

Künstliche Intelligenz fürchten oder beherrschen?

tamatzner.de

informatikethik.de



KI ist ein Mittel zum Zweck.

Leitfrage der Ethik: Was soll ich tun?

setzt Beschäftigung mit **Handlungen** voraus,
die bestimmten **Zwecken** dienen.

Chatbots

Was soll ich als Nutzer tun, um deren
Ergebnissen vertrauen zu können?

Zweck z.B.:
mich informieren
eine Hochzeitsrede vorbereiten
eine Seminaufgabe bearbeiten

Voraussagen menschlichen Verhaltens

Was soll ich als Betreiber tun, um ihnen
vertrauen zu können?

Zweck z.B. Voraussage der:
Rückfallgefahr von Straftätern
Erfolgsaussicht von Bewerbern
Bonität von Kreditnehmern

Der Einsatzzweck bestimmt das Fachgebiet.

Fachgebiete mit **nachweisbar** gültigen Regeln
Exakte Wissenschaften, Technik
weite Bereiche des Wirtschaftens

Behandlung seit jeher durch Menschen
seit 80 Jahren zunehmend durch Maschinen
auch bestimmte Zweige von KI
hohes Vertrauen

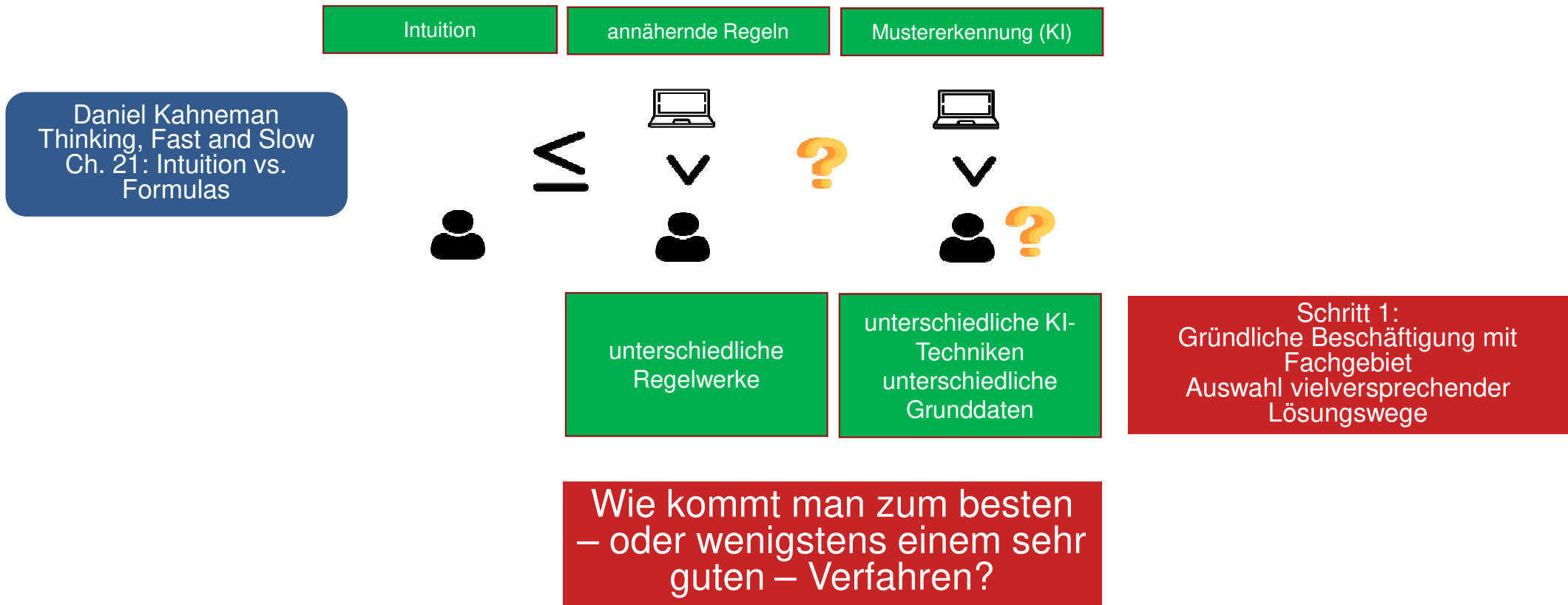
Wahrheit
Ja-Nein-Aussagen

Fachgebiete ohne nachweisbar gültige
Regeln
z.B. Voraussage menschlichen Verhaltens

Behandlung seit jeher durch Menschen
Vertrauen?
nun zunehmend durch Maschinen, auch KI
Wem sollen wir mehr vertrauen?

Wahrscheinlichkeits-
Aussagen

Lösungswege für Einsatzzweck ohne nachweisbare Regeln



ohm

Für schwierig algorithmisierbare Aufgaben reicht klassischer Test nicht aus.



^{ohm} Schwieriger Test gegen Wirklichkeit erfordert zusätzliche Maßnahmen.

- Historische Daten einholen.
- Im Prozess laufend Erfolgsdaten einholen.
- Veränderungen der Verteilung von Ein- und Ausgabedaten ermitteln und bewerten
- Abwägung Datenschutz gegen andere Menschenrechte.
- Mehrere Verfahren parallel betreiben und Ergebnisse speichern.
- Manchmal bewusst falsch handeln, um Feedback zu provozieren (u.U. moralisch verwerflich!)
- ??? – Forschungsarbeit für jeden Einsatzzweck!

Chatbots

fürchten

beherrschen

blind einsetzen

Betreiber und dessen Interessen
ermitteln

gegen Wirklichkeit testen
(vergleichsweise einfach möglich)

Gefahr: Chatbot gibt einseitige
Mainstream-Meinung wieder

nicht als einzige Quelle nutzen, wenn
wichtige Entscheidungen anstehen

KI

fürchten

beherrschen

Einsatzzweck unklar lassen

Einsatzzweck konkret fassen und moralisch bewerten

IT muss billig sein!
Erstbeste Lösung wählen.
Beim Test kann man sparen...
...erst recht, wenn System schon produktiv ist

Sorgfalt bei Entwurf und Test,
angemessen für Kritikalität

wenn schwierig gegen Wirklichkeit zu
testen, *ständige* Wachsamkeit und Test

Maschine vermenschlichen