TMMB-YÖTG Endüstri 4.0 ve Robotik Seminerleri

Insan-robot etkilesimi ve etik

Seda Göksu

24 Subat 2022

Içerik

- Temel kavramlar
- Endüstri 4.0, dijitallesme ve insan
- Robotik teknolojiler ve insan
- Insan-robot etkilesimi (human-robot interaction / human-robot collaboration)
- Yapay zeka ve robotik sistemlerde etik degerler (AI, robotics and ethics)
- Robotik teknolojilerde neden etik uygulamalara ihtiyacimiz var?
- Sonuç

Temel Kavramlar

**Human-robot interaction (insan-robot etkilesimi)**

1. human-robot collaboration
2. decision-making of robotics (cognitive aspects of robotics)

AI and ethics (yapay zeka ve etik)

1. ethical values and AI (trustworthy AI)
2. ethics by design
3. responsible and Green AI

Multidisciplinary research
(*Psychology, sociology, philosophy, anthropology, science and technology studies*)

3
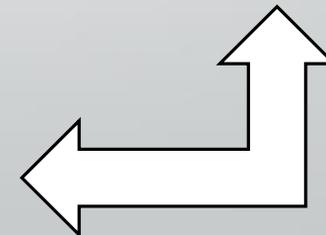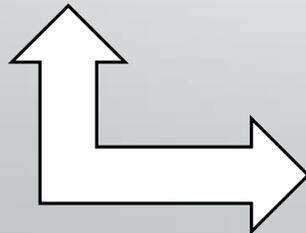
Endüstri 4.0, dijitallesme ve insan

- Advance production and operation services with digital technologies

- Production systems interact with their environment through data collection and information processing

- *Human factor* is one of the key elements in the whole process of digitization

- Sociotechnical system referring to social, technical, organizational and human elements
  (Weiss, A., Wortmeier, A.-K., & Kubicek, B. (2021). Cobots in Industry 4.0: A Roadmap for Future Practice Studies on Human–Robot Collaboration. IEEE Transactions on Human-Machine Systems, 51)

- New working ways and skills development for workers/employees

- A new generation of robots capable of performing all kinds of tasks based on learning capacity (cognitive aspects)

Robotik teknolojiler ve insan

- Robotization change organizations and jobs and, thus how employees/workers adapt to these changes

- Human-robot collaboration as a complex sociotechnical system

- How humans and robots mutually empathize for better communication in a natural way

- The expansion of cognitive abilities to robots. Recent research show that ***the future will be robots and humans working together*** *(The Impact of Robots on Productivity, Employment and Jobs positioning paper by the International Federation of Robotics, April 2017).*

Human-robot interaction

Human-centered aspects:

- manage safe, effective and pleasurable human-robot interaction

- cognitive and emotional aspects, empathy, intuitiveness, trustworthiness, privacy and security

- ethics in robotics technology design and responsive human-robot interfaces

- improve efficiency on collaborative tasks

Human-robot collaboration

How do we integrate human elements to the deployment of robotics technologies?

- *socially intelligent robots* to enable how these emerging technologies require understanding of how humans integrate visual and haptic information

- *human robotic collaborative tasks* (human robot intention awareness based on behavioral aspects (analysis on intuitiveness and responsive human-robot interfaces)

- human-robot collaboration in a processes of collaboration between robots and humans (*team dynamics perspective*)

- Implement social and cognitive psychology methods based on these behavioral aspects, *ethical elements and trustworthy concepts* (value-sensitive perspective on robotics technology design) of human-robot interaction

Use case 1 social robots

Use-case 1: humanoid robots for elderly care (assistive robots)

- smart solutions by developing a humanoid robot which serves as a personal assistant (care provider) for the elderly people to stimulate physical, cognitive activities and monitor health status of the elderly people or supports the care-givers

- **Implementation of human aspects**: inputs from social and cognitive psychology, identify *user needs and requirements* (e.g: there is a *vulnerable group* in this case), stimulate social interaction skills of robots, ethical and value-sensitive design (behavioral and social elements). It is essential to include moral values at the stage of design (*acceptability, trust of users and responsibility*)

Use case 2 collaborative robots

Use-case 2: collaborative robots in industry

- Support human-robot team, different types of risks in terms of security, ergonomic and psychosocial risks in production cycle requires *human-robot collaboration tasks*

- **Implementation of human aspects**: task sharing with a robot, effective execution of industrial activities and task completion, worker well-being (physically and mentally), adapting to a changing work environment in which people collaborate with robots (*from cooperation to collaboration based on common goals)*

Human-robot collaboration

Key points on human-robot collaboration

- Foster workers' willingness to collaborate with the robots
- Organizational context (work-related context) in human-robot collaboration
- Human-robot collaboration as a change process (before and after implementation)
- Human perception on robotic technologies (user needs and requirements / social acceptance / empathy / cognitive aspects)

AI and ethics

1.Ethical values and AI (trustworthy AI)

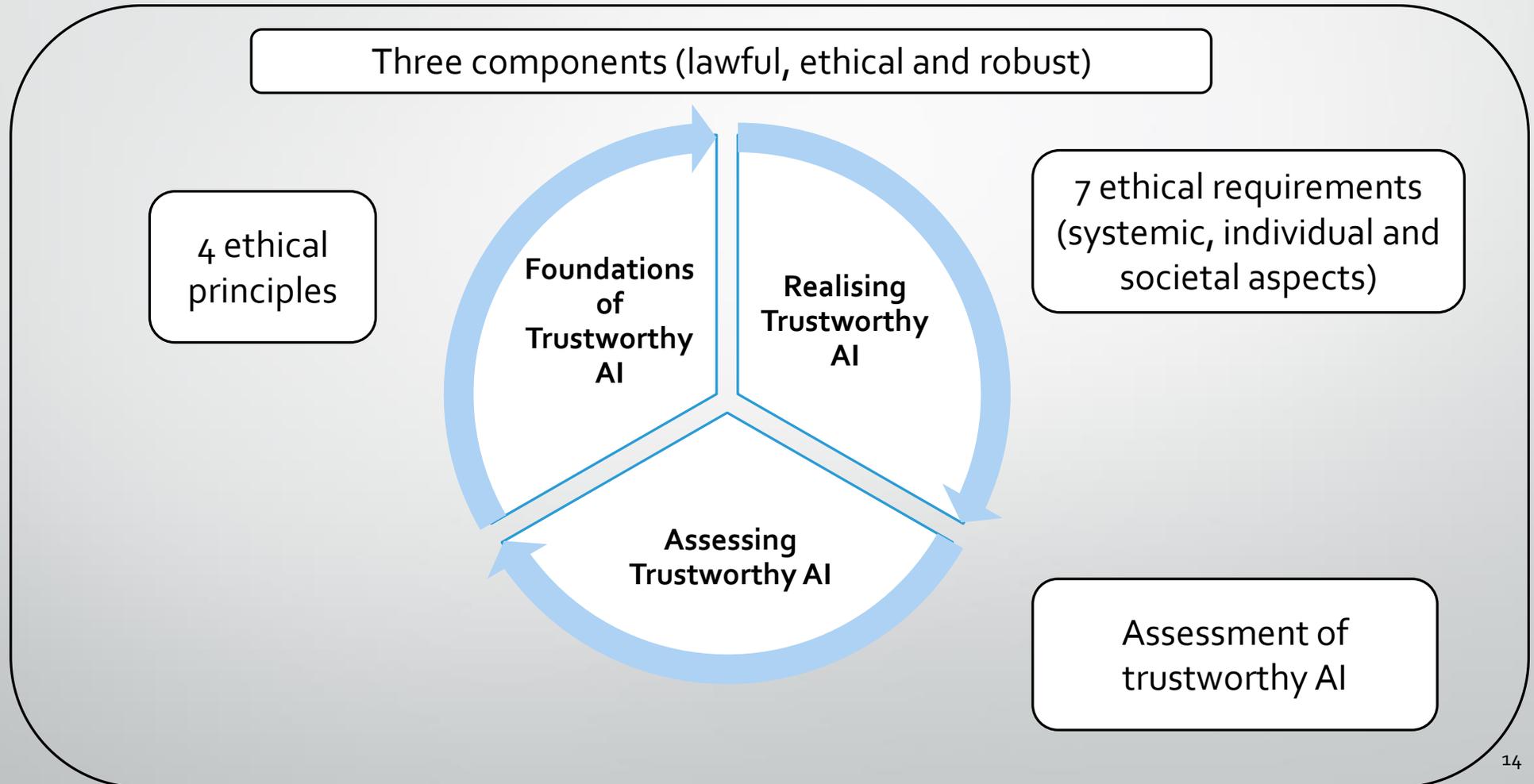2.Ethics by design

3.Responsible and Green AI

Ethical values

- **Ethics:** *referred to as the philosophical study of the concepts of moral right and wrong and any system or code of moral rules, principles, or values* (retrieved from https://www.britannica.com/topic/ethics-philosophy)

- Ethical values in AI*:* trustworthiness, accountability, explainability of AI, greener AI, societal and environmental well-being, social awareness, inclusiveness, gender and other forms of diversity, privacy, safety and reliability

- GDPR, open science and science integrity

- value-sensitive design approach and informed consent (explore moral values and principles)

Ethical values and AI

Different arguments on ethical principles/values and AI:

- Ethical principles: utilitarianism, diversity, autonomy and legacy (*Tegmark, Life 3.0, pp.269-279).*

- Ethics is not something that can be ticked off on a list (*Algorithm Watch, Germany, retrieved from* https://algorithmwatch.org/en/)

- Trustworthy AI applications should be developed and used in such a way that they respect *human autonomy, prevent harm, function fairly and intelligibly* (*EC Ethics Guideline for trustworthy AI).*

- Ethical analysis on the concepts of *agency, intelligence, responsibility, empathy and compassion* etc.

Trustworthy AI

Three components (lawful, ethical and robust)

4 ethical principles

Foundations of Trustworthy AI

Realising Trustworthy AI

Assessing Trustworthy AI

7 ethical requirements (systemic, individual and societal aspects)

Assessment of trustworthy AI

Source: *EC Ethics Guideline for trustworthy AI, retrieved from https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html)*

Trustworthy AI

- 3 components of trustworthy of AI (*lawful, ethical and robust)*

- 4 ethical principles *(respect for human autonomy, prevention of harm, fairness and explicability*)

- 7 requirements which refer to trustworthy AI in the context of AI systems' life cycle: (i) *human agency and oversight, (ii) technical robustness and safety, (iii) privacy and data governance, (iv) transparency, (v) diversity, non-discrimination and fairness, (vi) environmental and societal well-being and (vii) accountability* (EC Ethics Guideline for trustworthy AI, retrieved from https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html)

Implementation of trustworthy AI principles

- Involve all relevant stakeholders throughout the AI system's life cycle

- Trustworthy AI assessment list when developing, deploying or using AI systems, and adapt it to the specific use case in which the system is being applied

- Continuous process on the design, implementation and evaluation of solutions

- AI systems (***socio-technical systems***) should aim to improve human welfare and freedom

Use case on ethics and AI

Deployment of AR based technology:

- **Implementation of ethical values:** *value-sensitive co-design approach involving all relevant end-users (customers and producers), needs and requirements of users, trustworthy and accurate knowledge, informed consent of people who take part in the research, GDPR, privacy and integration of ethics by design approach from the beginning of the research to the implementation process (inclusion ethical principles of autonomy, accountability, acceptability and empathy)*

Ethical values in robotics

- User requirements and needs
- Ethical principles (human autonomy, empathy, social intelligence etc.) and value-sensitive design by technology
- Ethical Guidelines on robotics technologies (*Declaration of Helsinki)* and approval from Ethical Boards
- Social and behavioral model
- Trustworthiness, privacy, security and ethics in robotics technology design and intuitiveness and responsive human-robot interfaces

Social impact of AI and robotics technologies

Technology readiness level

- TRL 1 – Basic principles observed
- TRL 2 – Technology concept formulated
- TRL 3 – Experimental proof of concept
- TRL 4 – Technology validated in lab
- TRL 5 – Technology validated in relevant environment
- TRL 6 – Technology demonstrated in relevant environment
- TRL 7 – System prototype demonstration in operational environment
- TRL 8 – System complete and qualified
- TRL 9 – Actual system proven in operational environment

Societal readiness level

- SRL 1 – identifying problem and identifying societal readiness
- SRL 2 – formulation of problem, proposed solution(s) and potential impact, expected societal readiness; identifying relevant stakeholders for the project.
- SRL 3 – initial testing of proposed solution(s) together with relevant stakeholders
- SRL 4 – problem validated through pilot testing in relevant environment to substantiate proposed impact and societal readiness

A

- SRL 5 – proposed solution(s) validated, now by relevant stakeholders in the area
- SRL 6 – solution(s) demonstrated in relevant environment and in co-operation with relevant stakeholders to gain initial feedback on potential impact
- SRL 7 – refinement of project and/or solution and, if needed, retesting in relevant environment with relevant stakeholders

B

- SRL 8 – proposed solution(s) as well as a plan for societal adaptation complete and qualified
- SRL 9 – actual project solution(s) proven in relevant environment

C

A: problem-oriented R&D in the labs
B: test-redesign-validate with relevant stakeholders
C: actual use

Sonuç

- Technical and social impact of robotics technologies (technology readiness level + societal readiness level)

- Ethical principles and value-sensitive design approach

- Responsible research and innovation (Is your innovation good for society?)

- Multidisciplinary research / holistic approach on research and innovation activities

TESEKKÜRLER

Seda Göksu