Socially Responsible Perspectives on Data, Algorithms and Digitalization – Why Gender matters?

2022 DEWIS Symposium

Univ.-Prof. Dr. Carmen Leicht-Scholten

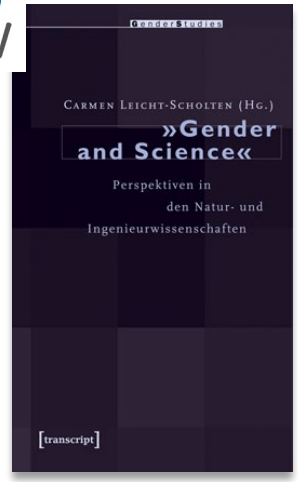
GDI Gender und Diversity in den Ingenieurwissen-schaften | RWTH AACHEN UNIVERSITY

Bridging Professorship Engineering & Humanities

VDI

RRI Hub
Responsible Research and Innovation Hub
Ein Projekt der RWTH

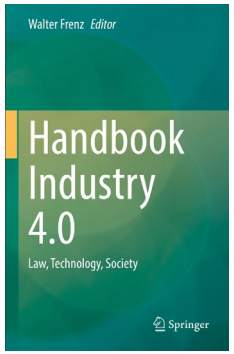
PROMOTIONS-KOLLEG NRW



RWTH FAKULTÄT für bauingenieurwesen
Vice Dean in Charge of Studies

Netzwerk Frauen- und Geschlechterforschung NRW

STIFTUNG Akkreditierungsrat

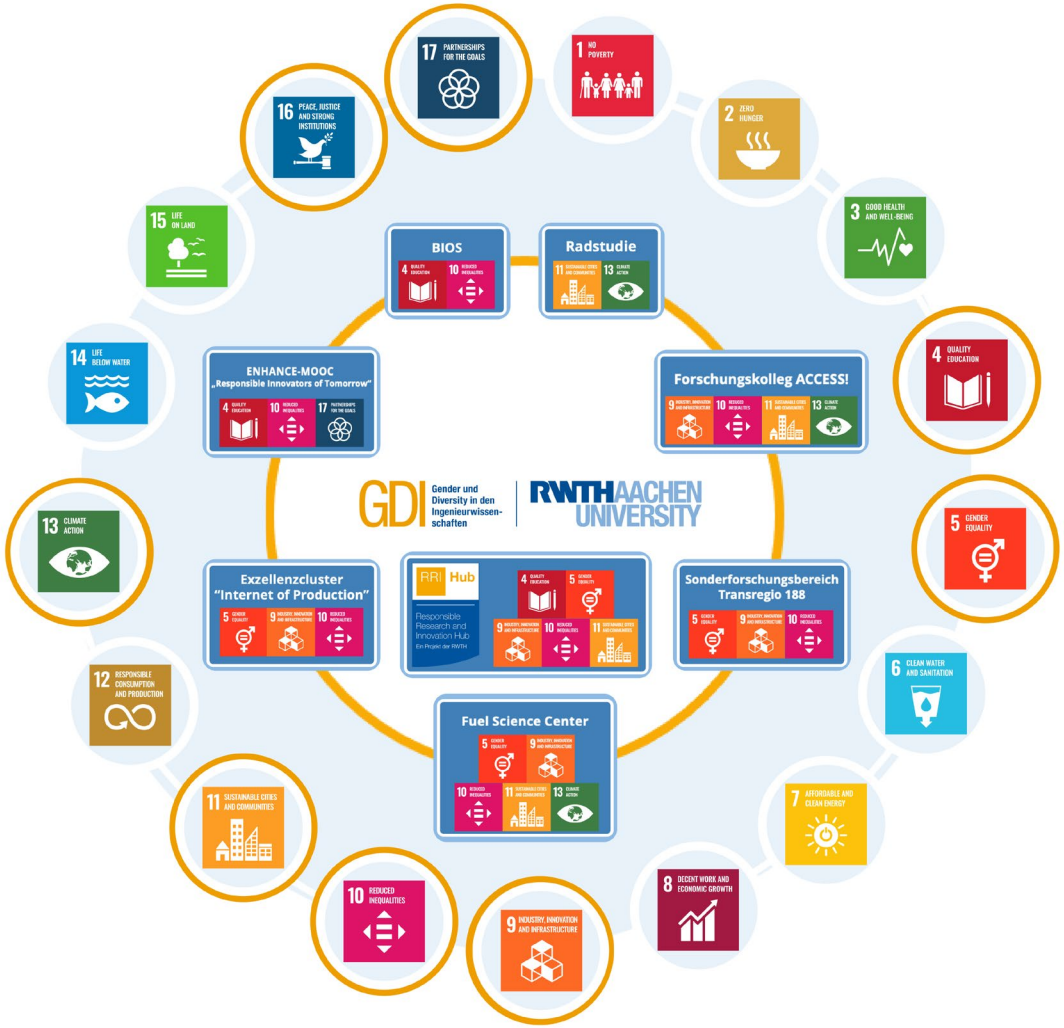


humtec | RWTH AACHEN UNIVERSITY

cdio™

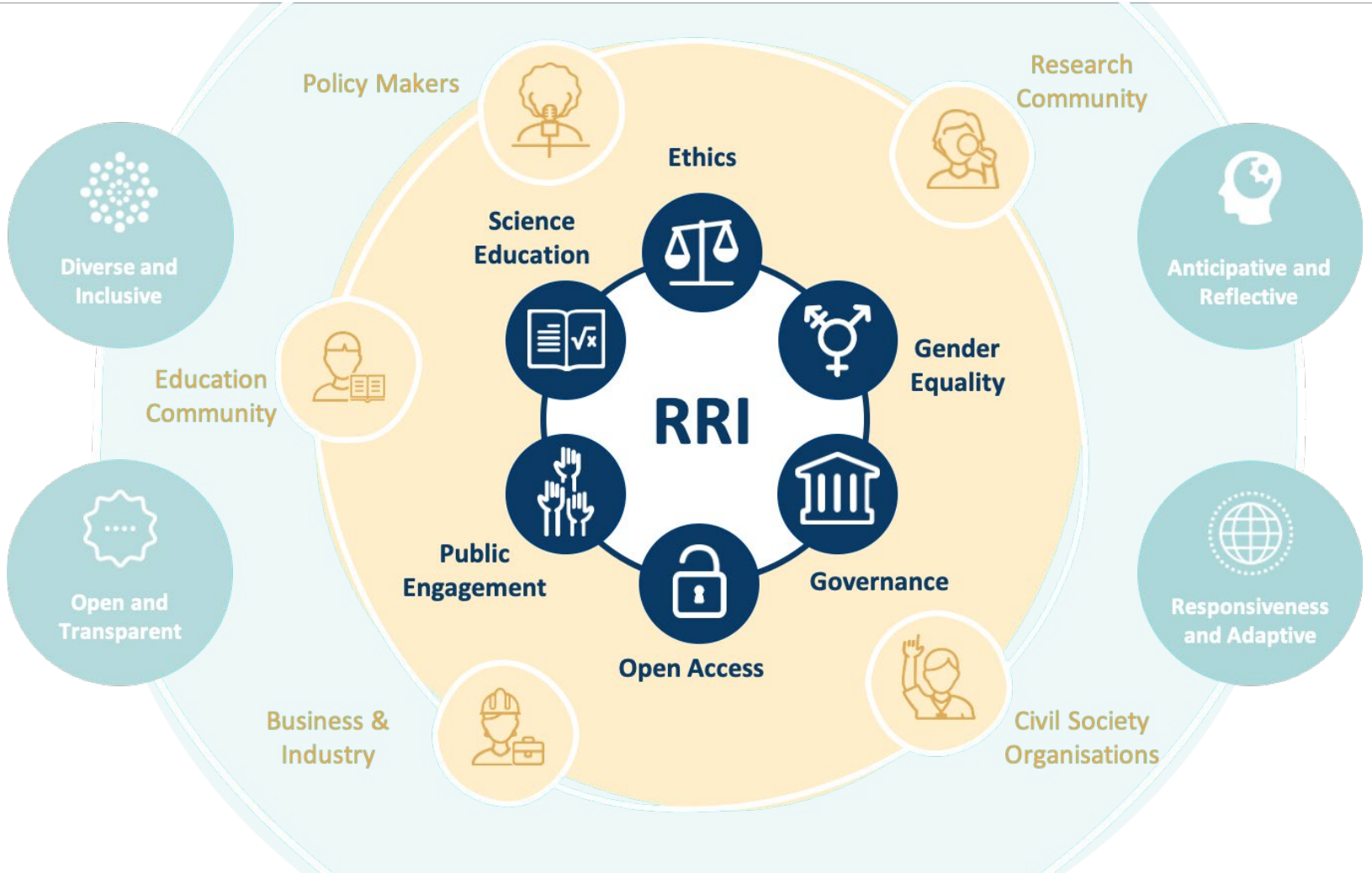
GDI | RWTH AACHEN UNIVERSITY

Fields of research



- Cultures of Science and Technology
- Participation and Change
- Engineering Education
- Diversity and Innovation

Responsible Research and Innovation



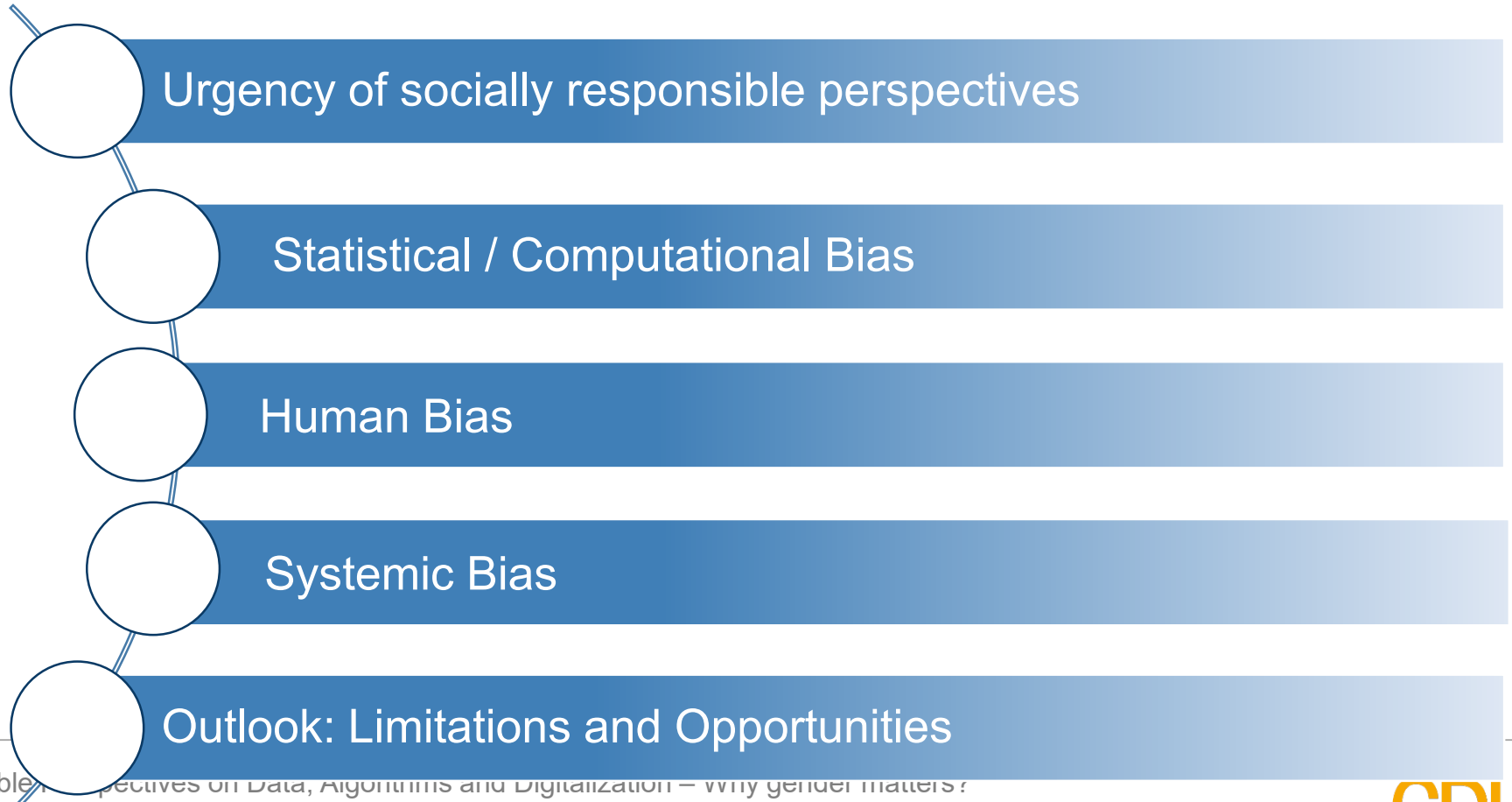
POLICY AGENDAS / KEY ISSUES

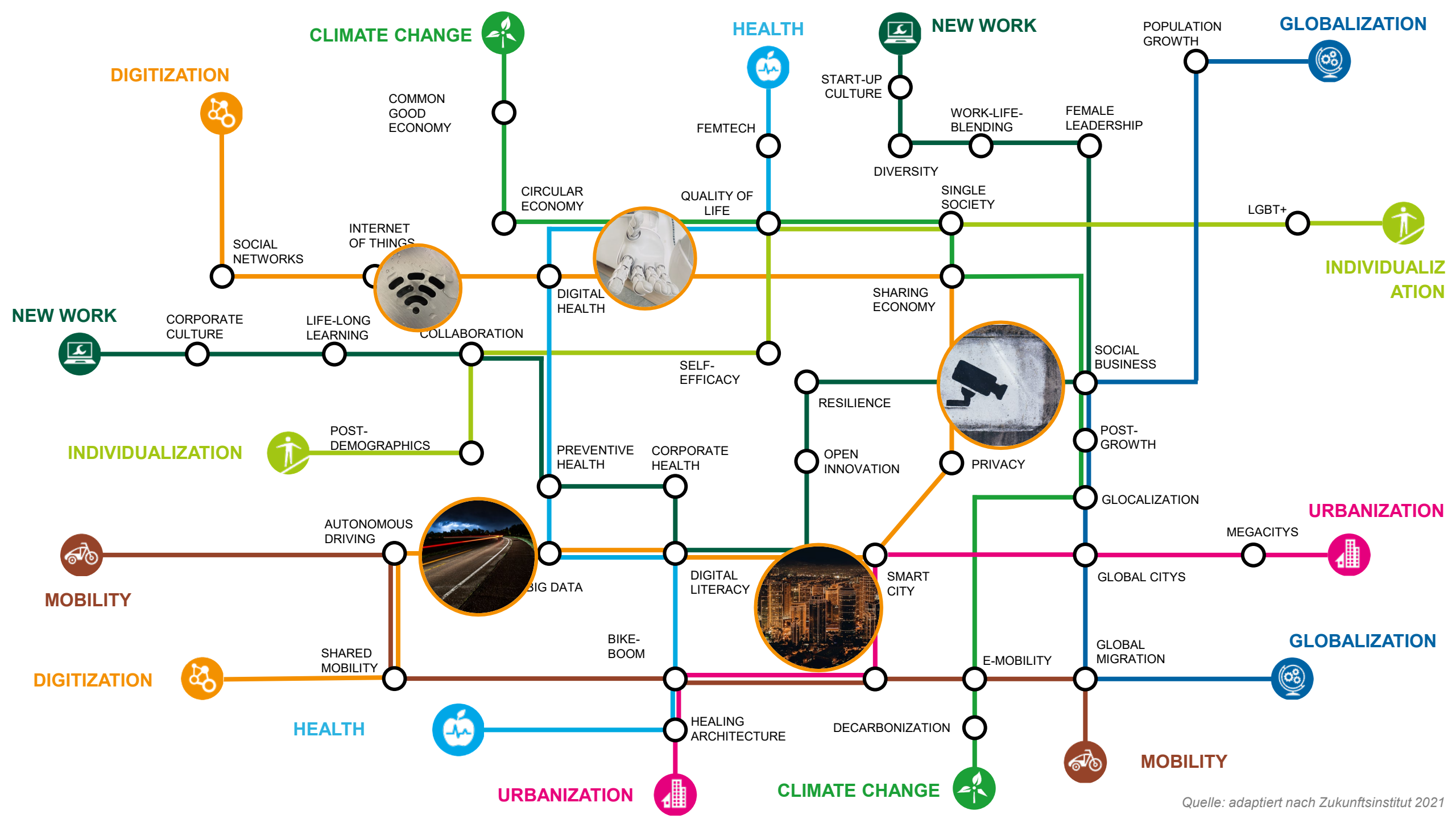
ACTORS

PROCESS DIMENSIONS

Quelle: Abgeleitetes Werk der Originalgrafik von www.rri-tools.eu CC BY-NC-SA 4.0

Content


Urgency of socially responsible perspectives
Statistical / Computational Bias
Human Bias
Systemic Bias
Outlook: Limitations and Opportunities



Quelle: adaptiert nach Zukunftsinstitut 2021

Diversity Attributes and Discimination in AI

Amazon Prime Free Same-Day Delivery – excluded minority neighborhoods

Racial Bias in predictive policing

Increased political polarization in social media

HR algorithm: hobby lacrosse became indicator of calculated success

Gender Bias in Online Ad Systems

Autocomplete on Google or Bing

Source: GDI 2021

Source: according to https://www.gardenswartzrowe.com/why-g-r

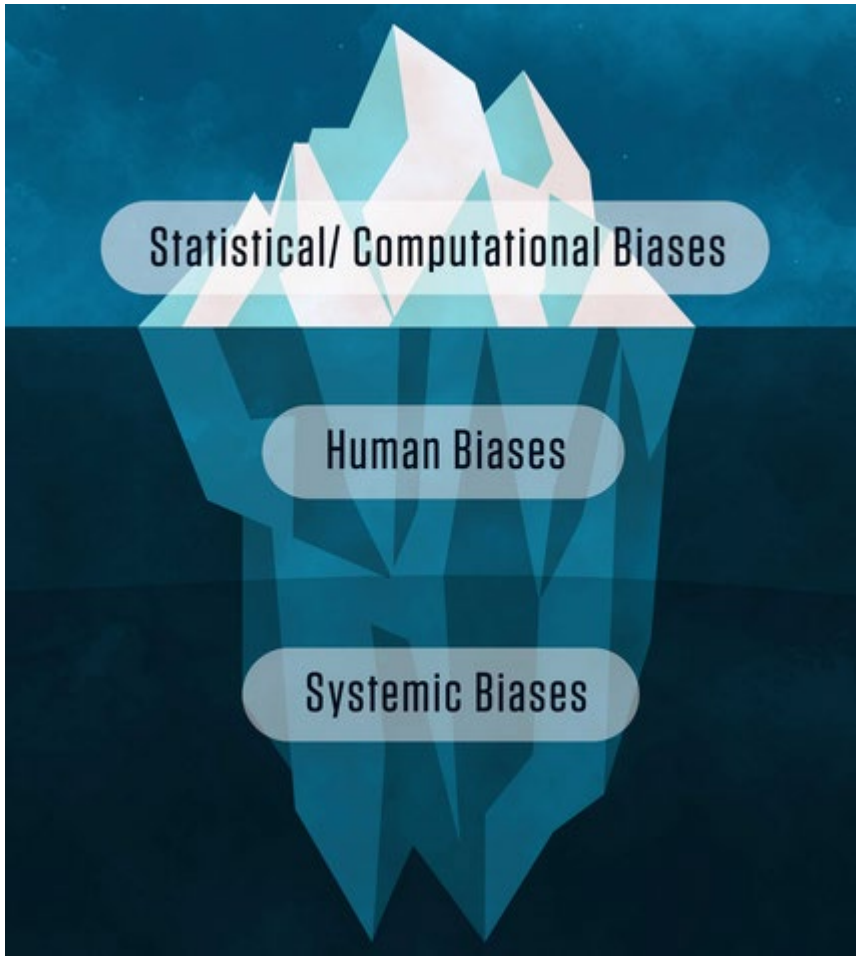
GDI | RWTH AACHEN UNIVERSITY

Causes of Algorithmic Bias

„**who** codes matters, **how** we code matters, **why** we code matters“

– Joy Buolamwini

Source: https://www.youtube.com/watch?v=UG_X_7g63rY



Source: https://www.nist.gov/news-events/news/2022/03/theres-more-ai-bias-biased-data-nist-report-highlights

Systemic Bias - Gendered Perspectives

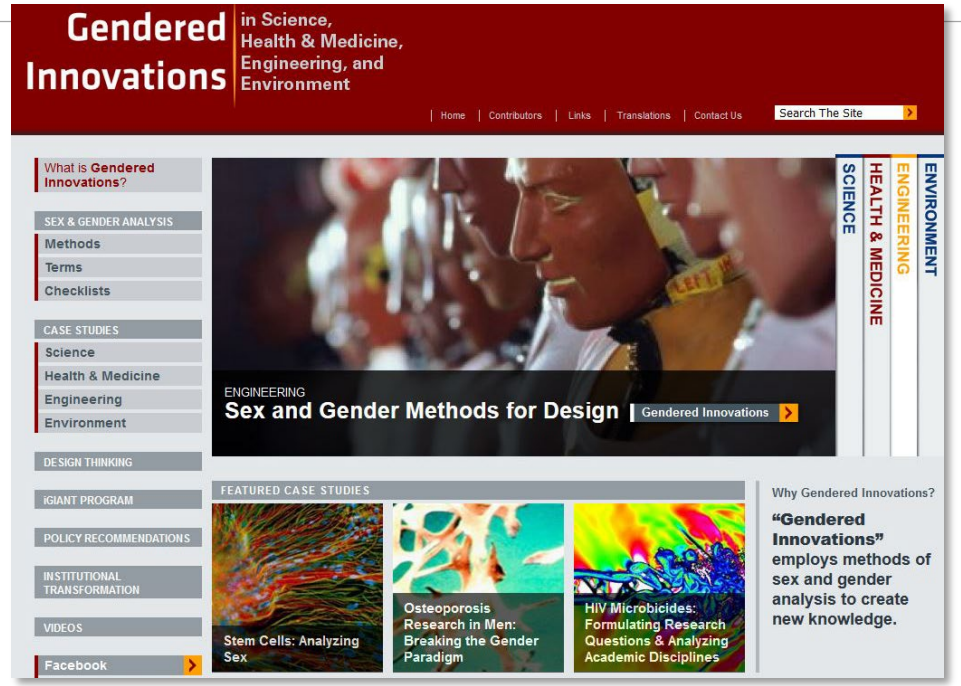
Situated knowledge

I. Knowledge is always partial

II. Knowledge is always involved in discourses of power

III. Reflection and responsibility

Donna Haraway



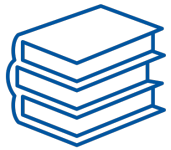
I. Fixing the Numbers

II. Fixing the Institution

III. Fixing the Knowledge

Quelle: https://genderedinnovations.stanford.edu/

Statistical/ Computational Bias



Observations or historical Data

Training Data may be biased
- Historical Bias, *e.g. Issues of Representativity*
- Measurement Bias, *e.g. How to infrastructure demands*

Detection of Regularities

Induction - inferring the future from the past
- Learning from unfair patterns in the past leads to repeating those unfair decisions in the future.

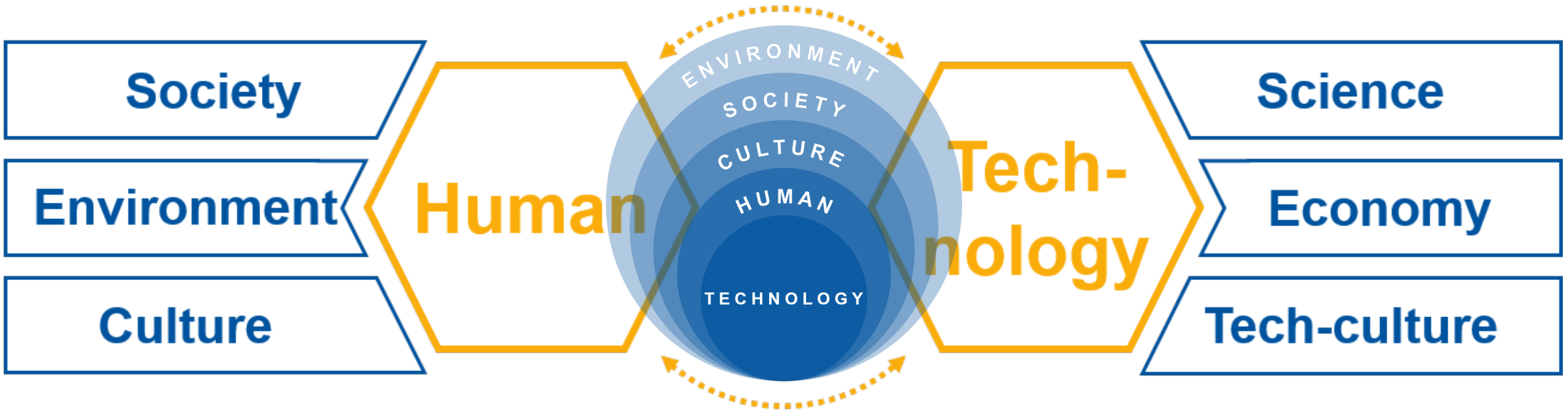
Prediction of Future Events

Predictions may dis-/ favor certain groups
Difficulty to quantify fairness: lack of consensus how to compute fairness, contradicting fairness metrics

→ Need for interdisciplinary approaches

→ Education in the field of ethics and gender/ and diversity, and intersectionality

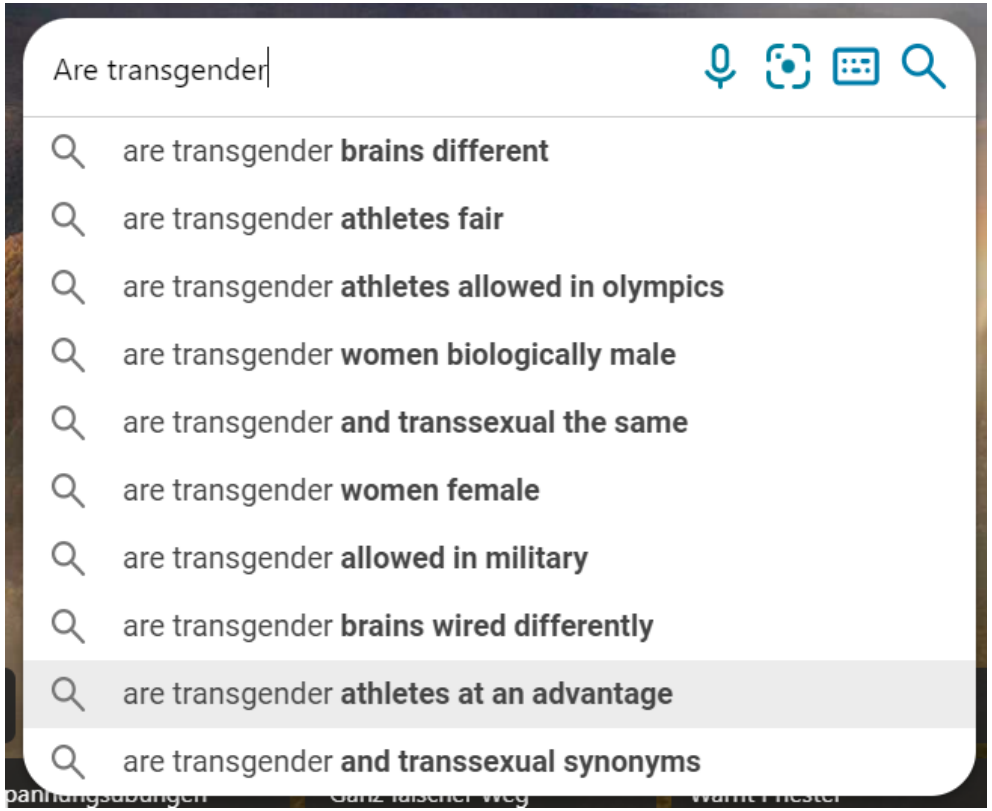
Status Quo



Source: GDI

Human Bias – Interaction

Google Search, 24th October 2022



Bing Search, 24th October 2022

Human Biases

Gender Marketing: Stereotypes may not match the Need of the Customers

Source: http://www.jeongmeeyoon.com/aw_pinkblue.htm

"For example, video game designers designed a game platform for girls in pink because that is what the parents (who purchase the game) perceived their girls wanted. The girls themselves preferred darker metallic colors (Rommes, 2006)."

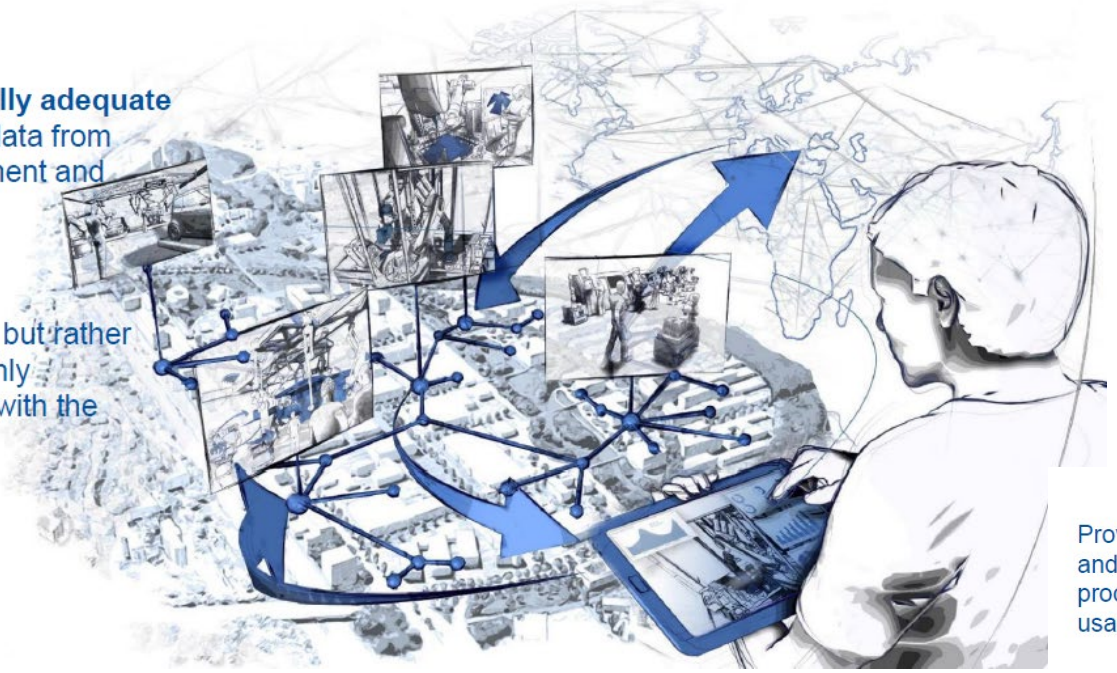
https://genderedinnovations.stanford.edu/terms/stereotypes.html

Cluster of Excellence - Internet of Production

Internet of Production
Cluster of Excellence EXC 2023

Providing **semantically adequate** and **context aware** data from production, development and usage in industry…

… not only one-time, but rather continuously and highly iterative in **real time** with the **adequate level of granularity**…



Providing **semantically adequate** and **context aware** data from production, development and usage in industry…

… not only one-time, but rather continuously and highly iterative in **real time** with the **adequate level of granularity**…

… for a new level of **cross-domain collaboration**.

Future Research: Living Lab on Gender-Responsive and Fair Algorithms

Development of interdisciplinary approaches for a discrimination-sensitive and gender-responsive use of algorithms through…

- … monitoring the impact of detailed explanations on the perception of fairness in AI,
- … evaluating methods of explanation regarding their suitability to make discrimination detectable for laypersons,
- … developing innovative strategies to integrate gender and other diversity dimensions into explainability tools.

Outlook: AI for social good (AI4SG)

> "**AI4SG** facilitates the attainment of **socially good outcomes** that were **previously unfeasible, unaffordable** or **simply less achievable** in terms of efficiency and effectiveness."
>
> – Cowls et al. (2021)

SDG 13 **Climate Action**

Technological improvement through AI

Enables …

- Smart Cities – e.g. by autonomous electrical vehicles
- Circular economies and efficient use of resources
- Demand response in the renewable energy sector

Vinuesa et al. (2020)

Outlook: AI for social good (AI4SG)

"**AI4SG** facilitates the attainment of **socially good outcomes** that were **previously unfeasible, unaffordable** or **simply less achievable** in terms of efficiency and effectiveness."

– Cowls et al. (2021)

SDG 1 **Reduce poverty**

Predict poverty by satellite imagery and machine learning

- Accurate and scalable method with low costs
- Track and target poverty
- Possible indicator for SDG 1 (new measurement)
- **Identify problems**
- **Create Actionable knowledge**

Jean et al. (2016)

Thank you for your attention!

I am looking forward to the discussion!

Univ.-Prof. Dr. Carmen Leicht-Scholten

GDI Gender and Diversity in the Engineering Sciences
Faculty of Civil Engineering
Director RRI Hub

RWTH Aachen University
Templergraben 55
52056 Aachen

www.gdi.rwth-aachen.de
www.rwth-aachen.de

GDI | RWTH AACHEN UNIVERSITY

Literature

Crawford, “The Hidden Biases in Big Data,” Harvard Business Review 1 (2013).

Cowls, J., Tsamados, A., Taddeo, M., & Floridi, L. (2021). A definition, benchmark and database of AI for social good initiatives. Nature Machine Intelligence, 3(2), 111-115.

Datta, Amit; Tschantz, Michael Carl; Datta, Anupam 2015. Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination. In Proc. Of PETS.

Haraway, Donna 1988: Situated Knowlege. The Science Question in Feminism and the Privilege of Partial Perspectives.

https://genderedinnovations.stanford.edu/terms/stereotypes.html

https://www.computerworld.com/article/3068622/amazon-prime-and-the-racist-algorithms.html

https://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html

Huszár, Ferenc; Ktena, Sofia Ira; O'Brien, Conor; Belli, Luca; Schlaikjer, Andrew; Hardt, Moritz (2022): Algorithmic amplification of politics on Twitter. In: *Proceedings of the National Academy of Sciences of the United States of America* 119 (1). DOI: 10.1073/pnas.2025334119

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. Science, 353(6301), 790-794.

Lambrecht, A. and Tucker, C. E. “Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads,” SSRN Journal, 2016. [Online]. Available: http://www.ssrn.com/abstract=2852260

Owen, R.; von Schomberg, R.; Macnaghten, P. (2021): An unfinished journey? Reflections on a decade of responsible research and innovation. Journal of Responsible Innovation, 8:2, 217-233. https://doi.org/10.1080/23299460.2021.1948789

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. Nature communications, 11(1), 1-10.

Yam, J., & Skorburg, J. A. (2021). From human resources to human rights: Impact assessments for hiring algorithms. Ethics and Information Technology, 23(4), 611–623. https://doi.org/10.1007/s10676-021-09599-7