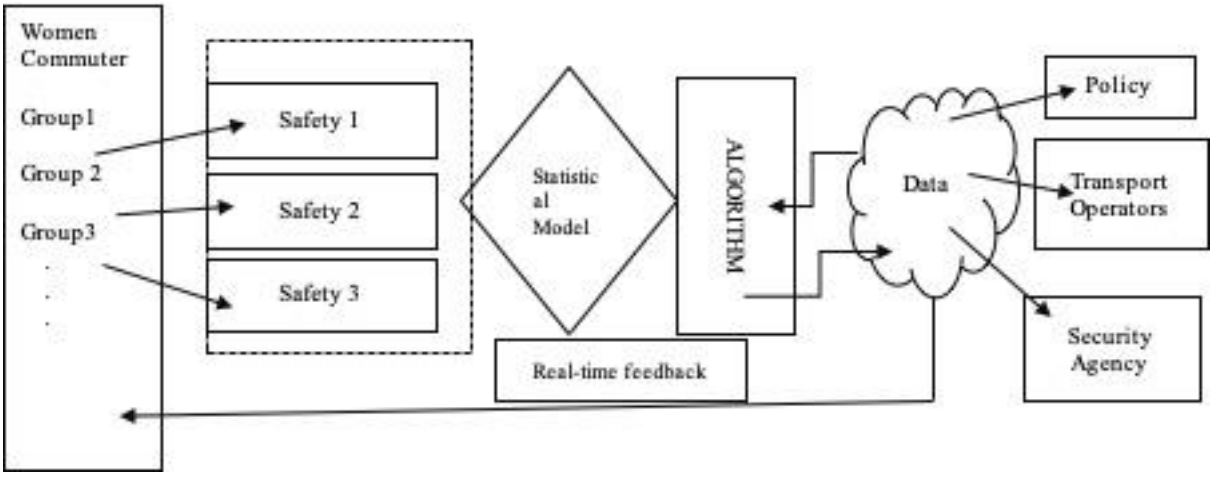


Figure 1: Proposed Pathway for Co-creating Responsible Algorithm for Women's Transit Safety

Many women respondents expressed an inability to do anything about the unsafe transport situation and feeling unsafe every day is very frustrating. Thus, providing a gateway to express feelings is a good starting point for ensuring the psychological safety of women commuters. The data cloud can be analyzed to provide inputs to security agencies in case of an emergency situation and identify the hotspots where women often feel unsafe while commuting. By analyzing the data, the transport managers can enhance their quality of service and attract more women commuters. In the long run, the data could be analyzed to find patterns and can provide critical policy inputs regarding women's safety in public transportation.

6 Discussion

It is clear from the field study and analysis of existing women's safety Apps that there is a knowledge gap between academic research (i.e. this study and other studies from different cities across the world) and understanding of women's safety among digital app developers. The academic research on women's safety offers deeper insights including highlighting the psychological aspects of safety which brings into question the effectiveness of the existing Apps. Thus, there is a need to translate academic knowledge into digital products with collaboration and consultation between academic institutions and technology development firms. Being responsive to the needs of women commuters such an approach has potential for economic as well as social sustainability. However, there are genuine concerns and limitations of the digital technology developers which we need to consider before proposing a methodological pathway for developing algorithms in an inclusive, participative, and responsible way. Currently, it is evident that there is unpreparedness in firms regarding upholding the moral overload of the digitization process. At this juncture, there is great potential for increasing the stakeholdership of academic institutions. Engagement with academic institutions before developing the logic or creating relevant data sets can help reduce the biases in the algorithms and lead toward a more inclusive digitization process.

References

Acilar, A., Sæbø, Ø., 2023. Towards understanding the gender digital divide: A systematic literature review. Glob. Knowl. Mem. Commun. 72, 233–249.

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y.K., D'Ambra, J., Shen, K.N., 2021. Algorithmic bias in data-driven innovation in the age of AI. Int. J. Inf. Manag.

Alvesalo-Kuusi, A., Malik, H.M., Viljanen, M., Lepinkainen, N., 2022. Dynamics of social harms in an algorithmic context. Int. J. Crime Justice Soc. Democr. 11, 182–195.

Bakon, K.A., Elias, N.F., Abusamhadana, G.A., 2020. Culture and digital divide influence on e-government success of developing countries: a literature review. Bakon Kinn Abass Cult. Digit. Divide Influ. E-Gov. Success Dev. Ctries. Lit. Rev. 15 1362–1378.

Bala, S., Singhal, P., 2018. Gender digital divide in India: A case of inter-regional analysis of Uttar Pradesh. J. Inf. Commun. Ethics Soc. 16, 173–192.

Buhmann, A., Fieseler, C., 2021. Towards a deliberative framework for responsible innovation in artificial intelligence. Technol. Soc. 64, 101475.

Ceccato, V., Gaudelet, N., Graf, G., 2022. Crime and safety in transit environments: a systematic review of the English and the French literature, 1970–2020. Public Transp. 1–49.

Chang, H.-C.H., Bui, M., McIlwain, C., 2021. Targeted Ads and/as Racial Discrimination: Exploring Trends in New York City Ads for College Scholarships. ArXiv Prepr. ArXiv210915294.

Choi, E.Y., Kanthawala, S., Kim, Y.S., Lee, H.Y., 2022. Urban/rural digital divide exists in older adults: Does it vary by racial/ethnic groups? J. Appl. Gerontol. 41, 1348–1356.

Choi, M., Cristol, D., 2021. Digital citizenship with intersectionality lens: Towards participatory democracy driven digital citizenship education. Theory Pract. 60, 361–370.

Dreyer, M., Chefneux, L., Goldberg, A., Heimburg, J. von, Patrignani, N., Schofield, M., Shilling, C., 2017. Responsible innovation: A complementary view from industry with proposals for bridging different perspectives. Sustainability 9, 1719.

Figueroa, C.A., Luo, T., Aguilera, A., Lyles, C.R., 2021. The need for feminist intersectionality in digital health. Lancet Digit. Health 3, e526–e533.

Gal, U., Jensen, T.B., Stein, M.-K., 2020. Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. Inf. Organ. 30, 100301.

KUMARI, S., SINGH, R., 2019. E-Mobility Through RRI to Achieve Social Sustainability: A Case Study of Women Commuters of Delhi, India, in: Proceedings of the STS Conference Graz 2019. Presented at the STS Conference Graz 2019, TU Graz, Graz. https://doi.org/10.3217/978-3-85125-668-0-24

Marjanovic, O., Cecez-Kecmanovic, D., Vidgen, R., 2022. Theorising algorithmic justice. Eur. J. Inf. Syst. 31, 269–287.

Missimer, M., Mesquita, P.L., 2022. Social Sustainability in Business Organizations: A Research Agenda. Sustainability 14, 2608.

Mogaji, E., Soetan, T.O., Kieu, T.A., 2020. The implications of artificial intelligence on the digital marketing of financial services to vulnerable customers. Australas. Mark. J. j-ausmj.

Mukerjee, N., 2019. Investigating women’s safety in New Delhi’s urban transport systems. Int. J. Soc. Sci. Econ. Res. 4, 3876–3886.

Organization, W.H., 2021. Ethics and governance of artificial intelligence for health: WHO guidance.

Qureshi, S., 2020. Why data matters for development? Exploring data justice, micro-entrepreneurship, mobile money and financial inclusion. Inf. Technol. Dev.

Steen, M., Timan, T., van de Poel, I., 2021. Responsible innovation, anticipation and responsiveness: case studies of algorithms in decision support in justice and security, and an exploration of potential, unintended, undesirable, higher-order effects. AI Ethics 1, 501–515.

Stilgoe, J., Owen, R., Macnaghten, P., 2013. Developing a framework for responsible innovation. Res. Policy 42, 1568–1580.

Ugwudike, P., 2022. Predictive algorithms in justice systems and the limits of tech-reformism. Int. J. Crime Justice Soc. Democr. 11, 85–99.