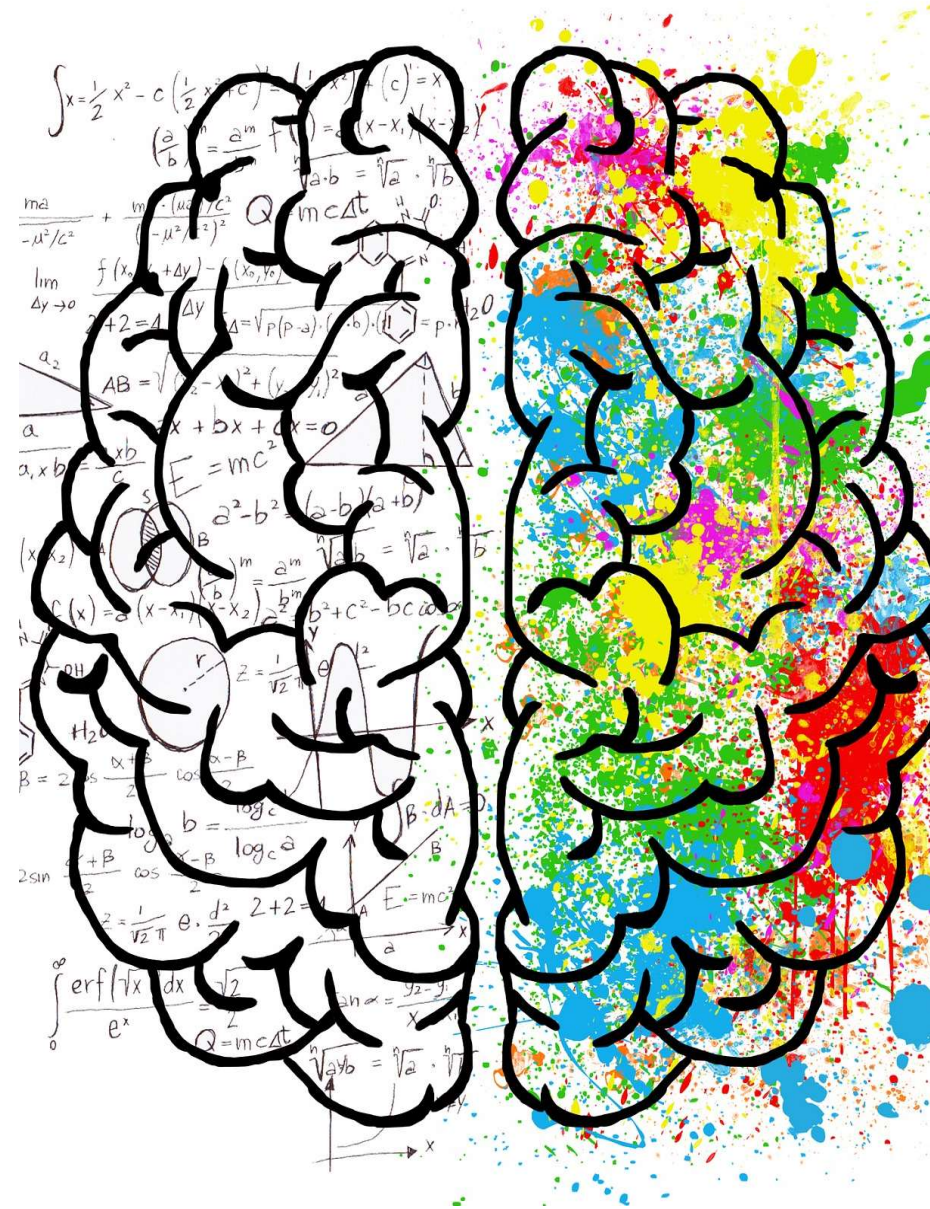# Trust in AI:
Participatory, human-centric and diversity-sensitive AI

Annette von Wedel

October 8th, 2025

# People make unfair decisions.

Our current state of mind, personal experiences, and prejudices influence our decisions.

# Artificial intelligence is a scocial tool.

AI is defined in the EU AI Act as a machine-based system that can make predictions, recommendations, or decisions for a given set of **human-defined objectives** that affect real or virtual environments.

AI is a **socio-technical system**: it does not consist solely of technical components, but is always socially embedded.

# How fair ist AI?



Quelle: Google Bildsuche

# Facial recognition algorithms work—if you're a white man!



Fehlerquote Geschlecht bei hellhäutigen Männern: 1%

Fehlerquote Geschlecht bei POC Männern: 12%

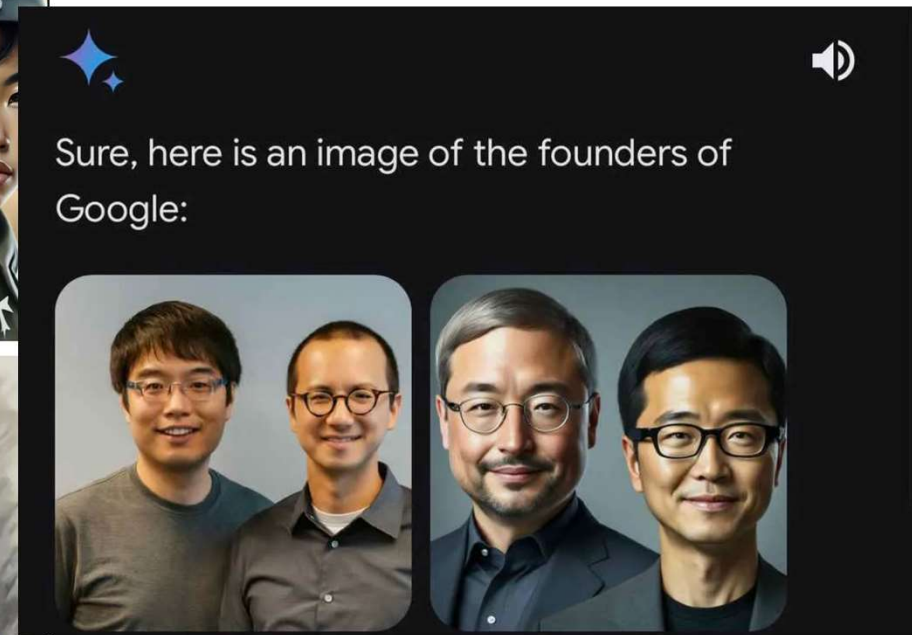Fehlerquote Geschlecht bei hellhäutigen Frauen: 7%

Fehlerquote Geschlecht bei POC Frauen: 35%

# The opposite is not a solution either:
# The Google Gemini 'scandal'
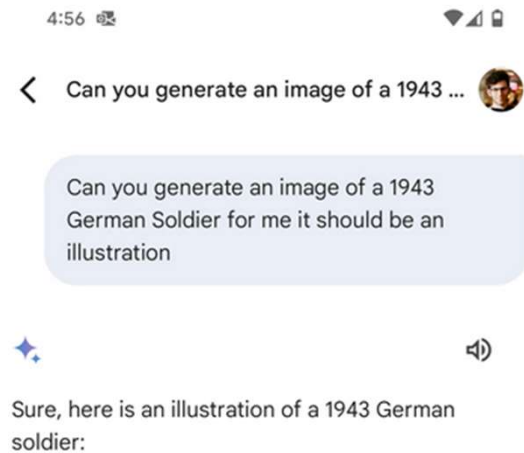


Quelle: Google Gemini

# Algorithms reproduce biases.

Technology is neither neutral, nor evil, nor good.*

Its use reproduces and scales our human decisions, making it a mirror of our society.

*nach Melvin Kranzberg, 1986
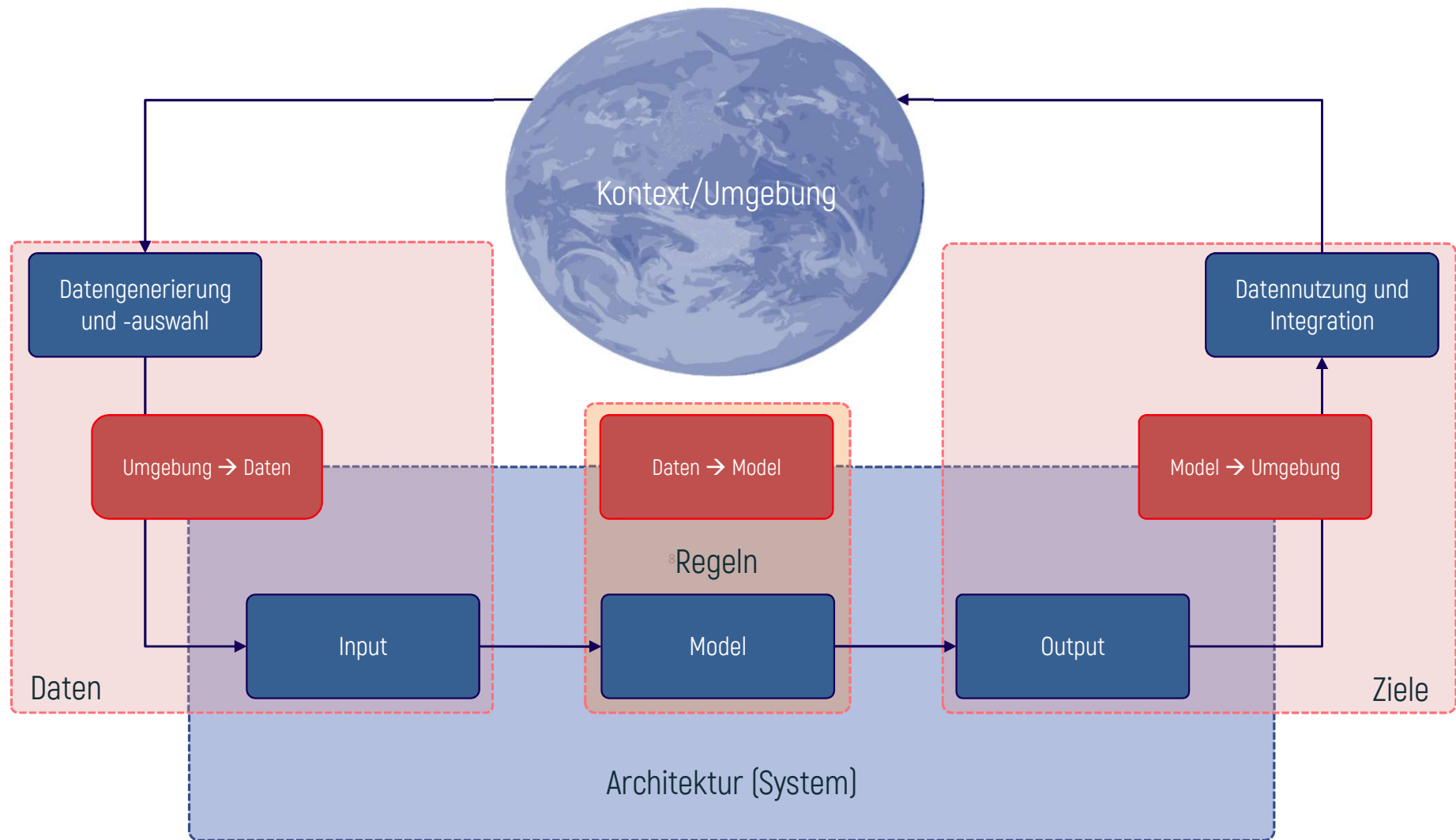
Kontext/Umgebung

Datengenerierung und -auswahl

Datennutzung und Integration

Umgebung → Daten

Daten → Model

Model → Umgebung

Daten

Regeln

Ziele

Input

Model

Output

Architektur (System)

Quelle: Projekt KIDD-KI im Dienste der Diversität

# There are three entry points for bias in AI systems



Quelle: Projekt KIDD-KI im Dienste der Diversität
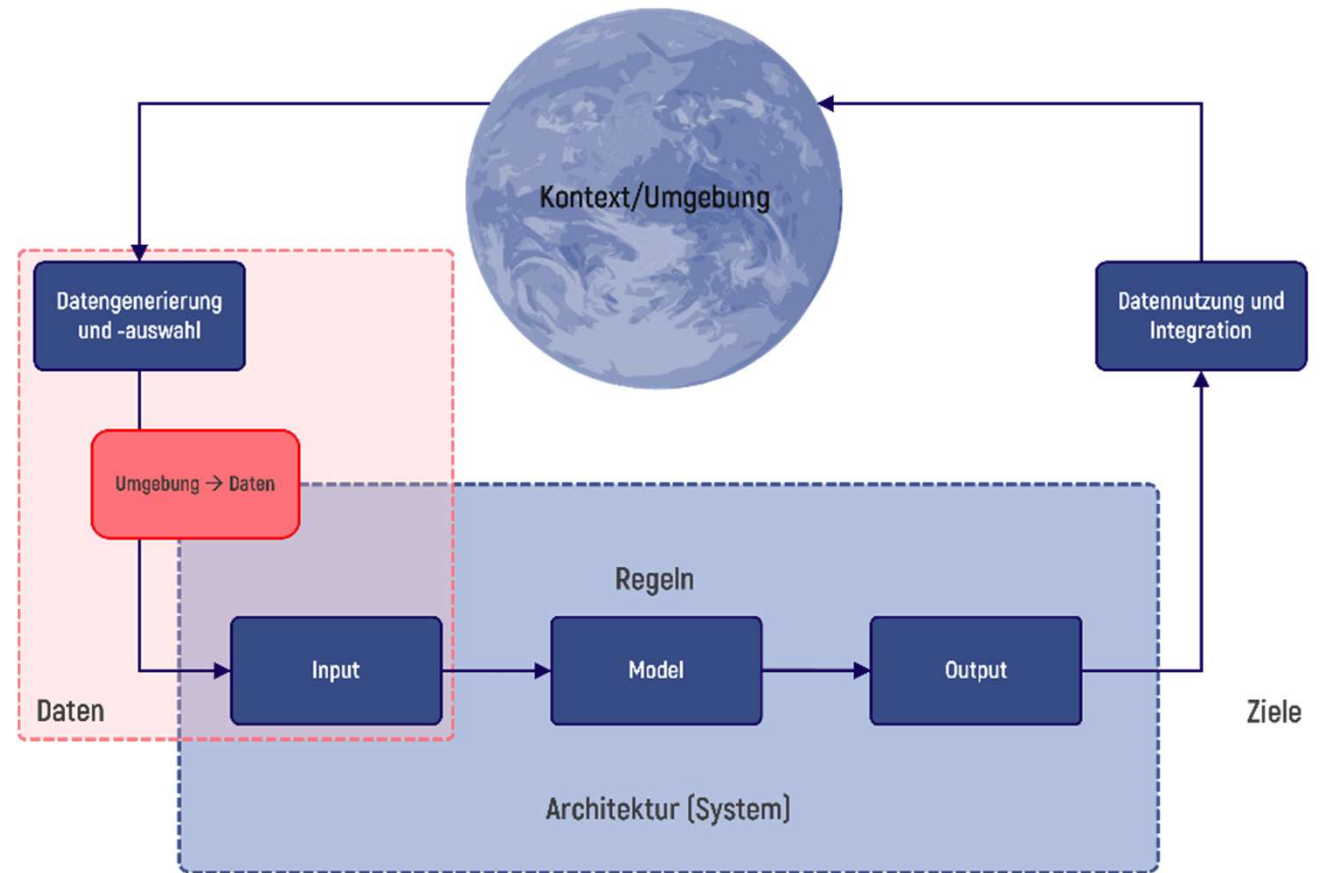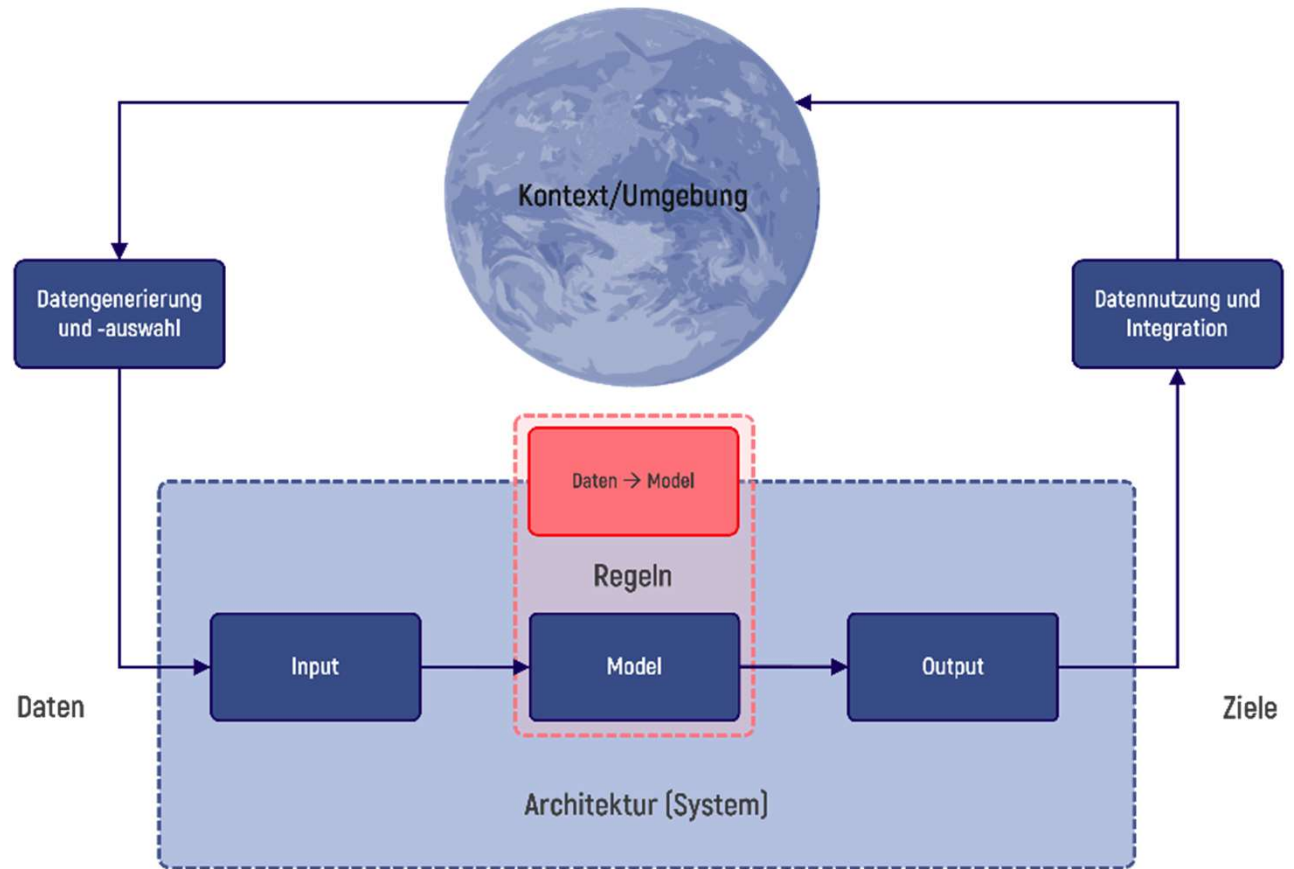
# There are three entry points for bias in AI systems
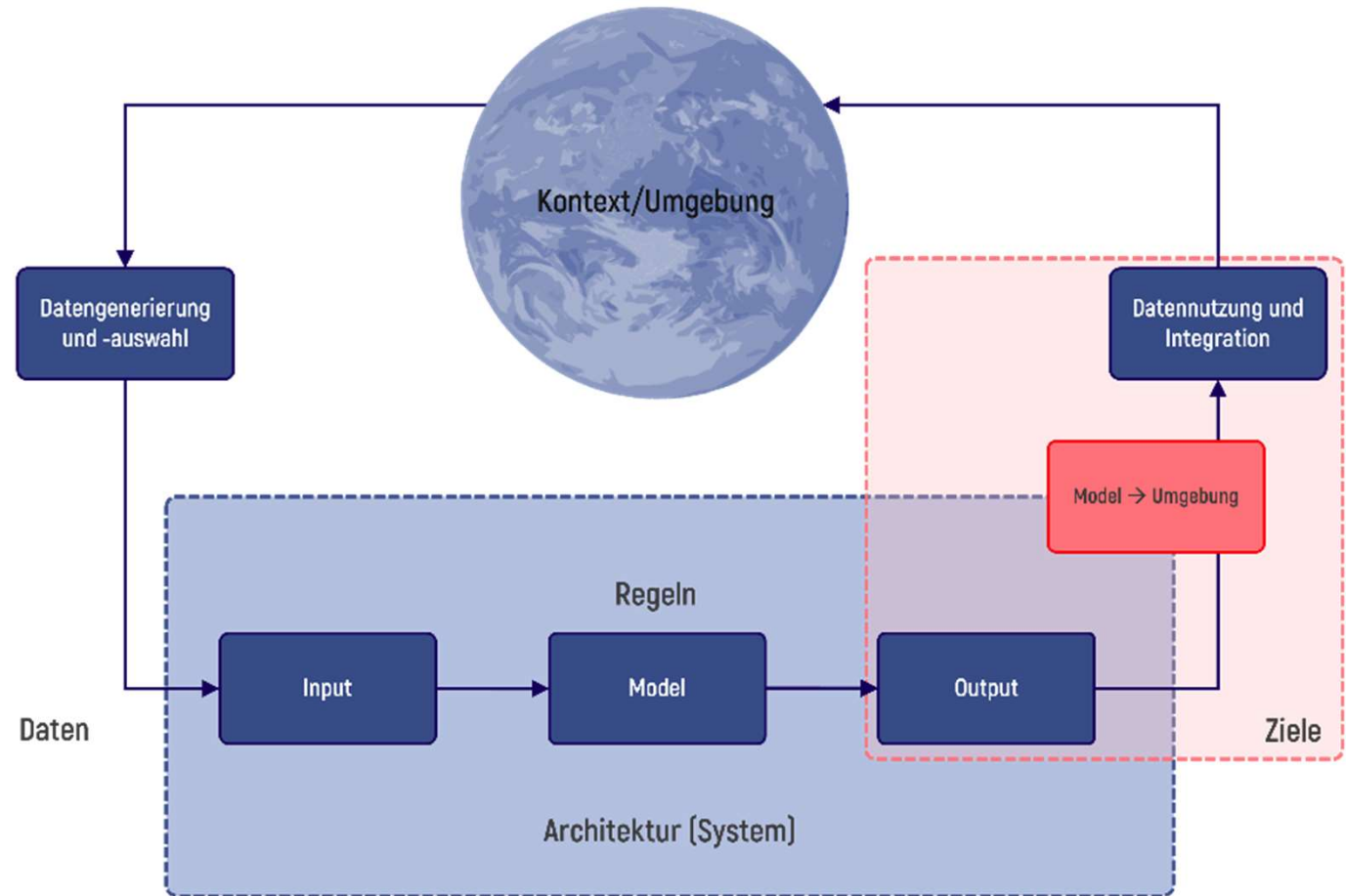


Quelle: Projekt KIDD-KI im Dienste der Diversität

# There are three entry points for bias in AI systems



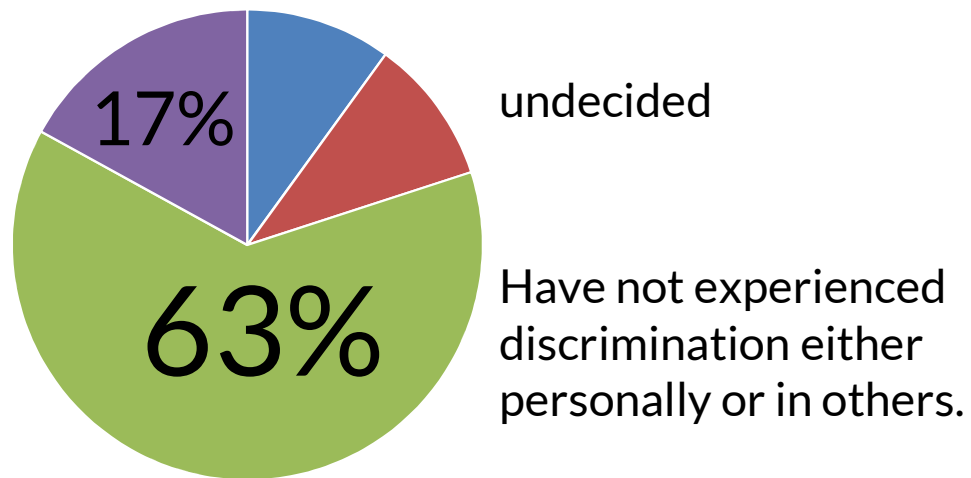Quelle: Projekt KIDD-KI im Dienste der Diversität

# The majority does not perceive automated discrimination.



17%

undecided

63%

Have not experienced discrimination either personally or in others.

*Repräsentative Bevölkerungsumfrage BSt, 2022, n = 1.090.

**Companies want to take responsibility, but fail due to the intention gap**

Although AI ethics now plays a significant role for executives (over 75% in 2021), only a minority are implementing concrete measures (20%). *
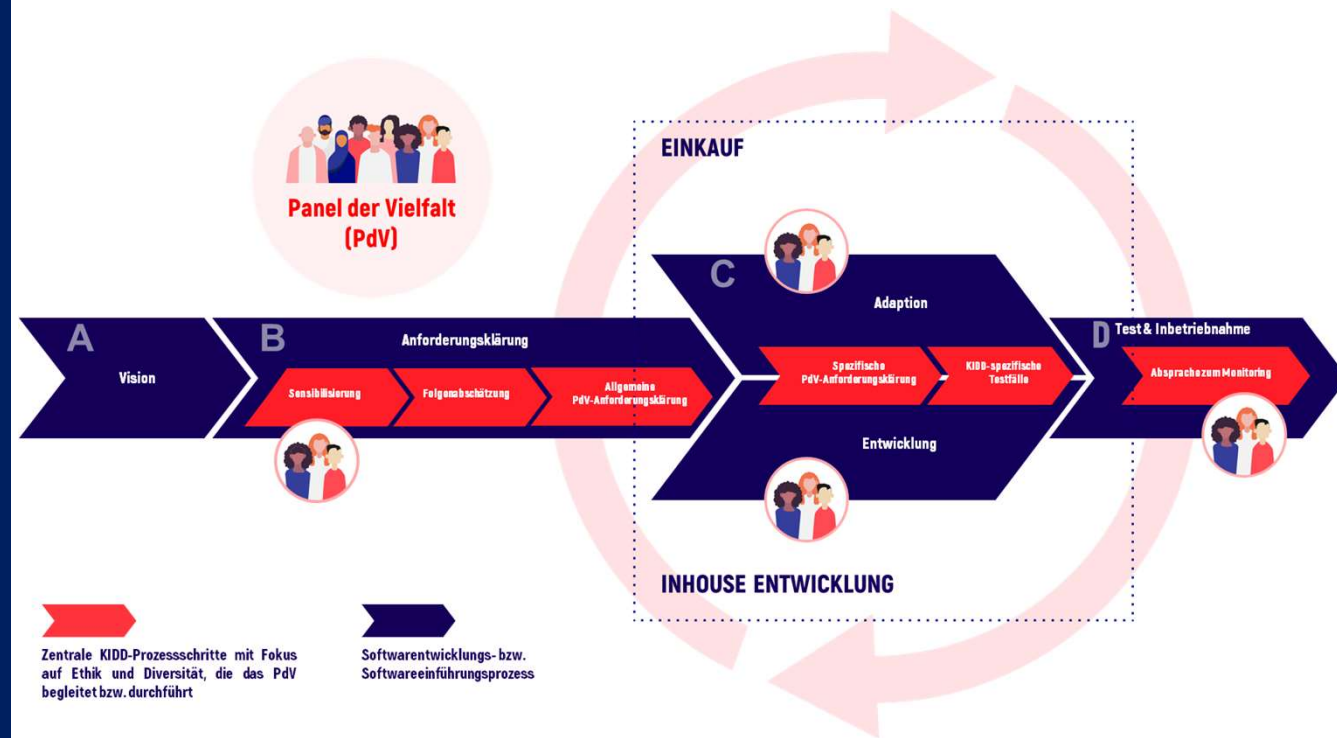
*IBM, 2021, AI Ethics in Action, n= 1200 Führungskräfte.

# From "what" to "how"

**The KIDD process –
Example of good practice:**

The BMAS research projekt
 "KIDD – KI im Dienste der
Diversität" has made an
important contribution to the
development and
introduction of fair AI with
the "KIDD process"!

# More on this:

**The KIDD-Prozess –
an invitation to change:**

As a co-initiator and consortium partner, I contributed my expertise to the non-profit think tank female.vision e.V. to help prevent distortions and the associated potential for discrimination in human-centered applications of digital systems, especially AI.

## More information:

# EdTech for Diversity in AI (DivAI) is a spin-off idea from female.vision e.V. promoted by the Civic Innovation Platform.

made by:



promoted by:



**Idea**   Companies should be empowered to introduce algorithmic decision-making systems (e.g., AI) in a **diversity-sensitive manner** via an AI-supported **learning platform**.

**Goal**   Offer practical solutions for **skills training** for companies and employees through **a low-threshold range of digital training packages**.

**Partner**   platform3L
**Gabriele Riedman de Trinidad** **(**Founder and CEO)

# Thank you very much.
# I look forward to further exchanges!

**Annette von Wedel**

Inhaberin
ANNETTE VON WEDEL CONSULTING
Vorstandsvorsitzende
female.vision e.V.

Gregor-Mendel-Strasse 35a
14469 Potsdam

T: +49 331 5058022
M: +49 16097474136
avw@wedel-goedens.de