



**“ETHICAL DATA MANAGEMENT  
FOR SMES”**

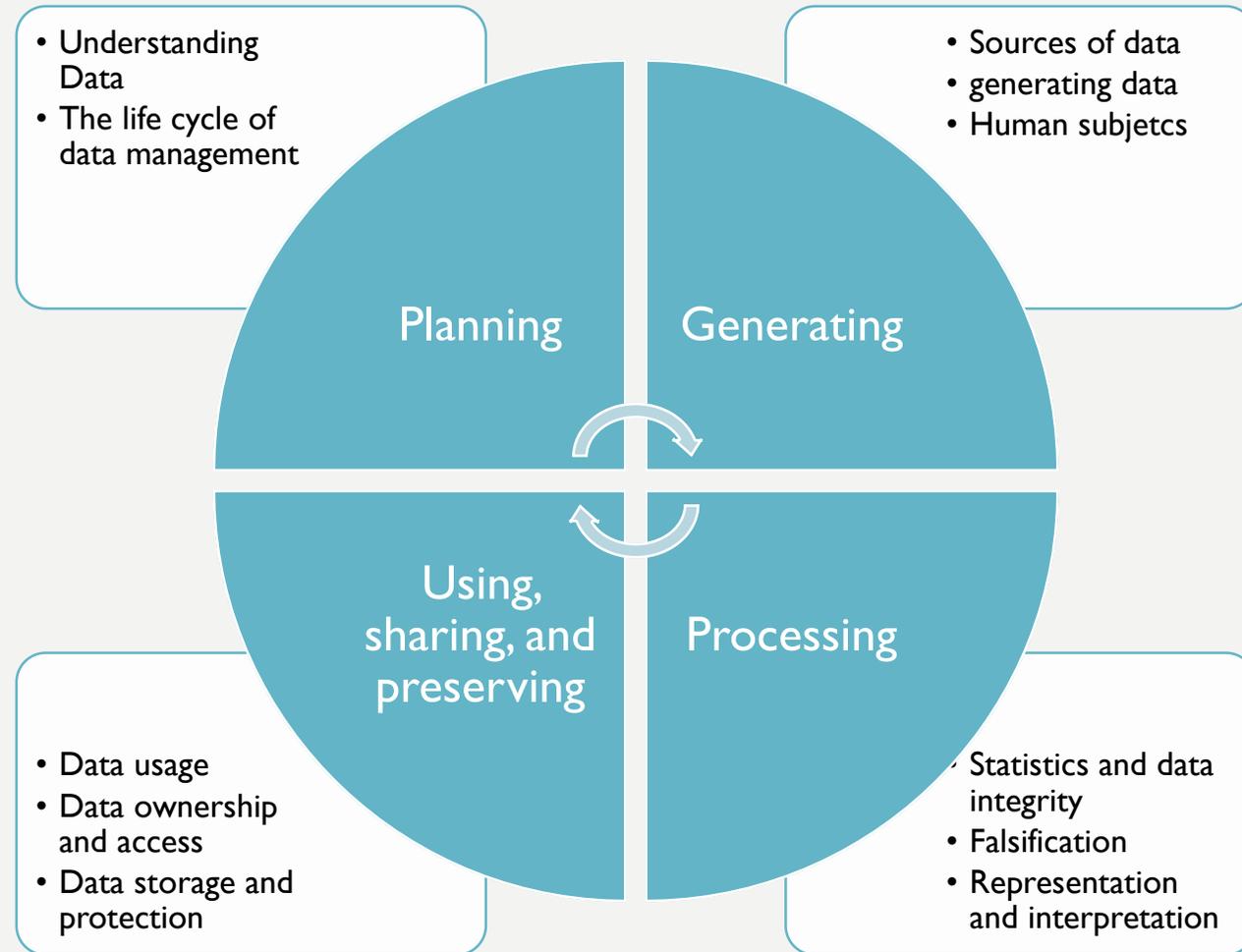
**KATHARINA MILLER - “SMEDATA II”**

**17 JUNE 2022**

# ETHICS OF DATA MANAGEMENT

- “Ecology of Data Management”
- Ethics in data management is not a schematic process;
- It is rather an ongoing process, a “life cycle”, in which SME’s are invited to question their activities from the very first moment on, when they start thinking and drafting their data processes.

# LIFE CYCLE OF DATA MANAGEMENT



# A HUMAN RIGHTS-BASED APPROACH TO DATA

LEAVING NO ONE BEHIND IN THE 2030  
AGENDA FOR SUSTAINABLE DEVELOPMENT



# HUMAN RIGHTS-BASED APPROACH TO DATA 1/5

- The **first principle is participation**. According to HRBAD, “participation of relevant population groups in data collection exercises, including planning, data collection, dissemination and analysis of data” is crucial.
- In the “Life cycle of data management” of any SME this principle affects the processes planning and generating. SMEs have to understand data and research even though some population groups may seem distant to them - such as for example marginalized groups - and they must include human subjects and generate data by relevant population groups.
- According to HRBAD, the participation can be achieved by:
  - Considering a range of processes that facilitate and encourage participation of relevant population groups.
  - Clearly communicating on the participatory processes and the outcomes of the exchanges to human subjects.
  - Including views of vulnerable or marginalized groups, and groups who are at risk of discrimination.

# HUMAN RIGHTS-BASED APPROACH TO DATA 2/5

- The **second HRBAD principle is data disaggregation**. Disaggregated data is data that has been broken down by categories, for example by vulnerable or marginalized groups, sex or gender, ethnical or racial origin. Disaggregated data can reveal inequalities and deprivations that may not be fully reflected in aggregated data. This principle affects the life cycle's processes planning, generating and processing.
- SMEs would have to integrate different categories in their data framework (planning process), they have to generate the different categories of data and do a correct interpretation to reveal to possible inequalities and deprivations.
- According to HRBAD, this can be adequately realized by:
  - Generating more detailed data than national averages. This is key in identifying and understanding inequalities.
  - Disaggregation of data according to characteristics identified in international human rights law.
  - Establishing alternate sampling and data collection approaches.
  - Recommended tool: UN Women (2021), Counted and Visible: Toolkit to better utilize existing data from household surveys to generate disaggregated gender statistics, in [https://unstats.un.org/unsd/statcom/52nd-session/documents/BG-31-Executive\\_Summary\\_Counted\\_and\\_Visible\\_Toolkit-E.pdf](https://unstats.un.org/unsd/statcom/52nd-session/documents/BG-31-Executive_Summary_Counted_and_Visible_Toolkit-E.pdf), page 4 (last access on 17.06.2022).

# HUMAN RIGHTS-BASED APPROACH TO DATA 3/5

- The third **HRBAD principle is self-identification**. According to HRBAD, “(f)or the purposes of data collection, populations of interest should be self-defining. Individuals should have the option to disclose, or withhold, information about their personal characteristics”.
- The principle affects especially the process “generating” within the life cycle. SMEs will have to carefully listen to the human subjects and how they want to define themselves. This should also be foreseen in the “planning” process within the life cycle. According to HRBAD, the principle of self-identification can be respected by:
  - It should be at the individual’s discretion how and which data about their personal characteristics they want to provide.
  - Data collection activities should be conducted in accordance with the human rights principle of ‘doing no harm’.

(Source: United Nations Human Rights Office of the High Commissioner (2018), A human-rights-based approach to data, page 11).

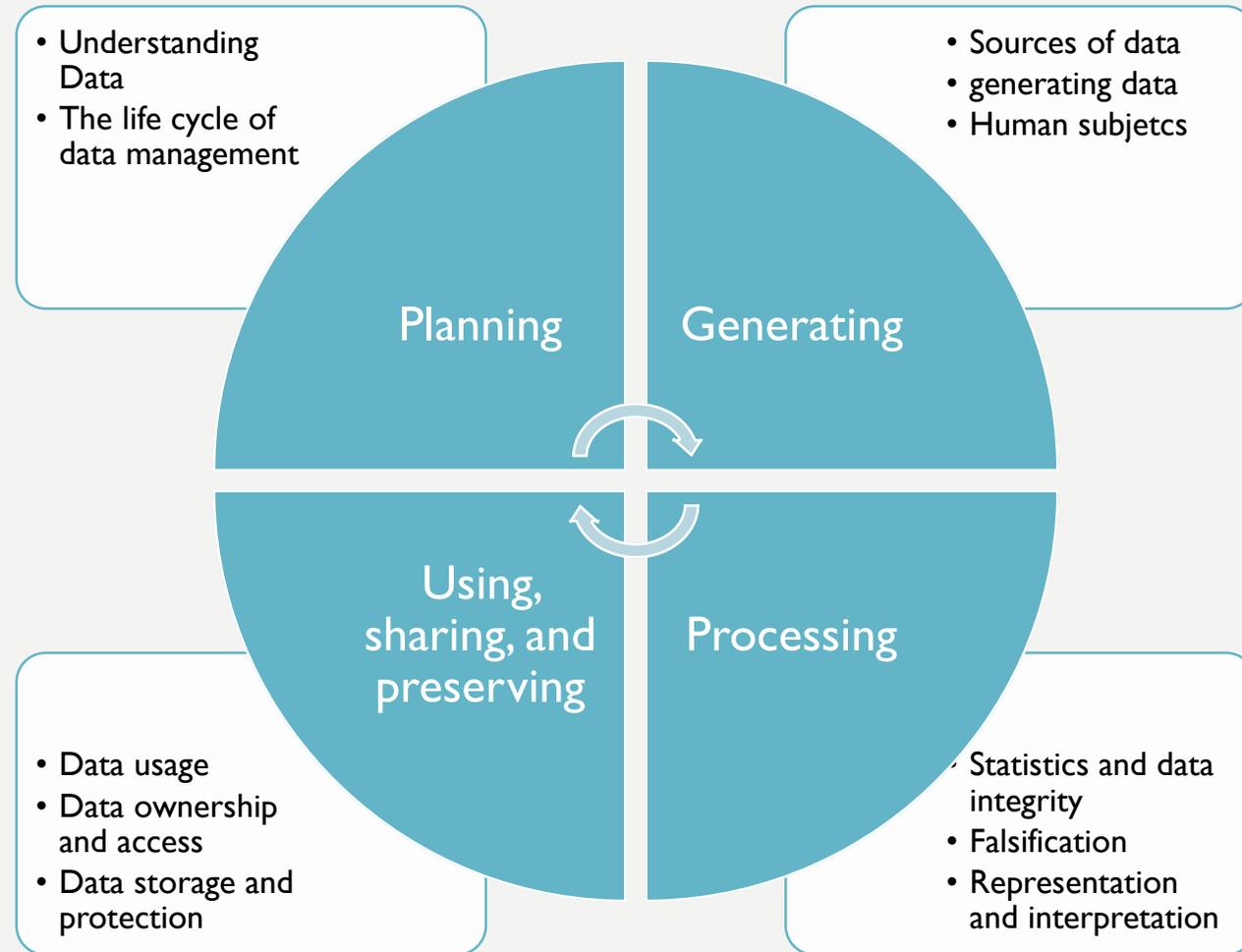
# HUMAN RIGHTS-BASED APPROACH TO DATA 4/5

- Another **HRBAD principle is transparency**. According to HRBAD, “(d)ata collectors should provide clear, openly accessible information about their operations, including research. design and data collection methodology. Data collected by State agencies should be openly accessible to the public”. This principle affects especially the processes processing and using, sharing and preserving within the life cycle.
- SMEs should act with data integrity and don’t falsify any data; they should furthermore make their data usage public and give access to the public. This is also in line with the FAIR principles (Findable, Accessible, Interoperable and Re-usable) for data management.
- According to HRBAD, this principle of transparency can be fulfilled by:
  - Making information about data collection publicly available.
  - Disseminating of data as quickly as possible after its collection.

# HUMAN RIGHTS-BASED APPROACH TO DATA 5/5

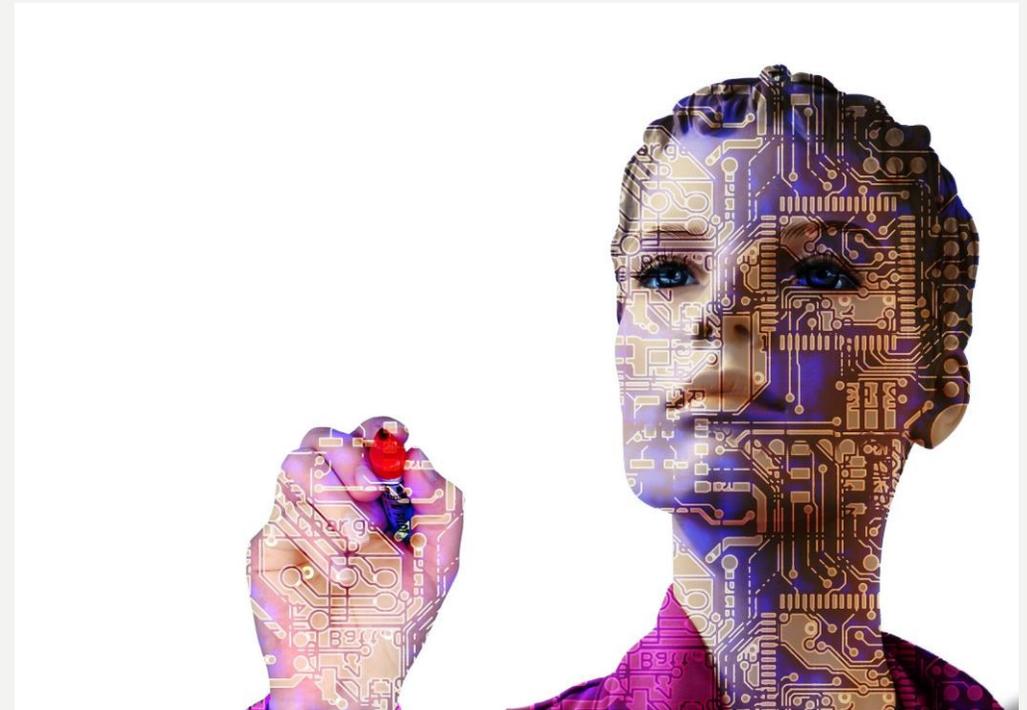
- The last HRBAD principle that we discuss is **accountability**. According to HRBAD, “(d)ata collectors are accountable for upholding human rights in their operations, and data should be used to hold States and other actors to account on human rights issues”.
- This principle affects the whole “Life cycle of data management” of SMEs, because SMEs are accountable for respecting, protecting and fulfilling human rights of the subjects at all times of the four processes.

# LIFE CYCLE OF DATA MANAGEMENT



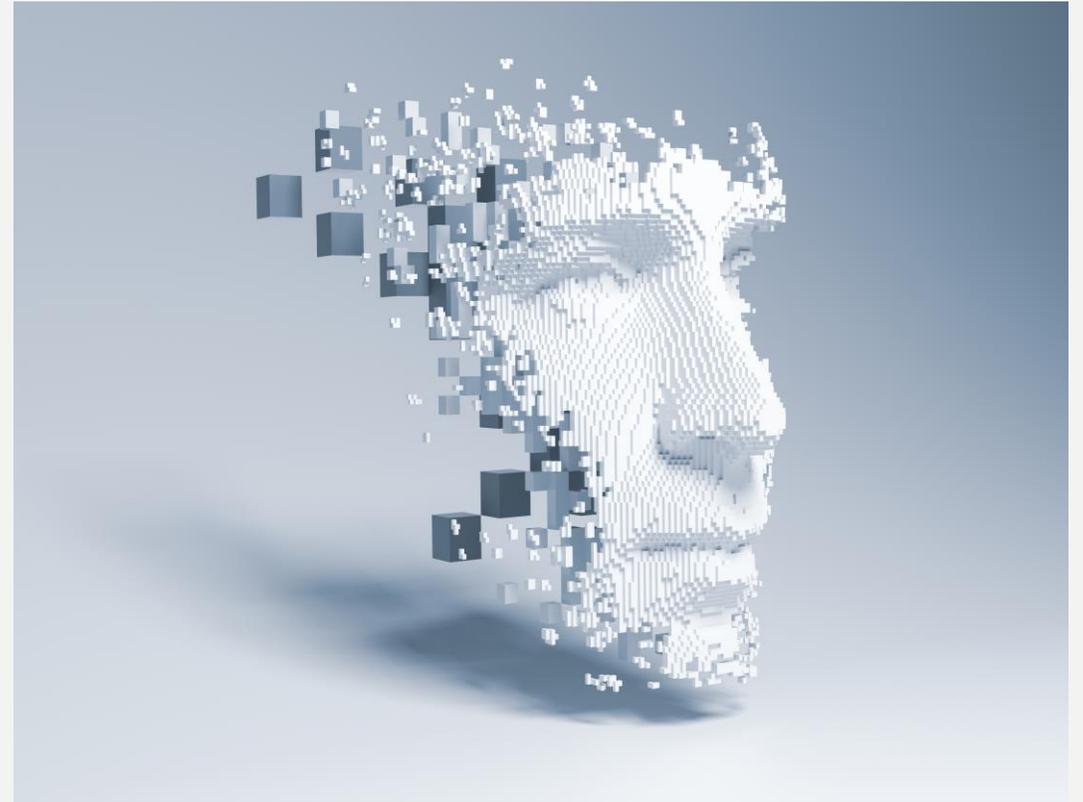
# CASE “AMAZON”

- The team had been building computer programs since 2014 to review job applicants’ resumes with the aim of mechanizing the search for top talent.
- But by 2015, the company realized its new system was not rating candidates for software developer jobs and other technical posts in a gender-neutral way.
- That is because Amazon’s computer models were trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Most came from men, a reflection of male dominance across the tech industry.



# CASE “ORACLE”

- The enterprise software company Oracle, through its omnibus Recruiting Cloud product, provides employers with predictions about the likelihood a candidate will accept a job offer, and what the employer can do to increase the candidate's chance of acceptance.
- The employer can adjust salary, bonus, stock options, and other benefits to see in real time how the prediction changes. The tool can update itself with employers' data about the outcome of previous offers and acceptances over time.
- Tools like this might amplify pay gaps for women and workers of color. Human resource data commonly include ample proxies for a worker's socioeconomic and racial status, which could be reflected in salary requirement predictions. In any case, offering employers highly specific insight into a candidate's salary requirements increases information asymmetry between employers and candidates at a critical moment of negotiation.



# CASE “HIREVUE”

- One prominent video interviewing company, HireVue, lets employers solicit recorded interview answers from applicants, and then "grades" these responses against interview answers provided by current, successful employees. More specifically, HireVue's tool parses videos using machine learning, extracting signals like facial expression and eye contact, vocal indications of enthusiasm, word choice, word complexity, topics discussed, and word groupings. It uses these signals to create a model that claims to capture relationships between interview responses and workplace performance, based on the employer's preexisting metrics.
  - Speech recognition software can perform poorly, especially for people with regional and nonnative accents.
  - Facial analysis systems can struggle to read the faces of women with darker skin.
  - Legitimacy of using physical features and facial expressions that have no credible, causal link with workplace success to make or inform hiring decisions.
  - Tests that have the effect of considering someone's immutable characteristics—even if they do so in a facially legal way—may violate expectations of dignity and justice, and prevent candidates from making a good-faith effort to demonstrate their suitability for a job.
  - Moreover, some worry that interviewees might be rewarded for irrelevant or unfair factors, like exaggerated facial expressions, and penalized for visible disability or speech impediments.





# Discussion!