Replicating Gender Bias from Above: Earth Observation, Machine Learning and SDG 5

Anthony Deen, Mateo Rojas Guerrero, Juan Enrique Bonilla Morales, Zinnya del Villar (Data-Pop Alliance)

Note: This report was partially based on a report of the ethical implications of ML4EO conducted for GIZ FAIR Forward by Data-Pop Alliance.

Abstract

With the increased application of Earth Observation (EO) data, combined with the higher spatial and temporal granularity of data and the potential benefits of Machine Learning (ML) algorithms, ethical concerns around EO have become more salient. In this science-policy brief, we examine the effects of EO data collected without a gender-inclusive perspective, and how that raw data (or processed by ML) may have negative implications for women. We will focus on three groups of women living in various states of vulnerability: those experiencing forced migration, those living in refugee camps, and smallholder farmers. The brief offers recommendations to relevant stakeholders –such as EO practitioners and NGOs utilizing this information– on possible mitigation strategies to ensure that the collection and use of EO data is done in a gender-inclusive manner, with the goal of designing more comprehensive and equitable interventions based on EO data.

Introduction & Context

The expansion and increased capabilities of EO technology has vast potential to help us better understand the world we live in and to advance progress towards achievement of the Sustainable Development Goals (SDGs). This includes both in the natural world (e.g. geographical mapping, natural disaster mitigation), as well as in the humanitarian sphere, where such knowledge can help organizations or governments designing humanitarian interventions (e.g. access to electricity, addressing food insecurity) to gain a fuller picture of the situation "on the ground" through satellite imagery.[i] Additionally, as more EO data is processed by ML algorithms, the increase in quality and scope of available data is also a cause for optimism.[ii]

Considering these possibilities, questions have also arisen around fairness and equity at all levels of the EO and ML data ecosystem, including who has access to these technologies, as well as around the collection, usage, and storage of the data. Ethical questions have also been raised regarding 'mapped' populations and their level of consent and participation in interventions designed on their behalf.[iii] While all of these are important concerns, this brief will focus on one aspect and its implications for EO (and ML-enhanced) data: gender-inclusivity. As ML has already been shown to replicate biases (or omissions) from the data it " learns" from[iv], the inclusion of gender-conscious data collection becomes even more pressing as these technologies are increasingly linked. It is important to note that while "gender" is a broad term encompassing a wide variety of identities, for the purposes of this brief we will use the term from the binary perspective of man and woman, considering the negative consequences for women.

SDG 5 calls for gender equality, including (but not limited to) an end to discrimination based on gender, freedom from gender-based violence, and equal access to technology.[v] These subcategories are addressed in this brief, with an explanation of how current EO practices and research may be failing to meet the standards of equality set forth by the SDGs, the consequences that can result, and recommendations to ensure that these frontier technologies are applied in a manner that is most inclusive and beneficial to all members of mapped populations.

Machine Learning and Earth Observation for SDG5 - Obstacles to Gender Equality

Bibliometric analysis[vi] is the quantitative study of bibliographic data and is used to identify relevant research trends related to a specific sector or topic. According to the analysis undertaken for this brief, which utilized R software to search the SCOPUS database[vii] , there is a lack of research related to our topic of concern. The analysis undercovered only 221 research articles mentioning "EO and SDGs" together from 2015-2021, a timeframe chosen to coincide with the publication of the SDGs in 2015 (Figure 1). Additionally, in the conceptual structure map made using author's keywords (which is more specific in terms of topics covered by the papers) the term "gender" does appear at all, indicating that while other SDG measurements are being considered with regards to EO, gender equality is not one of them. (Figure 2).

Figure 1. Annual Scientific Production

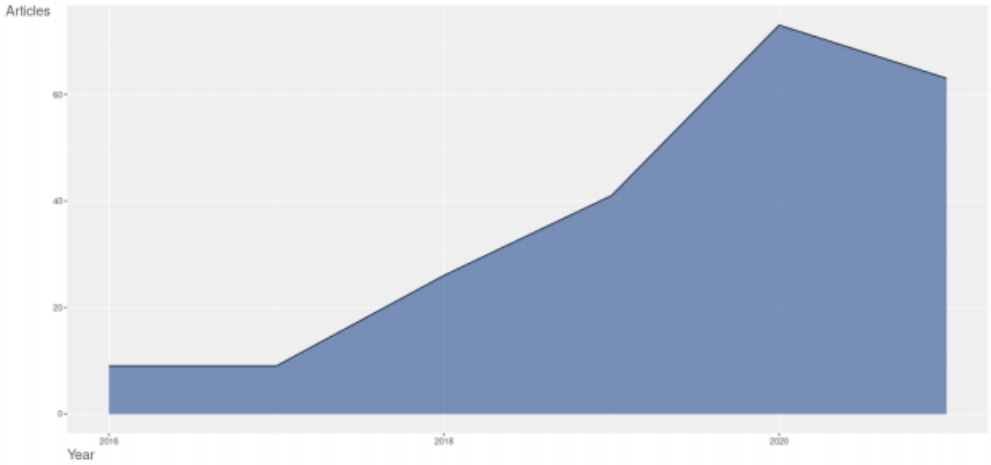
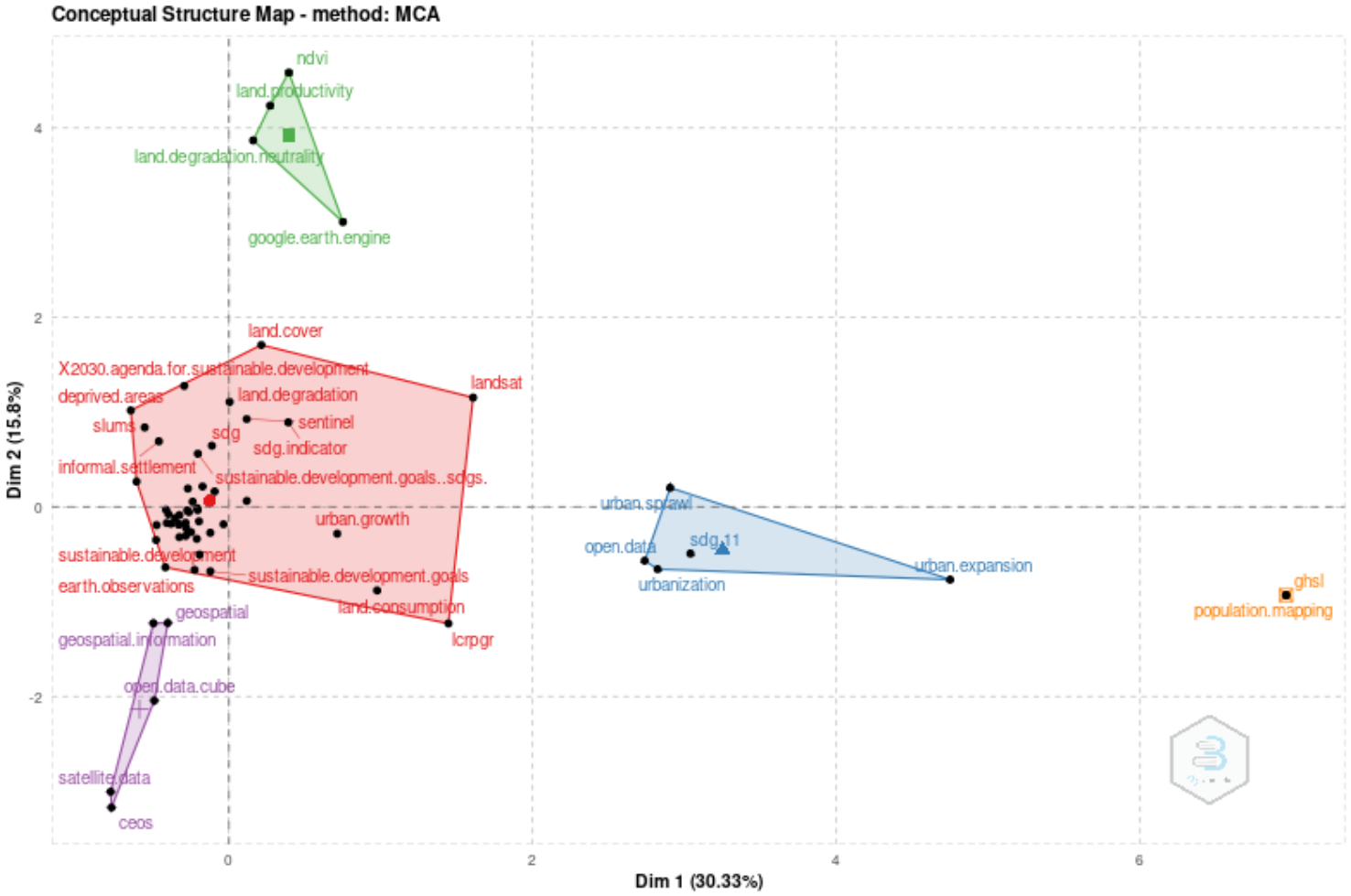


Figure 2. Conceptual Structure Map



As the bibliometric analysis revealed a lack of published research related to this topic, a more targeted review of individual research articles addressing EO and gender confirmed that gender perspectives were often excluded from the EO-related project design or data collection activities.[viii] For the purposes of this brief, "gender-inclusivity" is exemplified by the recognition of gendered implications when analyzing EO data. The most basic step towards this goal is the realization that there can be a "gendered" component to all aspects of

EO data, from mapped populations to inanimate objects. For example, if mapping a refugee camp, the number and spacing of toilet facilities may seem innocuous in relation to gender. However, research pointing to the risk of sexually-based violence (SBV) for women in camps accessing toilets indicates their location (remote) and prevalence have differential impacts based on gender.[1]

Additionally, the incorporation of a gender-inclusive lens from the beginning stages of the project design helps ensure the inclusion of these issues, which is especially pertinent for those designing humanitarian interventions based on EO data.

Use Cases

In order to demonstrate the saliency of this topic, we have identified several use cases based on real-world applications of EO output and the potential harms resulting from the lack of this perceptive. The first involves EO data tracking conflict situations likely to result in forced migration (refugees).[2] While this data is useful to monitor flows and prepare for arrival of these groups, without a gender-perspective the data can miss the heightened vulnerability of women migrants to a variety of abuses and violence.[3] The second involves mapping refugee camps[4] , in which lack of gender consideration can overlook factors that make living conditions (even more) more dangerous for women, such as the placement of toilets or shower facilities far from women's quarters, which increases the risk of assault or violence (especially at night). Finally, women smallholder farmers may experience discrimiation in agricultural insurance coverage and compensation, should firms rely solely on EO data[5] to determine restitution for crop failures due to disaster. Again, without knowing if the crops belong to women farmers, who often have less access to necessary farming equipment, they may incorrectly access damages and resultant compensation. It is important to note that with the application of ML, these omission biases would exclude important variables when constructing ML algorithms.

Subsequently, the biased data could lead to bias replication, as has happened with ML in the past.

Societal and Sector Challenges

The lack of a gender perspective in EO technology did not occur in a vacuum, and an examination of both societal factors and the sector itself provide evidence for how this situation developed. The societal challenges relate mostly to the skewed gender ratios in all areas of STEM, which are even more pronounced within the EO (and AI) communities.[6] Without the presence of women EO practitioners, it is not surprising that certain gendered aspects have been overlooked. However, these challenges also extend to the lack of clear ethical guidelines guiding the collection of EO data and, subsequently, ethical issues such as gender-mainstreaming being excluded from the design stage of projects. Taken together with a dearth of research related to ethical implications of EO, a clear picture emerges of challenges from outside of the sector that are nonetheless impacting the work being done within it.

Challenges to gender-inclusivity are also evident at the sector-level. Data collection (perhaps for the aforementioned reasons) is often undertaken without a view towards ethical implications or preventing biased output.[7] On the one hand, it is understandable that EO practitioners would view their output as agnostic or neutral, as they are literally images of what is happening on the ground. However, without the application of a gender-inclusive lens, this thinking can (and has) caused those using the data to miss certain "ground truths" about surveilled populations, such as the gender composition and differential needs of each gender within these groups.[8] The risk is that the EO community, in addition to development actors who design interventions based on their data, view EO images as the "whole truth", rather than a piece of the puzzle that must be complemented with additional data about mapped populations, in order to best meet their needs.

Public Policy Recommendations

Based on the information presented in this policy brief, recommendations on how to better incorporate gender-inclusive practices (and progress towards SDG 5) into EO collection and data usage are presented below.

EO Practitioners and NGOs

- Develop organizational frameworks to ensure that EO data is collected in an ethical manner, with specific consideration given to gender-inclusivity.
- Gender-inclusive considerations should be explicitly addressed in the development phase of new EO projects, and monitored throughout implementation. Organizations utilizing EO data

[1] Aubone, A., & Hernandez, J. (2013).
[2] Witmer, F. D. (2015).
[3] Kofman, E. (2019).
[4] Wickert, L. & Bogen, M. & Richter, M.. (2020).
[5] de Leeuw, J., et al (2014).
[6] Houser, K. (2018).
[7] Berman, G., de La Rosa S., and Accone, T. (2018).
[8] SERVIR-Mekong (2015).

should also ensure that this was done, and if not, develop their own strategies to make the data more reflective of gendered realities on the ground.

- If feasible, include consultation with gender experts to more fully understand the ever-evolving landscape of gender-inclusive practices and intersectionality. Smaller organizations (with less resources) could also gain knowledge through review of gender-inclusive literature before the implementation of new projects and as a best practice.
- Whenever possible, involve mapped populations in the development of humanitarian interventions using EO data. This is a best practice approach for all interventions, but especially so for EO-based work, for the reasons outlined in this study.
- Increased Diversity, Equity and Inclusion (DEI) practices within organizations to help counteract the current skewed gender ratio in the field of EO (and AI).

References

Anderson, K., Ryan, B., Sonntag, W., Kavvada, A. & Friedl, L. (2017) Earth observation in service of the 2030 Agenda for Sustainable Development, Geo-spatial Information Science, 20:2, 77-96, DOI: 10.1080/10095020.2017.1333230

Aubone, A., & Hernandez, J. (2013). Assessing refugee camp characteristics and the occurrence of sexual violence: a preliminary analysis of the Dadaab complex. Refugee Survey Quarterly, 32(4), 22-40.

Bajracharya, B., Irwin, D. E., Thapa, R. B., & Matin, M. A. (2021). Earth Observation Applications in the Hindu Kush Himalaya Region — Evolution and Adoptions. Earth Observation Science and Applications for Risk Reduction and Enhanced Resilience in Hindu Kush Himalaya Region: A Decade of Experience from SERVIR (pp. 1–22). Springer International Publishing. https://doi.org/10.1007/978-3-030-73569-2_1

Berman, G., de La Rosa S., and Accone, T. (2018) Ethical Considerations when Using Geospatial Technologies for Evidence Generation, Innocenti Discussion Paper, UNICEF Office of Research - Innocenti, Florence.

Derviş , H. (2019). Bibliometric analysis using Bibliometrix and R Package. Journal of Scientometric Research, 8(3), 156-160.

de Leeuw, J., Vrieling, A., Shee, A., Atzberger, C., Hadgu, K. M., Biradar, C. M., Keah, H., & Turvey, C. (2014). The potential and uptake of remote sensing in insurance : a review. Remote sensing, 6(11), 10888-10912. http://www.mdpi.com/2072-4292/6/11/10888

Goal 5 | Department of Economic and Social Affairs (un.org)

Hagen, E. (2020). The Ethical Mapping Guidelines: how not to map. EthicalGeo.

https://ethicalgeo.org/the-ethical-mapping-guidelines-how-not-to-map/

Houser, K. (2018). Report Finds Huge Gender Gap in Artificial Intelligence Workforce. Futurism. https://futurism.com/ai-gender-gap-artificial-intelligence

Kofman, E. (2019). Gendered mobilities and vulnerabilities: refugee journeys to and in Europe. Journal of Ethnic and Migration Studies, 45(12), 2185-2199.

Lary, D. J., Zewdie, G. K., Liu, X., Wu, D., Levetin, E., Allee, R. J., ... & Aurin, D. (2018). Machine learning applications for earth observation. Earth observation open science and innovation, 165.

Stephen R.J. Sheppard, Petr Cizek (2009),The ethics of Google Earth: Crossing thresholds from spatial data to landscape visualisation, Journal of Environmental Management, Volume 90, Issue 6,Pages 2102-2117,ISSN 0301-4797,https://doi.org/10.1016/j.jenvman.2007.09.012.

SERVIR-Mekong. 2015. Gender and GIS: Guidance Notes. Asian Disaster Preparedness Center: Bangkok, Thailand.

Wachter, S., Mittelstadt, B., & Russell, C. (2020). Bias preservation in machine learning: the legality of fairness metrics under EU non-discrimination law. W. Va. L. Rev., 123, 735.

Wickert, L. & Bogen, M. & Richter, M.. (2020). SUPPORTING THE MANAGEMENT OF HUMANITARIAN OPERATIONS CONCERNING MIGRATION MOVEMENTS WITH REMOTE SENSING. ISPRS -International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XLIII-B3-2020. 233-240. 10.5194/isprs-archives-XLIII-B3-2020-233-2020.

Witmer, F. D. (2015). Remote sensing of violent conflict: eyes from above. International Journal of Remote Sensing, 36(9), 2326-2352

[i] Anderson, K., et al (2017).
[ii] Lary, D. J., et al (2018).
[iii] Hagen, E. (2020).
[iv] Wachter, S., Mittelstadt, B., & Russell, C. (2020).
[v] Goal 5 | Department of Economic and Social Affairs (un.org)
[vi] Derviş , H. (2019).
[vii] Stephen R.J. Sheppard, Petr Cizek (2009).
[viii] Bajracharya, B., Irwin, D. E., Thapa, R. B., & Matin, M. A. (2021).