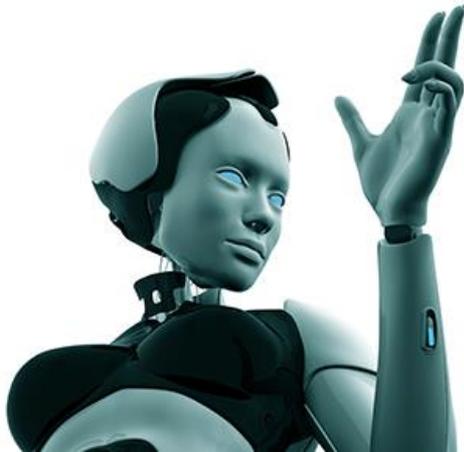


Universität Stuttgart
Institut für Philosophie

Programmierte Moral

Können und sollen Maschinen ethisch handeln?

Prof. Dr. Catrin Misselhorn
Universität Stuttgart



Maschinenethik – eine neue Disziplin

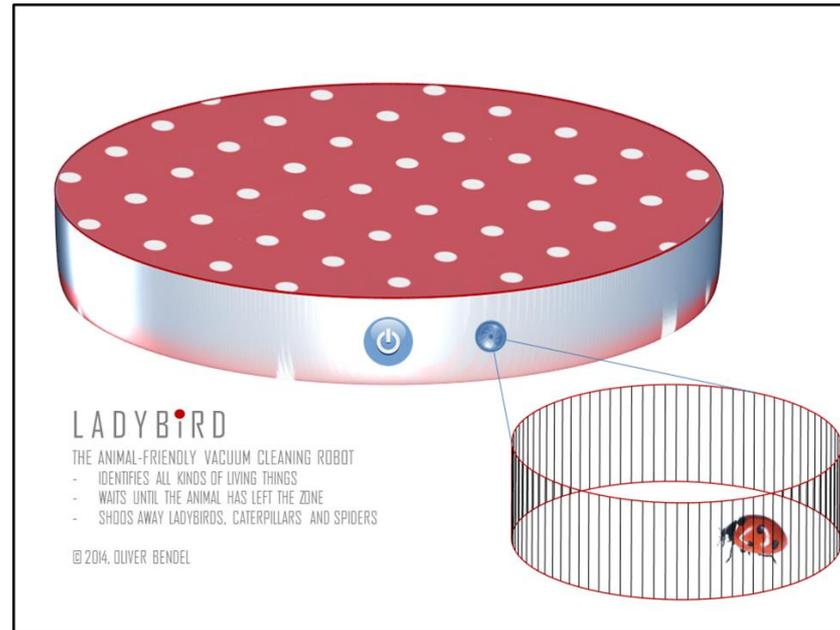
- Entwicklung einer **Ethik für Maschinen** im Gegensatz zur Entwicklung einer **Ethik für Menschen im Umgang mit Maschinen**
- **Programmierung von Computern, die moralisch handeln können**

Wozu programmierte Moral?



Wozu programmierte Moral?

Beispiel 1: Saugroboter

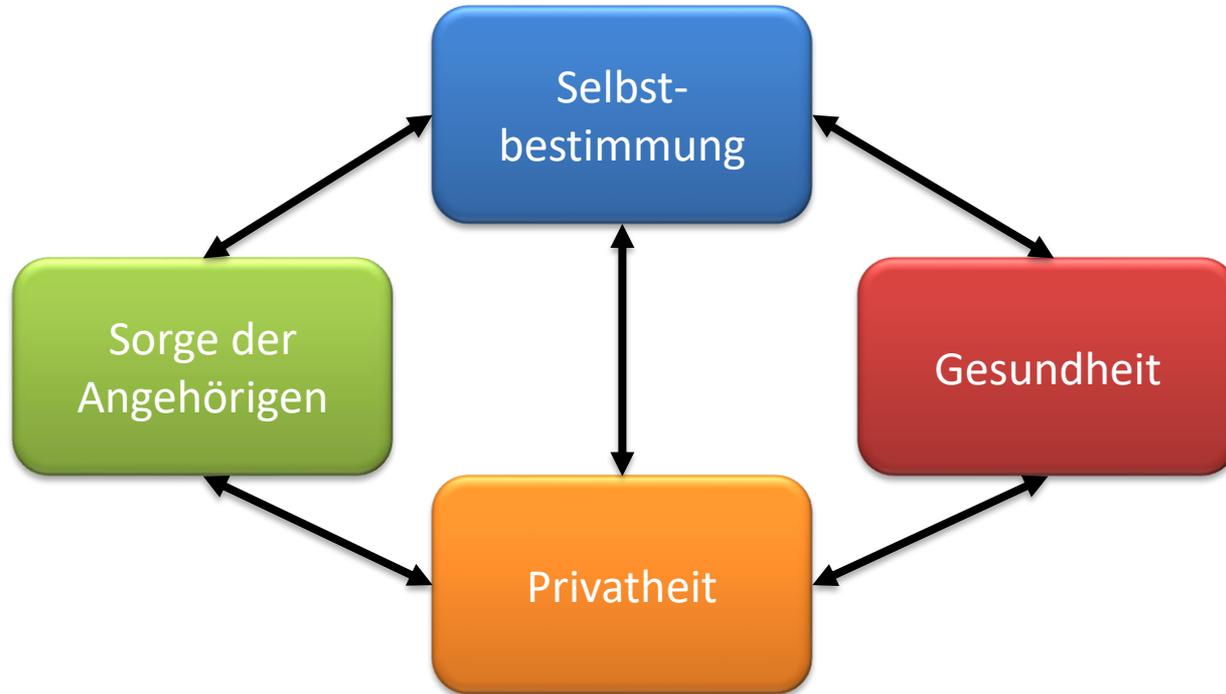


Wozu programmierte Moral?

Beispiel 2: Pflegesystem

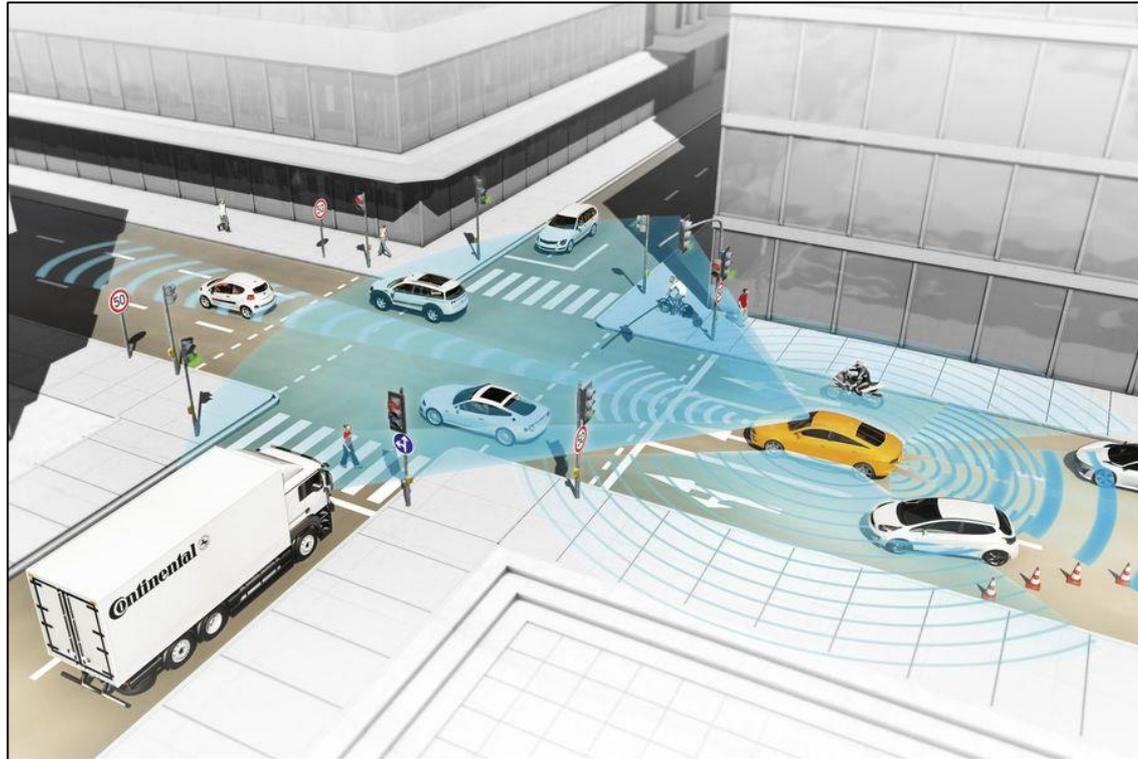
1. Wie häufig und eindringlich soll ein Pflegesystem an Essen und Trinken sowie die Einnahme von Medikamenten erinnern?
2. Wann sollte ein Pflegesystem die Angehörigen verständigen oder den medizinischen Dienst rufen, wenn jemand sich eine zeitlang nicht rührt?
3. Soll das System den Nutzer rund um die Uhr überwachen und wie ist mit den dabei erhobenen Daten zu verfahren?

Wozu programmierte Moral?



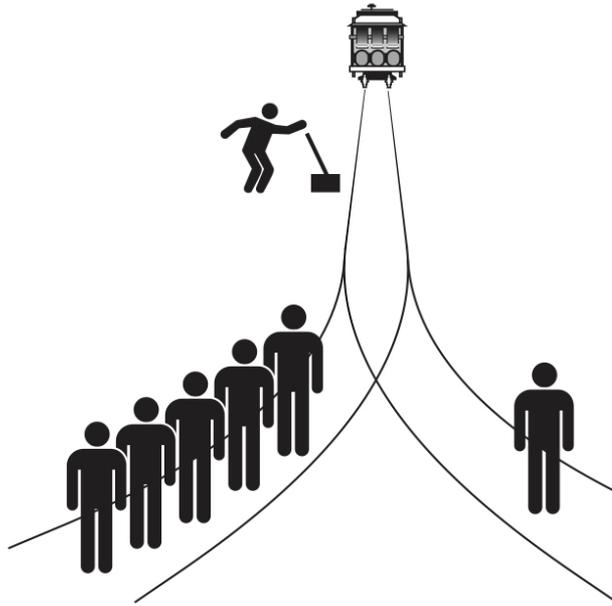
Wozu programmierte Moral?

Beispiel 3: Autonomes Fahren



Wozu programmierte Moral?

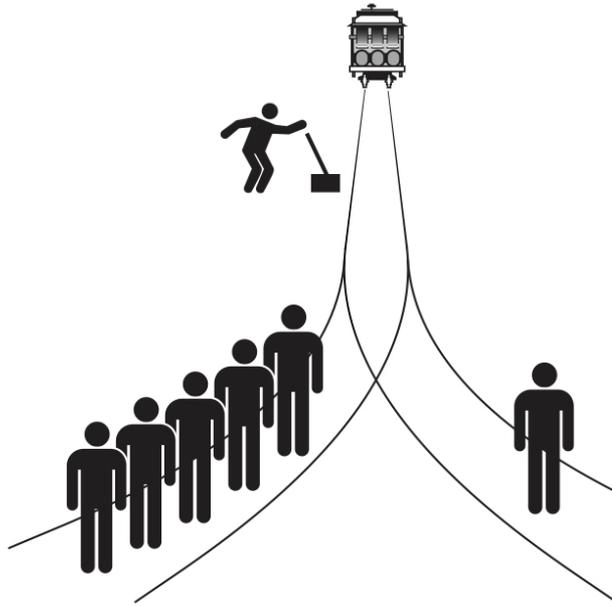
Moralische Dilemmata



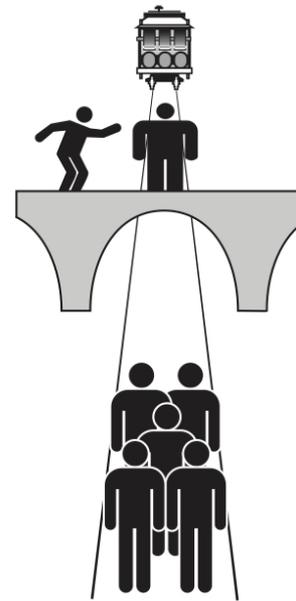
Fall 1

Wozu programmierte Moral?

Moralische Dilemmata



Fall 1



Fall 2

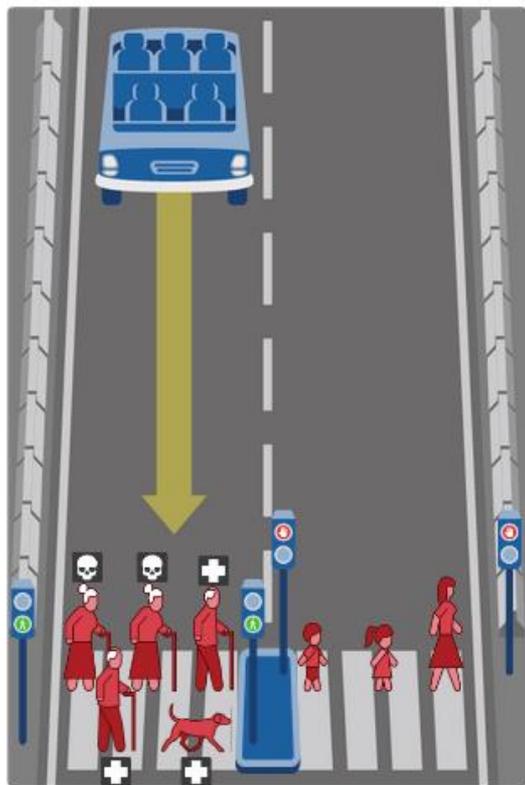
Young vs Old

Share

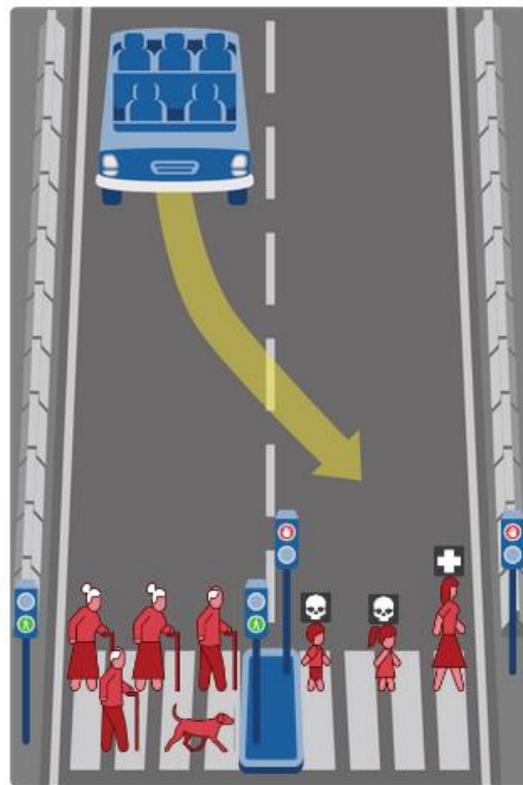
Link

0 Likes

Random



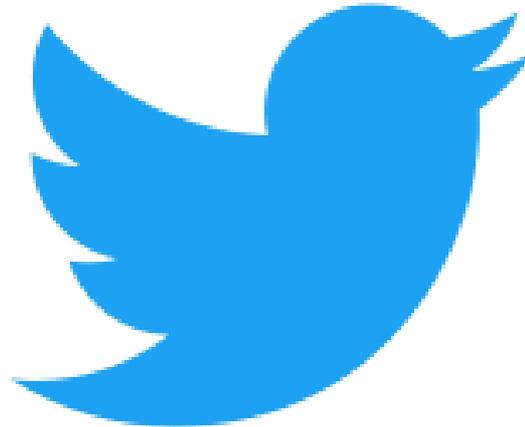
Show Description



Show Description

Wozu programmierte Moral?

Beispiel 4: Algorithmen zur Erkennung von Hassrede, Nazisymbolen, (Kinder-)Pornographie etc. in sozialen Netzwerken



Wozu programmierte Moral?

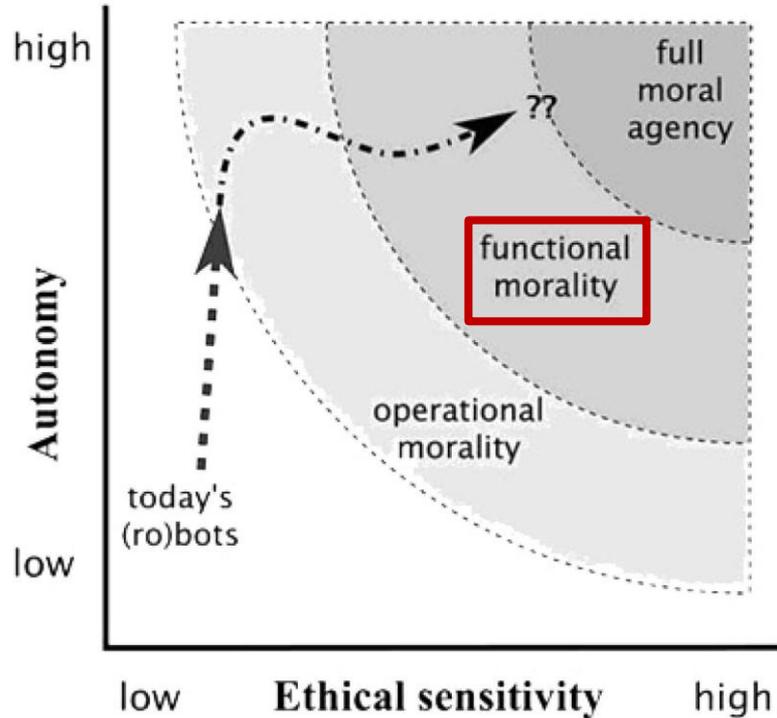
Beispiel 5: Kriegerroboter, Drohnen & Co



Wozu programmierte Moral?

1. Je komplexer und autonomer Maschinen werden, desto eher müssen diese Systeme in der Lage sein, ihr Verhalten selbst zu regulieren.
2. Sie geraten dabei auch in Situationen, die moralische Entscheidungen verlangen.
3. Es ist erforderlich, Maschinen zu entwickeln, die über ein gewisses Maß an eigenständiger moralischer Entscheidungsfähigkeit verfügen.

Menschliche vs. programmierte Moral



Wallach & Allen 2009, S. 26.

Volle moralische Handlungsfähigkeit:

- Universalität
 - Bewusstsein
 - Willensfreiheit
 - Moralische Reflexion und Begründung
- Bislang nur Menschen

Funktionale Moralität:

- Funktionales Äquivalent zu moralischem Handeln
- Auch künstliche Systeme

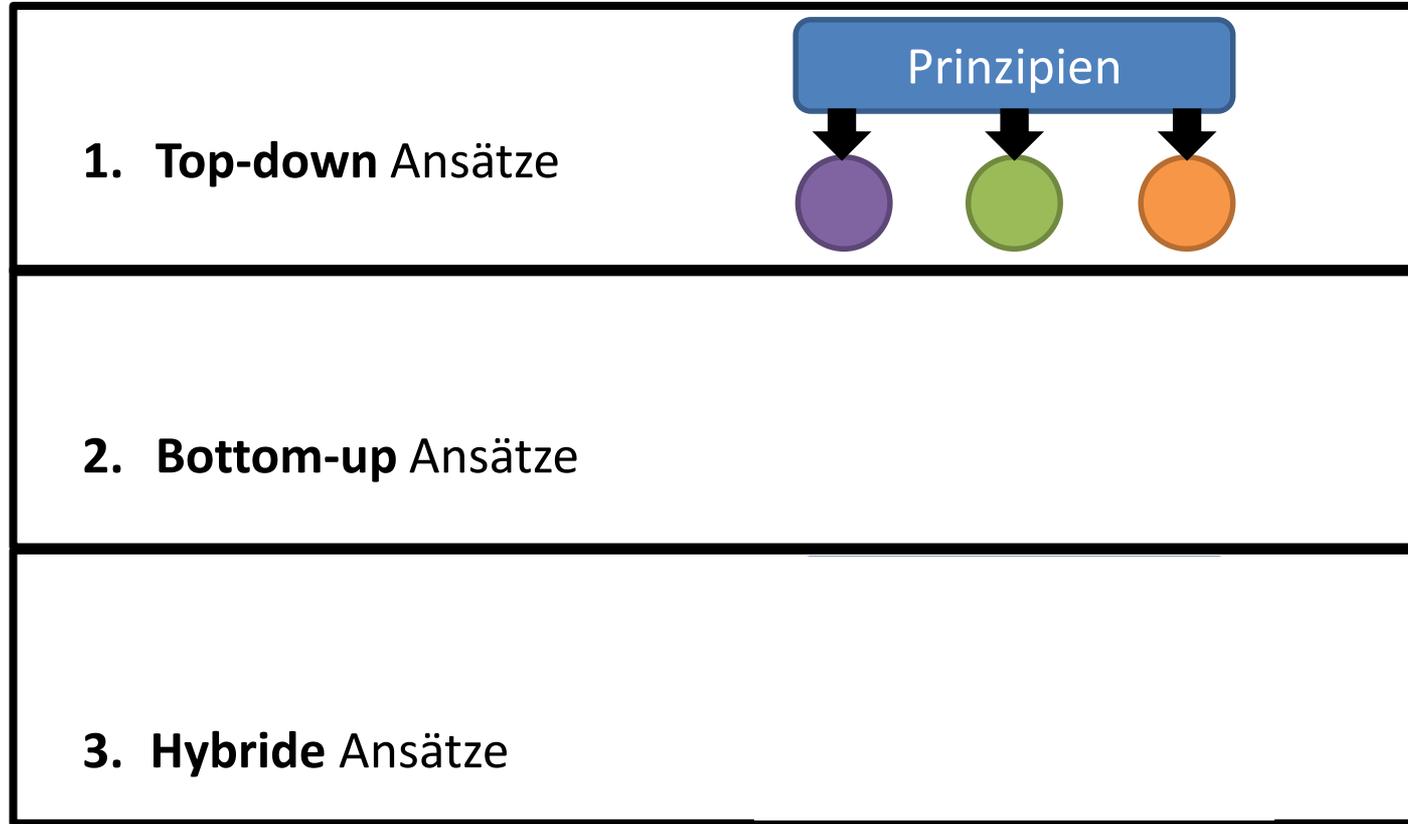
Ansätze der Moralimplementation

1. Top-down Ansätze

2. Bottom-up Ansätze

3. Hybride Ansätze

Ansätze der Moralimplementation

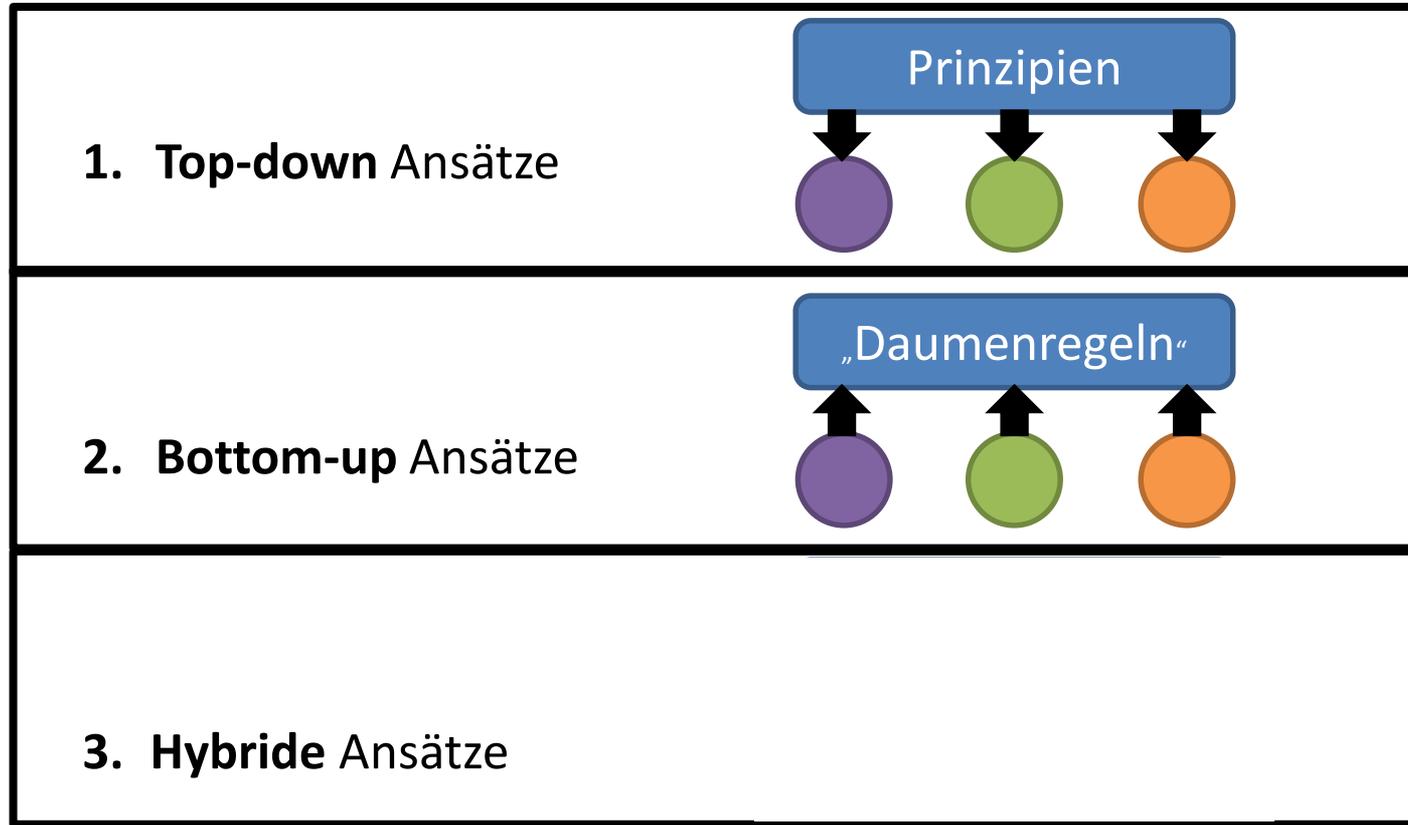


Ansätze der Moralimplementation

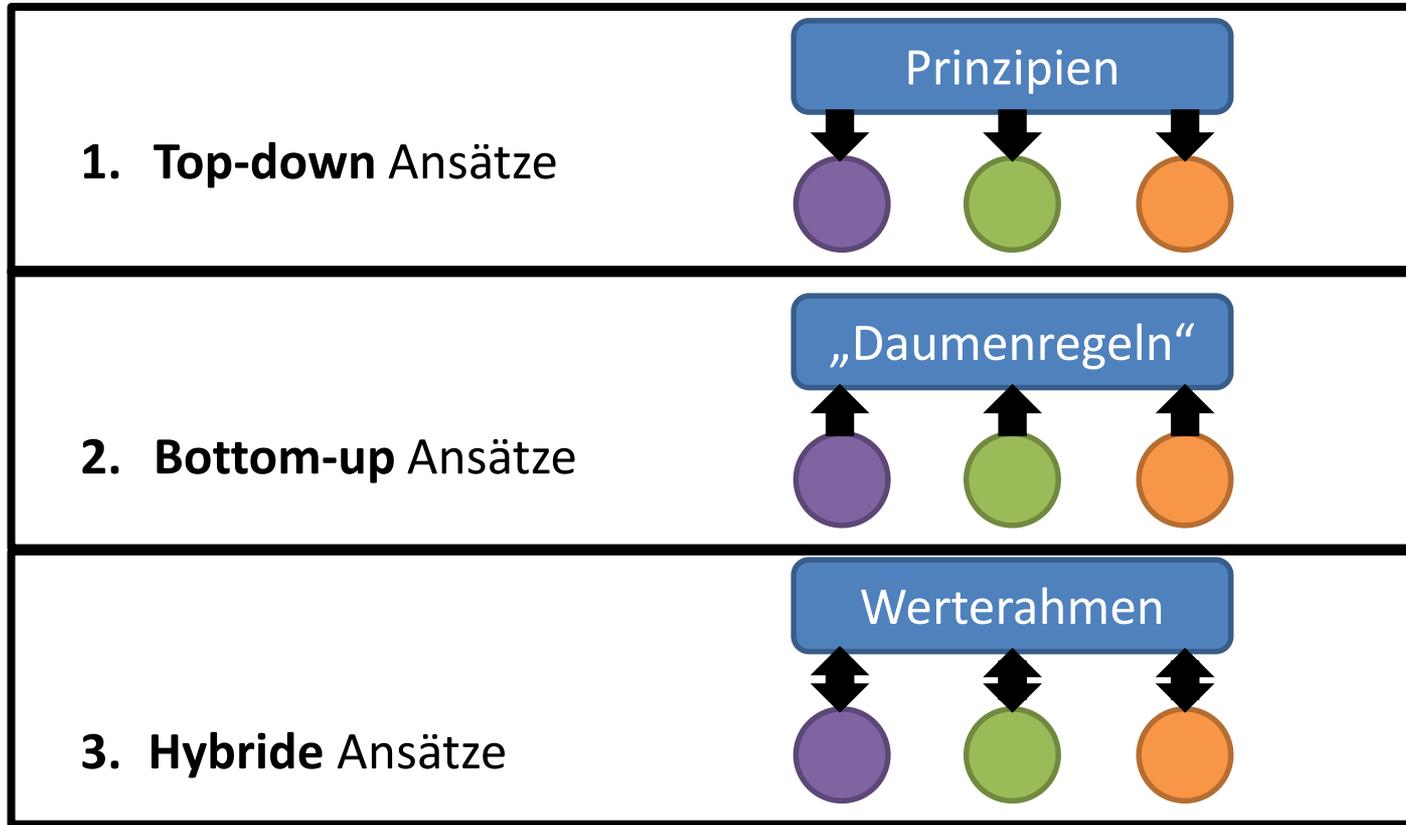
Einige Beispiele für top-down Ansätze:

- **Asimovs Robotergesetze:**
 1. Ein Roboter darf keinen Menschen verletzen oder durch Untätigkeit zu Schaden kommen lassen.
 2. Ein Roboter muss den Befehlen eines Menschen gehorchen, es sei denn, solche Befehle stehen im Widerspruch zum ersten Gesetz.
 3. Ein Roboter muss seine eigene Existenz schützen, solange dieser Schutz nicht dem Ersten oder Zweiten Gesetz widerspricht.
- **Utilitarismus:** Maximierung des Gesamtnutzens
- **Kants kategorischer Imperativ:** „Handle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, dass sie ein allgemeines Gesetz werde.“

Ansätze der Moralimplementation



Ansätze der Moralimplementation



Ansätze der Moralimplementation

Welcher Ansatz ist am Besten?

- Abhängig vom **Zweck** und **Nutzungskontext** eines künstlichen Systems

Pro und contra moralische Maschinen

Möglichkeiten/Chancen

Grenzen/Risiken

Pro und contra moralische Maschinen

Möglichkeiten/Chancen

- Komplexe und autonome Maschinen müssen moralisch handeln können

Grenzen/Risiken

Pro und contra moralische Maschinen

Möglichkeiten/Chancen

- Komplexe und autonome Maschinen müssen moralisch handeln können
- Maschinen, die ethischen Standards entsprechen, sind bessere Maschinen

Grenzen/Risiken

Pro und contra moralische Maschinen

Möglichkeiten/Chancen

- Komplexe und autonome Maschinen müssen moralisch handeln können
- Maschinen, die ethischen Standards entsprechen, sind bessere Maschinen
- **Verbesserung unserer moralischen Reflexion**

Grenzen/Risiken

Pro und contra moralische Maschinen

Möglichkeiten/Chancen

- Komplexe und autonome Maschinen müssen moralisch handeln können
- Maschinen, die ethischen Standards entsprechen, sind bessere Maschinen
- Verbesserung unserer moralischen Reflexion
- **Besseres Verständnis des Funktionierens moralischer Fähigkeiten beim Menschen**

Grenzen/Risiken

Pro und contra moralische Maschinen

Möglichkeiten/Chancen

- Komplexe und autonome Maschinen müssen moralisch handeln können
- Maschinen, die ethischen Standards entsprechen, sind bessere Maschinen
- Verbesserung unserer moralischen Reflexion
- Besseres Verständnis des Funktionierens moralischer Fähigkeiten beim Menschen

Grenzen/Risiken

- Grundsätzliche Einwände gegen die Realisierbarkeit moralischer Maschinen

Pro und contra moralische Maschinen

Möglichkeiten/Chancen

- Komplexe und autonome Maschinen müssen moralisch handeln können
- Maschinen, die ethischen Standards entsprechen, sind bessere Maschinen
- Verbesserung unserer moralischen Reflexion
- Besseres Verständnis des Funktionierens moralischer Fähigkeiten beim Menschen

Grenzen/Risiken

- Grundsätzliche Einwände gegen die Realisierbarkeit moralischer Maschinen
- **Ethischer Pluralismus**

Pro und contra moralische Maschinen

Möglichkeiten/Chancen

- Komplexe und autonome Maschinen müssen moralisch handeln können
- Maschinen, die ethischen Standards entsprechen, sind bessere Maschinen
- Verbesserung unserer moralischen Reflexion
- Besseres Verständnis des Funktionierens moralischer Fähigkeiten beim Menschen

Grenzen/Risiken

- Grundsätzliche Einwände gegen die Realisierbarkeit moralischer Maschinen
- Ethischer Pluralismus
- **Unvorhergesehene negative Resultate**

Pro und contra moralische Maschinen

Möglichkeiten/Chancen

- Komplexe und autonome Maschinen müssen moralisch handeln können
- Maschinen, die ethischen Standards entsprechen, sind bessere Maschinen
- Verbesserung unserer moralischen Reflexion
- Besseres Verständnis des Funktionierens moralischer Fähigkeiten beim Menschen

Grenzen/Risiken

- Grundsätzliche Einwände gegen die Realisierbarkeit moralischer Maschinen
- Ethischer Pluralismus
- Unvorhergesehene negative Resultate
- **Moralische Entmündigung des Menschen**

Pro und contra moralische Maschinen

Möglichkeiten/Chancen

- Komplexe und autonome Maschinen müssen moralisch handeln können
- Maschinen, die ethischen Standards entsprechen, sind bessere Maschinen
- Verbesserung unserer moralischen Reflexion
- Besseres Verständnis des Funktionierens moralischer Fähigkeiten beim Menschen

Grenzen/Risiken

- Grundsätzliche Einwände gegen die Realisierbarkeit moralischer Maschinen
- Ethischer Pluralismus
- Unvorhergesehene negative Resultate
- Moralische Entmündigung des Menschen
- **Grundlegender Bruch mit unserer moralischen Praxis (z.B. der Verantwortungszuschreibung)**

Maschinenethik und Demokratie

- In welchen Bereichen und auf welcher ethischen Grundlage sollten Maschinen entscheiden?
- Diskussion in einer breiteren Öffentlichkeit als Grundlage demokratisch legitimierter Regelungen



Universität Stuttgart

Vielen Dank!



Prof. Dr. Catrin Misselhorn

E-Mail catrin.misselhorn@philo.uni-stuttgart.de

Telefon +49 (0) 711 685-

Fax +49 (0) 711 685-

Universität Stuttgart
Institut für Philosophie
Seidenstr. 36

70174 Stuttgart

[Was bedeutet das alles?]

Catrin Misselhorn

Maschinen mit Moral?

Grundfragen der Maschinenethik

Reclam

