

Artificial Moral Agents (AMAs)

Dr. Michael Nagenborg
Institut für Philosophie, Uni Karlsruhe
Mitarbeiter EU-Projekt „Ethicbots“

www.michaelnagenborg.de

Emerging Technoethics of Human Interaction with Communication, Bionic, and Robotic systems

Emerging Technoethics of Human Interaction
with Communication, Bionic, and Robotic systems

Artificial Moral Agents (AMAs)

Ursprung: „Machine Ethics“

Allen, Colin; Wendell Wallach; Iva Smit: *Why Machine Ethics?*
In: IEEE Intelligent Systems, July/August 2006, 12-17.

Worum es nicht geht...

- Künstliche Personen
- „Böse“ oder „gute“ Programme (Floridi)

Was ist ein „Artificial Agent“?

Programm

- Werkzeug
- Nutzung
- Interaktiv

Software Agent

- autonomer Partner
- Sozial
- Delegation
- Pro- und reaktiv

„Autonom?“

- Selbstkontrolle und -steuerung
- Ergebnisse nicht vorhersehbar, aber *nicht* willkürlich (regelgeleitet)

Beispiel:

Google

Autonomie

- Vollständige autonom (= Person)
- ...
- **Hinsichtlich einiger Aspekte autonom**
- ...
- „tut was selbstständig“

Autonomie

- Vollständige autonom (= Person)
- ...
- **Hinsichtlich einiger Aspekte autonom**
- ...
- „tut was selbstständig“

Artificial Moral Agents (AMAs)

- AA
- „moralische Normen“ (i. w. S.)
- Kann diese nicht ändern

Beispiel:

Robocop

Beispiel:

Google.de

Beispiel: Privacy

- „Privacy“ wird interaktiv ausgehandelt.
- „Privacy“ wird durch soziale Normen geschützt.
- Diese Normen sind „moralisch“ (i. w. S.).

Beispiel:

Zwei Personen sprechen mit einander.
Dritte kommt hinzu.

Kriterien der Beurteilung

- Kontext des Gespräches

Kriterien der Beurteilung

- Kontext des Gespräches
 - Konferenz?
 - Am Arbeitsplatz?
 - In der Öffentlichkeit? (Welche Öffentlichkeit?)

Kriterien der Beurteilung

- Kontext des Gespräches
- Beziehung der Personen untereinander
- Kultureller Background
- Details der Interaktion etc.

Kurzum: Ganz schön kompliziert das Ganze...

Kurze Geschichte des „Datenschutz“

- Vom „right to be let alone“ zum „Datenschutz“
 - Verhältnis: Bürger / Staat
 - Schutz von „sensiblen Daten“
 - Mythos der trivialen Daten

Kurze Geschichte des „Datenschutz“

- „Datenschutz“ und „e-Commerce“
 - Verhältnis: Konsumenten / Firmen
 - Vertrauenswürdiger Umgang mit „sensiblen Daten“
 - Rechtlicher Datenschutz
 - Technischer Datenschutz (z. B. P3P)

Kurze Geschichte des „Datenschutz“

- P3P (www.w3c.org/p3p)
 - Maschinenlesbar
 - Entlastung der Nutzer (Lessig)
 - Privacy as Property (Lessig)
 - Keine Integration von Mechanismen zur Kontrolle und Durchsetzung der Regeln

AMAs und PETs

- Semantic Web (-> Agenten)
- Policy Aware Web
- Privacy and Identity Management for Europe
 - Schwerpunkt: E-Commerce
 - Agentenbasiert
 - Beachtet z. B. EU-Richtlinien (Kontrolmechanismen)

AMAs und PETs

Fragen:

- Wie zuverlässig sind die „Agenten“?
- Welcher „Moral“ folgen sie?
- Wie werden Personen im Netz abgebildet?
- ... und wer entscheidet darüber?

Kurzum:

- Ich bin mißtrauisch...
- ... und ihr?