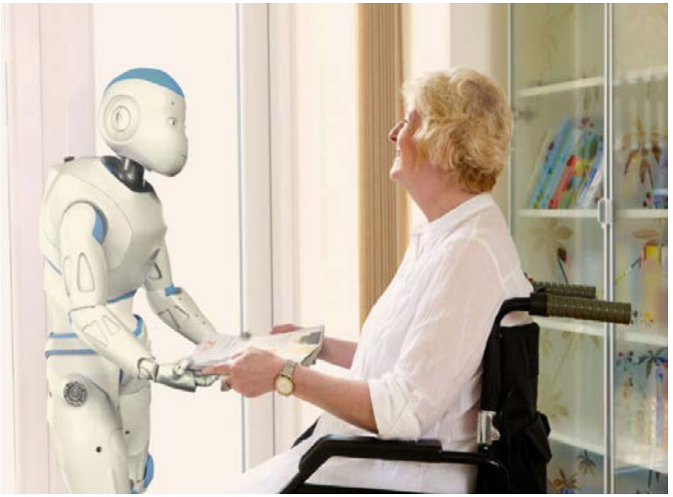
INTERACTING WITH AI
MODULE 3

Session 2, November 3, 2020

Amela Karahasanović, SINTEF and UiO

SINTEF

SUSTAINABILITY
& DESIGN LAB



Module 3 Living and working with AI

Objectives

Understanding of challenges related to use of AI infused systems in everyday life and at work

- How to evaluate them?
- **When and how to use them?**
- What do we know about living and working with them?

Module 3 Overview

- Evaluation of interaction with AI [27th of October]
- **Human - AI partnership [3rd of November]**
- Lessons learned from studies of human – AI interaction [10th of November]

Plan for today

- Human-robot teams
- Task distribution between humans and AI
- Levels of automation
- Human-in-the-loop and situation awareness
- Explainable AI

Your turn

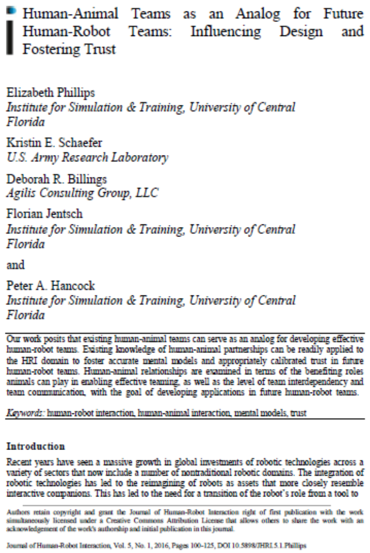
- Which tasks do YOU think could be done by AI?
- Which tasks do YOU think should not be done by AI?

- 5 minutes, group discussions

Roles of robots

Roles of robots *(Philips at al. 2016, Human-Animal Teams as an Analog for Future Human-Robot Teams: Influencing Design and Fostering Trust, Journal of Human-Robot Interaction, Vol. 5, No. 1, 2016, pages 100-125)*

- Peoples mental models of robots doesn't fit reality leading to distrust or discounting using the automated systems
- A robot is a team member instead of a tool



Human-Animal Teams as an Analog for Future Human-Robot Teams: Influencing Design and Fostering Trust

Elizabeth Phillips
Institute for Simulation & Training, University of Central Florida

Kristin E. Schaefer
U.S. Army Research Laboratory

Deborah R. Billings
Agilis Consulting Group, LLC

Florian Jentsch
Institute for Simulation & Training, University of Central Florida

and

Peter A. Hancock
Institute for Simulation & Training, University of Central Florida

Our work posits that existing human-animal teams can serve as an analog for developing effective human-robot teams. Existing knowledge of human-animal partnerships can be readily applied to the HRI domain to foster accurate mental models and appropriately calibrated trust in future human-robot teams. Human-animal relationships are examined in terms of the benefiting roles animals can play in enabling effective teaming, as well as the level of team interdependency and team communication, with the goal of developing applications in future human-robot teams.

Keywords: human-robot interaction, human-animal interaction, mental models, trust

Introduction

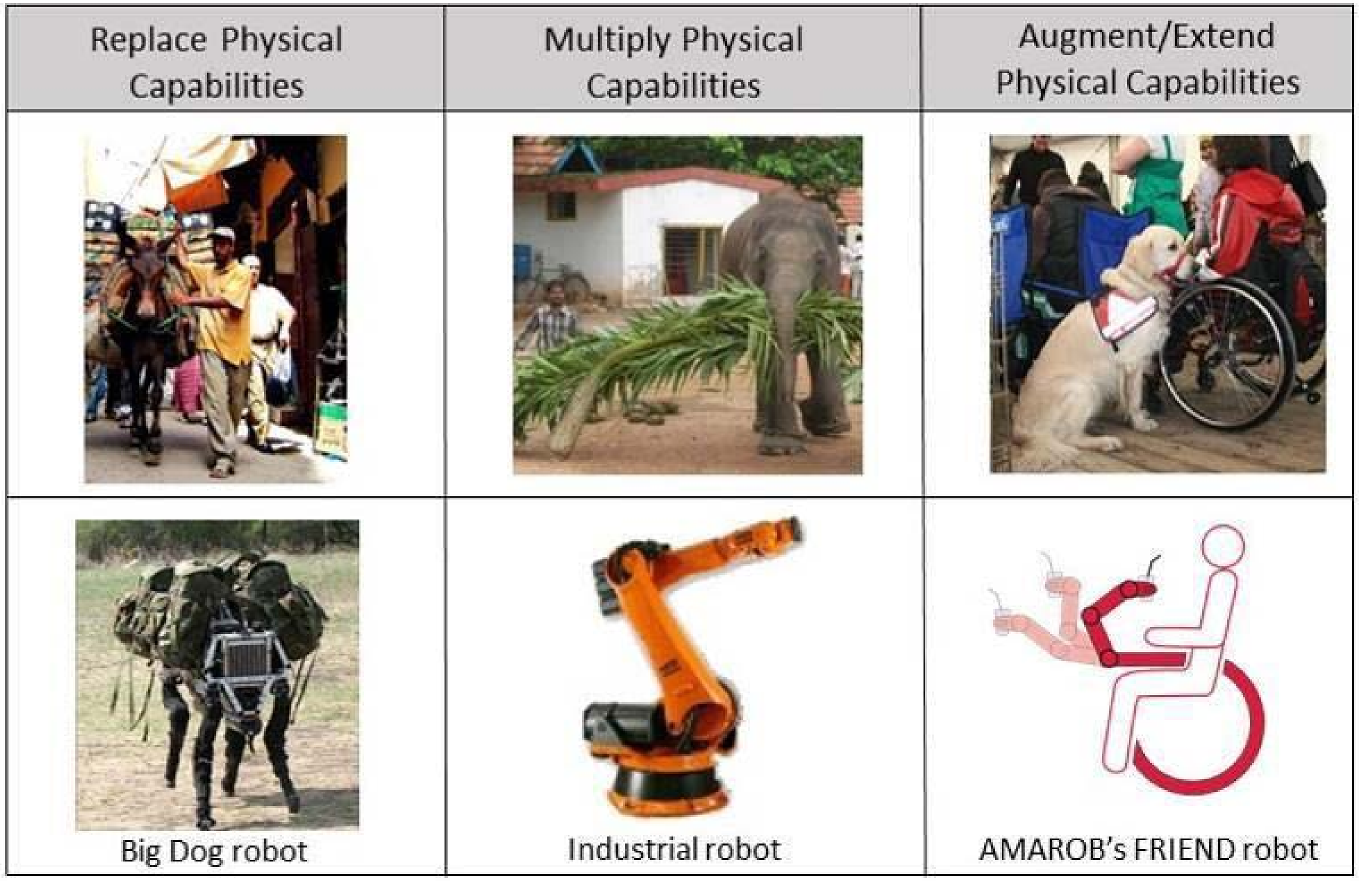
Recent years have seen a massive growth in global investments of robotic technologies across a variety of sectors that now include a number of nontraditional robotic domains. The integration of robotic technologies has led to the reimagining of robots as assets that more closely resemble interactive companions. This has led to the need for a transition of the robot's role from a tool to

---

Journal of Human-Robot Interaction, Vol. 5, No. 1, 2016, Pages 100-125, DOI 10.5898/JHRI.5.1.Phillips

Physical benefits

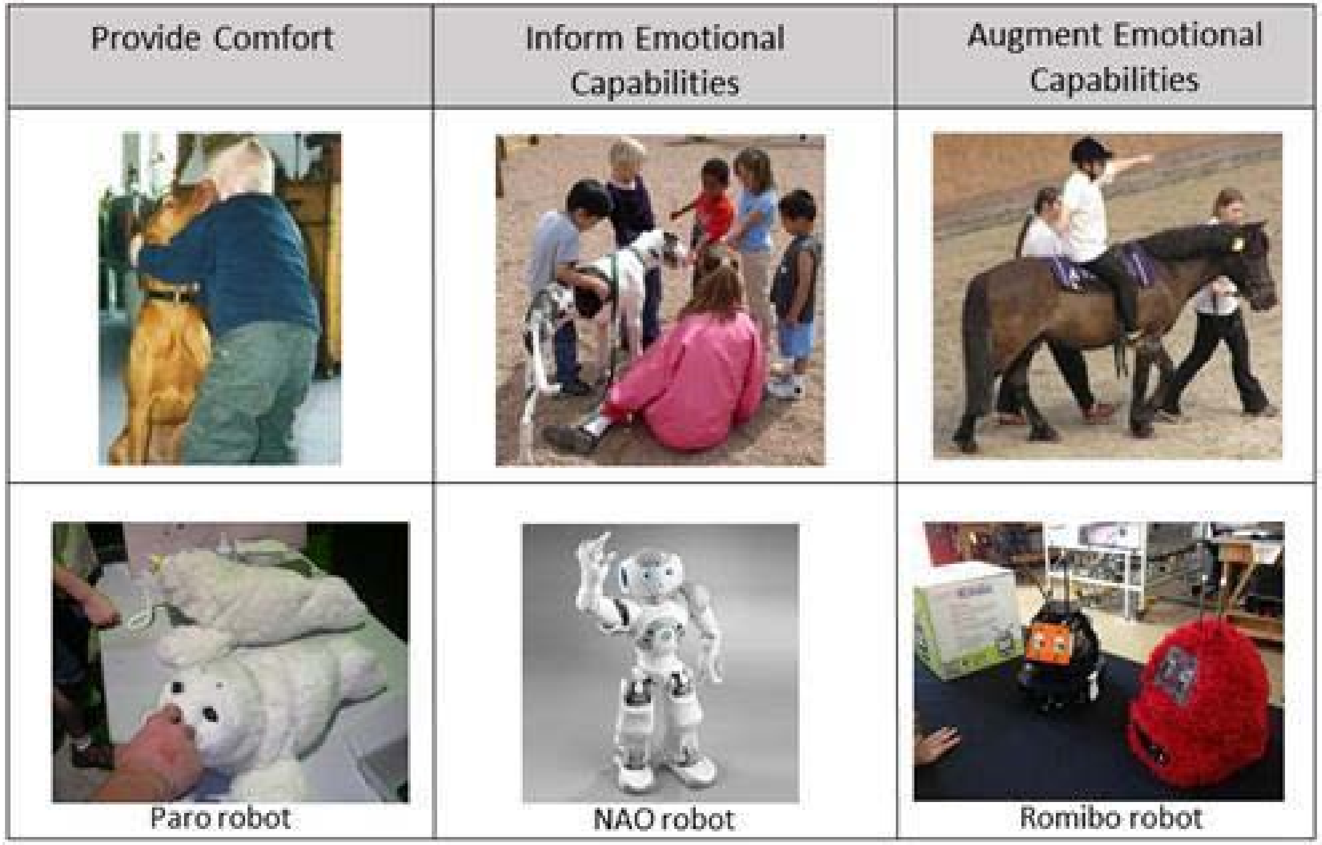
- Replace physical capabilities
  - Big Dog robot (Boston Dynamics)
- Multiply physical capabilities
  - Industrial robot arm
- Augment/Extend physical capabilities
  - AMAROB's Functional robot arm with user-friendly interface for disabled people (FRIEND) for people with skeletal-muscular disorders



Replace Physical Capabilities	Multiply Physical Capabilities	Augment/Extend Physical Capabilities
Big Dog robot	Industrial robot	AMAROB's FRIEND robot

Emotional benefits

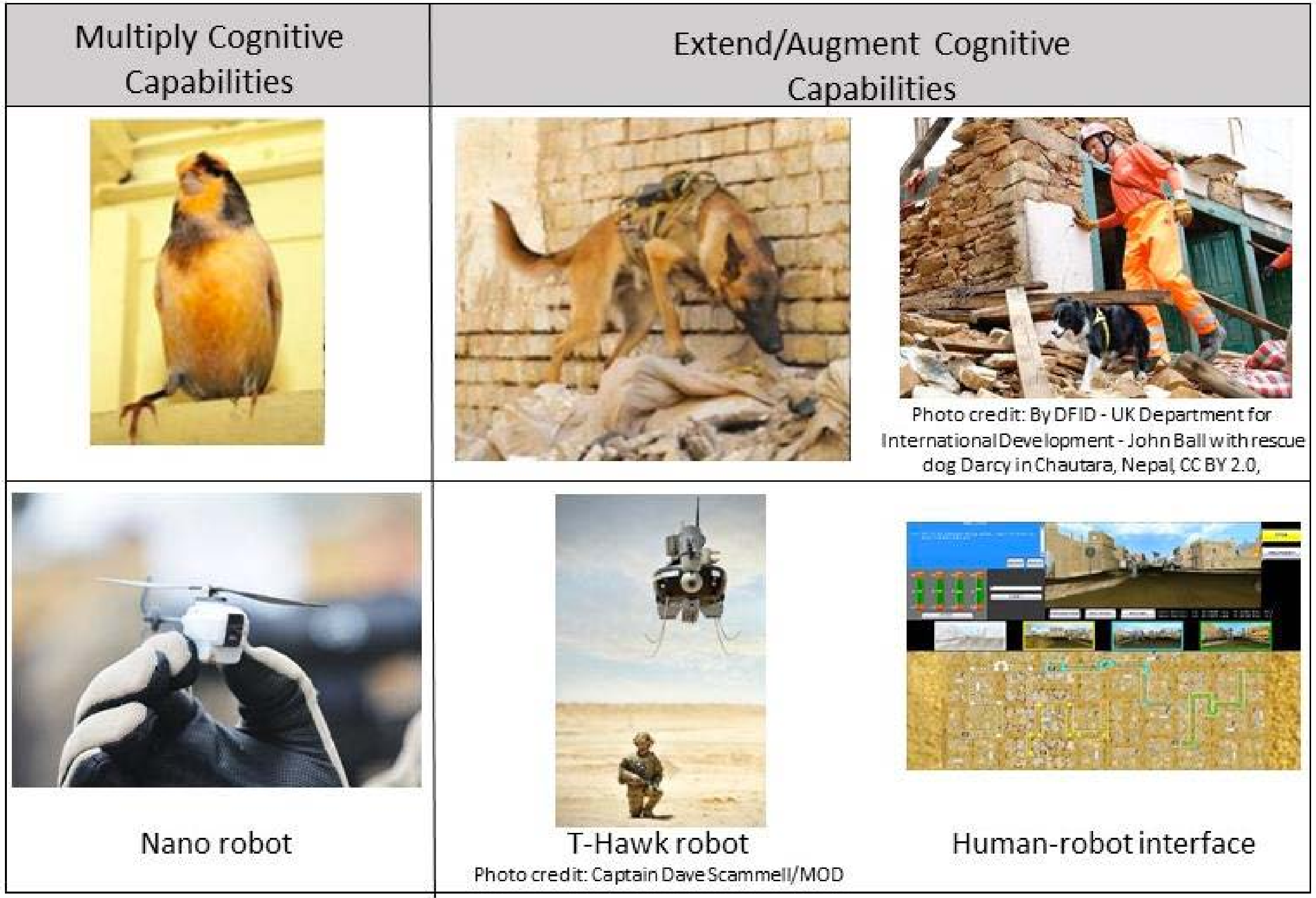
- Provide comfort
  - Paro therapeutic robot - responds as if it is alive, moving its head and legs, making sounds, imitates the voice of a real baby harp seal.
- Inform/Augment emotional capabilities
  - NAO robot, Romibo robot
  - Teaching social skills
  - Therapeutic horse riding (autism)



Provide Comfort	Inform Emotional Capabilities	Augment Emotional Capabilities
Paro robot	NAO robot	Romibo robot

Cognitive benefits

- Multiply cognitive capabilities
    - Nano Unmanned Aerial Vehicle collect additional sensory information
- Extend cognitive capabilities
    - Robots helping in nuclear disaster
- Human-dogs narcotics search team
    - Reciprocal interdependencies: handler provides search guidelines, dog provides sensory alerts, team provides location of narcotics



Multiply Cognitive Capabilities	Extend/Augment Cognitive Capabilities	
		Photo credit: By DFID - UK Department for International Development - John Ball with rescue dog Darcy in Chautara, Nepal, CC BY 2.0,
Nano robot	T-Hawk robot Photo credit: Captain Dave Scammell/MOD	Human-robot interface

How to allocate tasks to humans and AI?

Approaches to task allocation

- Ad hoc allocation
- Formal and balanced approach including comparative assessment of human and machine performance; using KPIs (performance, situation awareness, costs, cognitive task load, trust, human tendency for boredom, keeping skills, recovery from system failure, team dynamic …) and knowledge about man/machine capabilities
  - Take into account political, ethical and legal reasons

Approaches to task allocation

Sheridan and Verplank's 10 level Autonomy scale

- Level 1 – humans take all decisions
- Level 2 – Computer aids in highlighting key information on screen or decluttering irrelevant information
- Level 3 - System gathers key information and integrates
- Level 4 - Computer aids in doing each action as instructed
- Level 5 - Computer completely carries out singular or sets of tasks commanded by human
- Level 6 - Computer and human generate decision options, human decides and carries out with support
- Level 7 - Computer generates recommended options, human decides (or input own choice) and system carries out
- Level 8 – informs the human only if asked
- Level 9 – informs the human only if the computer decides to
- Level 10 – the computers acts autonomously ignoring the human

Sheridan, T.B., Verplank, W., Human and Computer Control of Undersea Teleoperators, MIT, 1978

Adaptive Automation

Decision of when to pass control from automation to the human based on:

- A consistent time interval
- The occurrence of critical events
- Detection of human performance below a certain criterion level
- Use of psychophysiological monitoring to detect losses of arousal or other cues of poor performance (e.g., loss of consciousness)
- The use of models of human performance to predict the best times to intervene

Your turn – group work 10 - 15 minutes

A company introduces an intelligent agent (robot) that will take care of recruitment and hiring new employees. Describe the functionality that such agent should have. Which task could it perform? When doing this consider not only current technology but also technology that will come in relatively near future (5-10 years). Then write two scenarios where this agent have two different automation levels. In the first scenario the agent will have a level of automation 6 or 7. In the second scenario the agent will have a level of automation between 8, 9 or 10. Describe which tasks the agent perform and which tasks should humans perform. What are advantages and disadvantages related to this task distribution? What are the possible problems that might occur? How to overcome them?

Approaches to task allocation

Sheridan and Verplank's 10 level Autonomy scale

- Level 1 – humans take all decisions
- Level 2 – Computer aids in highlighting key information on screen or decluttering irrelevant information
- Level 3 - System gathers key information and integrates
- Level 4 - Computer aids in doing each action as instructed
- Level 5 - Computer completely carries out singular or sets of tasks commanded by human
- Level 6 - Computer and human generate decision options, human decides and carries out with support
- Level 7 - Computer generates recommended options, human decides (or input own choice) and system carries out
- Level 8 – informs the human only if asked
- Level 9 – informs the human only if the computer decides to
- Level 10 – the computers acts autonomously ignoring the human

Sheridan, T.B., Verplank, W., Human and Computer Control of Undersea Teleoperators, MIT, 1978

Other classifications

Levels of autonomy for self-driving cars

5. **Full autonomy**: equal to that of a human driver, in every driving scenario.

4. **High automation:** Fully autonomous vehicles perform all safety-critical driving functions in certain areas and under defined weather conditions.

3. **Conditional automation:** Driver shifts “safety critical functions” to the vehicle under certain traffic or environmental conditions.

2. **Partial automation:** At least one driver assistance system is automated. Driver is disengaged from physically operating the vehicle (hands off the steering wheel and foot off the pedal at the same time).

1. **Driver assistance:** Most functions are still controlled by the driver, but a specific function (like steering or accelerating) can be done automatically by the car.

0. **No Automation:** Human driver controls all: steering, brakes, throttle, power.

(from Shneiderman, 2020)

Two-dimensional framework with the goal of Trusted, Reliable & Safe systems

Control

Human	**HUMAN MASTERY** **Bicycle** **Piano playing** ***Car 1980***	**TRUSTER, RELIABLE & SAFE SYSTEMS** **Elevator** **Camera** ***Car 2040***
Computer	**Clock, mousetrap**	**COMPUTER CONTROL** **Airbag deployment, pacemakers** ***Car 2020***
	Low	High

Automation

(Shneiderman, 2020)

AI in Complex Time-Critical domains

Decision making in complex time-critical domains

Steag

Type of problems where decision support systems can be useful and ways for involving humans

- Combinatorial problems; large search space – humans can help pruning a decision tree
- Visual problems; abstract problems that might be represented visually: image classification, geographical clustering; visual presentation help humans
- Computationally intensive problems; humans can guide computation and weight cost/benefits of further computations
- Heuristic-heavy problems; humans help selecting heuristics

Malasky, J.S, Human Machine Collaborative Decision Making in a Complex Optimization System, MIT, 2003

What is the problem?

A simple error made large deviations from the intended path possible. ….example, an American Airlines flight crashed in the mountains of Colombia in 1996 killing all aboard due to a mixture of programming error, overreliance on the automation, and poor feedback depriving the pilots of an understanding of what the automation was doing

(Endsley, Mica R. Designing for Situation Awareness:

An Approach to User-Centered Design, Second Edition CRC Press. 2011)

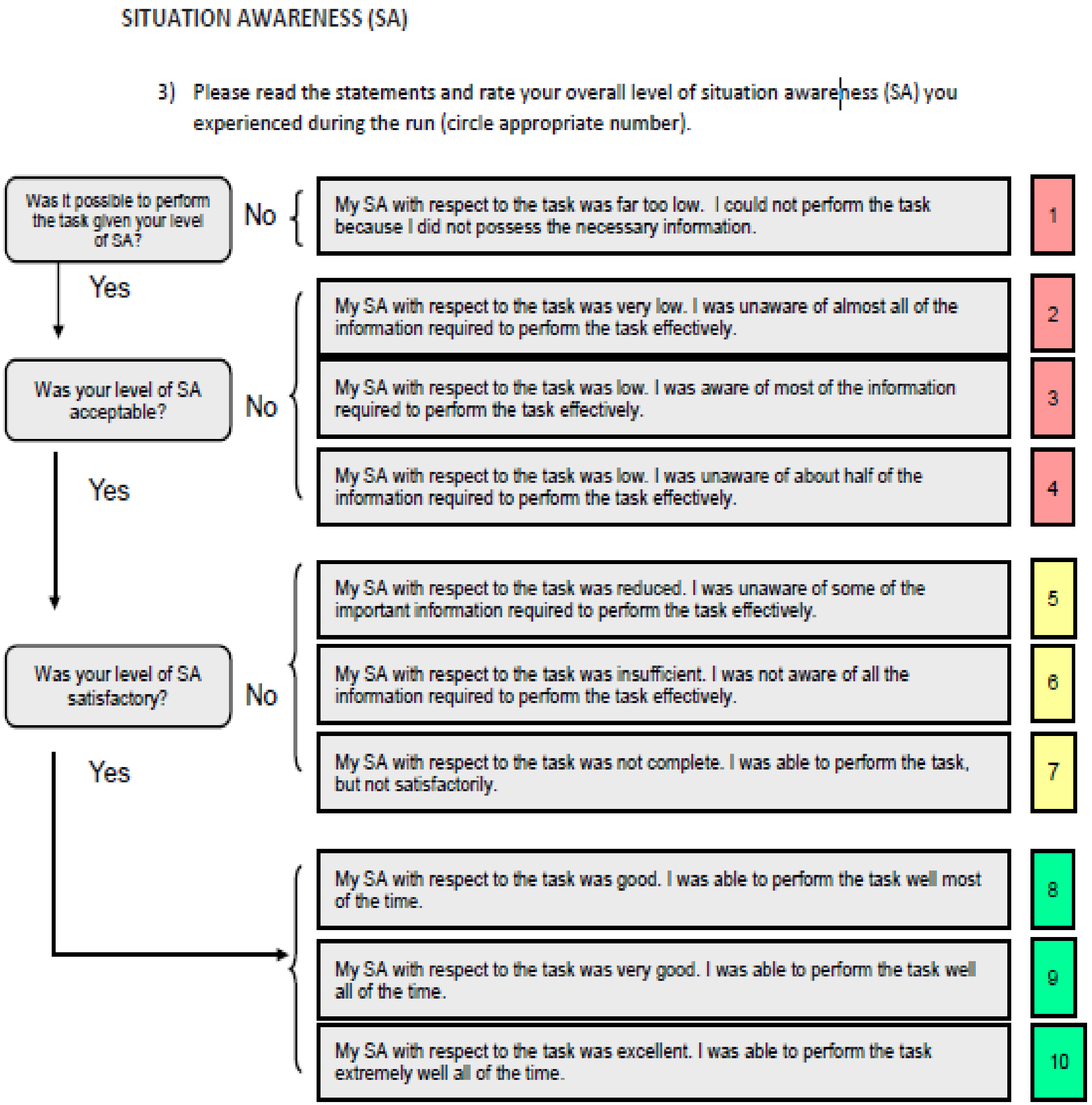
How to assure good collaboration between AI and humans?

Situation awareness

“the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future” (Endsley, 2011).

Important in complex time-critical domains

- Important with or without automation
- Different levels of automaton lead to different challenges
- Out-of-the-loop syndrome – what if automation fails, boredom, fatigue
- Loosing the competence and ability to adequately react in non-nominal situations



SITUATION AWARENESS (SA)

3) Please read the statements and rate your overall level of situation awareness (SA) you experienced during the run (circle appropriate number).

Was it possible to perform the task given your level of SA?

No

My SA with respect to the task was far too low. I could not perform the task because I did not possess the necessary information.	1

Yes

Was your level of SA acceptable?

No

My SA with respect to the task was very low. I was unaware of almost all of the information required to perform the task effectively.	2
My SA with respect to the task was low. I was aware of most of the information required to perform the task effectively.	3
My SA with respect to the task was low. I was unaware of about half of the information required to perform the task effectively.	4

Yes

Was your level of SA satisfactory?

No

My SA with respect to the task was reduced. I was unaware of some of the important information required to perform the task effectively.	5
My SA with respect to the task was insufficient. I was not aware of all the information required to perform the task effectively.	6
My SA with respect to the task was not complete. I was able to perform the task, but not satisfactorily.	7

Yes

My SA with respect to the task was good. I was able to perform the task well most of the time.	8
My SA with respect to the task was very good. I was able to perform the task well all of the time.	9
My SA with respect to the task was excellent. I was able to perform the task extremely well all of the time.	10

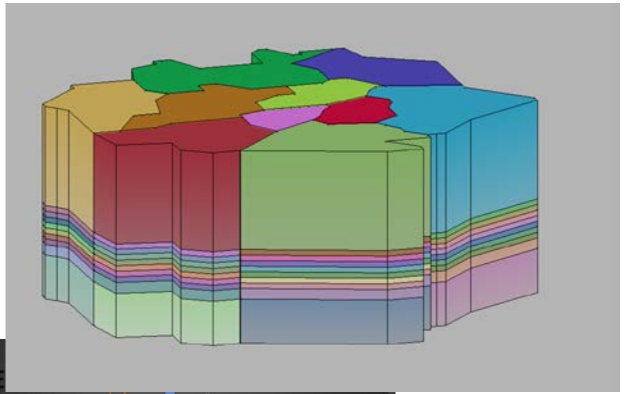
Your turn – 10 minutes

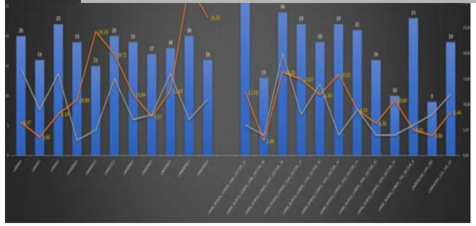
GO TO https://www.topspeed.com/car-games/car-games/octane-racing-ar183400.html

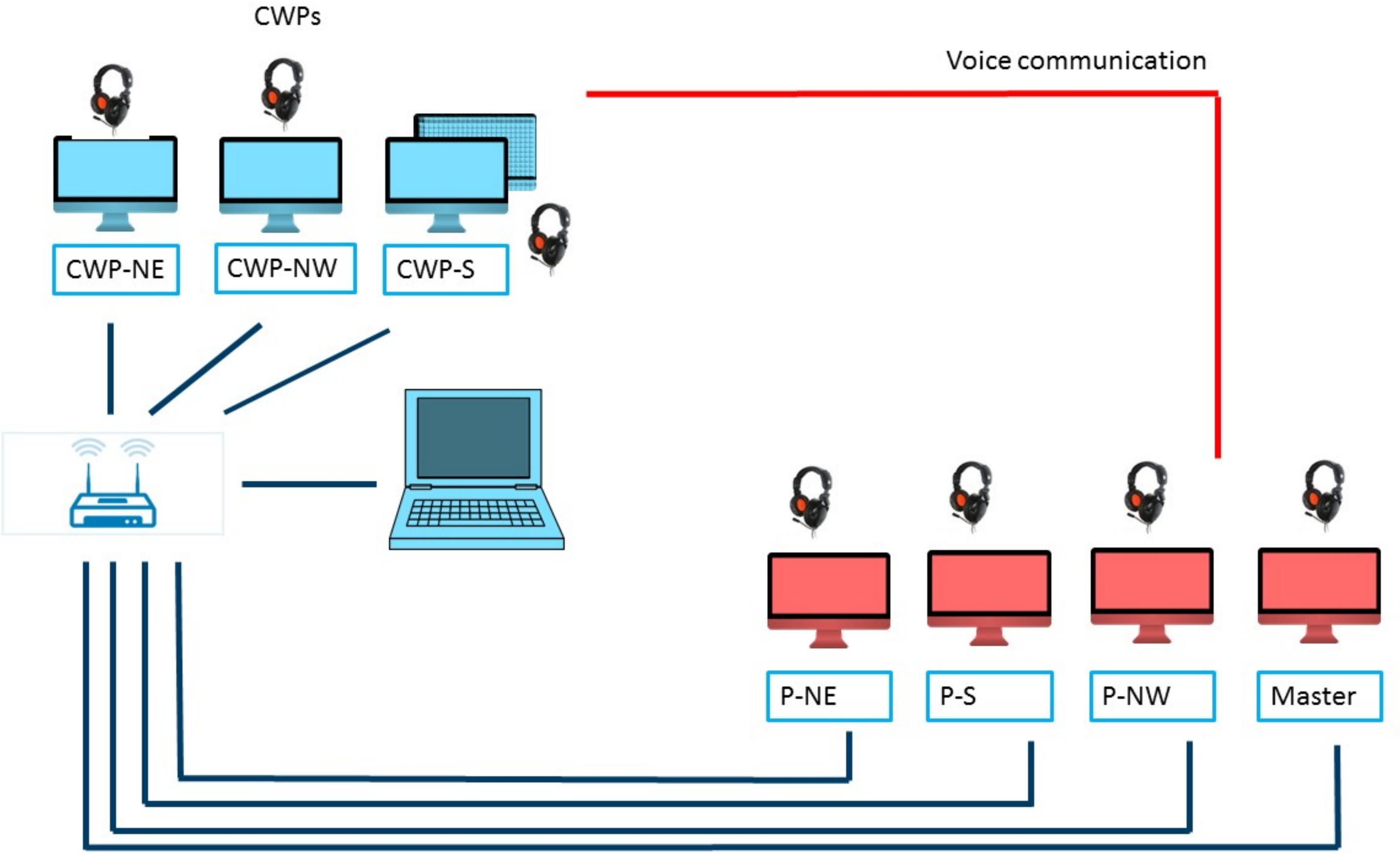
and play for a while

- Evaluate situation awareness using the given questionnaire

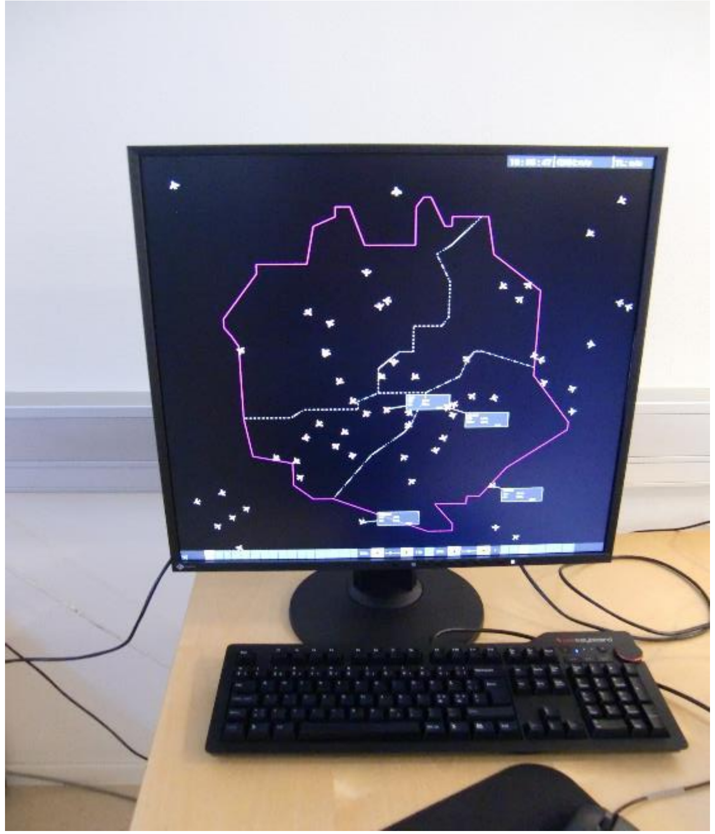
ATM example
Evaluation of DAC with ATCOs



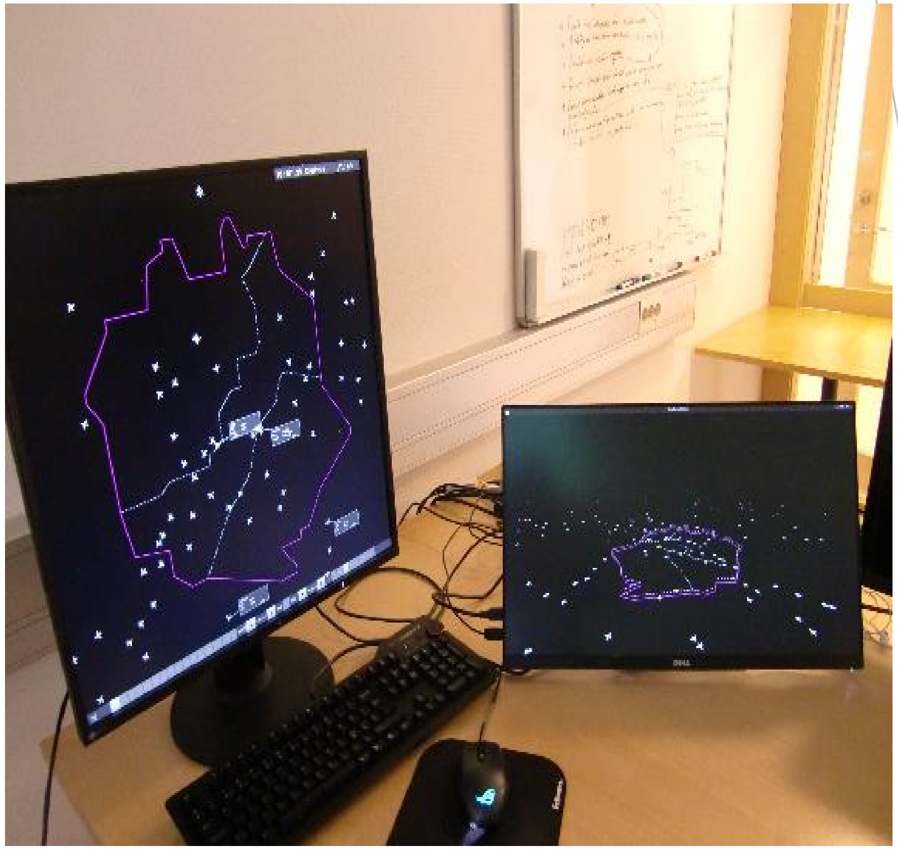




Layout of the exercise



ATCOs CWP static approach

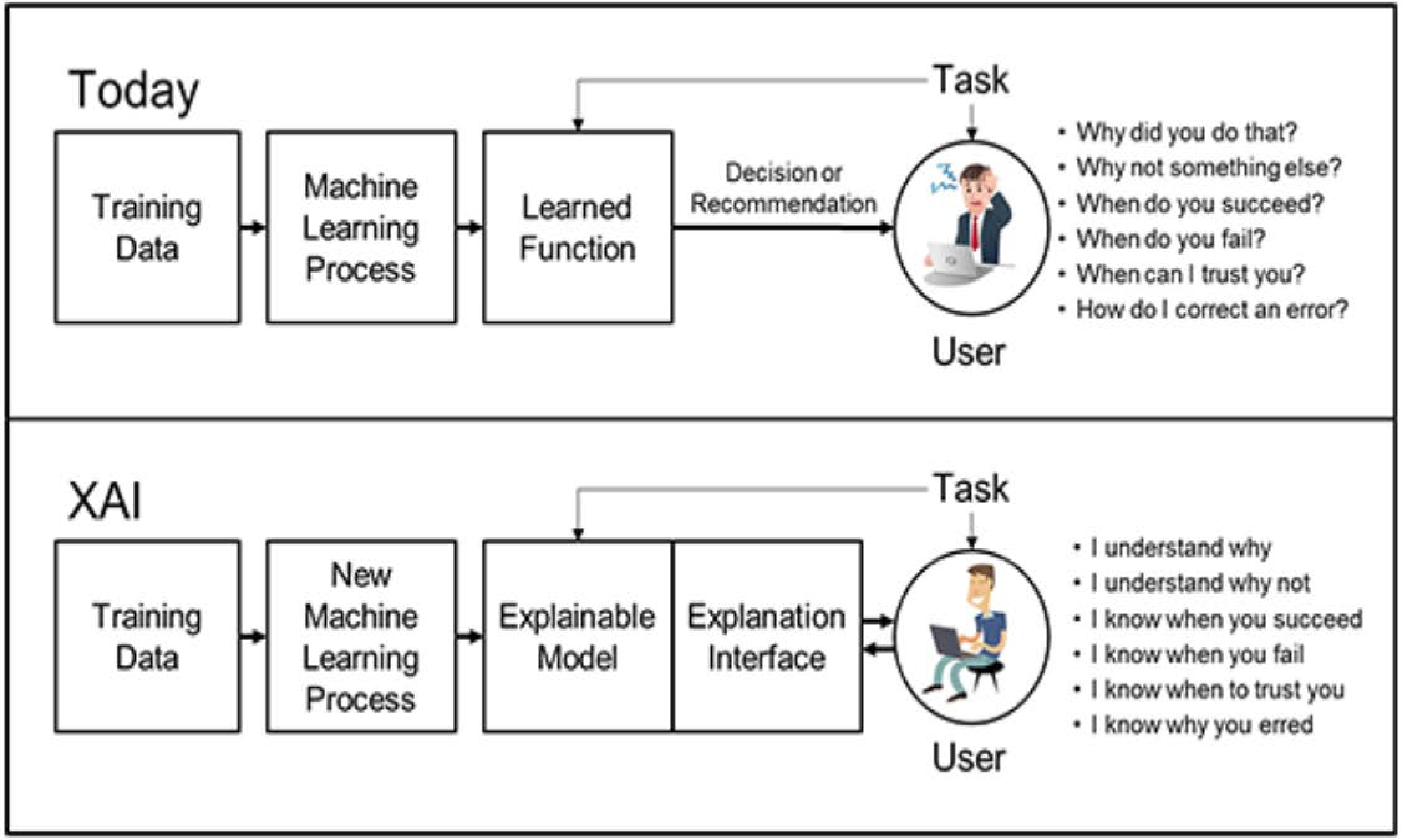


ATCOs CWP DAC

Data collection

- Observations
- Log files (simulator, UI)
- Screen captures
- Video recording of screens
- Interviews (audio records)
- Audio records of the communication between the ATCOs and pilots
- Questionnaires

Explainable AI



(from: https://www.darpa.mil/program/explainable-artificial-intelligence)

- Transparency: We have a right to have decisions affecting us explained in a language we understand
- Bias: How can we ensure that AI system has not learned a biased view of the world?
- Fairness: Can we verify the fairness of decisions?
- Safety: Can we gain confidence in reliability of our AI system?

Module 3

Literature

Buolamwini, J. and Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, in PMLR 81:77-91 http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

De-Arteaga, M., Fogliato, R., and Chouldechova. A., 2020. A Case for Humans-in-the-Loop: Decisions in the Presence of Erroneous Algorithmic Scores. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–12. DOI: https://doi.org/10.1145/3313831.3376638

Hernández-Orallo, Evaluation in artificial intelligence: from task-oriented to ability-oriented measurement, J. Artif Intell Rev (2017) 48: 397. https://dl.acm.org/doi/10.1007/s10462-016-9505-7

****

Endsley, Mica R. Designing for Situation Awareness: An Approach to User-Centered Design, Second Edition CRC Press. 2011 (chapters 2 and 10)

Hosanagar, K. A human's guide to machine intelligence, Viking, 2019 (chapters 7- 10)

****

iHUMAN documentary by Tonje Hessen Schei
https://tv.nrk.no/program/KOID75003817

Group assignment Deadline – the final report deadline

(new) Appendix 3: Evaluation - Evaluation plan, findings and reflections. Each group is to plan the evaluation of their own chatbot or a publicly available chatbot of their choice. The evaluation should include an evaluation using the guidelines for Human-AI Interactions and an abusability test. Briefly describe the subject and the scope of the evaluation, the evaluation plan, your findings, and lessons learned. Approx. 3 pages.

Individual assignment Deadline – the final individual report deadline

Human AI collaboration

Philips at al. (2016) give a taxonomy and examples of human-robots collaboration. Choose 2-3 examples, describe their levels of autonomy as described in Shneiderman (2020) and reflect on advantages and disadvantages if we decrease/increase their current level of autonomy. Reflect on their current and needed explainability (Hagras, 2018; Smith-Renner et al. 2020).