Industrial Artificial Intelligence and its Ethical and Governance Challenges

Shanghai, September 18, 2019

Prof. Dr. Ulrike Reisach (冉珊鹤)

Münchener Kreis for the Digital Transformation

Neu-Ulm University of Applied Sciences, Information Management Department

www.ulrike-reisach.de/en/profile/

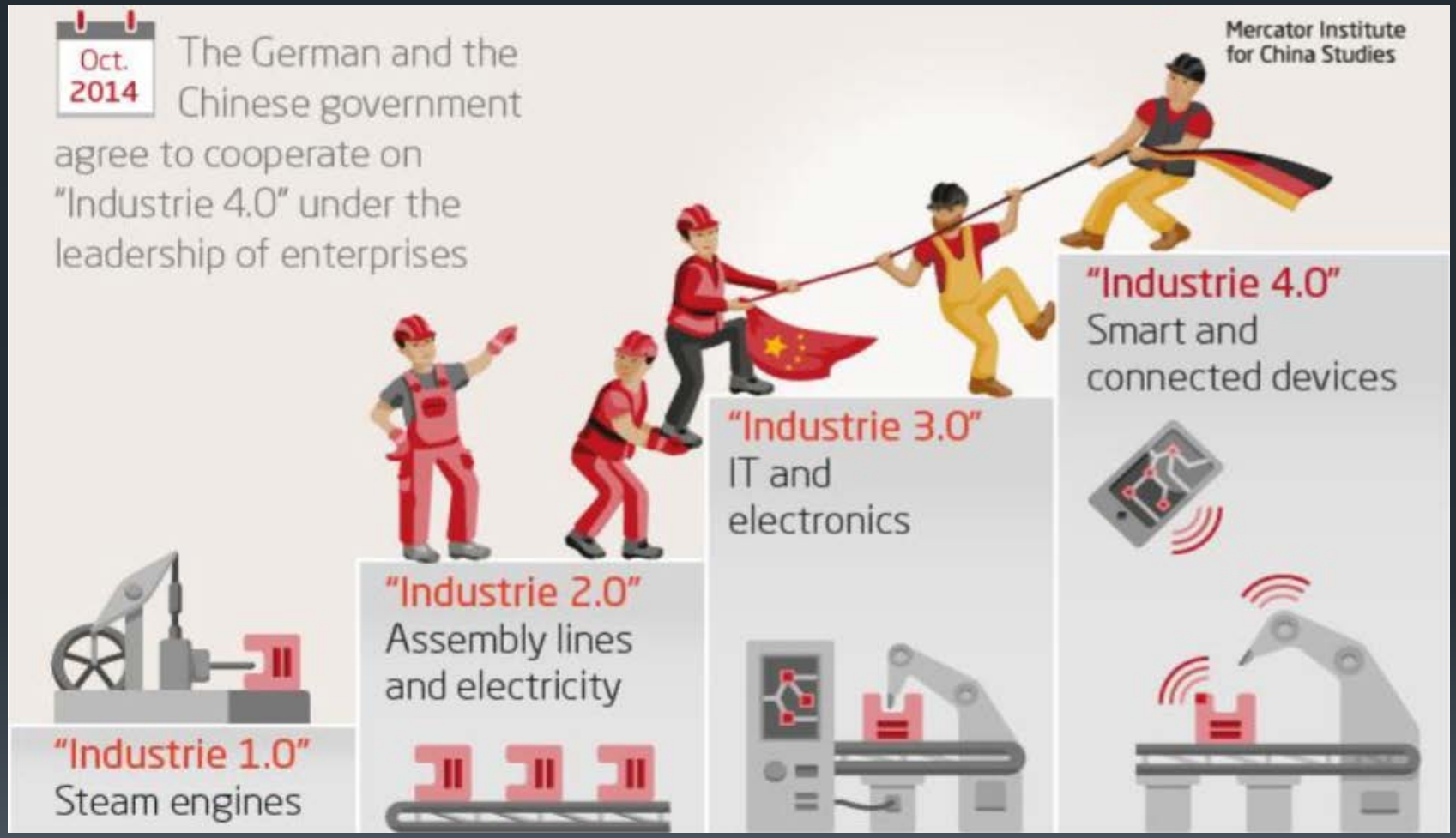
ulrike.reisach@hs-neu-ulm.de

Industrial Artificial Intelligence
and its Ethical and Governance Challenges

"Confronting quantitative methodologies with qualitative questions" (J. van Dijck, 2014)

1. **Industrial AI – Optimizing value added processes**
2. **Key ethical issues of Industrial AI**
3. **Examples from different industries**
4. **Knowledge based on experience and effective teamwork**
5. **Ethics Codices for machine learning**

Industry 4.0



Source: Wübbeke, Jost und Conrad, Björn (2015) Industrie 4.0: Will German Technology Help China Catch Up with the West? In: Mercator Institute for China Studies: China Monitor, Issue 23, 14. April 2015. https://www.merics.org/sites/default/files/2017-09/China_Monitor_23_Industrie40_DE.pdf

Industry 4.0 and AI (1)

Optimizing value added processes / Increasing operative efficiency

Disruption	Explanation	Industry case
Big Data	Rise in data volumes, computational power, and **connectivity**; Cloud solutions; Horizontal and vertical **system and networks integration,** cross-company and **cross-border**	Collaborative **B2B Platforms**; **Sensors** for analyzing and optimizing processes, operating conditions and customer experience
Advanced Analytics and business-intelligence capabilities	Optimized **supply chain management** through forecasting Decrease of ramp up time and scrap rates, shorter development time and **less production cost** **IoT:** Devices and unfinished products are enriched with **embedded computing**	**Prediction and digital twins** e.g. for mainte-nance, performance, customer demand/needs/behaviors Savings through supply chain and stock management **Simulation/testing** of machine settings and product quality,new materials

Artificial Intelligence (AI) -What's in it for Germany and its Industrial Sector?, BCG: Embracing Industry 4.0 and Rediscovering and PWC (2017) Digital factories 2020 – shaping the future of manufacturing.

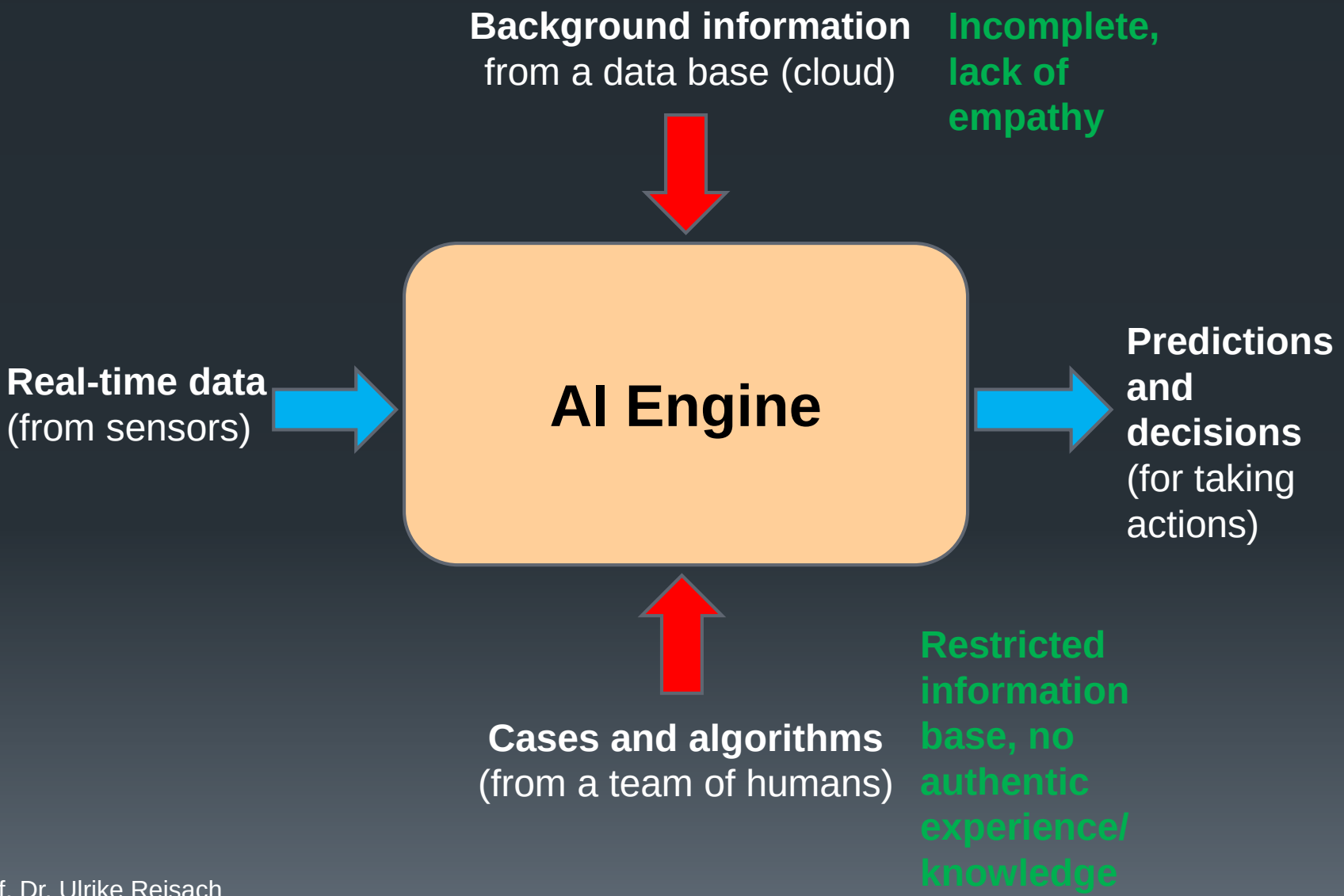
Industry 4.0 and AI (2)

Optimizing value added processes / Increasing operative efficiency

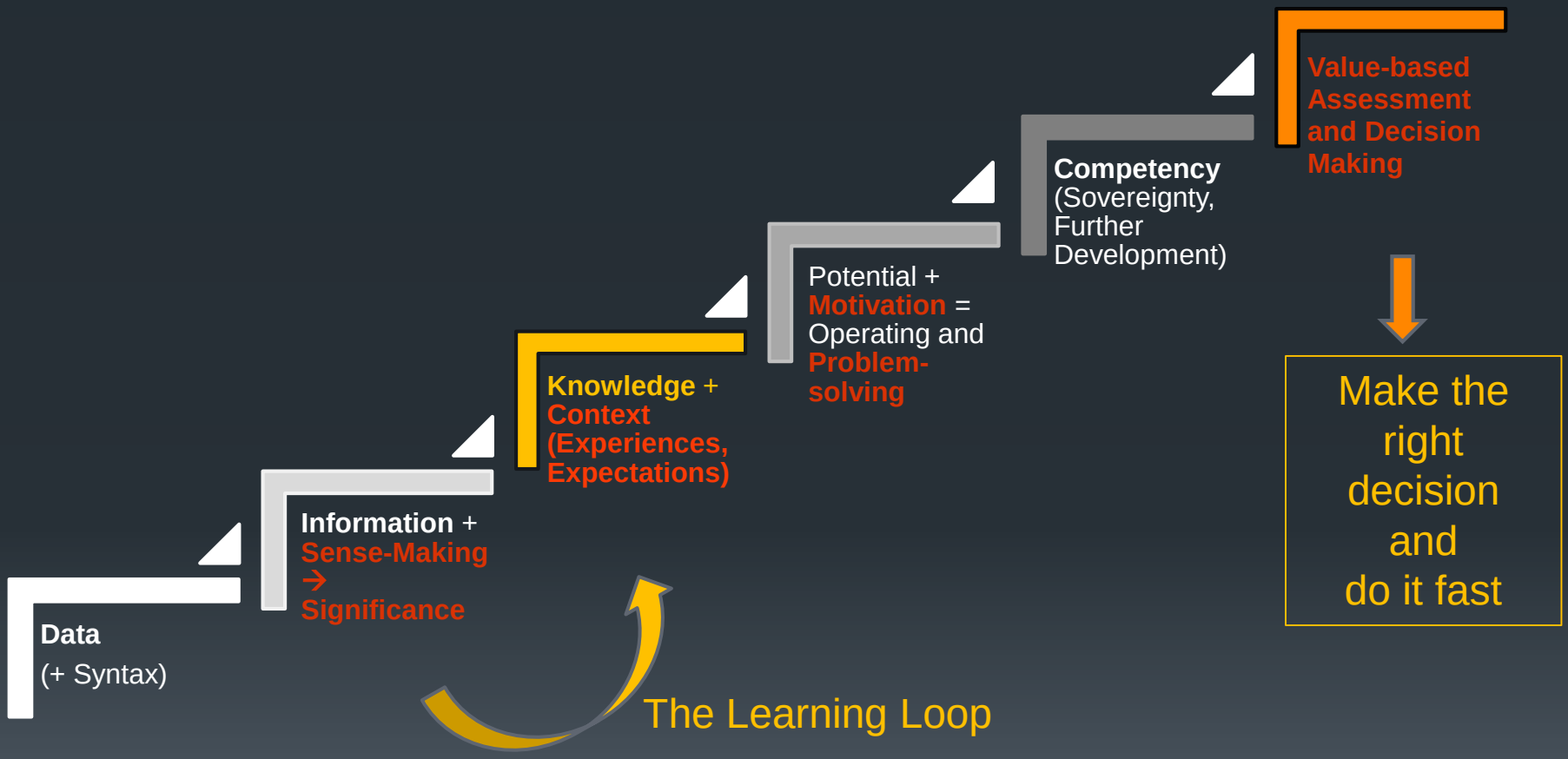
Disruption	Explanation	Industry case
Human-machine interfaces	New forms of human-machine interaction, e.g. through **Collaborative and context-aware robots** Machine Learning for **High-performance R&D projects**	Touch interfaces, Augmented-Reality Systems, **Autonomous agents,** e.g. (forklift) trucks, cobots, complex autonomous systems Time-to market acceleration
Digital-to physical transfer	Improvements in transferring digital instructions to the physical world: **Digital modelling and simulation of new materials and designs**	**Virtual factory planning,** Augmented Reality, Drones; **Advanced robotics;** Additive manufacturing, 3-D printing

UReisach based on Baur, C and Wee, D. (2015) Manufacturing‘s next act, in: McKinsey, June 2015,
Mc Kinsey (2017) Smartening up with Artificial Intelligence (AI) -What’s in it for Germany and its Industrial Sector?,
BCG: Embracing Industry 4.0 and Rediscovering
PWC (2017) Digital factories 2020 – shaping the future of manufacturing.

Artificial Intelligence - simplified



Knowledge Acquisition and Decision Making



In extention of Erlach, C./Orians, W./Reisach, U.: Wissenstransfer bei Fach- und Führungskräftewechsel, München 2013, Kap. 3, in Abwandlung von North, K:. Wissensorientierte Unternehmensführung. Wertschöpfung durch Wissen. Wiesbaden 2005

Key ethical issues of Industrial AI

1. **Purpose**

   "Is the purpose we put into the machine the purpose which we really desire?" (Norbert Wiener, 1940)
2. **Value alignment**
3. **The off switch problem**

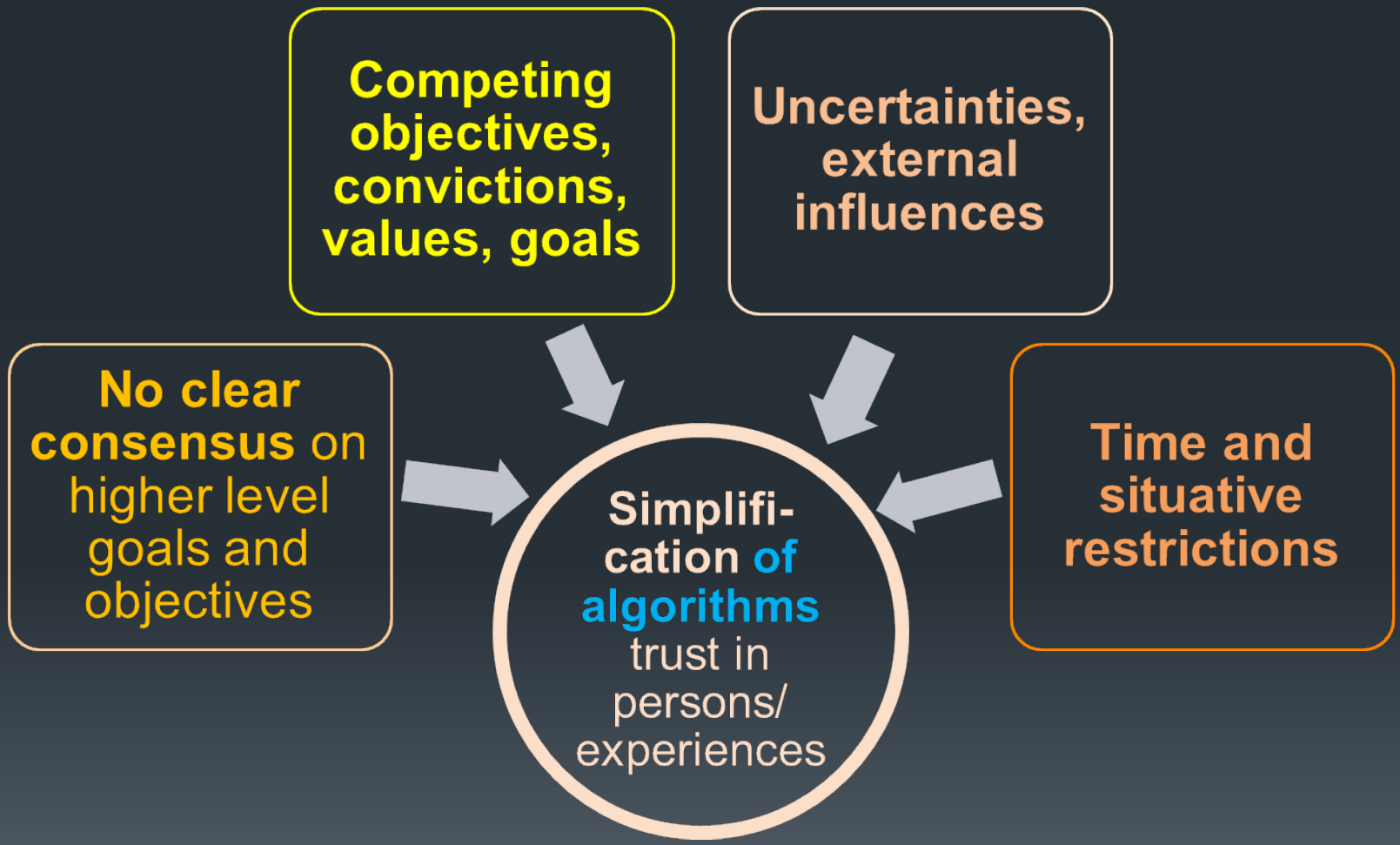
Consequences:

- Cooperative reinforcement learning better than copying human behavior
- Context is essential
- Ethical problems cannot be resolved by algorithms: dialogue required

Source for 1-3: Russell S.: Three principles of creating safer AI

The Psychological Side of Decision Making

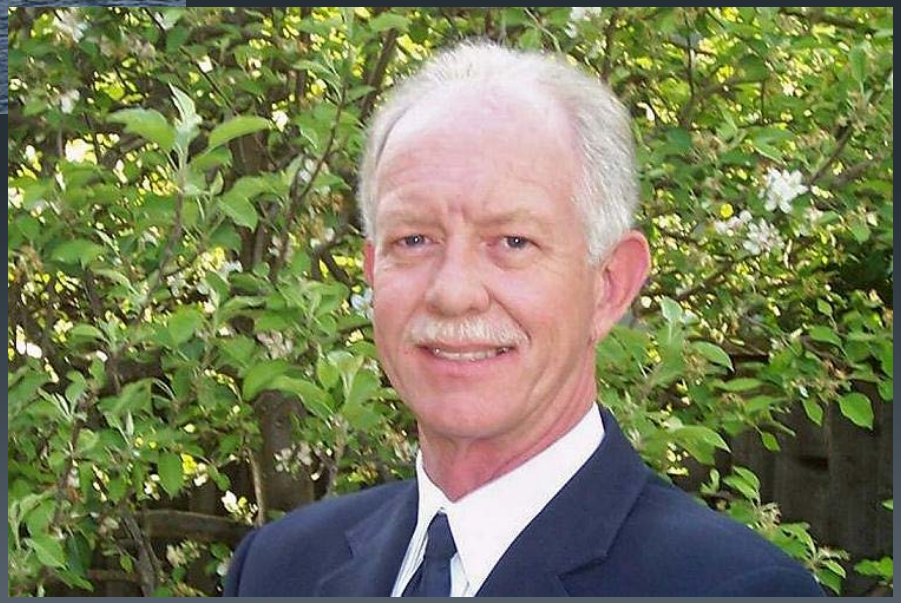
of **coders**, especially if they are acting within groups



Reisach, following ideas Roth, G (2017) Personality, decisions and behavior, pp. 117, Gigerenzer, G (2007) Gut decisions. The intelligence of the unconscious and the power of intuition, and Kleinberg, J. and Mullainathan S. (2019) Simplicity Creates Inequity.

Emergency landing on the Hudson River

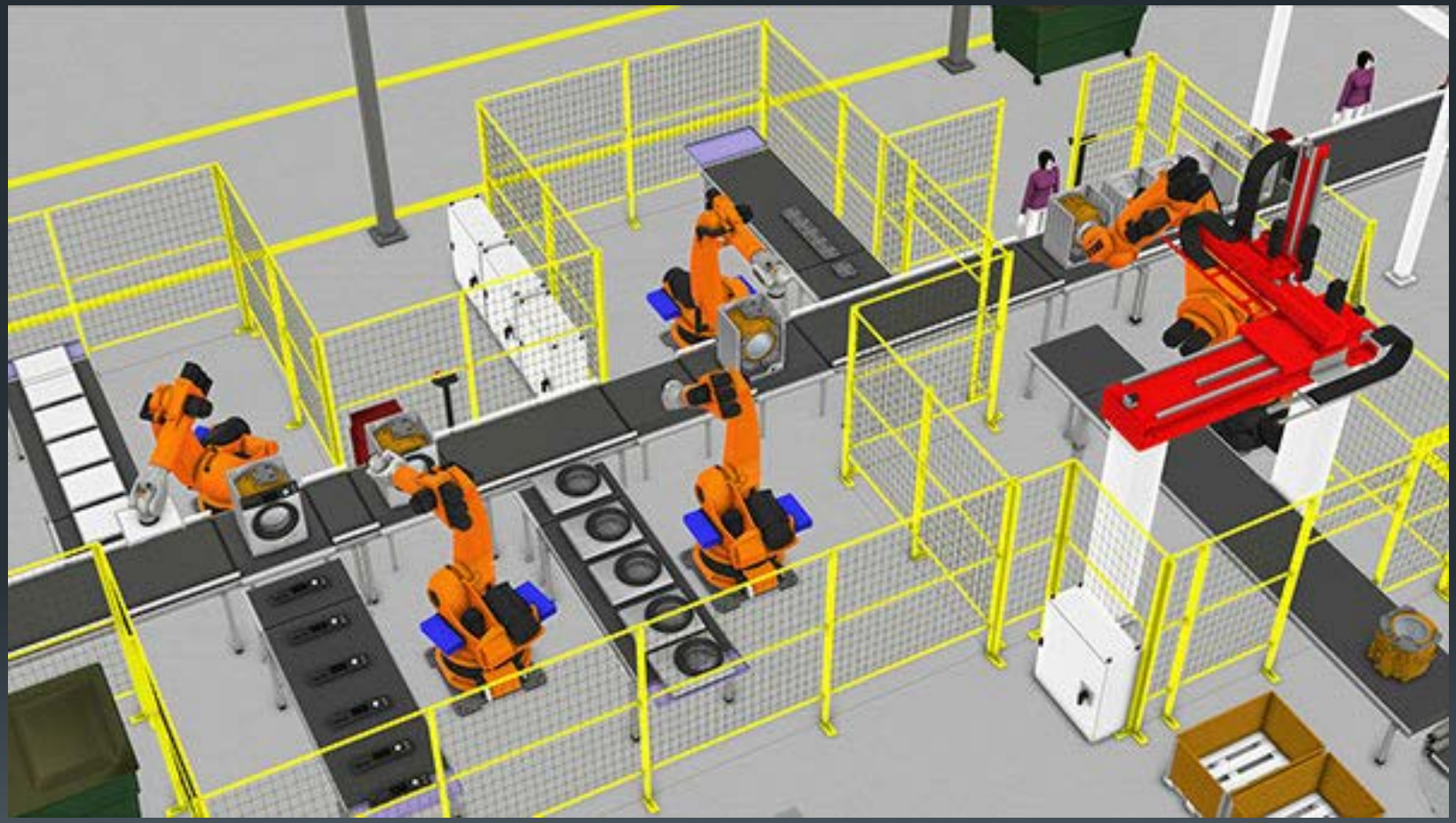
January 2009: Saving of 155 passengers after a collision with a bird swarm.



Pilot Chesley Sullenberger,
58 Jears old,
Pilot of the Airbus A 320

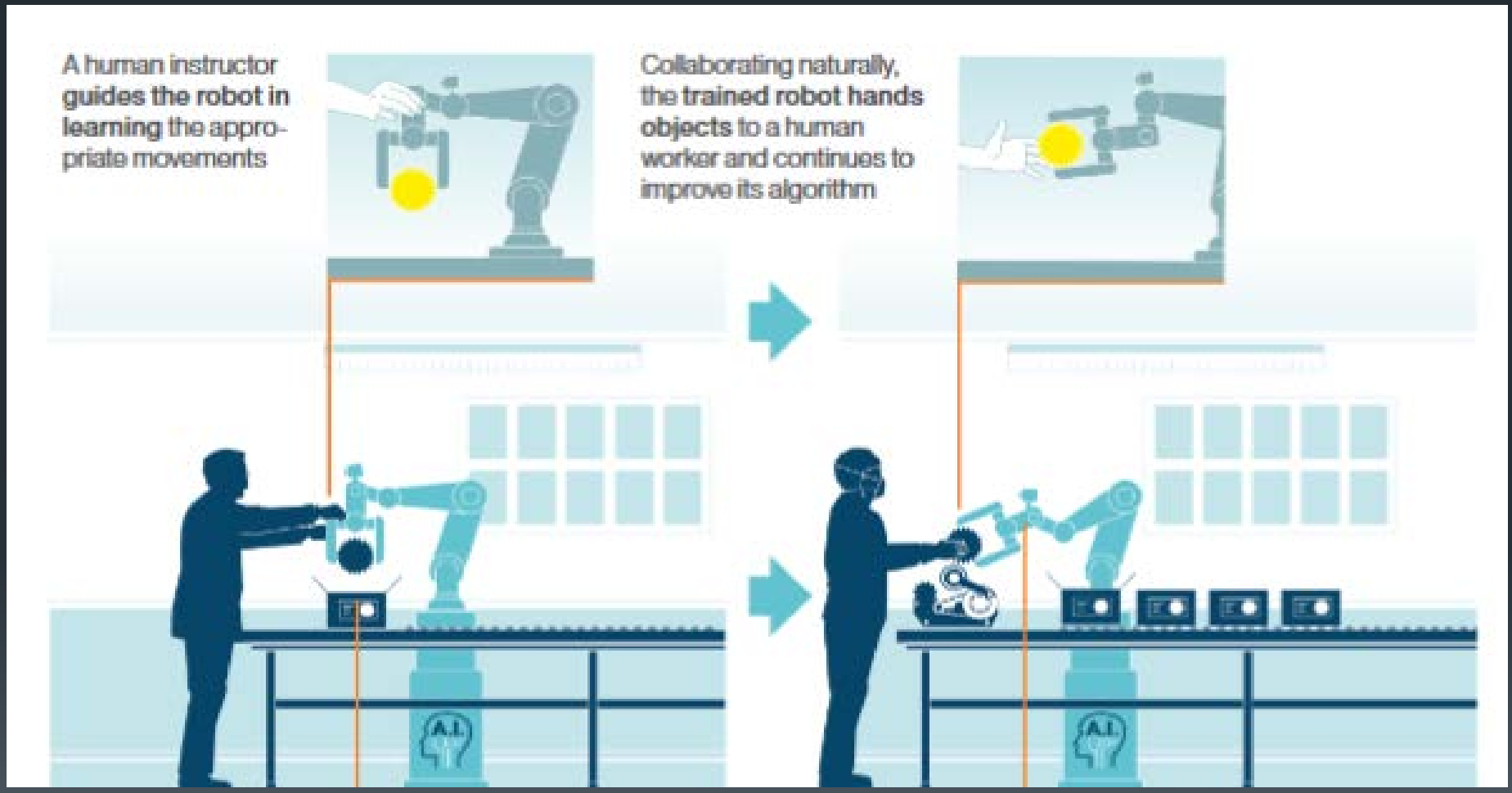
Bildquellen: Reuters nach Spiegel online, http://www.spiegel.de/panorama/flugzeugunglueck-in-new-york-airbus-notlandung-im-hudson-river-vor-manhattan-alle-passagiere-gerettet-a-601567.html , 25.04.2013

Simulation and modelling increases the capacity and flexibility of a high-end assembly line, while reducing costs by 15%



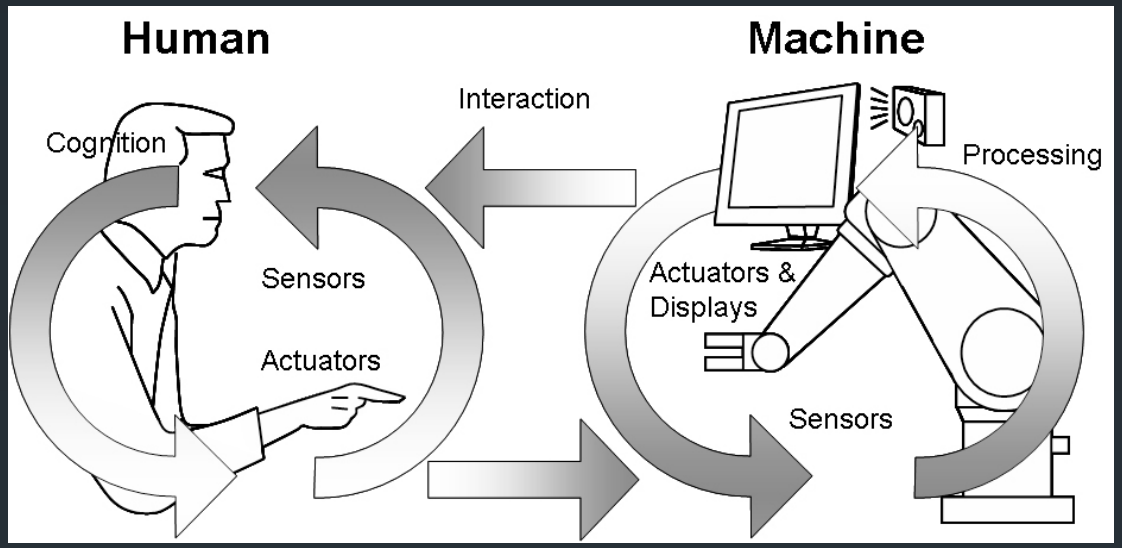
Source: Robotics.org, Case Study Midea, using Visual Components. https://www.robotics.org/content-detail.cfm/Industrial-Robotics-Case-Studies/Midea-Case-Study/content_id/8108

Co-operative reinforced learning



https://www.mckinsey.com/~/media/McKinsey/Industries/Semiconductors/Our%20Insights/Smartening%20up%20with%20artificial%20intelligence/Smartening-up-with-artificial-intelligence.ashx

Ethically sensitive: Human-Machine Interaction



1. Human-robot collaboration: mutual understanding?
2. Remote monitoring and control: accepted?
3. Digital HR and performance management/control
4. Automation of knowledge work: innovativity?
5. Sensors in the human and customer sphere

Safety, User-Friendliness and Ergonomics

MRK-Systeme GmbH, Augsburg, picture International Federation of Robotics https://ifr.org/case-studies/industrial-robots

Scientific Management Reloaded?

Scientific Management (Frederick Winslow Taylor, 1911) suggested that almost every act of the worker would have to be preceded by one or more preparatory acts of management, thus separating the planning of an act from its execution.

The problem:
Humans would be trained and guided as if they were machines.

Human-relations Theory (Elton Mayo)
"Instead of treating the workers as an appendage to 'the machine', the Hawthorne experiments brought to light ideas concerning motivational influences, job satisfaction, resistance to change, group norms, worker participation, and effective leadership.

→ Pay sufficient attention to people and the deep sentiments and relationships connecting them.

Sources: Taylor, FW: Scientific Management. London, Routledge 2003 https://doi.org/10.4324/9780203498569
Sonnenfeld, J.A.: Shedding Light on the Hawthorne Studies, Journal of Occupational Behavior, Vol. 6, 1985, p. 125

BMW FIZ Research and Engineering Center, Munich

… has been a pacesetter in the automobile industry with having all of those involved in product development across all areas of expertise are united under one roof. **Fostering innovation by involving employees.**

Source: BMW Blog, https://www.bmwblog.com/2017/10/07/bmw-investing-400-million-euros-expansion-fiz-munich/

What keeps societies/institutions/teams together?

1. Social Relations
- Social networking
- Trust in people
- Acceptance of diversity

2. (Personal) **Connectedness**
- Identification with the society
- Trust in institutions
- Perception of Fairness

3. Focus on the common good
- Solidarity and helpfulness
- Respect for societal rules
- Civic participation

Effective teamwork is the basis for most modern technical or engineering environments:

a. Cross-functional; complementary personal skills
b. Feedback and ideas, respect and mutual learning
c. Communication, trust and support
d. Collective skills, experience and knowledge
e. Problems are tackled faster

→ **Improved productivity**

Left side adapted from Walkenhost, P (2018) What Holds Asian SocietiesTogether? In: Bertelsmann Endowment, Guetersloh.
Right side: Reisach based on Wilson, J.R. & Whittington, C.M. AI & Soc (2001) 15: 58. https://doi.org/10.1007/BF01205738

Beijing AI Principles https://www.baai.ac.cn 2019/5/28

Research & Development

- Do Good, For Humanity, Be Responsible: conform to human values and the overall interests of humankind
- Control Risks: **safety for the external environment**
- Be Ethical: Improve **transparency**, explainability, predictability
- Be Diverse and Inclusive
- Open and Share

Use

- Use Wisely and Properly
- Informed Consent
- **Education and Training** („stakeholders shold be able to receive“)

Governance

- Optimizing Employment: Explorations on human-Ai coordination
- Harmony and Cooperation, Adption and Moderation
- Subdivision and Implementation: more specific guidelines
- Long-term planning, research on potential risks

OECD Principles on AI

1. AI should benefit people and the planet by driving **inclusive growth**, sustainable development and well-being.
2. AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, and they should include appropriate safeguards – for example, enabling human intervention where necessary – to **ensure a fair and just society.**
3. There should be **transparency** and **responsible disclosure** around AI systems to ensure that people understand AI-based outcomes and can challenge them.
4. AI systems must function in a robust, secure and safe way throughout their life cycles and **potential risks should be continually assessed and managed.**
5. Organizations and individuals developing, deploying or operating AI systems should be held **accountable** for their proper functioning in line with the above principles.

Source: OECD Principles on AI, Paris May 22, 2019. https://www.oecd.org/going-digital/ai/principles/

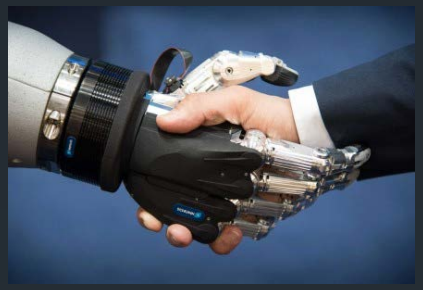
Reflectivity in algorithm design and implementation

Reflectivity is an approach close to critical thinking and consequentialism, deeply rooted in the humanities, especially in philosophy and psychology.

- **Interdisciplinary thinking, considering the contextual framework** such as cultural, political and societal conditions and short- plus long-term developments;
- **Social skills like empathy and compassion** (A. Schweizer) in order to identify with another person's situation and r**ecognizing human motivations**, rational or irrational interests;
- A transgression of the own self and an **interest in living together harmoniously: Consideration of the broader societal impact** (Confucius, Consequentialism, M. Weber);
- **Balancing legitimate self-interest and legitimate societal interest = corporate and societal governance,** no purely mathematical calculation or optimization (not like J.S. Mill's utilitarism).

Discussion: Autonomous driving, Medicine testing, Emergency rescue …

Thanks for your kind attention!



Prof. Dr. Ulrike Reisach (冉珊鹤)

www. ulrike-reisach.de/en/profile
www.ulrike-reisach.de/china
www.hs-neu-ulm.de/en/ulrike-reisach

ulrike.reisach@hs-neu-ulm.de

Cobot-Source: Deutsche Messe, July 06, 2019
Emperial Palace, Beijing: UReisach

References (A-I)

Aldewereld, H., Dignum, V., & Hua Tan, Y. (2014) Design for values in software development. In J. van den Jeroen, P. E. Vermaas, & I. van de Poel (Eds.), *Handbook of ethics, values, and technological design.* Dordrecht: Springer.

Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014) Power to the people: The role of humans in interactive machine learning. *AI Magazine, 35*(4), 105–120.

Baur, C and Wee, D. (2015) Manufacturing‘s next act, in: McKinsey, June 2015, , https://www.mckinsey.com/business-functions/operations/our-insights/manufacturings-next-act

BCG: Embracing Industry 4.0 and Rediscovering. https://www.bcg.com/capabilities/operations/embracing-industry-4.0-rediscovering-growth.aspx

Bergmann, Gustav/Daub, Jürgen: Systemisches Innovations- u. Kompetenzmanagement, 2. ed., Wiesbaden 2008

Brundage et al. (2018) The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, Oxford, February 2018. https://img1.wsimg.com/blobby/go/3d82daa4-97fe-4096-9c6b-376b92c619de/downloads/1c6q2kc4v_50335.pdf.

European Data Protection Supervisor: Declaration on Ethics and data protection in AI. https://edps.europa.eu/data-protection/our-work/subjects/artificial-intelligence_en

Floridi, L (2013): Distrubuted Morality in an Information society. Sci Eng Ethics 19-727-743, DOI: 10.1007/s11948-012-9413-4

IEEE (2018) Global Initiative on Ethics of Autonomous and Intelligent Systems

References (M-Z)

Mc Kinsey (2017) Smartening up with Artificial Intelligence (AI) -What's in it for Germany and its Industrial Sector?
https://www.mckinsey.com/~/media/McKinsey/Industries/Semiconductors/Our%20Insights/Smartening%20up%20with%20artificial%20intelligence/Smartening-up-with-artificial-intelligence.ashx

PWC (2017) Digital factories 2020 – shaping the future of manufacturing.
https://www.pwc.com/gx/en/industries/industrial-manufacturing/publications/digital-factories-2020.html

Rahwan, I (2018) Society-in-the-loop: programming the algorithmic social contract. In: Ethics Inf Technol (2018) 20: 5. DOI https://doi.org/10.1007/s10676-017-9430-8

Russell, JS and Norvig, P. (2018) Artificial Intelligence. A modern Approach, 3rd. ed., Chapter 26: Philosophical Foundations and 26.3: The Ethics and Risks of Developing AI (pp1034-1039)

Solms, M./Turnbull, O.: The Brain and the Inner World. An introduction to neuroscience and subjective experience, Karnac Books, London 2002

Standing Committee of the One Hundred Year Study of Artificial Intelligence (2016) Artificial intelligence and life in 2030. Technical report, Stanford University.

Van Dijck, J (2014) Datafication, dataism and dataveillance. Big Data between scientific paradigm and ideology. In: Surveillance & Society, 12(2): 197-208.
www.academia.edu/26648535/Datafication_dataism_and_dataveillance_Big_Data_between_scientific_paradigm_and_ideology

Zenglein, MJ and Holzmann (2019) Evolving Made in China 2025. China's industrial policy in the quest for global tech leadership. In: Mercator Institute for China Studies, Merics Papers on China 08/July 2019, Berlin
https://www.merics.org/sites/default/files/2019-07/MPOC_8_MadeinChina_2025_final_3.pdf