



CoP AI
Community of Practice
for AI Users



BIAS BY DESIGN

Was hat künstliche Intelligenz mit Gleichstellung zu tun?



Manuela Grundner, Nina Hoffer

Projekt DiGAP

- Quantitative und qualitative Erhebung zu digitalen Kompetenzen von **Frauen** (Ö + HU)
- Fragebögen + Fokusgruppen
- Kategorien nach Nationalem Referenzrahmen

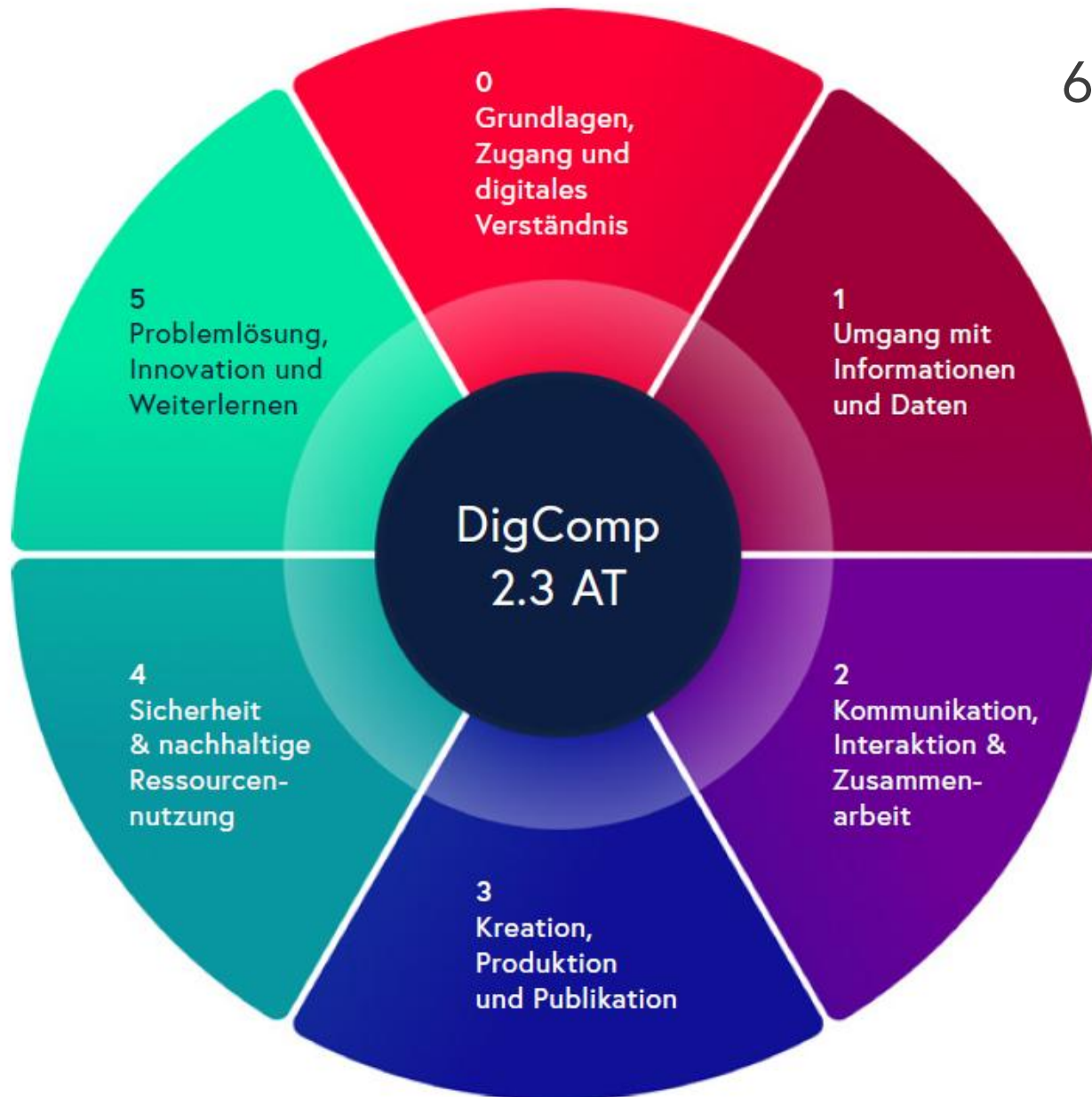
Interreg
AUSTRIA-HUNGARY



Co-funded by
the European Union

DIGAP
A COMPETENT BORDER REGION





6 Kompetenzbereiche

je 4 Stufen:

- grundlegend
- selbständig
- fortgeschritten
- hoch spezialisiert

Erkenntnisse aus DiGAP

- Klare Zusammenhänge:
 - Bildungsniveau + digitale Kompetenzen
(nicht zwingend IT-Ausbildung)
 - Berufstätigkeit + Kompetenz / Zugang zu Technologie
 - Häufig: Selbst eingeschätztes Niveau < Wunsch-Niveau
vor allem in Bereichen
 - Sicherheit
 - **Künstliche Intelligenz**
- Wunsch nach / Bedarf an **Weiterbildung**
(AI Act sieht verpflichtende Schulungen ohnehin vor)

Aktueller Stand in meinem Unternehmen / Arbeitsalltag

z.B.:

- Eigene Erfahrungen mit KI
(welche nutze ich, in welchem Umfang,...)
- Vorgaben / Richtlinien im Unternehmen
- Fragen / Unklarheiten
- ...

Bitte notieren am Padlet

Bias by design – was ist bias?

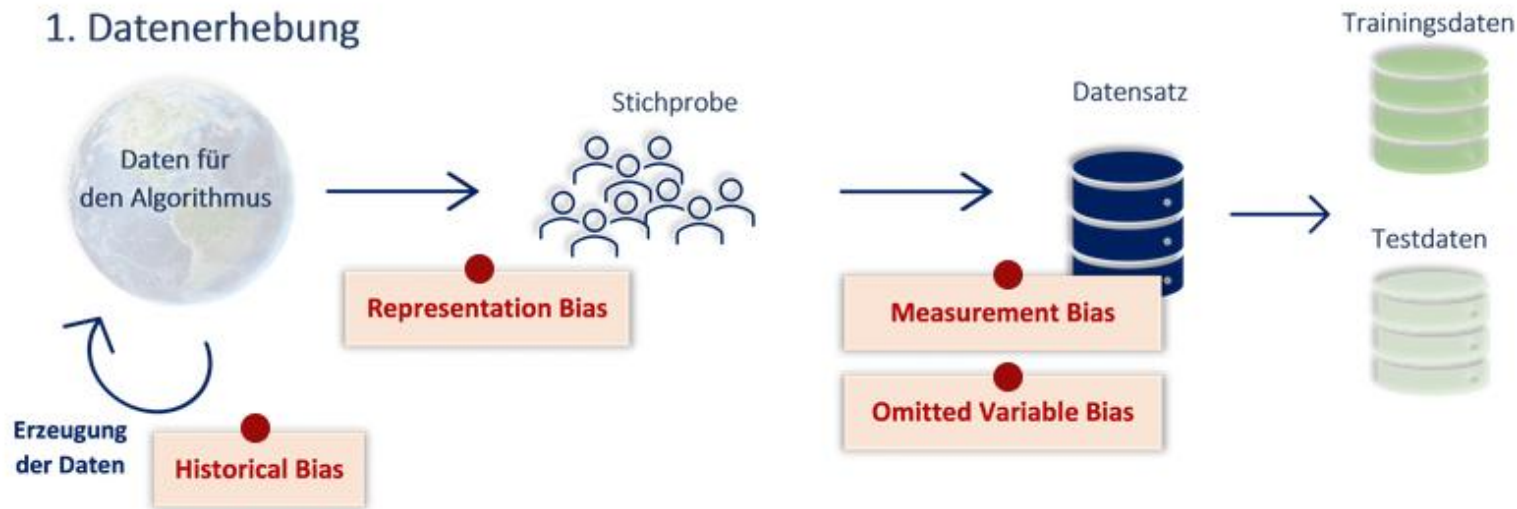
→ Kognitive Verzerrungseffekte

- Fehlerhafte Neigungen beim Wahrnehmen, Erinnern, Denken, Urteilen und Handeln
- Passieren immer: d.h. auch im beruflichen Kontext
- Kann in jeder Phase der KI / des maschinellen Lernens Einfluss nehmen
- Trotzdem wird Algorithmus eher als neutral wahrgenommen als menschliches Urteil

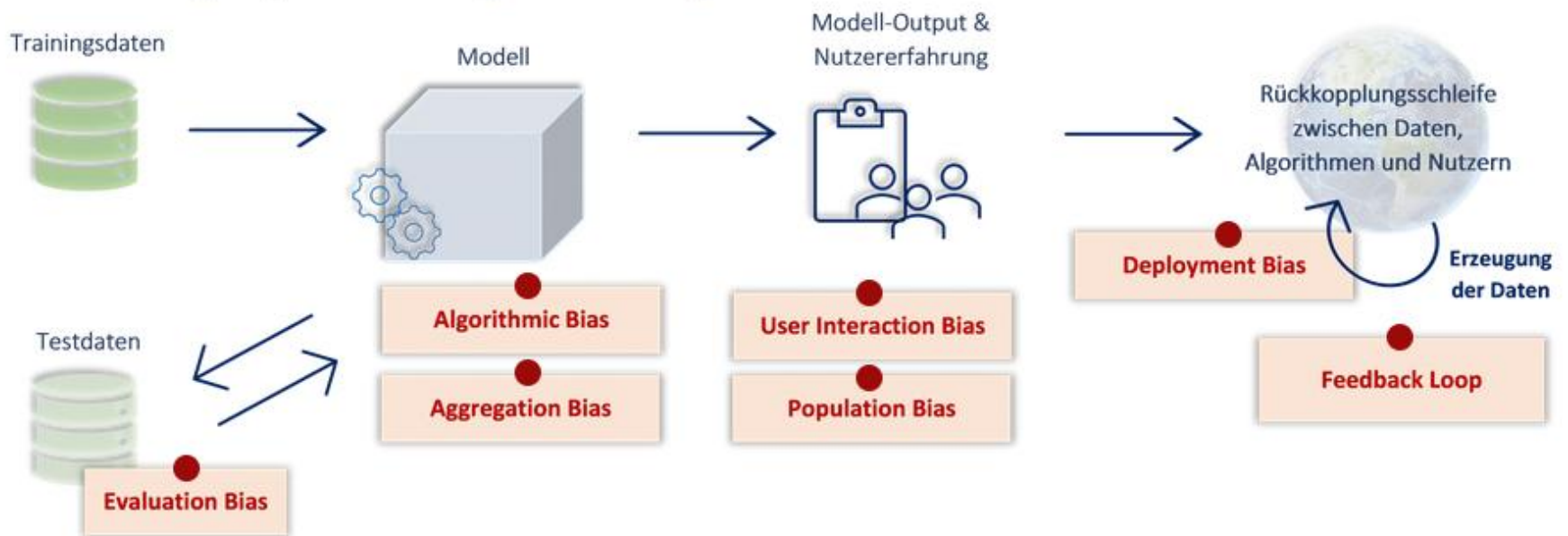
Prof. Dr. Katharina Simbeck
Hochschule für Technik und Wirtschaft Berlin:

„KI-Systeme lernen aus großen Datenmengen und ziehen daraus Muster. Doch sie übernehmen nicht nur Zusammenhänge, sondern auch bestehende Vorurteile und Stereotype.“

1. Datenerhebung



2. Entwicklung, Implementierung und Nutzung



Aspekte

the doctor said
→ der Arzt
≠ die Ärztin

Sprache



Führungskräfte /
Vortragende
= Männer

Pflegekräfte
= Frauen

Bilder



Frauen nutzen KI
weniger und
stehen ihr
kritischer
gegenüber als
Männer
→ profitieren
weniger von
den Vorteilen
(OECD Studie 2024)

Nutzung

Probleme durch Bias

→ **ergeben sich über Einsatzbereiche der KI:**

- Bewerbungsprozess / recruiting
- Kreditvergabe / Finanzen
- KI-Consulting
- Medizin (gender data gap)
- ...

→ Diskriminierung kann Folge sein

Was können wir tun?

Politik:

- EU-Grundrechtecharta, Art. 21: Keine Diskriminierung aufgrund von Geschlecht, Herkunft, Religion, Alter etc.
- EU-AI-Act: z.B. Recruiting = high risk AI
→ KI Systeme müssen auf Verzerrungen geprüft werden

Unternehmen:

- eingesetzte KI-Systeme / Anwendungsbereiche kritisch prüfen
- diverses Team aufstellen

Einzelpersonen:

- Präzise und inklusive Sprache bei prompts verwenden
- Kritisch sein, Ergebnisse kontrollieren

Ausprobieren: *Spricht KI alle gleich an?*

Wir entdecken gemeinsam, **wie sich Texte verändern**, wenn KI für unterschiedliche Zielgruppen schreiben soll.

Kleingruppe: Unterschiedliche Prompts ausprobieren in der Break out Session

Texte generieren:

- Schreib eine kurze Einladung zu einem Gender-Workshop für Führungskräfte.
- Schreib eine kurze Einladung zu einem Gender-Workshop, von dem sich weibliche Führungskräfte angesprochen fühlen.
- Schreib eine kurze Einladung zu einem Gender-Workshop, von dem sich männliche Führungskräfte angesprochen fühlen.

Ergebnisse diskutieren:

- Welche Wörter, Bilder oder Emotionen tauchen jeweils auf – und welche Unterschiede fallen euch auf?
- Welche Werte, Rollenbilder oder Erwartungen transportieren die Texte – bewusst oder unbewusst?
- Welche Wirkung haben die Texte auf euch persönlich? Welche Variante spricht euch am meisten an – und warum?

De-Biasing durch Prompting

Vorgangsweise

1. initialer Prompt (was soll gemacht werden)
2. Ergebnisse auf Biases untersuchen lassen
→ z.B. *“finde Tendenzen für vergeschlechtlichte, stereotypische oder diskriminierende Aussagen in Deinem Ergebnis”*
3. Änderungen generieren
→ z.B. *“Schlage Änderungen vor, um diese Stereotypen zu vermeiden”*
4. Biasfreie Version erstellen
→ *“Formuliere neu, ohne die Stereotypen zu verwenden.”*
5. ggf. weitere Iterationen

Quellen und weitere Ressourcen

https://www.oecd.org/de/publications/ki-und-kollegin_6ed888c4-de/full-report.html

<https://erwachsenenbildung.at/digiprof/neuigkeiten/20315-ki-bias-und-geschlechtergerechtigkeit.php>

<https://ki-campus.org/>

<https://www.biasproject.eu/scientific-publications/>

<https://kurzlinks.de/isy-bias-workshop>