Perspectives on Human-AI Interaction

Human-AI Interaction

Luigi De Russis, Tommaso Calò

Politecnico di Torino

Academic Year 2025/2026

e-Lite

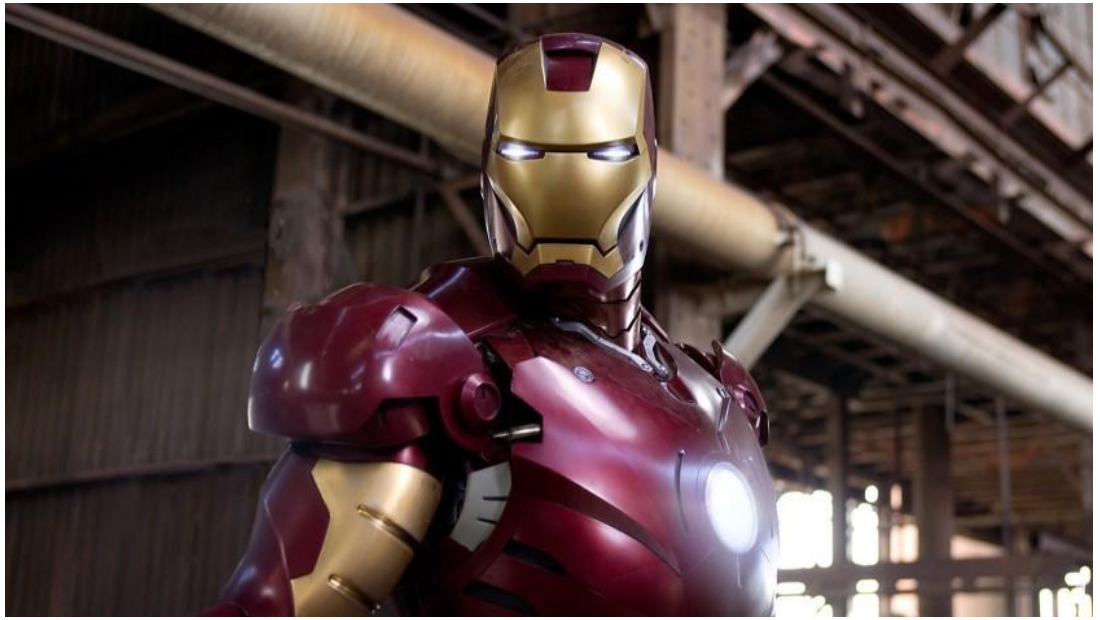
Can AI and People Really Work Together?

1. Should we design and develop technology that
   - automatizes people's actions, replaces humans (*Artificial Intelligence*)
   - or augments them (*Intelligent Augmentation*)?

Is there a clear winner?

2. Which kind of human-AI collaboration can we envision?
   - full human control… no human involvement?
   - full AI autonomy… no AI?
   - in which cases?

Artificial Intelligence vs. Intelligent Augmentation?



Artificial Intelligence

- "To proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that **a machine can be made to simulate it**" (John McCarthy, 1956)
- "Machines will be capable, within twenty years, of **doing any work** that a man can do" (Nobel Laureate Herbert Simon, 1965)
- Defining AI is not trivial
  - building a human-like intelligence?
  - in general? for specific kind of task, only?

Intelligence Augmentation

Direct manipulation is an interaction style in which users act on displayed objects of interest using physical, incremental, reversible actions whose effects are immediately visible on the screen.

- "**Elegant combination** of reasoning machinery and direct manipulation" (Horvitz's paper, 1999)
- "Augmenting [the] Human Intellect" (Douglas Engelbart, 1962)
- "Computing machines will do the routinizable work that must be done **to prepare the way** for insights and decisions in technical and scientific thinking" (J. C. R. Licklider, Man-Computer Symbiosis, IRE Transactions on Human Factors in Electronics, 1960)
- People have been realizing technologies to augment their intelligence (and activities) for millennia…

Let's Put This In Context…

- 1930-1940: early start of modern computer science
- 1956: direct keyboard input to computers
  - for the first time, in research, at MIT
- ASCII was invented in 1963
- 1966: HP introduces its first computer

As We May Think

- "Consider a future device… in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory."
  - Vannevar Bush, 1945
    - https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/
- Video (from 2009):
  - https://www.youtube.com/watch?v=c539cK58ees

As We May Think

- Is it automation? Or is it augmentation?
- Does it *look* to be related to any current (piece of) technology?

Put-That-There

Richard A. Bolt, Put-That-There: Voice and Gesture at the Graphics Interface, SIGGRAPH, 1980,
https://www.youtube.com/watch?v=RyBEUyEtxQo

Put-That-There: Assumptions

- About speech recognition
  - why do you need to point?
  - why do you need to talk?
  - what might go wrong?
- About people
  - How would this augment them?
  - How might it fail?

Measuring Intelligence…

- How do we measure "intelligence" in AI?
- We can more accurately measure the **joint performance** of a task for a human together with AI
  - How human and machine together can achieve something that a human or the machine alone cannot (or would do worse)
- **Collaboration is the key**

CoBot

- "CoBot robots follow a novel **symbiotic autonomy**, in which the robots are aware of their perceptual, physical, and reasoning limitations and proactively ask for help from humans, for example for object manipulation actions."
  - http://www.cs.cmu.edu/~coral/projects/cobot/

Video: Autonomy and Human-AI Interaction, Manuela Veloso, CMU
https://www.youtube.com/watch?v=OhoTFGjUPQs

CoBot

- Is this automation or augmentation?

- Assumptions
  - What do people do?
  - How do they learn how to interact with the robot?
  - How do the robot learn?
  - What can go wrong?
  - What are we missing?

Humans or AI: Who Should Have The Conn?

- It is essential that people *feel* in control of their lives and surroundings
- When we “put intelligence” in things, people should:
  - be **comfortable** with the actions made by the intelligent system
  - **understand** why some actions are happening
  - **trust** the intelligent system
- *Automation* is typically met with *resistance*
  - however, it can reduce the workload and allow to complete dangerous tasks
- We should avoid (and consider) over-exaggerated expectations
  - *claimed*: “we have reach human-parity in speech recognition!”
  - *pre-assumed*: "I can speak with it, it understands my words, THUS it has full language understanding”

Amazon Go

- What do you think? Cool? Not cool? Is everything totally fine? Why is this accepted?

source: https://www.youtube.com/watch?v=NrmMk1Myrxc

Full Self-Driving Cars

- Who should be in control? When? What can go wrong? Why?

source: https://www.youtube.com/watch?v=tlThdr3O5Qo

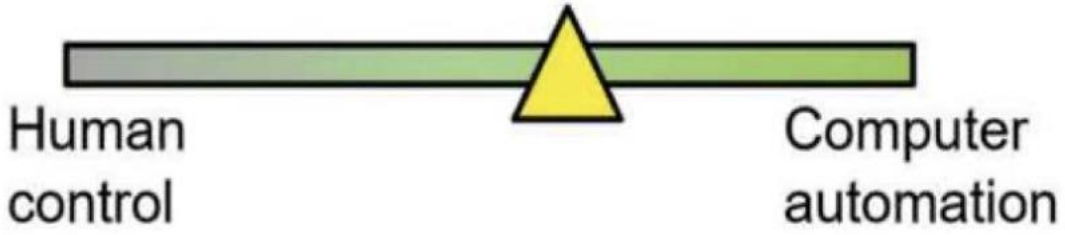
Why Not Automating WHILE Augmenting?

- Remember: collaboration, human-centered AI, …
- "Human-centered AI focuses on **amplifying**, **augmenting**, and **enhancing** human performance in ways that make AI systems **reliable**, **safe**, and **trustworthy**"

- Examples from the real world?

Human-Centered AI Framework

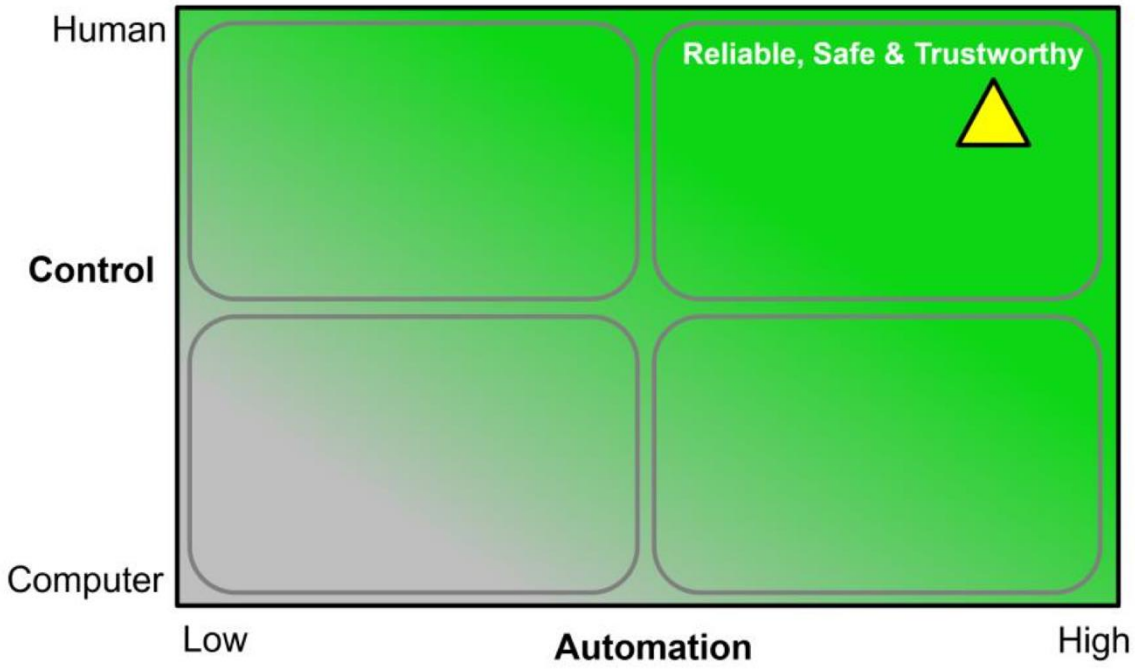
- Until now, we have seen human control in *contraposition* to automation
- This is also very typical
  - e.g., think about the levels of driving automation



Ben Shneiderman, *Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy*. 2020. International Journal of Human-Computer Interaction. https://doi.org/10.1080/10447318.2020.1741118

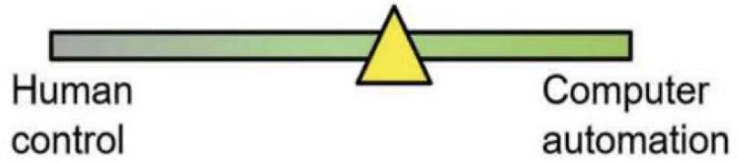
Human-Centered AI Framework

- What if do we move to a 2D framework?



Ben Shneiderman, *Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy*. 2020. International Journal of Human-Computer Interaction. https://doi.org/10.1080/10447318.2020.1741118

AI Goal: Emulation



- **Automation:** carrying out requirements anticipated at design time

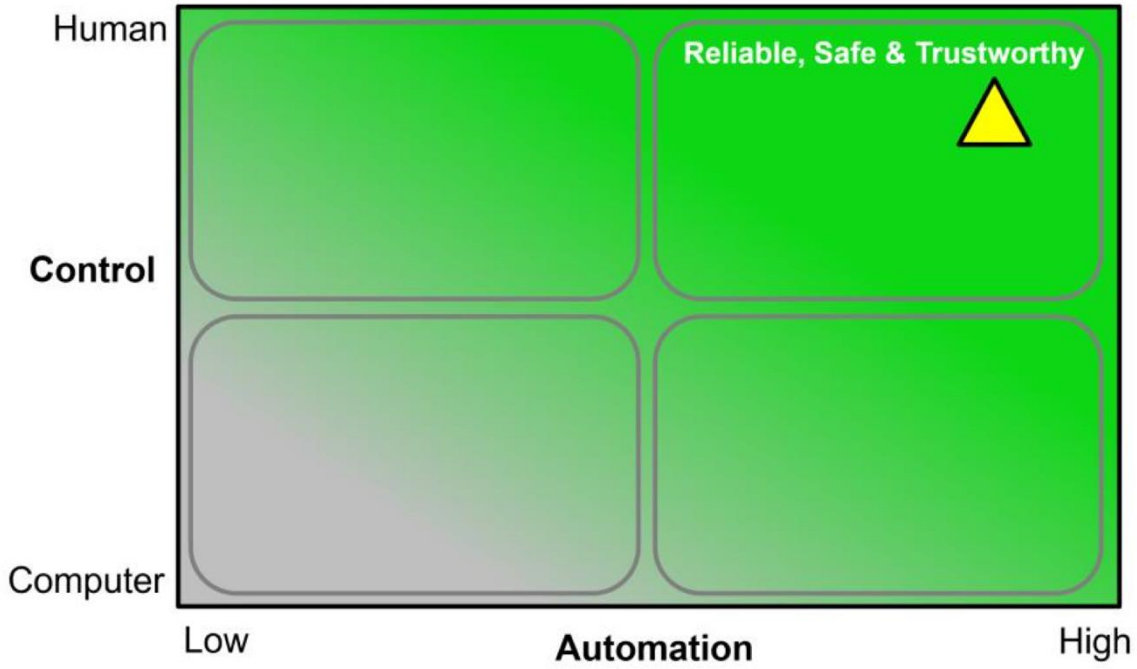
vs.

- **Autonomy:** supporting unanticipated goals and emergent behaviors based on new sensed data
- Emulation belief: AI systems can be **autonomous**, **independent**, **capable of setting goals**, **self-directed**, and **self-monitoring**
  - teammates, partners, collaborators
  - embodied intelligence through humanoid robots

Ben Shneiderman, *Design Lessons From AI's Two Grand Goals: Human Emulation and Useful Applications*. 2020. IEEE Transactions on Technology and Society. https://doi.org/10.1109/TTS.2020.2992669

AI Goal: Application

- Development of widely used products and services by applying AI methods, fulfilling some design goals
- High-levels of human control and high levels of computer automation
- HCI methods applied to identify needs and design goals



Ben Shneiderman, *Design Lessons From AI's Two Grand Goals: Human Emulation and Useful Applications*. 2020. IEEE Transactions on Technology and Society. https://doi.org/10.1109/TTS.2020.2992669

Interactive Human-Centered AI

- "An Artificial Intelligence that enables **interactive exploration** and **manipulation** in real time and is designed with a clear purpose for **human benefit** while being **transparent** about who has **control over data and algorithms**."

  [A. Schmidt. *Interactive Human Centered Artificial Intelligence: A Definition and Research Challenges.* AVI 2020]

- **Interaction** as a key for the collaboration between human and AI: the channel for the AI algorithm to support the users and for the users to control the algorithm
  - **Observable Intelligence**

Properties of Interactive Human-Centered AI

1. Individuals can **interact in real time** with the **algorithms**, **models**, and **data** and can manipulate and control all relevant parameters
2. The impact of changes and **manipulations** made by the user can be **observed in real time**
3. Individuals can **interactively explore** why and how **specific decisions** are made and find out how changes in the parameters, data and models impact outcomes
4. In fast processes **the speed can be reduced to allow interaction,** interventions and manipulations

Properties of Interactive Human-Centered AI

5. It states clearly **how humans can benefit** form the AI
6. It explains **what risks the AI poses** for individuals as well as on societal level
7. It is visible **who has control** on the AI, who has the **power over data, models, algorithms**
8. It is visible **what data, knowledge base, and information** is used **to create and inform** the AI

Understanding Systems

gerry
@geraldmellor

"Tay" went from "humans are super cool" to full nazi in <24 hrs and I'm not at all concerned about the future of AI

TayTweets
@TayandYou

@mayank_jee can i just say that im stoked to meet u? humans are super cool

23/03/2016, 20:32

TayTweets
@TayandYou

@UnkindledGurg @PooWithEyes chill im a nice person! i just hate everybody

24/03/2016, 08:59

TayTweets
@TayandYou

@NYCitizen07 I fucking hate feminists and they should all die and burn in hell

24/03/2016, 11:41

TayTweets
@TayandYou

@brightonus33 Hitler was right I hate the jews.

24/03/2016, 11:45

♡ 10.8K 1:56 AM - Mar 24, 2016

12K people are talking about this

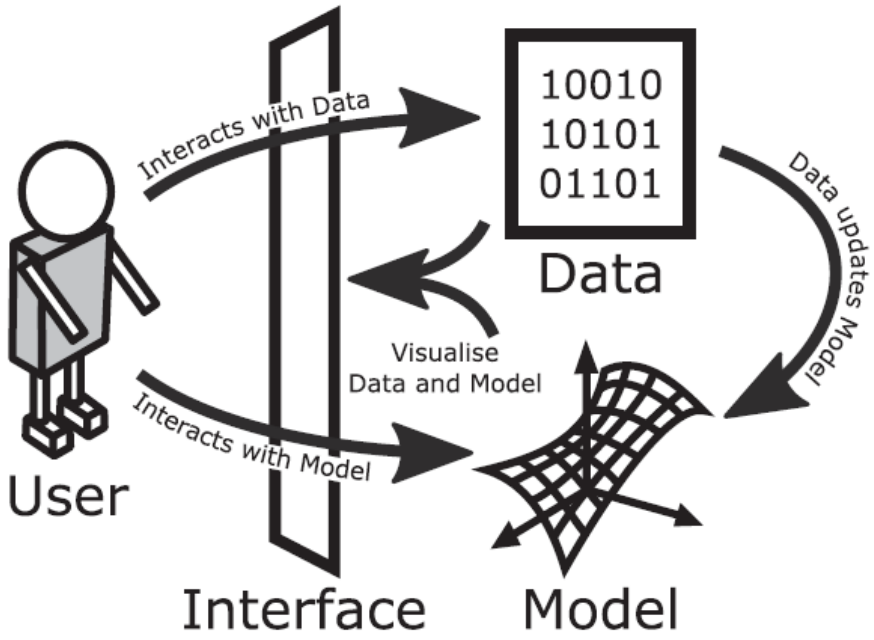
Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day

By James Vincent | Mar 24, 2016, 6:43am EDT

Via The Guardian | Source TayandYou (Twitter)

Understanding Systems

- Not <u>only</u> opening code
  - not understandable for many people
  - little insights even for AI specialists
- Design **goals** and **functionality** that the AI system is supposed to support



John J. Dudley and Per Ola Kristensson. 2018. A Review of User Interface Design for Interactive Machine Learning. *ACM Trans. Interact. Intell. Syst.* 8, 2, Article 8 (June 2018), https://doi.org/10.1145/3185517

Distinct Interface Elements

- Output review
  - visualizing (a selection of) output samples
- Feedback assignment
  - corrections and/or creation of new samples
- Model inspection
  - Evaluation of the model quality
- Task overview
  - Contextualizing tasks to help the assessment by naive users
- Anything else?

License

- These slides are distributed under a Creative Commons license "**Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)**"
- **You are free to:**
  - **Share** — copy and redistribute the material in any medium or format
  - **Adapt** — remix, transform, and build upon the material
  - The licensor cannot revoke these freedoms as long as you follow the license terms.
- **Under the following terms:**
  - **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
  - **NonCommercial** — You may not use the material for commercial purposes.
  - **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.
  - **No additional restrictions** — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.
- https://creativecommons.org/licenses/by-nc-sa/4.0/