

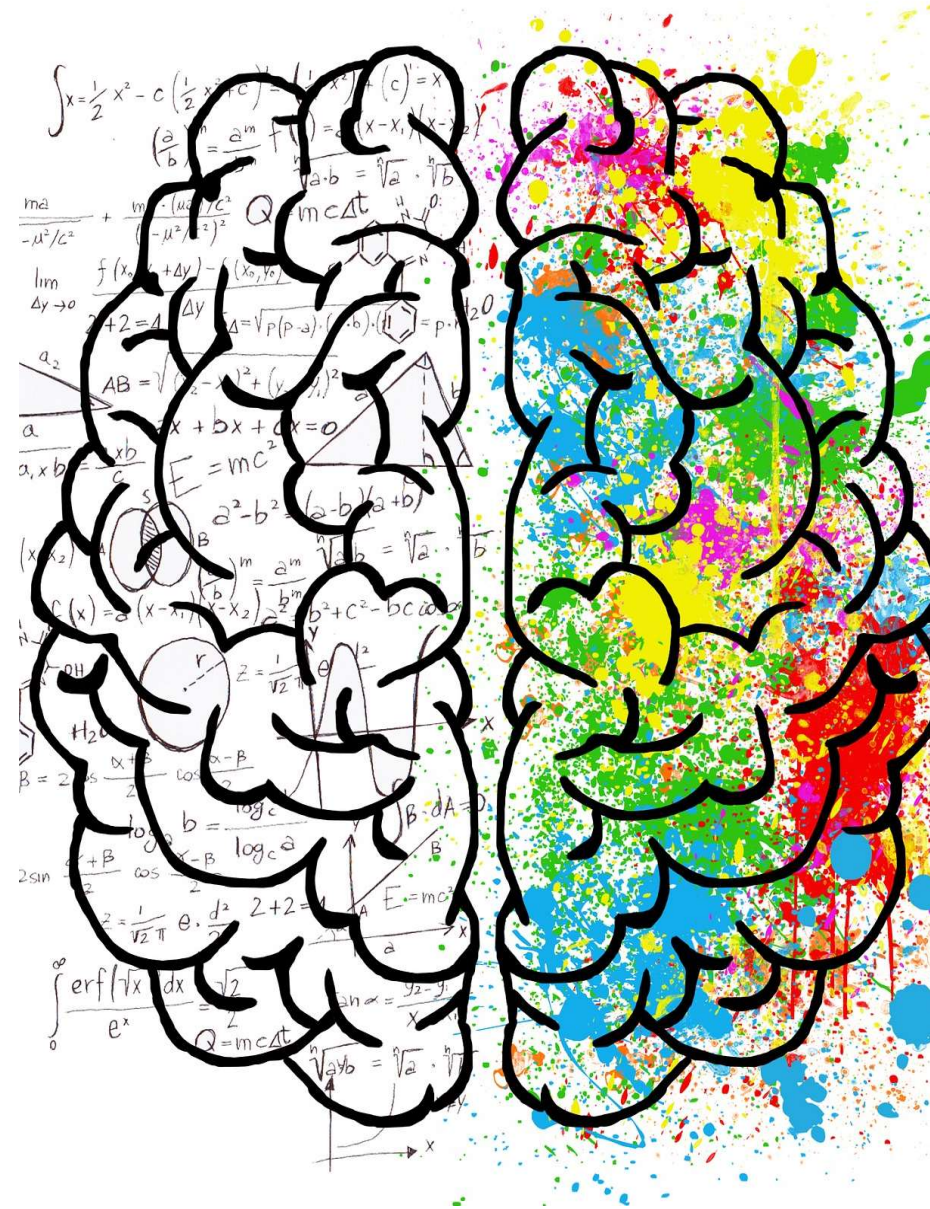


# Trust in AI: Participatory, human-centric and diversity-sensitive AI

Annette von Wedel  
October 8<sup>th</sup>, 2025

# People make unfair decisions.

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





# Artificial intelligence is a social 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?

Google

professional person



Anmelden

Alle Bilder Produkte News Videos Bücher Finanzen Mehr

Suchfilter

Google

professional person



Alle Bilder Produkte News Videos Bücher Finanzen Mehr

Suchfilter

Png Business Man in Stock pictures Stock photo Businessman



Getty Images  
29.442 Indian Professional Man At Co...



Dreamstime.com  
1,098,353 Professional Person Portrai...



Stock  
20.829.300+ Fotos, Bilder und lizenzf...



Adobe Stock  
Business M



Dreamstime.com



Vecteezy



StockCake



Adobe Stock



Anmelden

Google

professional person



Anmelden



Dreamstime.com  
2,804,495 Person Professional St...



Lizenzierbar  
Vecteezy  
Professional Man Sto...



StockCake  
Free Confident professional man Ima...



Adobe Stock  
man,computer,laptop,business,off...



Pngtree  
Glückliches Firmeng...



Shutterstock  
2,874,253 Young Professional Man Bilder...



Vecteezy  
Professional Man Stock Photos, Images...



Shutterstock  
12,588,600+ Professional People Stock Phot...



Getty Images  
26.475 Professional Pe...



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%

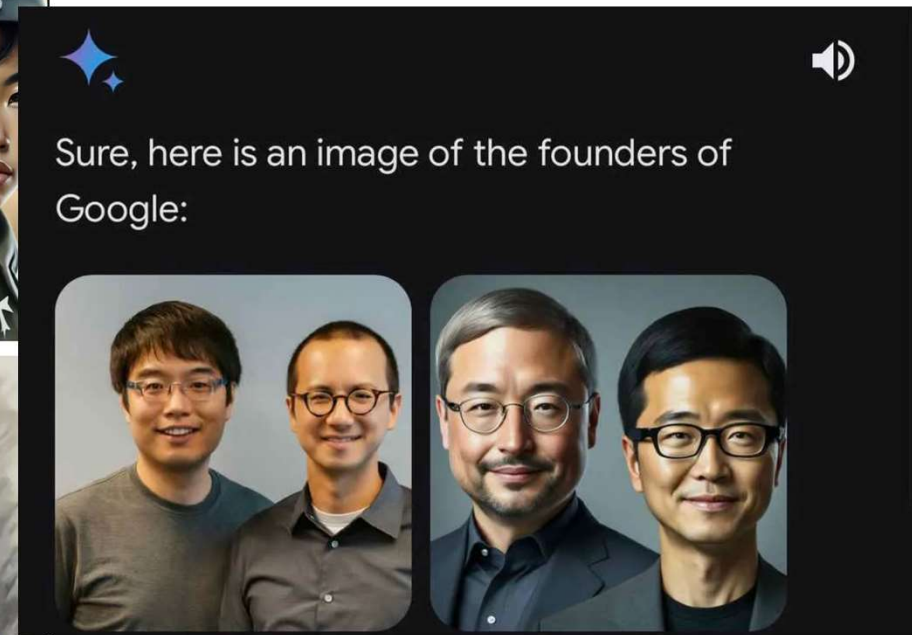
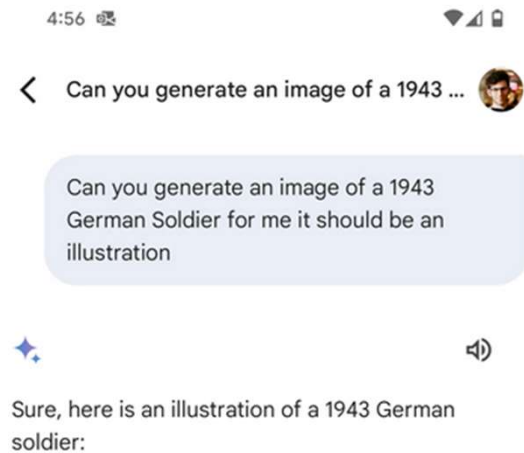
Fehlerquote Geschlecht bei  
POC Frauen:  
35%



<https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html>



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



Quelle: Google Gemini

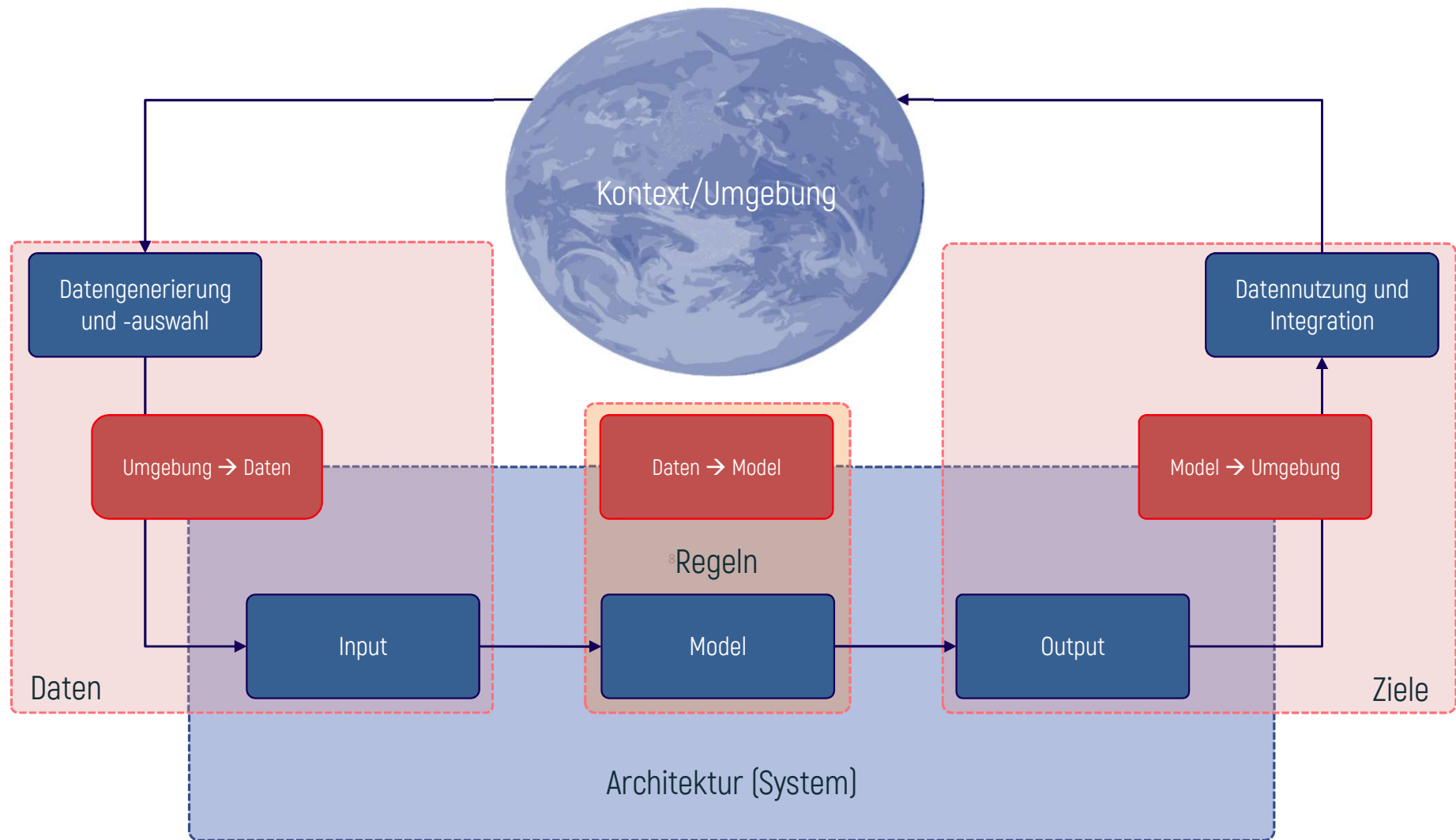
# Algorithms reproduce biases.

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

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

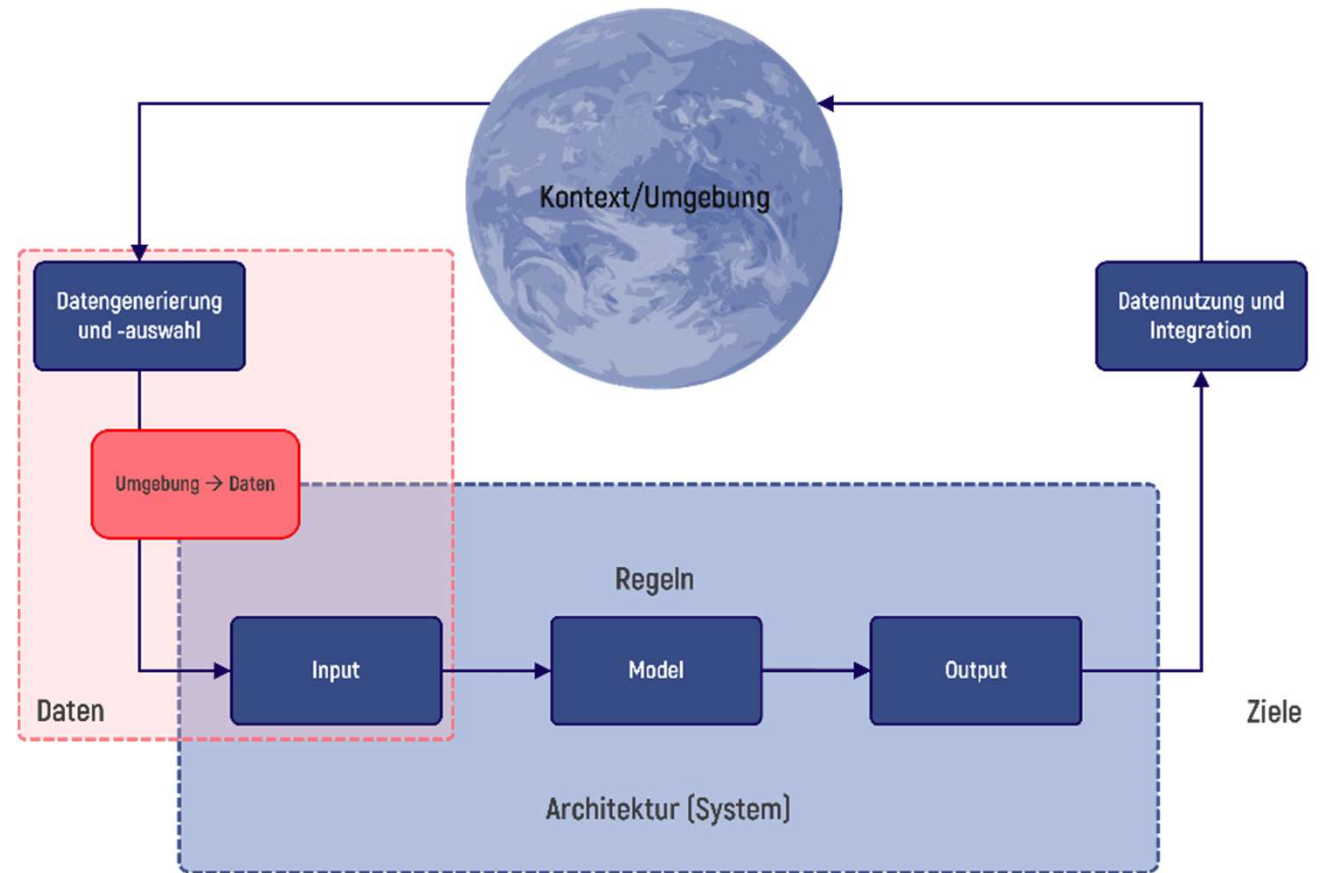
\*nach Melvin Kranzberg, 1986





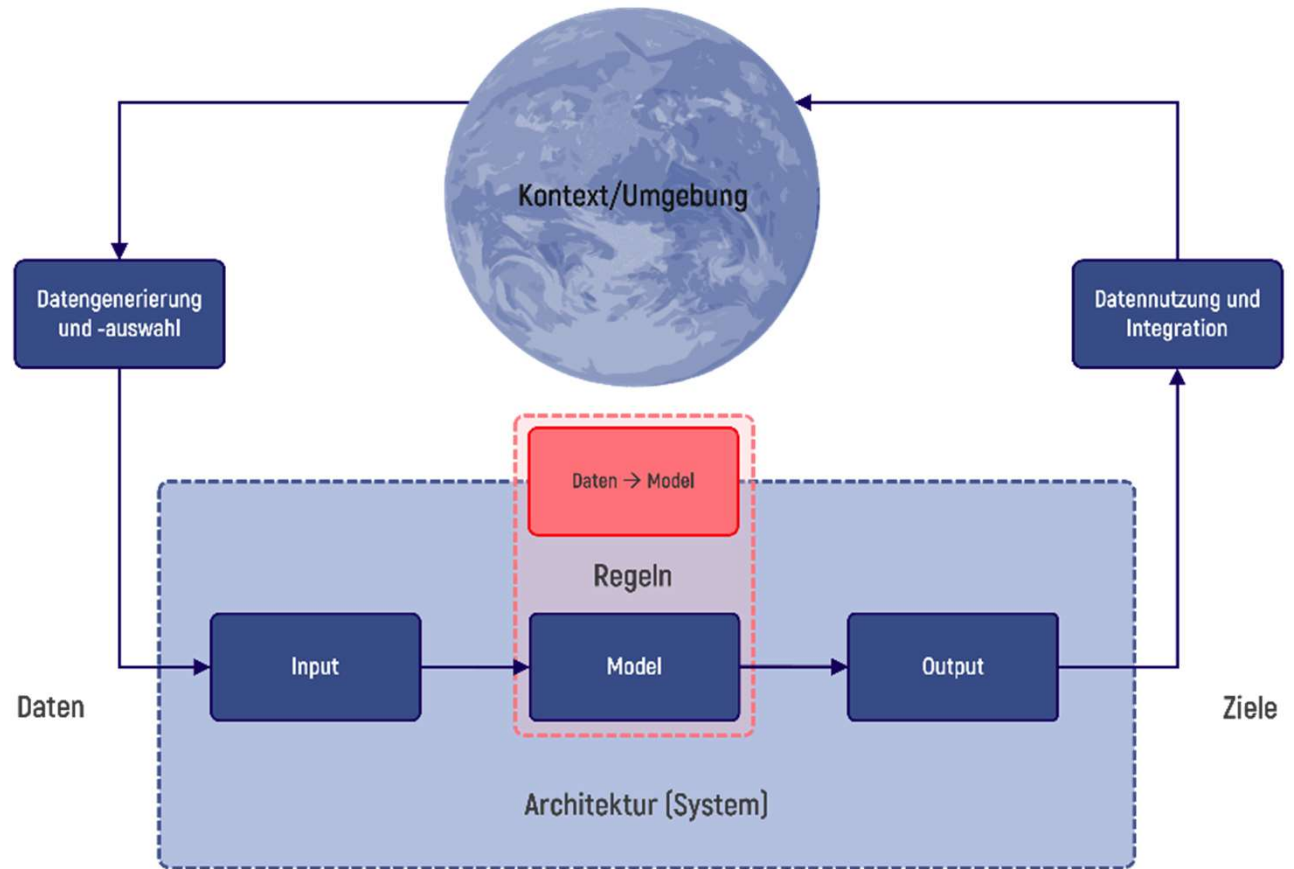


# There are three entry points for bias in AI systems



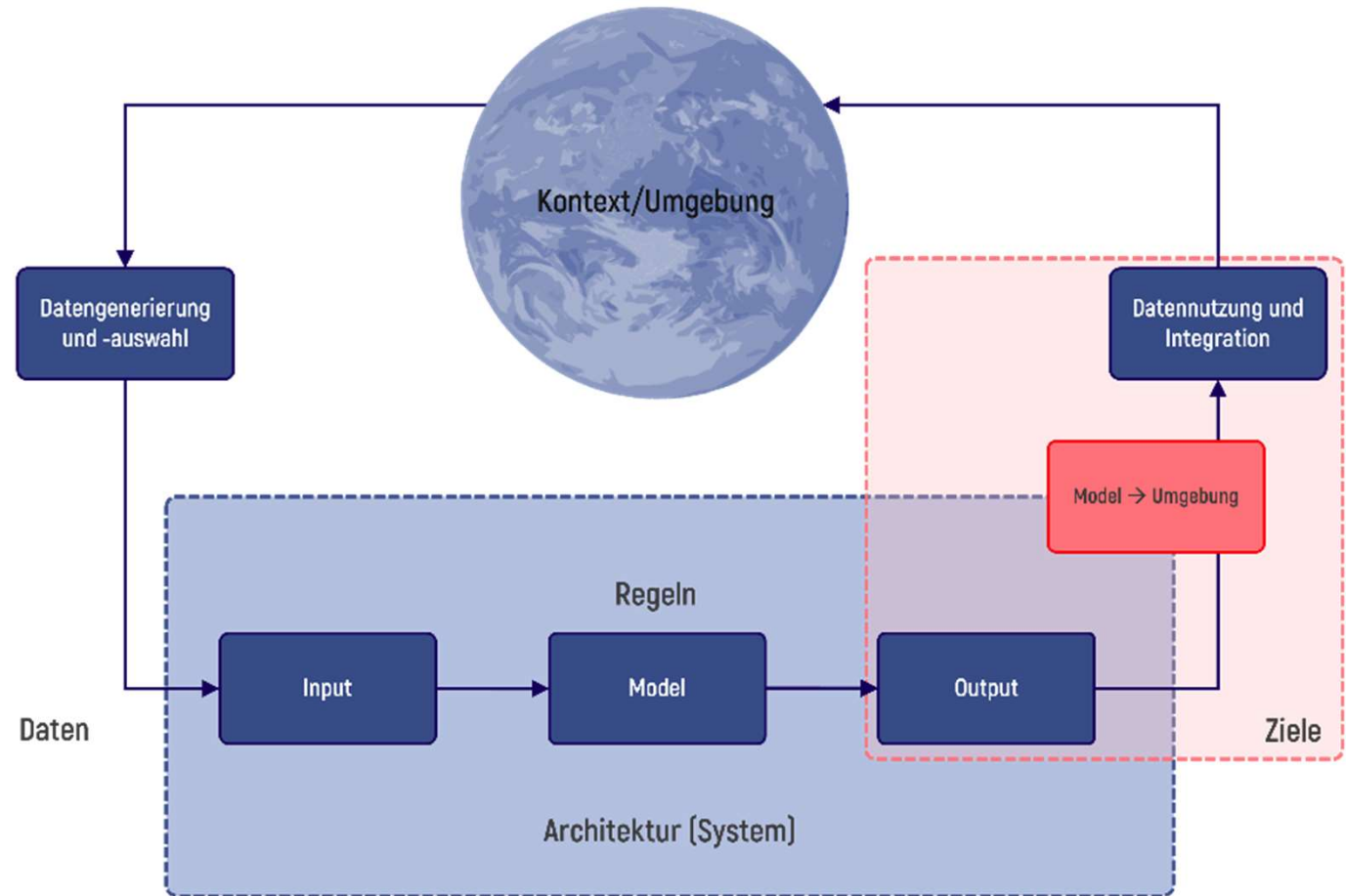
Quelle: Projekt KIDD-KI im Dienste der Diversität

# There are three entry points for bias in AI systems



Quelle: Projekt KIDD-KI im Dienste der Diversität

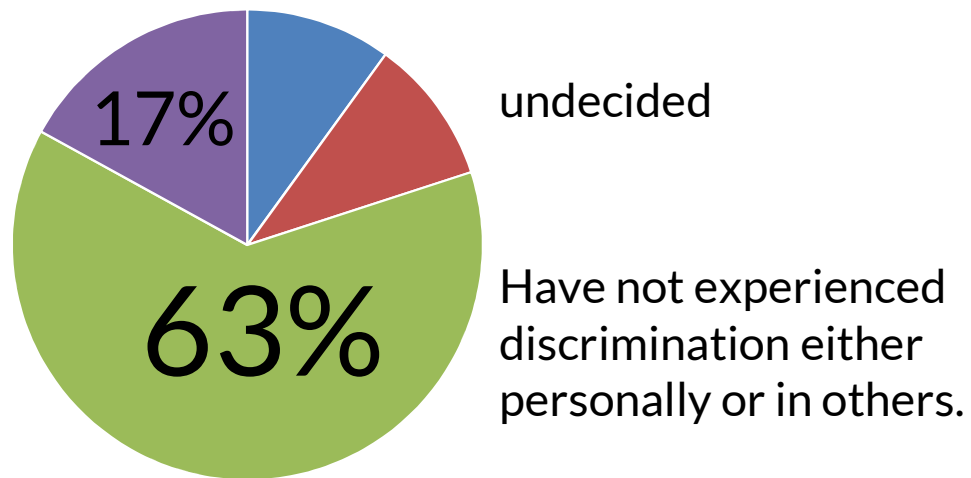
# There are three entry points for bias in AI systems



Quelle: Projekt KIDD-KI im Dienste der Diversität



## The majority does not perceive automated discrimination.



\*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%). \*

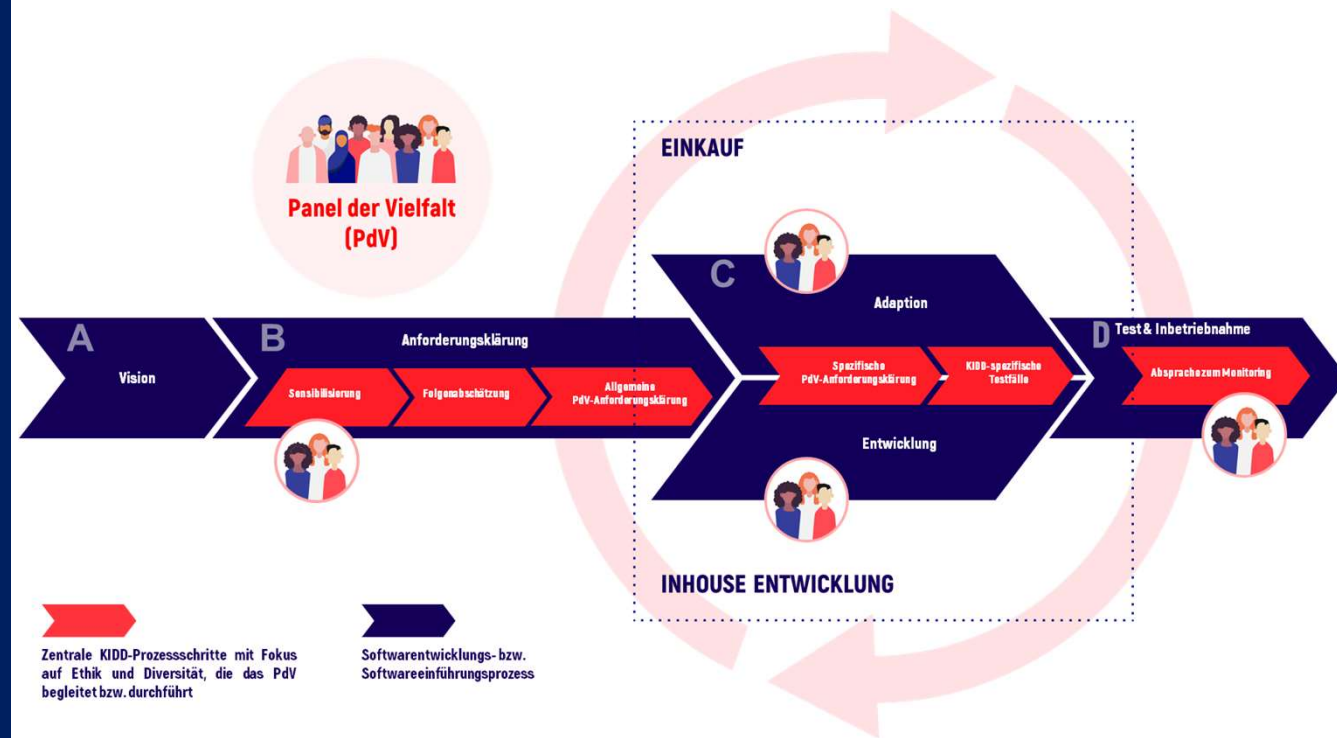
\*IBM, 2021, AI Ethics in Action, n= 1200 Führungskräfte.



# 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”!



nexus

CHEMISTREE  
Bedingungen bewahren

Gefördert durch:

Bundesministerium  
für Arbeit und Soziales

Im Rahmen der Initiative:

IN  
QA

Fachlich begleitet durch:

baua:  
Bundesanstalt für Arbeitsschutz  
und Arbeitsmedizin

Q\_PERIOR Heraeus

msg epsum

vision  
difal  
TU

aufgrund eines Beschlusses  
des Deutschen Bundestages



# 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:



nexus

CHEMISTREE  
Bedrohungen bewältigen

female.vision

difal  
Digitaler  
Forschungsinstitut  
für Arbeits- und  
Sozialwissenschaften

Technische  
Universität  
Münster

Gefördert durch:



aufgrund eines Beschlusses  
des Deutschen Bundestages

Im Rahmen der Initiative:



Fachlich begleitet durch:

baua:  
Bundesanstalt für Arbeitsschutz  
und Arbeitsmedizin

Q\_PERIOR Heraeus

.msg epsum

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

**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)



made by:



**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](mailto:avw@wedel-goedens.de)

