UiO : Department of Informatics
University of Oslo

Tutorial: Ethical and Technical Challenges and Considerations of Robots and Systems

Session I: Ethical and technical perspectives

Jim Torresen
Research group Robotics and Intelligent Systems (ROBIN)
Department of Informatics and RITMO
University of Oslo, Norway
Email: jimtoer@ifi.uio.no

2022 IEEE/RSJ International Conference on Intelligent Robots and Systems ***(IROS-2022)***

IROS KYOTO 2022

UiO : University of Oslo

The Research Council of Norway

AINEWS

Zoom receives backlash for emotion-detecting AI

By Ryan Daws | May 19, 2022 | TechForge Media
Categories: Applications, Artificial Intelligence, Companies, Ethics & Society, Face Recognition, Privacy, Voice Recognition,

Zoom has caused a stir following reports that it's developing an AI system for detecting emotions.

The system, first reported by Protocol, claims to scan users' faces and their speech to determine their emotions.

- Scan users' faces and their speech to determine their emotions.
- Zoom detailed the system more in a blog post last month (April 2022). The company says 'Zoom IQ' will be particularly useful for helping salespeople improve their pitches based on the emotions of call participants.

Tutorial Agenda – Ethical perspectives and technical challenges and opportunities with care robots

1. Introduction and motivation
2. Ethical challenges and perspectives
3. How to address ethical considerations in research including examples
4. Ethical considerations in own research at University of Oslo
5. Future opportunities in ethics-related research

UiO : Department of Informatics
University of Oslo

Collaboration on Intelligent Machines Norway (COINMAC)

←→ US, Brazil and Japan

Project manager: Jim Torresen

- Short term mobility stay for meetings
- Sharing and development of curriculum and teaching material for courses
- Long term mobility stay (students)
- Intensive course/student workshop
- Workshop/conference org. activities
- New collaborations/guest lectures (can be in Canada, South-Korea, ++)

The Research Council of Norway

Funding: *INTPART 2017-2024*

Diku Norwegian Agency for International Cooperation and Quality Enhancement in Higher Education

https://www.mn.uio.no/ifi/english/research/projects/coinmac/index.html

Why do we not want AI and robots?

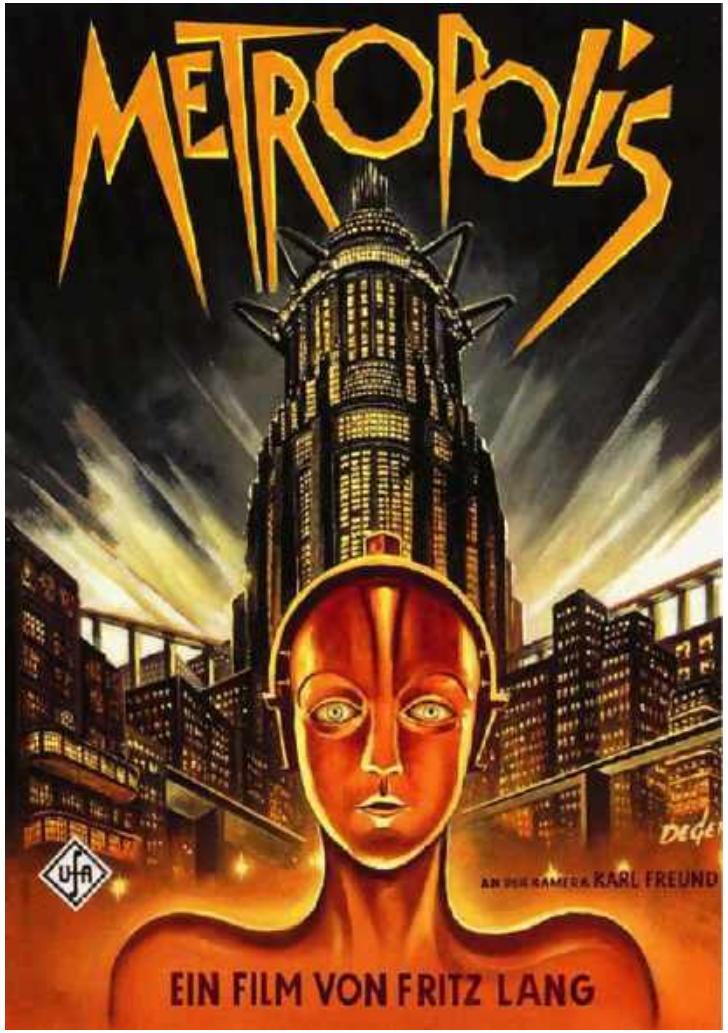
- Privacy: Collects a lot of **data** that **can be published or misused**?
- Safety and security: Can **harm us physically** (and mentally)?
- Society: **Colder society**?
- Future jobs: **Taking away our jobs**?

Photo by Philipp Katzenberger on Unsplash

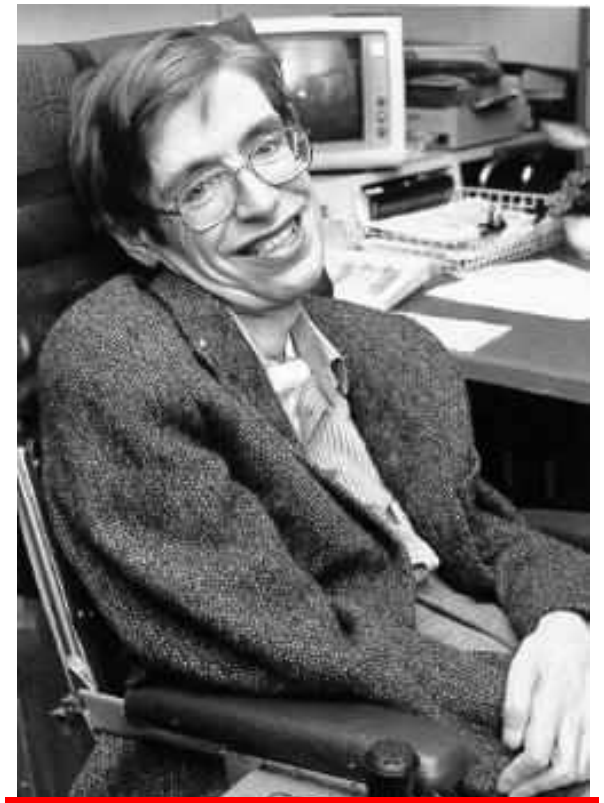
Why AI and robots?

- We can do (mental and physical) **work with less effort**
- Reduced **manual routine work**
- Improve dignity by making us **more independent** of the help of others
- Giving us **better health and longer life**

Is Terminator Coming Soon?



http://digitalresult.com/view/28-terminator-desktop-wallpapers



“Humans, limited by slow biological evolution, couldn’t compete and would be superseded by A.I.”

AI is our “biggest existential threat”

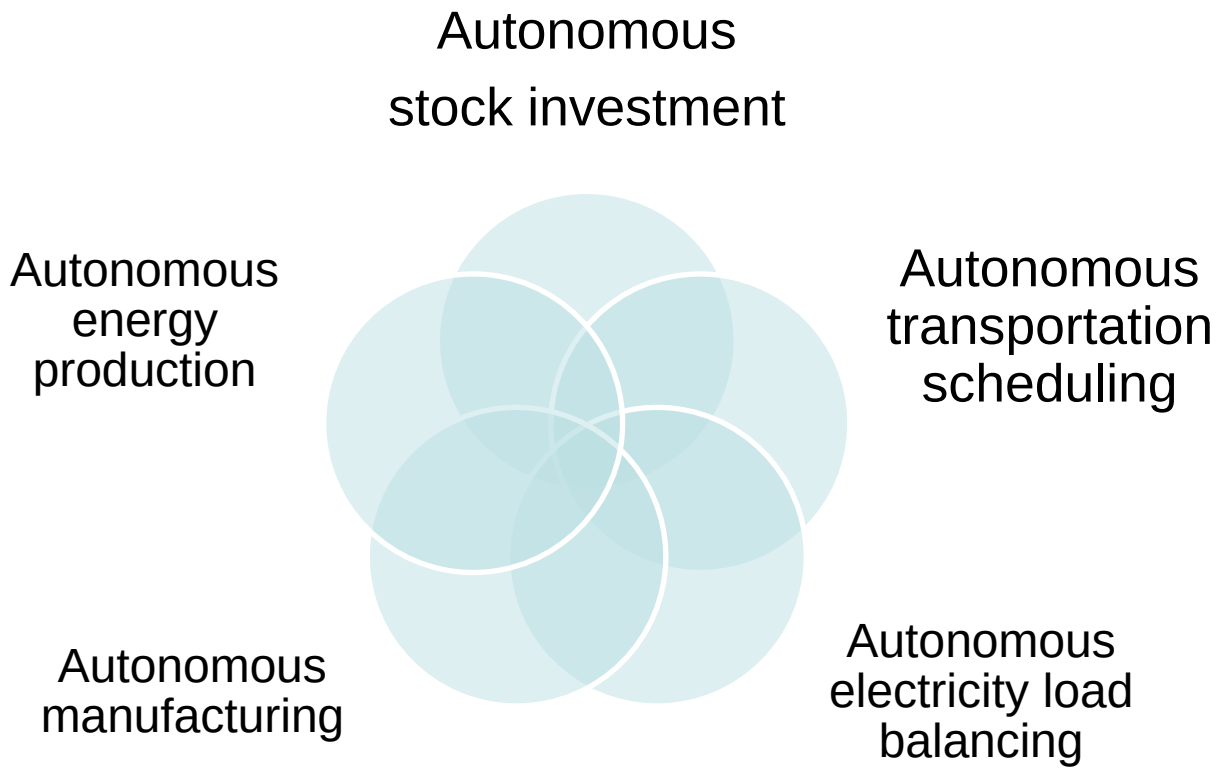
"I am in the camp that is concerned about super intelligence."

Ethical Risks of Developing AI Systems

- **Jobs:** People may become unemployed because of automation.?
- **Jobs:** We get too much free time.?
- **Technology:** Loosing human skills?
- **Technology:** Artificial intelligence can be used for destructive and unwanted tasks.?
- **Technology:** Successful KI can lead to the extinction of mankind?

Jim Torresen (2018). A Review of Future and Ethical Perspectives of Robotics and AI. *Frontiers in Robotics and AI*.

Future Scenario with Autonomous Interacting AI Systems



Wallach, Wendell og Allen, Colin. 2009. *Moral Machines: Teaching Robots Right from Wrong* New York: Oxford University Press.

Ethical Countermeasures

- **Designers, procurers and users need to be aware** of possible ethical challenges that should be considered
  - e.g. avoiding misuse and allowing for human inspection of the functionality
- The **systems** should **themselves** be able to **do ethical decision making** to reduce the risk of unwanted behavior
  - Decide when a human is to be contacted or the machine should stop

Jim Torresen (2018). A Review of Future and Ethical Perspectives of Robotics and AI. *Frontiers in Robotics and AI*.

Machine Ethics: Moor´s four categories of ethical agency

1) *Ethical Impact Agents:* Any machine that can be evaluated for its ethical consequences.

2) *Implicit Ethical Agents:* Machines that are designed to avoid unethical outcomes.

3) *Explicit Ethical Agents:* Machines that can reason about ethics.

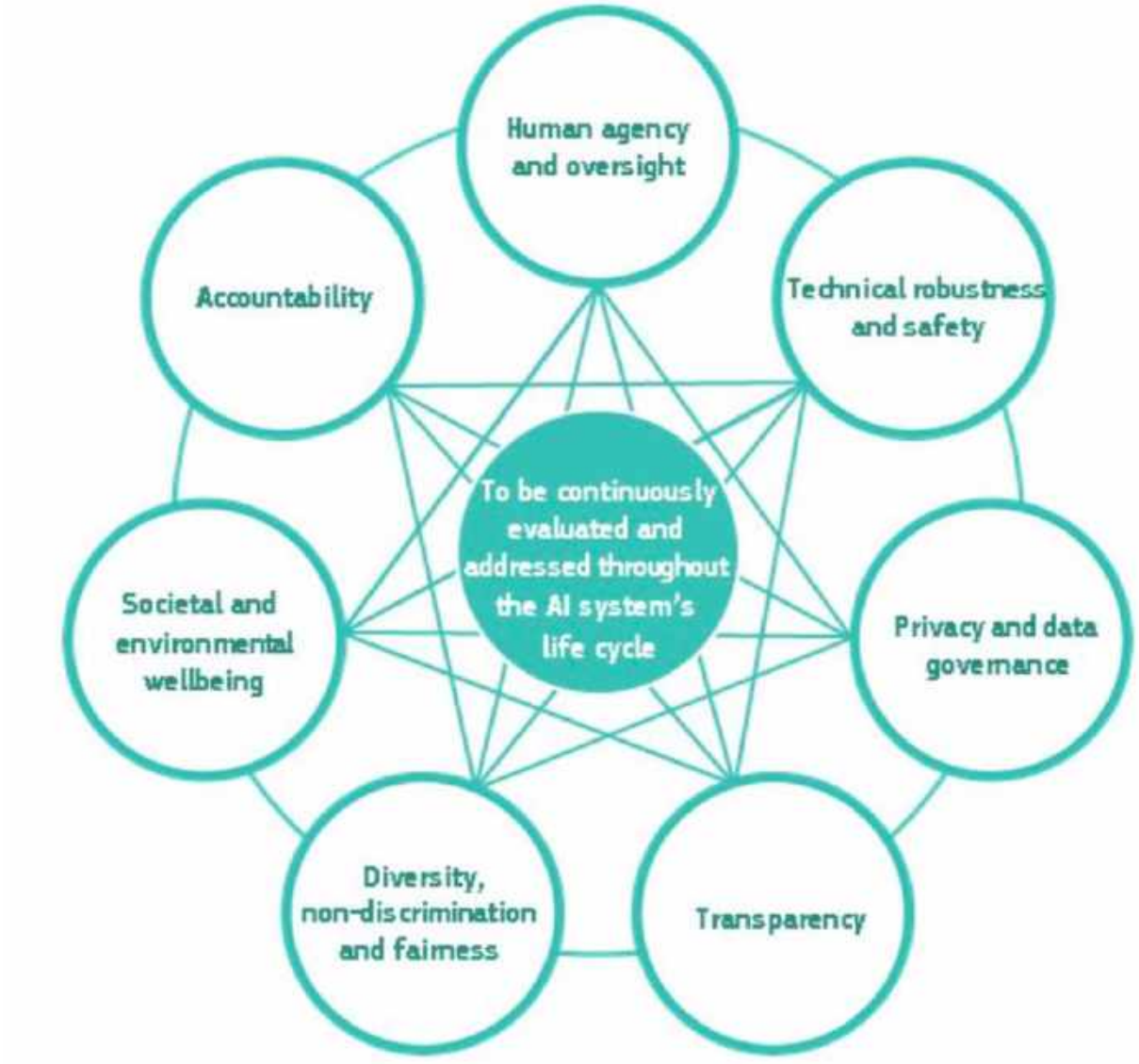
4) *Full Ethical Agents:* Machines that can make explicit moral judgments and justify them.

A. F. Winfield, K. Michael, J. Pitt and V. Evers, "Machine Ethics: The Design and Governance of Ethical AI and Autonomous Systems [Scanning the Issue]," in *Proceedings of the IEEE*, vol. 107, no. 3, pp. 509-517, March 2019, doi: 10.1109/JPROC.2019.2900622.

J. H. Moor, “The nature, importance, and [33] difficulty of machine ethics,” *IEEE Intell. Syst.*, vol. 21, no. 4, pp. 18–21, Jul./Aug. 2006.

International Work on Ethics and AI

- **IEEE** Standards Association
  - Prioritizing Human Wellbeing with Artificial Intelligence and Autonomous Systems
- **ISO**/IEC JTC 1/SC 42 Artificial intelligence
  - Standardization in the area of Artificial Intelligence
- **EU** High-Level Expert Group on Artificial Intelligence
  - 52 experts on Artificial Intelligence, comprising representatives from academia, civil society, as well as industry.
  - support the implementation of the European strategy on Artificial Intelligence.
  - recommendations on future-related policy development and on ethical, legal and societal issues related to AI, including socio-economic challenges.

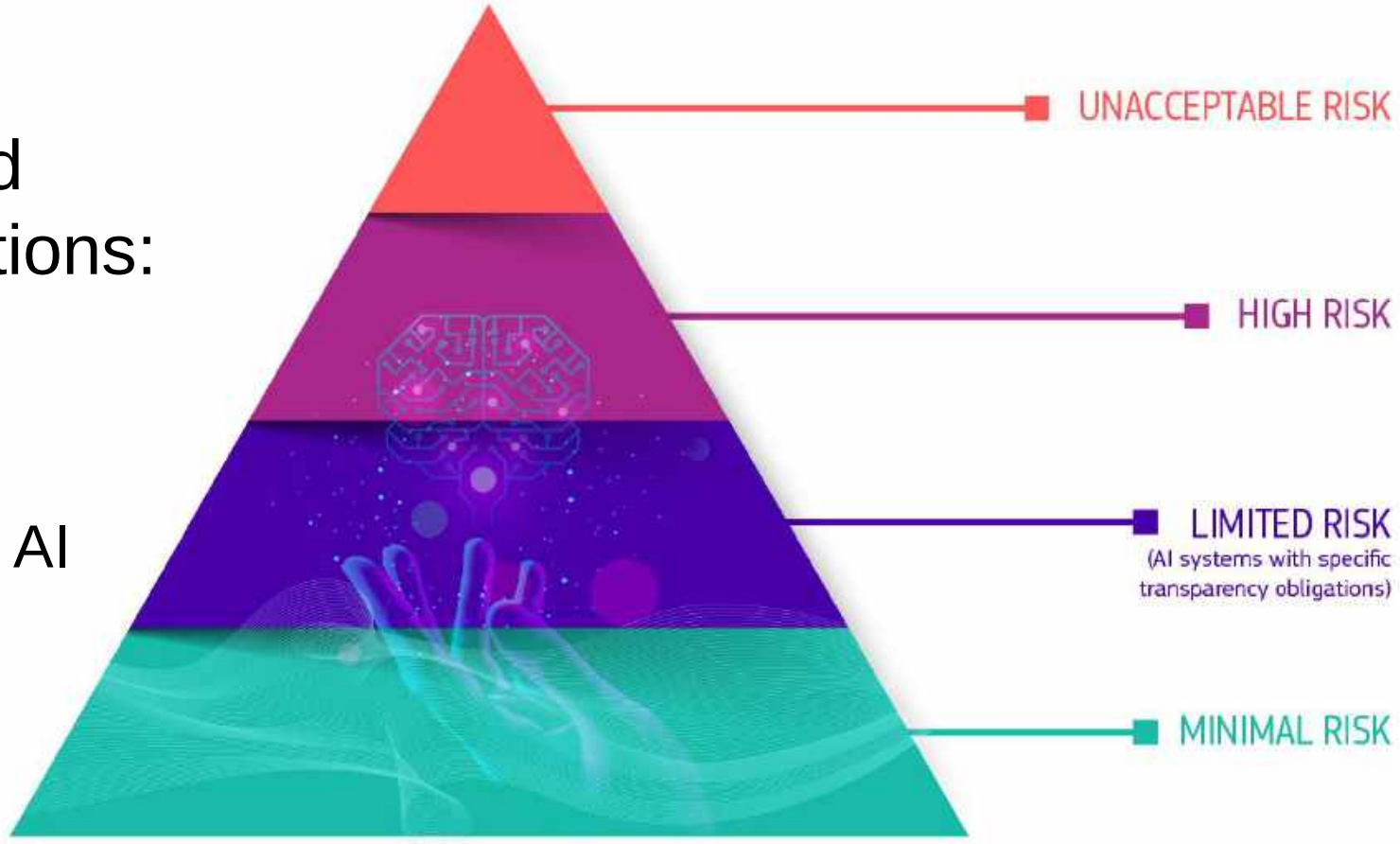

Human agency and oversight
Technical robustness and safety
Privacy and data governance
Transparency
Diversity, non-discrimination and fairness
Societal and environmental wellbeing
Accountability
To be continuously evaluated and addressed throughout the AI system's life cycle


INDEPENDENT
HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE
SET UP BY THE EUROPEAN COMMISSION
AI
ETHICS GUIDELINES FOR TRUSTWORTHY AI

EU´s risk-based approach to AI

- European Commission has proposed harmonised rules regarding AI applications:

- **EU AI Act Proposal**:
  - Prohibition of unacceptable AI practices
  - Regulation of high-risk AI systems



Source:
https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_en

UiO : Department of Informatics
University of Oslo

The Norwegian National RESEARCH ETHICS COMMITTEES

How to address ethical considerations in research: NENT AI Research Ethics Considerations Report

NENT – The National Committee for Research Ethics in Science and Technology (**Norway**)

- **Input** from relevant **academic/research institutions** involved in artificial intelligence research
- Review of **international and national reports** and guidelines
- Given by NENT in 2019 (Norwegian). An English translation published in October 2020.

Look at the report? Google for "statements AI NENT"

Characteristics of AI – NENT AI report considerations

1. AI mimics, **replaces and extends human intelligent action** and human decision-making and assessment (four considerations)
2. AI has **numerous applications** (two considerations)
3. AI uses and generates **big data** (three considerations)

UiO : Department of Informatics
University of Oslo

Ethical considerations in own research at Univ of Oslo

University of Oslo, Norway

UiO : Department of Informatics
University of Oslo

Robotics and Intelligent Systems (ROBIN)

Jim Tørresen
Professor, Group leader

Mats Høvin
Assoc. Prof.

Kyrre Glette
Professor

Kai Olav Ellefsen
Assoc. Prof.

Yngve Hafting
Ass. Prof.

Vegard D Søyseth
Principal Engineer

Adrian Bergflødt
Assistant Engineer

Postdocs:
Frank Veenstra (IFI)

Ulysse Côté-Allard (PIRC)

Diana Saplacan
(VIROS)

Adjunct positions (20%):

Alexander Wold (assoc.prof.)
Ole Jakob Elle (Prof.)
Roar Skogstrøm (lecturer)
Ståle Skogstad (assoc.prof.)
Tønnes Nygaard (lecturer)

PhD students
(ROBIN main superv.)**:**

Adel Baselizadeh
Benedikte Wallace (RITMO)
Bjørn Ivar Teigen (DOMOS)
Bjørn Thor Jonsson (RITMO)
Emma H Stensby
Farzan M. Noori
Ivar-Kristian Waarum (NGI)
Julian Fuhrer (RITMO)
Katrine Nergård
Marieke van Otterdijk
Mojtaba Karbasi (RITMO)
Mateusz Wasiluk (BioAI)
Shin Watanabe
Tom Frode Hansen (NGI)

Students: Bachelor ~200; Master: ~60
Robotics and Intelligent Systems program

Students hired on hourly basis:
Magnus E. Seierstad

Visiting researchers

https://www.mn.uio.no/ifi/english/research/groups/robin

UiO : Department of Informatics
University of Oslo

RITMO Centre of Excellence for Interdisciplinary Studies in Rhythm, Time and Motion grant 262762 (2017-2027)

- The center will study the **perceptual, cognitive** and **acting mechanisms** underlying our ability to experience rhythm and act rhythmically.
- Interdisciplinary **collaboration** between **musicology, psychology, computer science** and **robotics.**
- Machine learning and robotics to be applied

https://www.uio.no/ritmo/english

The Research Council of Norway

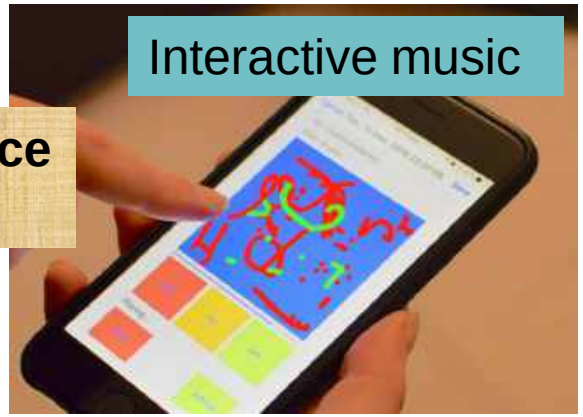
Norwegian Centre of Excellence

UiO : Department of Informatics
University of Oslo

Robotics and Intelligent Systems (ROBIN) research group

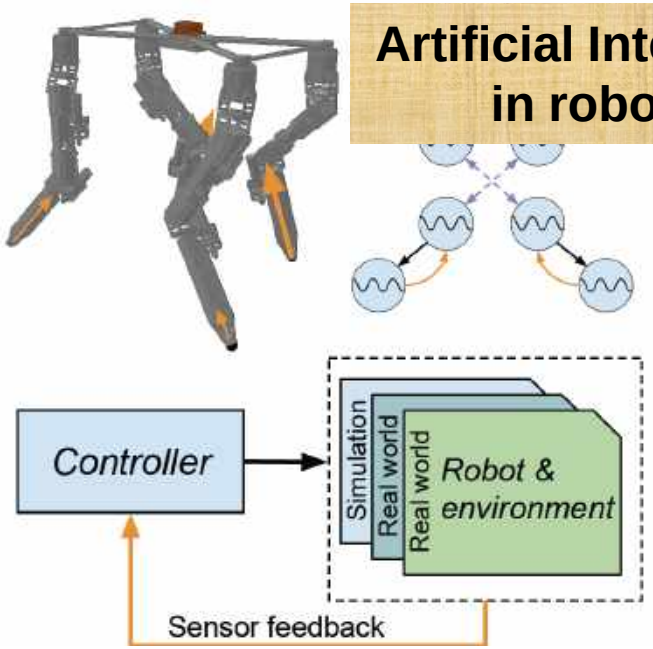
Adaptive and autonomous mental health treatment

Interactive music

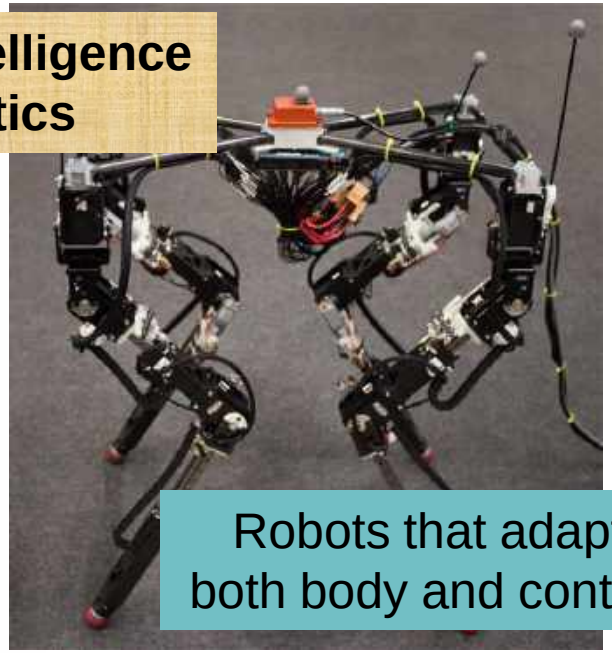


Artificial Intelligence in smartphones

Robots that assist people living at home



Artificial Intelligence in robotics



Robots that adapt both body and control

ROBIN Research Projects and Centre
Funded by the Research Council of Norway

- **INtroducing personalized TReatment Of Mental health problems using Adaptive Technology** (INTROMAT, 2016-2021, LightHouse project)
- **Multi-sensor Elderly Care Systems/Robots** (MECS, 2015–2021, IKTPLUSS)
- **Vulnerability in the Robot Society** (VIROS, 2019-2023, IKTPLUSS)
- **Predictive and Intuitive Robot Companion** (PIRC, 2020-2025, IKTPLUSS)
- **Centre of Excellence for Interdisciplinary Studies in Rhythm, Time and Motion** (RITMO, 2017-2027, CoE)

The Research Council of Norway

UiO : Department of Informatics
University of Oslo

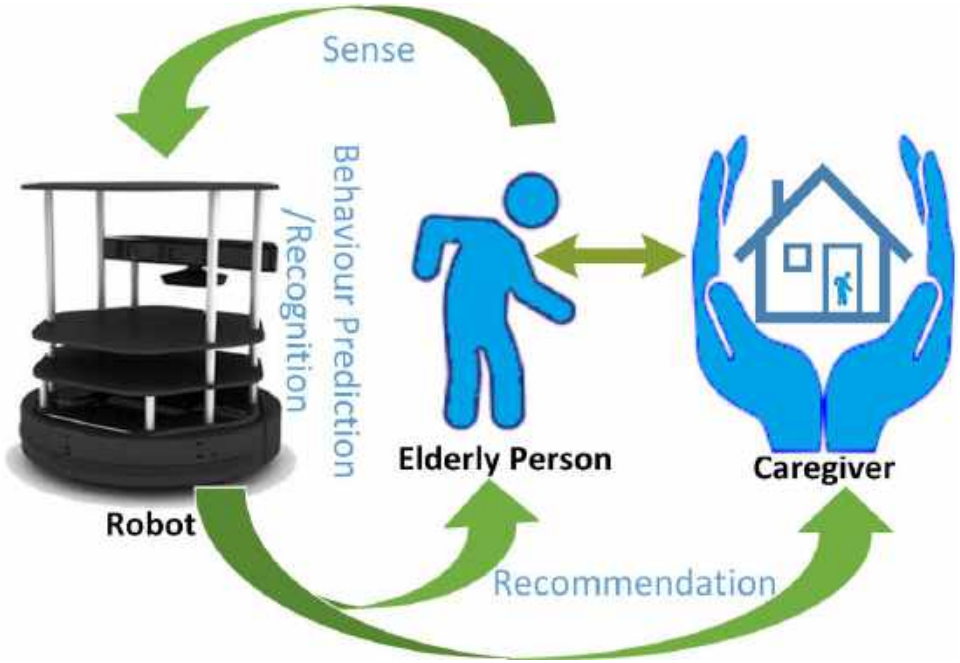
MECS
Multimodal Elderly Care Systems

MECS: Multi-sensor Elderly Care Systems

Research Council of Norway grant 247697 (2015–2021)

https://www.mn.uio.no/ifi/english/research/projects/mecs

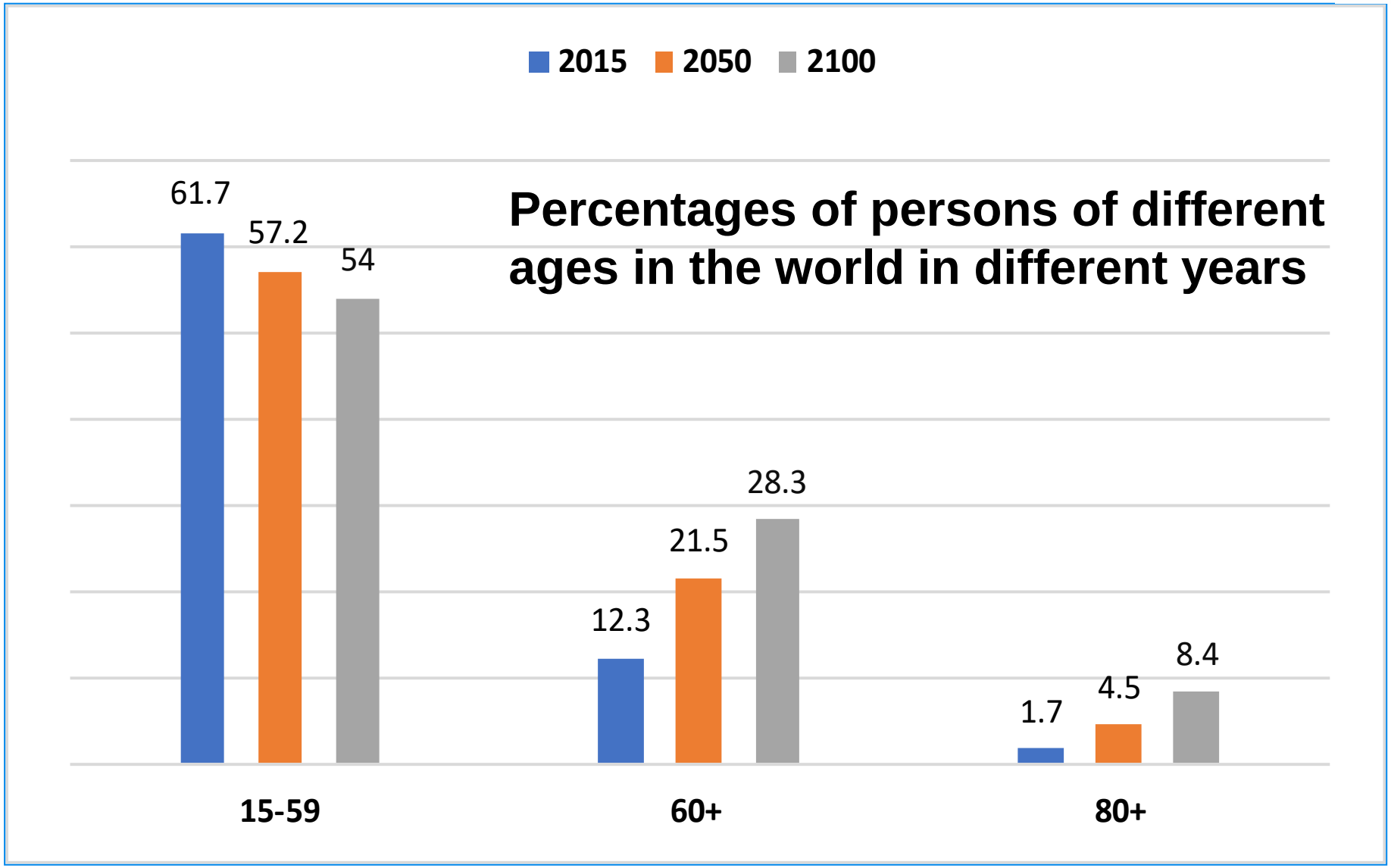
Goal: Create and evaluate multimodal mobile **human supportive systems** that are able to **sense, learn and predict future events**.



Funding: *FRINATEK*
Research Council of Norway

The Research Council of Norway

Web page: Google for ”MECS IFI”



United Nations (2015) World population ageing. United Nations, New York.

Robots and Older People

Would we like to be surrounded by robots rather than humans?

Robots and Older People

Would we like – with some help from robots – **to be independent with regards to our key needs like personal care, eating and transportation?**

Myself: *Yes, but my preference would be impacted by the ease of use and performance and to some extent the look of the robot.*

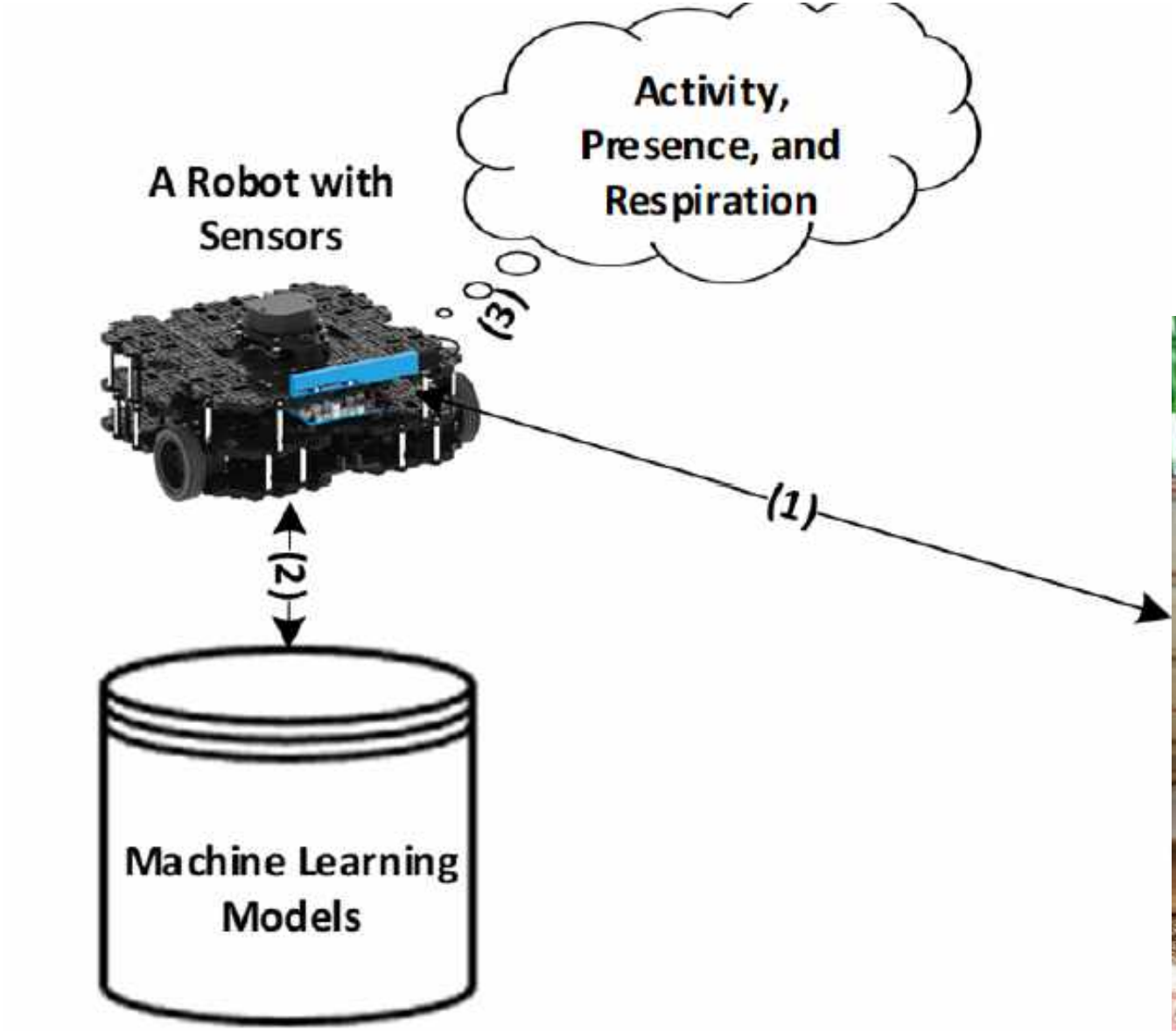


Image by rawpixel.com

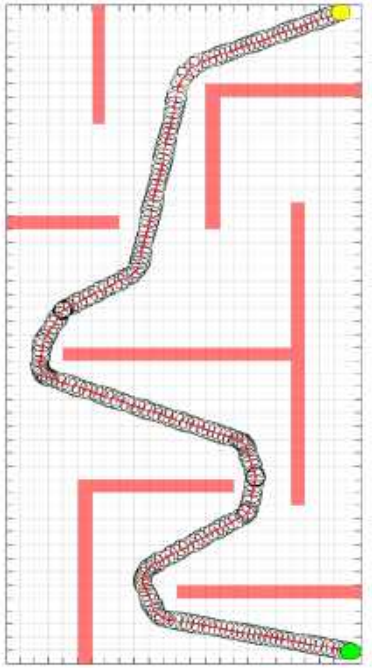
MECS
Multimodal Elderly Care Systems

MECS Research

Diana Saplacan

Rebekka Soma

Trenton Schulz



User needs and preferences

+ Master students

Robot sensing

Robot control

Apply sensors that provides non/less-intrusive sensing

Navigation without a map

Farzan M. Noori

Md. Zia Uddin

Weria Khaksar

User Centered Design – Participatory Design

- involve real users in **actual use contexts** (home of older people)
- focus on behavior and **satisfying the needs** and desires of the users
- achieve improvements through **iterative testing and improvement**

- Oslo municipality elderly care facility: **Kampen Omsorg +**
- **Vitalis home** for elderly in Eindhoven in the Netherlands

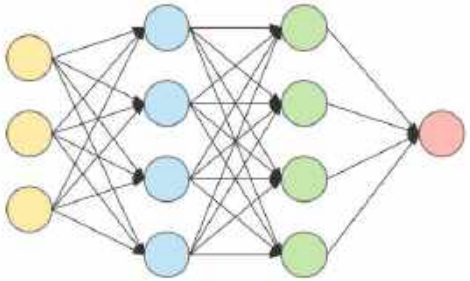

Dutch national TV (NPO)

Ethical Concerns: 1.Privacy 2.Security 3.Safety

ROBOT = Sense → Think → Act

Sense → Sensors

Think → Artificial Intelligence



Act → Motors + Mechanics

Top image: www.robotpark.com

Ethical Concerns: 1. Privacy

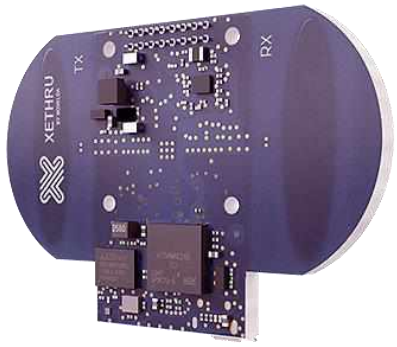
- Challenge 1: Balance the **privacy of the elderly** against the **needs for data** collection for having an efficiently functioning elderly care systems.
- Challenge 2: Protection of sensitive data to **avoid unwanted distribution and misuse** of such data.
- Mitigation:
  - **Sensor type:** Use sensors collecting less privacy related information (person identifiable vs bio-signals/medical diagnostic)
  - **Sensor data processing**: Process data locally rather than sending sensor data over Internet

Privacy challenges / paradoxes

- **How much data to collect:** The more data that are used during research and development, the less data is needed when a system is to be applied.
- **What kind of sensors:** The more sensors, the more accurate and safer the robot will be.
- **Consent**: A personalised robot will act better and safer: the more you know about a friend, the more you tend to adapt to the person. (ref. web page cookie preferences will give easier use and more personalised web page view)

Images by **rawpixel.com**

The MECS project has explored a number of novel non/less-intrusive sensors



Novelda XeThru

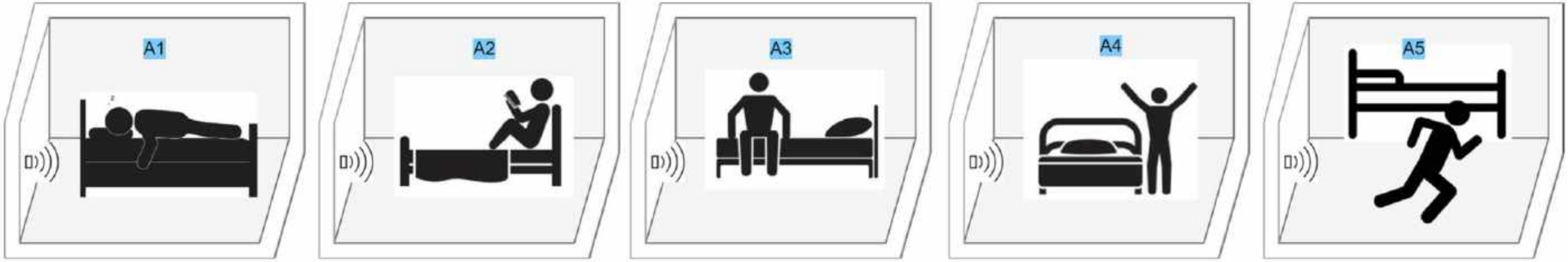
FLIR Thermal Camera

Md Zia Uddin; Weria Khaksar & Jim Tørresen (2018). Ambient Sensors for Elderly Care and Independent Living: A Survey. *Sensors*.

Non/less-Intrusive Sensing (thermal camera)

Md. Zia Uddin and Jim Torresen. A Deep Learning-Based Human Activity Recognition in Darkness, The 9th IEEE Colour and Visual Computing Symposium 2018 (CVCS 2018), Sept. 19-20, 2018

Ultra-Wideband (UWB) Radar-Based Activity Recognition



- **LSTM-based activity recognition** approach performed better than conventional approaches, with an **accuracy of 99.6%**.
- We applied 5-fold cross-validation to test our approach.

F. M. Noori, M. Z. Uddin and J. Torresen, "Ultra-Wideband Radar-Based Activity Recognition Using Deep Learning," in *IEEE Access*, vol. 9, pp. 138132-138143, 2021, doi: 10.1109/ACCESS.2021.3117667.

Ethical Concerns: 2. Security

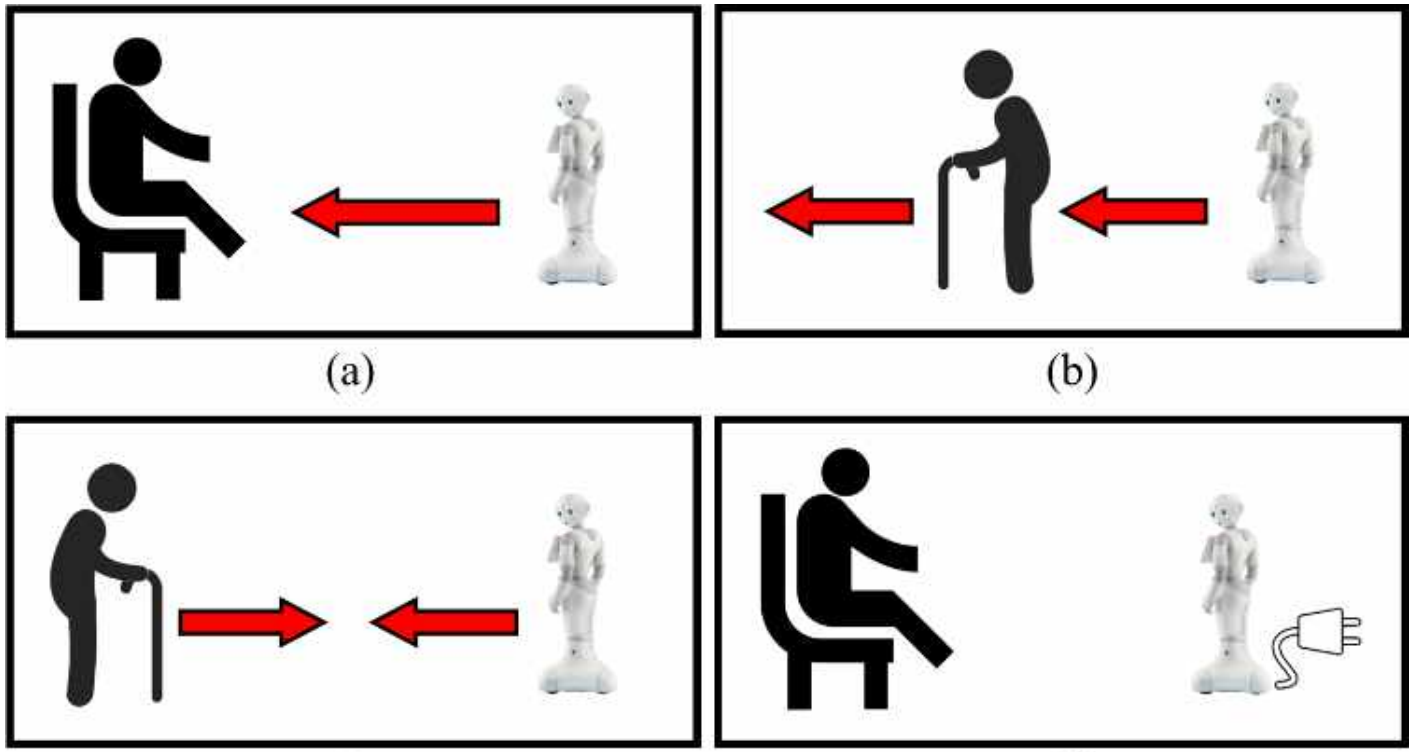
- Concern 1: **Sensing** – possible theft and unwanted distribution of sensor data from a robot.
- Concern 2: **Control** – risk of misbehaviour of the robot in similar ways as computers can be attacked with malware.
- **Mitigation** 1: Regular security measures with **passwords and authentication**
- **Mitigation** 2: Add an **external user assessment** module that can consider the current context (ref. ethical reasoning engine)

Ethical Concerns: 3. Safety

- Challenge: Robots getting **physically much closer to humans** than what we are used to. => Can hit us unintentionally or hurt us through un-authorized access
- Trade-off between robot size, performance and safety
- Mitigation:
  - Equip robots with **soft material**
  - Provide a self/user-aware adaptable system that can **learn about the user´s daily activities** and preferences
  - **Explainable AI/transparent systems** to be able to correct for unwanted or harming behavior

Farzan Majeed Noori, Michael Riegler, Md Zia Uddin, and Jim Torresen. 2020. Human Activity Recognition from Multiple Sensors Data Using Multi-fusion Representations and CNNs. ACM Trans. Multimedia Comput. Commun. Appl. 16, 2, Article 45 (June 2020),

User studies on robot and user encounters

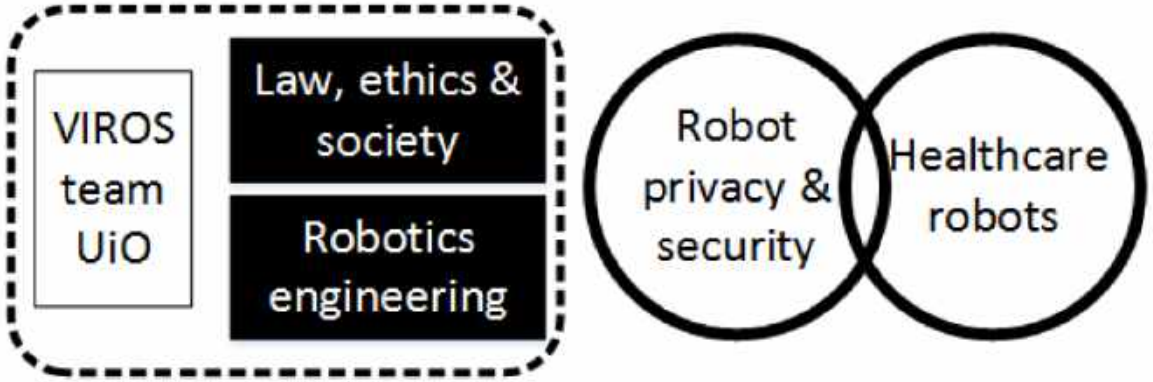


Khaksar, W.; Neggers, M.; Barakova, E.; and Torresen, J., "*Generation Differences in Perception of the Elderly Care Robot*," *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, 2021, pp. 551-558

UiO : Department of Informatics
University of Oslo

VIROS: Vulnerability in the Robot Society (2019-2023)

Research Council of Norway grant 288285



Dep. of Private Law +
Dep. of Informatics and other depts/partners

Goal:
Develop technology and proposals for **regulatory measures** to reduce vulnerabilities regarding robotics. **Focus on privacy, security and safety**, particularly in healthcare contexts.

Funding: *IKTPLUSS, Research Council of Norway*

The Research Council of Norway

https://www.jus.uio.no/ifp/english/research/projects/nrccl/viros/index.html

UiO : Department of Informatics
University of Oslo

RITMO Centre of Excellence for Interdisciplinary Studies in Rhythm, Time and Motion grant 262762 (2017-2027)

- The center will study the **perceptual, cognitive** and **acting mechanisms** underlying our ability to experience rhythm and act rhythmically.
- Interdisciplinary **collaboration** between **musicology, psychology, computer science** and **robotics.**
- Machine learning and robotics to be applied

https://www.uio.no/ritmo/english

The Research Council of Norway

Norwegian Centre of Excellence

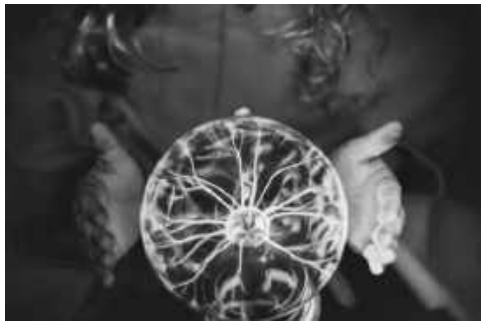
Human – Robot Interaction
Slow Versus Safe Robot

- sloppy vs too slow
- must have general capabilities

UiO : Department of Informatics
University of Oslo

Predictive and Intuitive Robot Companion (PIRC) (2020-2025)

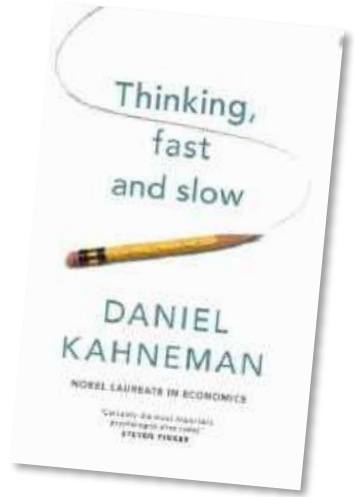
Research Council of Norway grant 312333



TIAGo mobile robot assistant

Goal: Build **models** that **forecast** future events and **respond dynamically by psychology-inspired computing**:

- Apply recent models of **human prediction** to perception-action loops of future intelligent robot companions.
- Include mechanisms for **adaptive response time** from quick and intuitive to slower and well-reasoned
- **Applications**: Physical rehabilitation and home care robot support for older people.



Funding: *IKTPLUSS, Research Council of Norway*

The Research Council of Norway

Future work and opportunities in ethics related research

- **privacy**
  - work with sensors collecting less privacy related data
  - work on algorithms for local /edge computing (rather than cloud computing)
- **security**
  - external user assessment module that can consider the current context
- **safety**
  - user-aware systems
  - explainable AI/transparent systems to be able to correct for unwanted or harming behavior
- **potential lack of contact with other humans**
  - **human dignity** impacted by our **independence**
  - politicians and society decide on the development as much as technology providers
  - researchers can contribute to a **balanced public discussion**

We should focus as least as much on improved quality of life as reducing the cost by the technology being developed

Questions or Comments?

Make contact: jimtoer@ifi.uio.no

www.jimtoer.no