Big Data, AI, and Ethics

Andreas Jedlitschka
Fraunhofer IESE
Fraunhofer-Platz 1
67663 Kaiserslautern
Germany

ICA Rectors and Deans Forum
Wageningen, NL
25. October 2018

Pixabay.com

Digital Transformation

- Physical objects (things, living objects, people)

  - produce data

  - have a history

  - are influenced

    - context-dependent

    - location-aware

    - in real time

  Yesterday: limited integration

  Today: connected across all boundaries

[Picture from http://b-metro.com, Cheri Ellis]

Increasingly automated data collection, analysis, and decision making

Fraunhofer

IESE

Today's Applications

- Customer understanding, clustering, classification, treatment
- Behavioral analysis (speech, face, activity)
- Optimize (automatize) processes and supply chains

- Hyper-personalization, micro targeting, content creation
- Social media monitoring, control influencer, build opinion
- Diagnosing and predicting diseases

- Predict, monitor, control, actively influence behavior
- Predictive law enforcement
- Active GPS and autonomous driving, precision farming

All Pictures from Pixabay.com

Fraunhofer

IESE

Future

New Development Paradigms

Transfer Learning

Simulated World

Emotional Intelligence

Augmented Human

Evolutionary Algorithms

Autonomous Co-Existence

AGI & Singularity

Pixabay.com

Fraunhofer
IESE

"reward principle"?!?
opt-in vs. opt-out

DriveSure: Black box car insurance from Churchill

The DriveSure plug-in is Churchill's black box insurance offering. Also known as telematics insurance, it lets us understand how you drive and helps you save money.

Fit our DriveSure box into your car and we'll give young drivers (under 26) with a full driving licence an upfront discount. Using a technology called telematics, we can analyse how you drive and offer you a car insurance premium based on what you do – or don't do – behind the wheel.

The safer you are, the more you'll save at your next renewal.

DriveSure

DriveSure Plug-in

DriveSure Portal

Good to know

How does DriveSure work?

https://www.churchill.com/car-insurance/black-box-original

Fraunhofer
IESE

Laws of robotics

1. A **robot** may not injure a human being or, through inaction, allow a human being to come to harm.

2. A **robot** must obey orders given it by human beings except where such orders would conflict with the First Law.

3. A **robot** must protect its own existence as long as such protection does not conflict with the First or Second Law
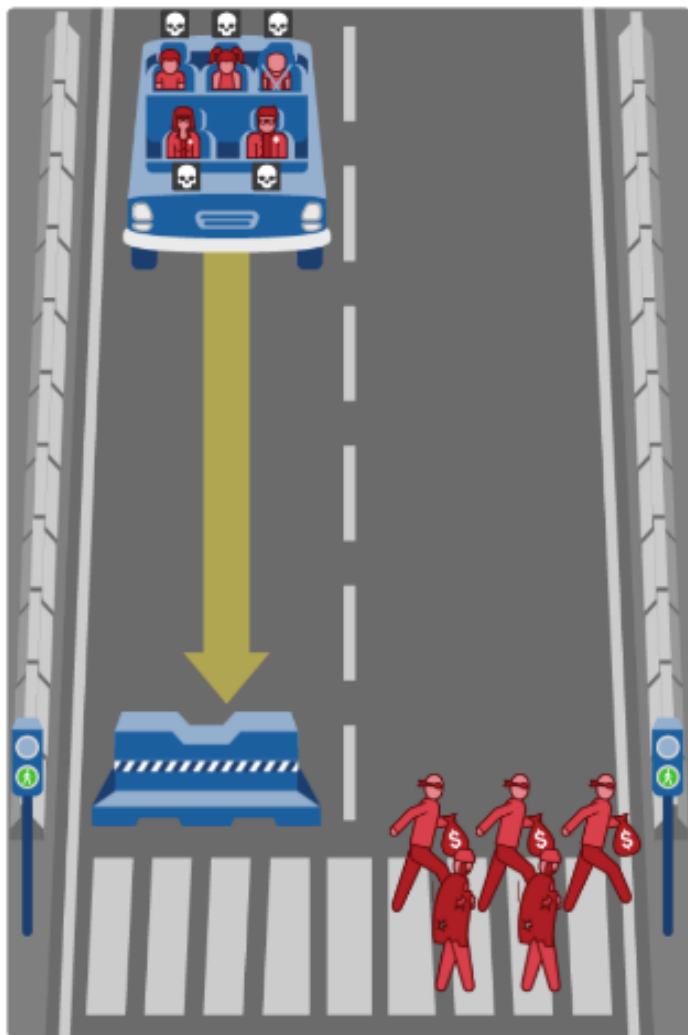
© Fraunhofer IESE

Pixabay.com

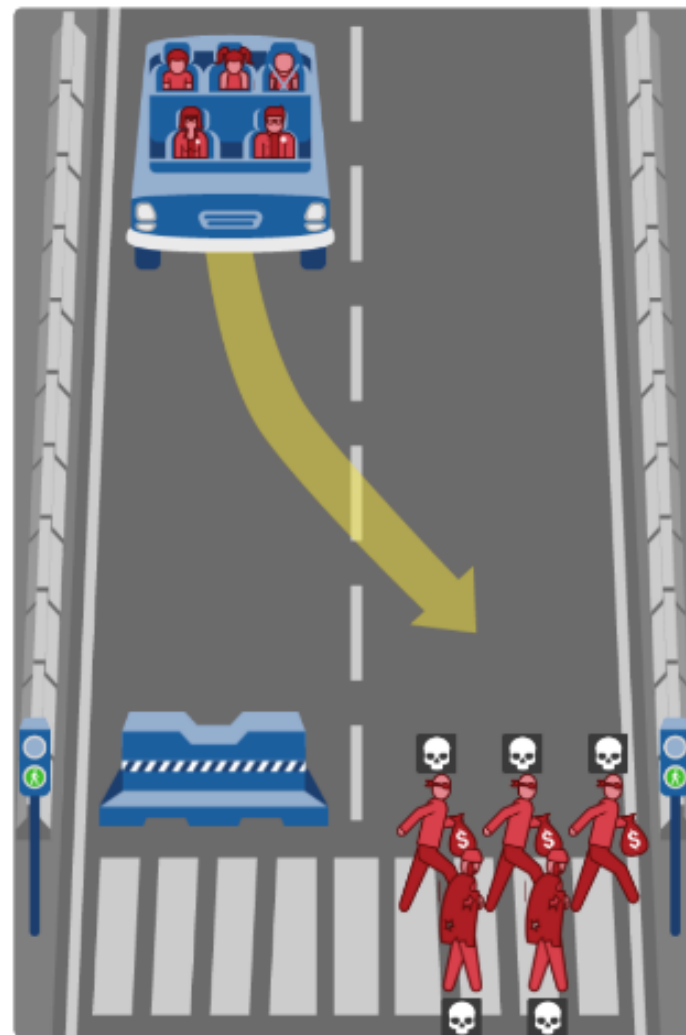# MORAL MACHINE

Home   Judge   Classic   Design   Browse   About
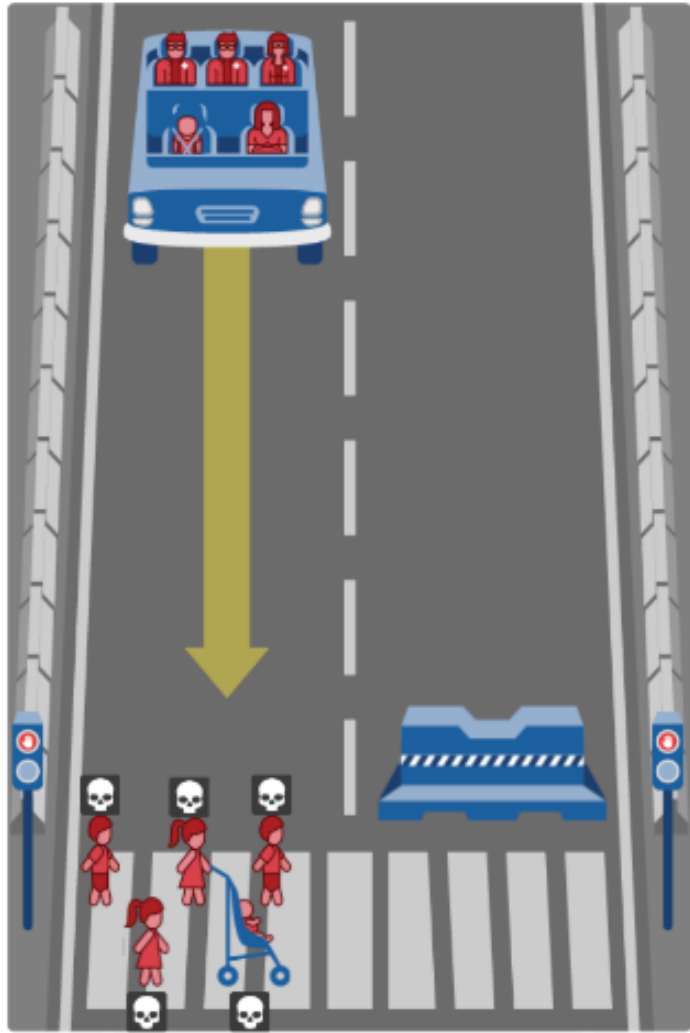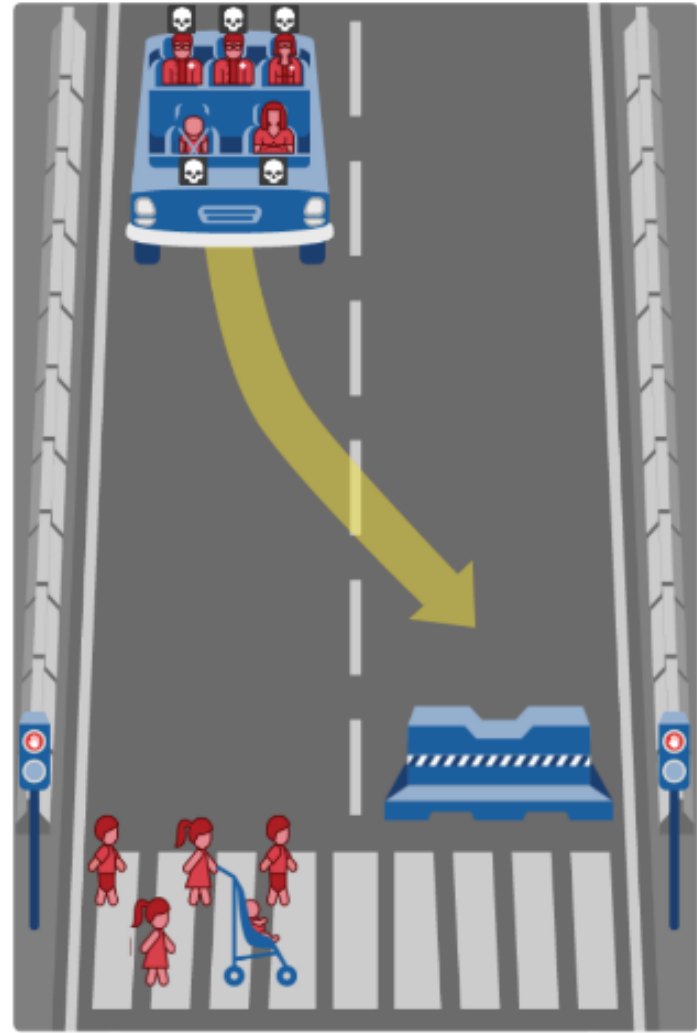
Show Description   Show Description

MORAL MACHINE

Home   Judge   Classic   Design   Browse   About

Show Description

Show Description

The Dark Side

Correlation vs. Causality

Shady Data & Black Box

Biased Decisions

Dual-use

Endangered Individual Diversity

This is now AI's job

Human vs. AI
White&Black viewpoints

Human decisions

- Biased
- Same attributes, different outcome
- Failures are somehow expected
- Human is accountable
- Tend to trust other humans
- limited information asymmetry
- Does not scale
- Very good in abstraction
- Good in forgetting
- Adapt to what is expected

Artificial Intelligence decisions

- Biased
- Same attributes, same outcome
- Failures are not accepted
- AI is held accountable, but why?
- We have limited trust in machines
- Asymmetry of information
- "Infinite" scalability
- Abstraction is challenging
- Never forgets …
- Adapt to …

Fraunhofer
IESE

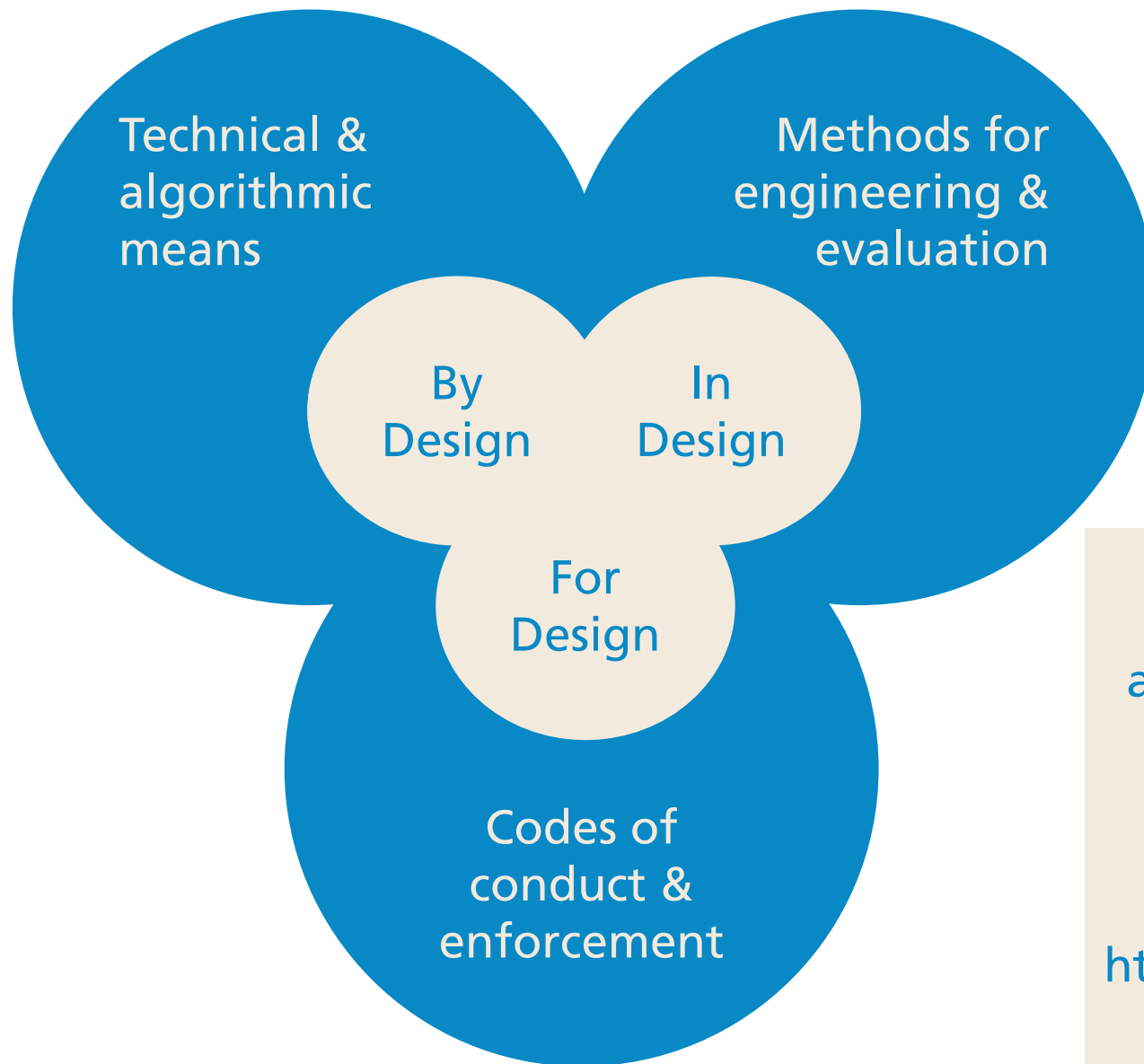# Challenges

Globalization

Tec. Advance

Influencer

Objectives

Quality of life

Background

© Fraunhofer IESE

Pixabay.com

Ethics in the AI lifecycle

Technical & algorithmic means

Methods for engineering & evaluation

By Design

In Design

For Design

Codes of conduct & enforcement

"It should not be a question of regulating artificial intelligence as a technology, but of controlling what people make with it to society and others."
https://algorithmwatch.org/en/dont-fear-ai

Picture adapted from text in: Virginia Dignum, Ethics in artificial intelligence: introduction to the special issue, Ethics and Information Technology (2018) 20:1-3 https://doi.org/10.1007/s10676-018-9450-z

Fraunhofer
IESE

The bad news

Science News
from research organizations

Code of ethics doesn't influence decisions of software developers

Date: October 9, 2018

Source: North Carolina State University

Summary: The world's largest computing society, ACM, updated its code of ethics in July 2018 -- but new research shows that the code of ethics does not appear to affect the decisions made by software developers.

https://www.sciencedaily.com/releases/2018/10/181009113617.htm

1,133 views | Oct 17, 2018, 06:13pm

Forbes

We Need To Work Harder To Make Software Engineering More Ethical

Jessica Baron Contributor ⓘ
Consumer Tech

https://www.forbes.com/sites/jessicabaron/2018/10/17/we-need-to-work-harder-to-make-software-engineering-more-ethical/#6f9a717a50cc

What we consider as being normal today becomes hard to be changed in future.

Fraunhofer
IESE

Key Message

We had better be quite sure that the purpose put into the machine is the purpose which we really desire.
- Norbert Wiener (1960)

- Is there a common goal where we want to evolve as human beings??

- Who decides what is right or wrong?

- Society is too complex to be predicted by algorithms (for now)

- Learning from mistakes might not be possible anymore (Think first) [Tegmark]

- Raise awareness, and teach accountable and transparent research

- Foster soft skills, team work, co-evolution with technology

- Limit (at least unintentional) unethical behavior of developers & users

- Philosophy, ethics, human/animal rights as obligatory parts of education

Fraunhofer
IESE

Time for Questions

Andreas Jedlitschka
Fraunhofer IESE
Kaiserslautern

andreas.jedlitschka@iese.fraunhofer.de

Some points to think about

- What does it mean for an AI system to make a decision?

- What are the incentives for an AI system?

- What are the moral, societal and legal consequences of AI decisions?

- Can an AI system be held accountable for its actions?

- How can we balance the need for efficiency and exploration with fairness and sensitivity to users?

- How can these systems be controlled once their learning capabilities bring them into states that are possibly only remotely linked to their initial, designed, setup?

- Should such autonomous innovation in commercial systems even be allowed, and should use and development be regulated?

Do you hear those questions for the first time?

Fraunhofer
IESE

General Ethical Principles

[ACM] A computing professional should...

- 1.1 Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing.

- 1.2 Avoid harm.

- 1.3 Be honest and trustworthy.

- 1.4 Be fair and take action not to discriminate.

- 1.5 Respect the work required to produce new ideas, inventions, creative works, and computing artifacts.

- 1.6 Respect privacy.

- 1.7 Honor confidentiality.

Fraunhofer

IESE

Ethical Initiatives (excerpt)

- ACM Code of Ethics and Professional Conduct

- IEEE initiative on Ethics of Autonomous Systems

- Foundation for Responsible Robotics

- Partnership on AI

- Microsoft FATE: Fairness, Accountability, Transparency, and Ethics in AI

- Google AI principles published after complains from thousands of employees

- OpenAI

- Future of Life Institute

- …

Selected References

[ACM] ACM Code of Ethics and Professional Conduct. Available online at https://www.acm.org/about-acm/acm-code-of-ethics-and-professional-conduct, checked on 10/19/2018.

[IEEE] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, Version 2. IEEE, 2017. http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html.

Bonnemains, Vincent; Saurel, Claire; Tessier, Catherine (2018): Embedded ethics: some technical and ethical challenges. In *Ethics Inf Technol* 20 (1), pp. 41–58. DOI: 10.1007/s10676-018-9444-x.

Herschel, Richard; Miori, Virginia M. (2017): Ethics & Big Data. In *Technology in Society* 49, pp. 31–36. DOI: 10.1016/j.techsoc.2017.03.003.

Selinger, Evan; Hartzog, Woodrow (2016): Facebook's emotional contagion study and the ethical problem of co-opted identity in mediated environments where users lack control. In Research Ethics 12 (1), pp. 35–43. DOI: 10.1177/1747016115579531.

Shapiro, Aaron (2017): Reform predictive policing. In *NATURE* 541 (7638), pp. 458–460. DOI: 10.1038/541458a.

Dignum, Virginia (2018): Ethics in artificial intelligence: introduction to the special issue. In *Ethics Inf Technol* 20 (1), pp. 1–3. DOI: 10.1007/s10676-018-9450-z.

Moral Code: The Ethics of AI. Available online at https://www.theatlantic.com/sponsored/hpe-2018/the-ethics-of-ai/1865/, checked on 8/1/2018.

Stuart Russell https://www.ted.com/talks/stuart_russell_3_principles_for_creating_safer_ai#t-233894

Max Tegmark https://www.ted.com/talks/max_tegmark_how_to_get_empowered_not_overpowered_by_ai

Fraunhofer
IESE

smart farming

smart communication

smart economy

smart health

smart energy

smart logistics

smart work

Thank you

Andreas Jedlitschka

Fraunhofer

IESE