

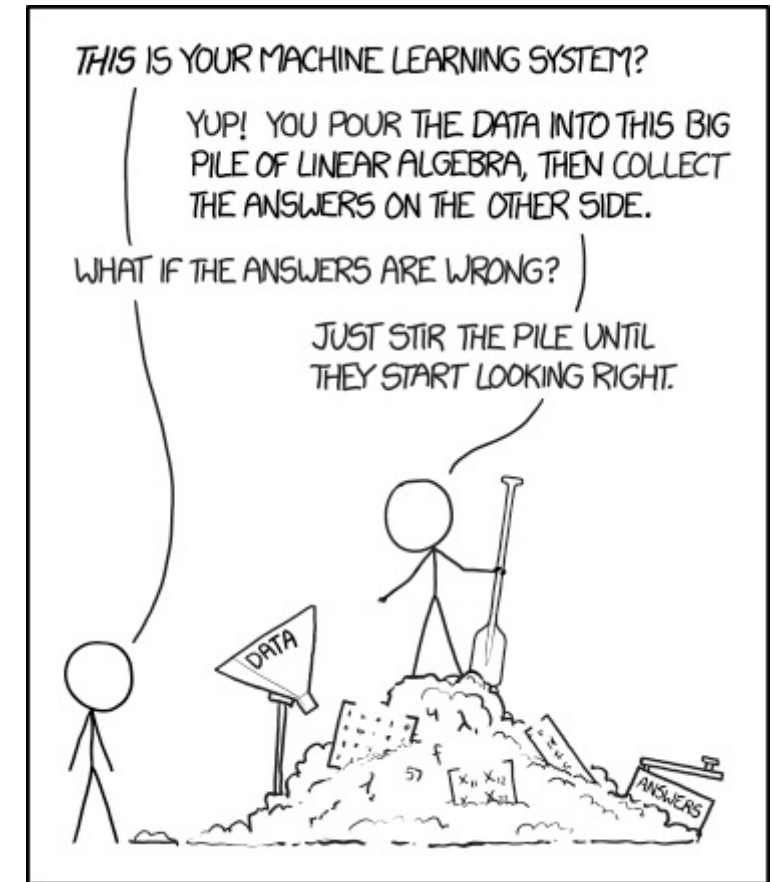
Human-AI Interaction

Introduction

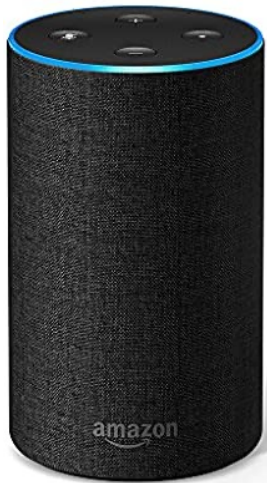
Luigi De Russis, Alberto Monge Roffarello

Let's Try to Set a Baseline...

- Classification? Clustering?
- Classification vs. regression?
- Unsupervised vs. supervised learning?
- Ontology?
- Cold start problem?
- Precision and recall?
- Expert Systems?



AI is everywhere!

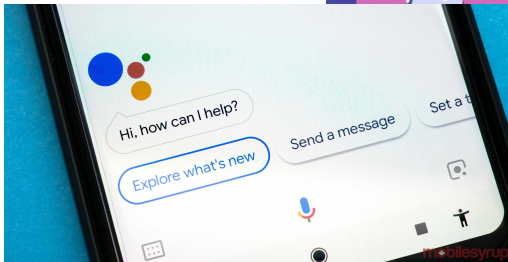


Chapter 1
What is AI?

Section	Exercises
I. How should we define AI?	1/1
II. Related fields	0/2
III. Philosophy of AI	0/1

Chapter 2
AI problem solving

Section	Exercises
I. Search and problem solving	0/1
II. Solving problems with AI	0/1
III. Search and games	0/1



Chapter 5
Neural networks

Section	Exercises	Section	Exercises
I. Neural network basics	0/1	I. About predicting the future	0/1
II. How neural networks are built	0/2	II. The societal implications of AI	0/1
III. Advanced neural network techniques	0/4	III. Summary	0/1

Create ML File Edit View Window Help

Project	Input	Accuracy	
AnimalClassifier	12	100%	97%
Model Sources	Classes	Training	Validation

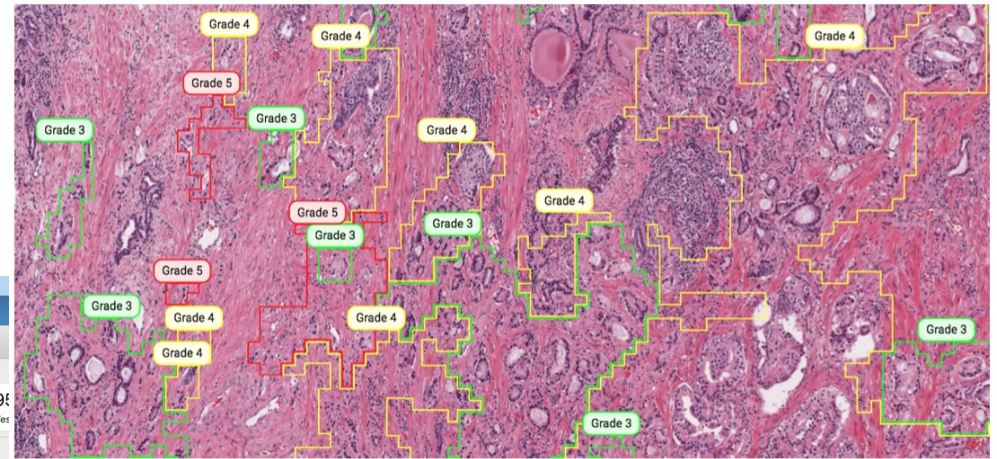
AnimalClassifier

Size 181 KB
Model Name AnimalClassifier
Author John Applesped
License License
Description A machine learning model that has been trained on a dataset of safari animals.

Drag or Add Files

Select File to Preview Output

Training completed after 32 seconds — today at 1:12 AM



AI is everywhere!

- When it "works", it's *great!*
- When it "fails", it does it *spectacularly...*
 - Tesla Smart Summon,
<https://www.youtube.com/watch?v=VbVoTK-IMoo>
 - Alexa,
<https://www.youtube.com/watch?v=QFpUN3kYTDA>



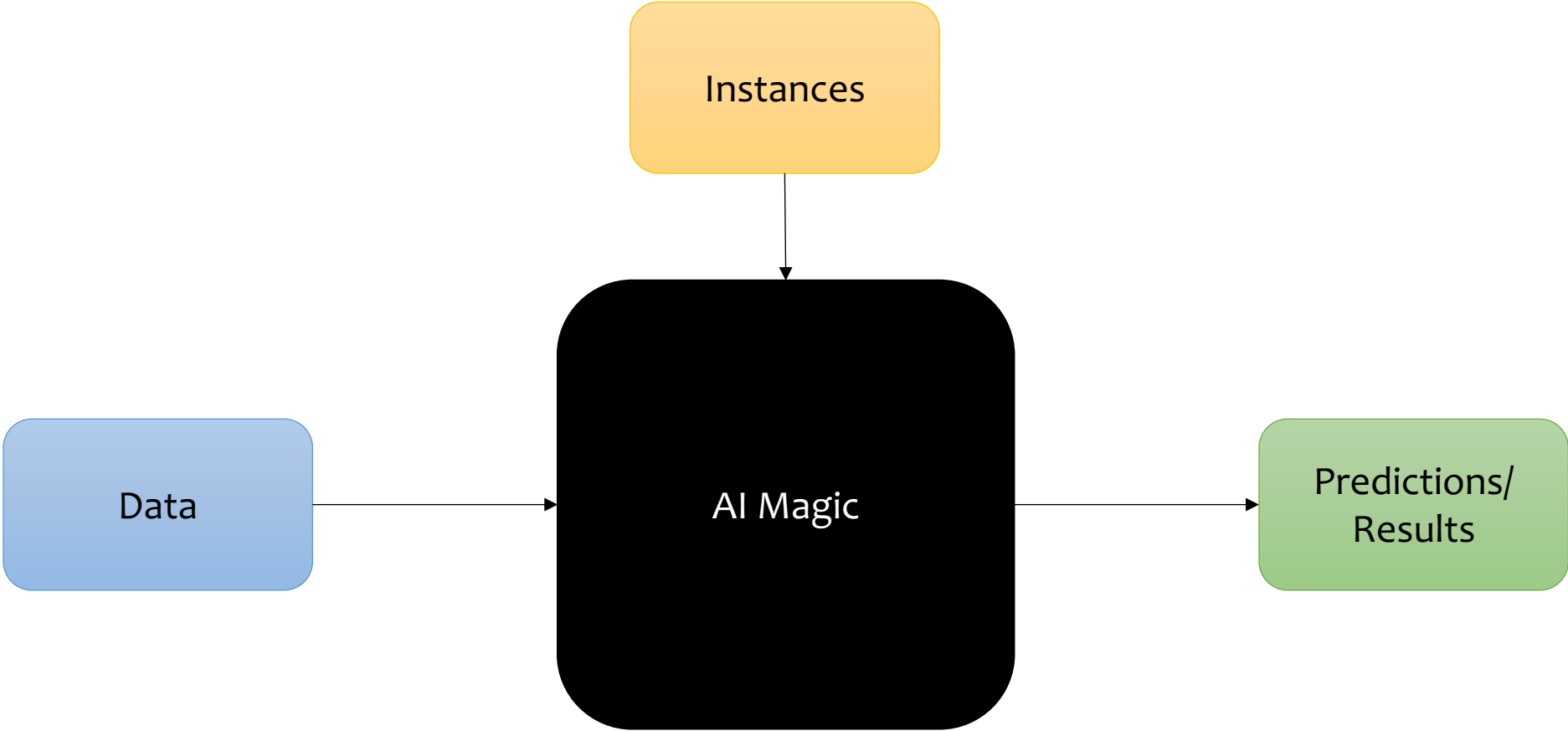
AI is everywhere!

- ... and/or it is **very** problematic, for instance:
 - *"IBM boasted that its AI could 'outthink cancer.' Others say computer systems that read X-rays will make radiologists obsolete..."*
 - *"Systems developed in one hospital often flop when deployed in a different facility.
Software used in the care of millions of Americans has been shown to discriminate against minorities.
And AI systems sometimes learn to make predictions based on factors that have less to do with disease than the brand of MRI machine used, the time a blood test is taken or whether a patient was visited by a chaplain."*

[source: <https://www.scientificamerican.com/article/artificial-intelligence-is-rushing-into-patient-care-and-could-raise-risks/>]

- Why?

A Possible Reason: The Typical Approach



Motivation

- Most AI/ML courses consider "user interfaces" or humans as an *afterthought*, near the end
 - several times they do not even think about "humans" 😞
 - they focus on algorithms/models, basically
- Why do not consider people from the *beginning*, and along the design, algorithmic choices, ... in an *iterative* way?!

Ultimately, AI Systems Are...

- Designed by *humans*
- To solve a problem framed by *humans*
- With *humans* taking specific choices (e.g., which algorithm to use)
- Evaluated and tested by *humans*
- With an outcome for *humans* (often)
- Presented to *humans* with a user interface (sometimes)

Algorithms As The (Main) Answer?

- Algorithms are **not always** the "answer"
 - for instance: if you go to Netflix for the first time, what should it recommend you watch?
 - this is the *cold start problem*, and it is not really and fully solved
 - algorithmically speaking, at least

⇒ A **suitable** user interface is **critical** to overcome some limitations!

- Keeping people involved and considering them since the beginning is **fundamental!**



Challenges

- How to ensure that people use AI-powered interfaces and systems with "joy" rather than "frustration"?
- How can we design and evaluate human-centered AI systems?
- How can we avoid (or minimize) problems, failures, ethical issues, ... in AI systems?

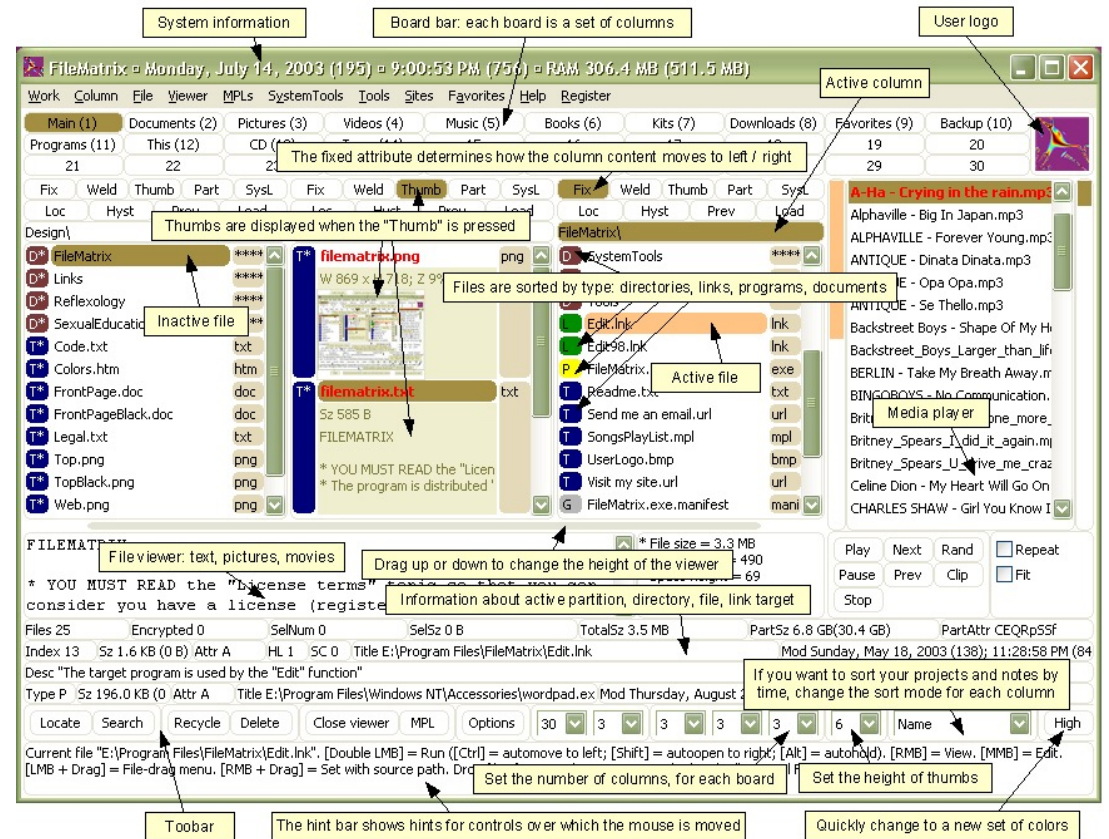
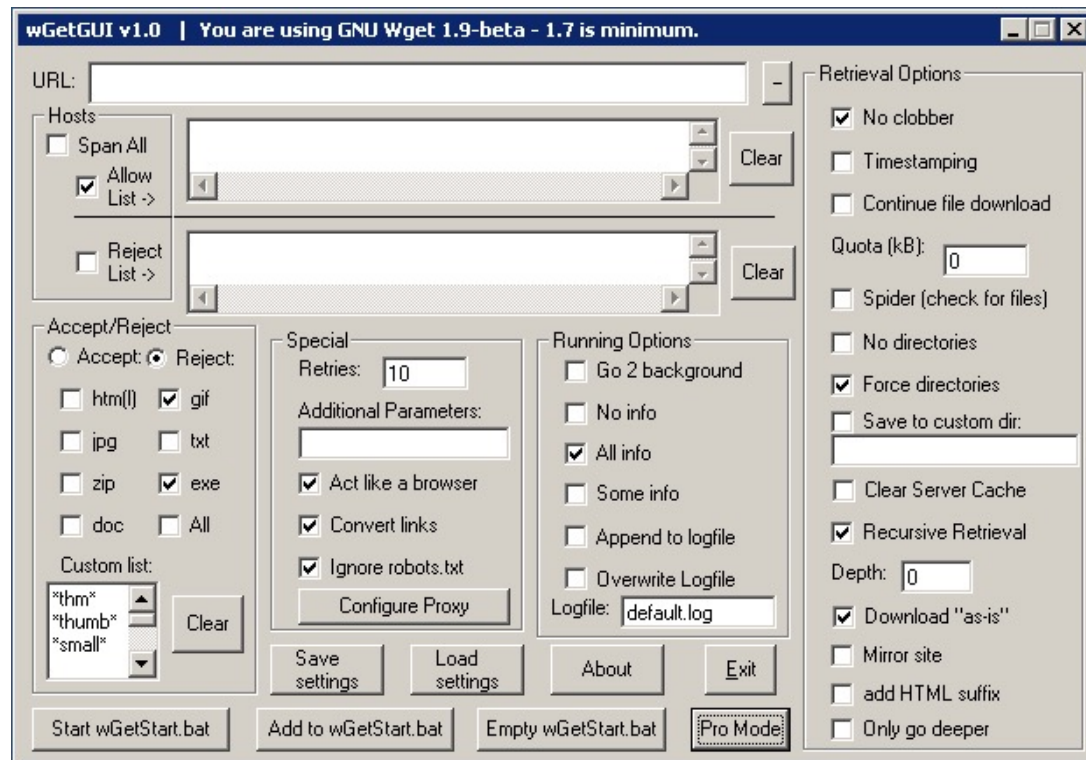
People & Computers

"The two hardest problems in computer science are: (i) people, (ii), convincing computer scientists that the hardest problem in computer science is people, and, (iii) off by one errors."

Prof. Jeffrey P. Bigham, 2018

<http://www.cs.cmu.edu/~jbigham/>

You Know... Developers...



https://thedailywtf.com/articles/Classic_WTF_-_Enter_the_Matrix

Human-AI Interaction: a Ph.D. Course

- *"Both [AI and HCI] explore the nexus of computing and intelligent behavior."*
[source: Jonathan Grudin, "AI and HCI: Two Fields Divided by a Common Focus", 2009,
<https://doi.org/10.1609/aimag.v30i4.2271>]
- We will build upon both disciplines.

What is *different* in interactive AI systems?

What is Different in Interactive AI Systems?

- AI-based systems are typically performed under **uncertainty**
 - often producing false positives and false negatives
 - they can be **incorrect**
- They may demonstrate unpredictable behaviors that can be *disruptive, confusing, offensive*, and even *dangerous* for users



What is Different in Interactive AI Systems?

- Before designing an AI system, we should ask:
 - what problems should be solved?
 - which AI approaches match human expectations given a problem?
 - which problems can be solved well enough for a particular use case?

Human-AI Interaction: a Ph.D. Course

- Great interest in **research!**
 - Human-centered AI, eXplainable AI, ...
 - you are "training" as researchers, after all
- The course will give some starting points and directions
 - research-based
 - if you want, you can go deep on different topics
 - general principles and ideas still apply

Some Other Relevant Courses @ PoliTo

- AI
 - Machine Learning and Artificial Intelligence, 6 credits, M.S. in Computer Engineering
 - Mimetic Learning, Ph.D. course
- HCI
 - Human-Computer Interaction, 6 credits, M.S. in Computer Engineering
 - Human-Machine Interaction, Ph.D. course

What Do I Mean For AI, here?

- Umbrella Term
 - Machine Learning, Knowledge Representation, Evolutionary Algorithms, ...
- Various Application Areas
 - Computer Vision, Natural Language Understanding and Processing, ...
- "Computers doing things that we expect people to be able to do"
 - Recognize if a photo contains a chair
 - Compute directions from here to Ikea
 - Infer that a chair is a piece of furniture
 - Recommend a movie

What Do You Mean For AI?

- "That one nerd-thing that is **thrilling** and **disturbing** at the same time" (x3)
- "Well-written algorithms that seem to be **intelligent**" (x3)
- "A clever organization of **deterministic operations** directed by a cleverer randomness"
- "An algorithm able to manage something **never seen before** (the opposite of being deterministic), also something able to pretend (very well) to **understand human behaviors**"
- "Machines doing stuff and **solving problems** on their own "
- "A bunch of math for **cool applications**"
- "**Smart** machines capable of performing **any tasks**"
- "Something able to **discover the unknown**"
- "An attempt to **replicate** how humans think" (x2)
- "An invention **as big as the fire** in the Stone Age", "A child **prodigy**"
- "A nice trick to **increase productivity and efficiency** in all walks life.", "Tool to **enhance life**"
- "Like a human being but **better** (hoping it **doesn't kill us all**)"

What Do You Mean For AI?

We have a winner! 😊

"Surely not the Ryanair helping bot"

What Did Your Colleagues Mean For AI? (2020)

- "Computers **taking decisions** as they are 'thinking'" (x2)
- "A machine to **answer questions** in a reasonable way"
- "[A system] **self-conscious, explainable** and show **creative** behaviors"
- "A tool that emulates the capacity of the humans to **make decisions**" (x2)
- "Human **empowerment** through intelligent **data processing**", "**Supportive** intelligence for the human being"
- "Algorithms great for **solving some/hard problems**" (x2)
- "**Transferring** human **intelligence** into machines"
- "Machines capable of **reasoning**"
- "A tool to let humans be humans by **delegating tasks** to machines"

Course Logistics

"Teaching Philosophy"

- Put persons first!
 - different backgrounds and expectations in this room
 - how to do something "good" for all of you?
- Interactivity
- Learn by doing, do by learning
 - mix of lectures, "practical" exercises, and readings
 - programming included!
- To learn something, teach it
 - Panel and workshop-style sessions

About You (hello!)

- 52 (enrolled) students
 - 18 from the Ph.D. in Computer and Control Engineering
 - 11 from the Ph.D. in Electrical, Electronics and Communications Engineering
 - 5 from the Ph.D. in Management, Product, and Design
 - 4 from the National Ph.D. in Artificial Intelligence
 - 3 from the Ph.D. in Mechanical Engineering
 - 3 from the Ph.D. in Bioengineering and Medical-Surgical Sciences
 - 3 from the Ph.D. in Metrology
 - 2 from the Ph.D. in Energetics
 - 2 from the Ph.D. in Chemical Engineering
 - 1 from the Ph.D. in Aerospace Engineering
 - 1 from the Ph.D. in Civil and Environmental Engineering

About You (hello!)

- Different (research) interests
 - IT + Medical Sciences,
 - Sustainable Data-Driven Design,
 - Computer Vision,
 - HW Testing,
 - Graph Neural Networks,
 - Robotics,
 - Decision-Support Systems,
 - Virtual Reality,
 - Swarms and Drones,
 - Wireless Communication,
 - Trustworthy AI,
 - NLP,
 - Software Engineering, ...

About Us





- Luigi De Russis
 - Assistant Professor (RTDb)
 - Department of Control and Computer Engineering
 - Research Topic: HCI in Complex Settings
 - luigi.derussis@polito.it
- Alberto Monge Roffarello
 - Postdoc
 - Department of Control and Computer Engineering
 - Research Topic: Digital Wellbeing
 - alberto.monge@polito.it

Course Topics

- Introduction to Human-AI Interaction
- Trade-offs and perspectives in Human-AI Interaction
 - Augmenting or replacing people?
 - Direct manipulation or agents?
- Designing and evaluating human-centered AI systems
 - Guidelines and methods
 - Data, bias, and trust
- Paradigms for Human-AI Interaction
 - Smart interfaces and conversational agents
- Hands-on sessions: design and prototyping a conversational agent

Course Information

- Material
 - <https://elite.polito.it> -> Teaching -> Current Courses -> 01UJUIU – Human-AI Interaction
 - short link: <http://bit.ly/polito-hcai>
 - Slides, exercises, etc.
- Students are encouraged to attend the classes with their laptops, to work on the proposed exercises

Date	Type	Details	Time	Room	Video
12/01/2022	Lecture	Introduction	14:30-17:30	7S	
18/01/2022	Exercise	Reading panels	09:00-13:00	7S	-
21/01/2022	Lecture	Perspectives on Human-AI Interaction	09:00-12:00	29B	
25/01/2022	Lecture	Paradigms for Human-AI Interaction	09:00-10:30	7S	
	Exercise	Journey Map	10:30-12:00	7S	-
01/02/2022	Lecture	Designing and Evaluating Interactive AI Systems	09:00-11:00	7S	
	Exercise	Design & Evaluation Workshop	11:00-13:00	7S	-
03/02/2022	Exercise	Case study: designing and implementing a conversational assistant	09:00-13:00	7S	-
09/02/2022	Exercise	Case study (cont'd).	09:00-13:00	29B	-
		Final presentation.			

The Plan: Overview



- 7 classes
 - around 50% interactive lectures and 50% exercises
- 3-4 hours per class
- Schedule
 1. 12/01/2022 h. 14:30-17:30, room 7S
 2. 18/01/2022 h. 9:00-13:00, room 7S
 3. 21/01/2022 h. 9:00-12:00, room 29B
 4. 25/01/2022 h. 9:00-12:00, room 7S
 5. 01/02/2022 h. 9:00-13:00, room 7S
 6. 03/02/2022 h. 9:00-13:00, room 7S
 7. 09/02/2022 h. 9:00-13:00, room 29B

The (Tentative) Detailed Plan

Class	Type	Topic	Teacher(s)
1	Lecture	Course introduction, logistics, introduction to Human-AI Interaction.	Luigi
2	Exercise	Reading Panels. Finalize group formation.	Luigi
3	Lecture	Perspectives on Human-AI Interaction.	Luigi
4	Lecture	Paradigms for Human-AI Interaction.	Alberto
	Exercise	Journey Map.	
5	Lecture	Designing and Evaluating Interactive AI Systems.	Alberto
	Exercise	Design & Evaluation Workshop.	
6	Exercise	Case Study: designing and implementing a conversational assistant.	Alberto
7	Exercise	Case Study (cont'd).	Alberto
	Exercise	Case Study (cont'd): Final Presentation.	Alberto, Luigi

Exam

Four practical activities, to be carried out **in class**.

1. Readings Panels (individual, next week) -> *to be prepared before the class!*
2. Journey Map (in group, class #4)
3. Design and Evaluation Workshop (in group, class #5)
4. Case Study Prototype and Presentation (in group, classes #6-7)

To **pass** the exam:

- 3 activities completed with success
- **MERIT** with all four activities (successfully) done

Our COVID-19 Times...

- The preferred way to follow the course is *in person*.
- However, "life happens":
 - Lectures will be video-recorded and shared after each class.
 - Group exercises can have *hybrid* groups, with at least one person of the group in the room.
 - Try hard to be in the class for the individual exercise (i.e., the panels). Send me an email if you **cannot really** be there but you want to join remotely.
- We will reserve the possibility to shift the course entirely online, if needed...

TODO: Group Composition

- Form a group for the upcoming activities
- **3-5 people** per group
- Fill up this spreadsheet:
 - <https://docs.google.com/spreadsheets/d/1JrluovlsTPnMV33Wp6joUOi6IUeBbw246-h8Qu4MDfA/>
- **Deadline:** Friday 21st of January 2022

About Programming...

- Do you know "enough" programming?
- You need to know some Python (preferably)
 - other languages may be ok (e.g., JavaScript, Java, ...)
- Needed for the **case study**, only
 - We will provide examples and projects to get started with (~1)
 - We will be here (*obviously!*)

Questions?

I Have Some Questions For You...

- I am a ML expert, a smart home enthusiast, and I applied AI in my home
- After an *adequate* period of data collection about my habits at home, I wrote a ML system to automatize my most frequent habits
- For instance, the AI detected that:
 - almost every morning, Mon-Fri, I wake up at 6:30
 - then, I turn on the light
 - I open the window for around 10 minutes
 - I start my coffee machine
- The system automatically executes these steps

I Have Some Questions For You...

- Is it a good problem to solve?
- Does it solve the "morning routine" totally?
- What can go wrong?
- Any failures and possibility to recover?
- Better ways to do this?
- ...

Fundamentals

Human-Computer Interaction and Human-Centered AI

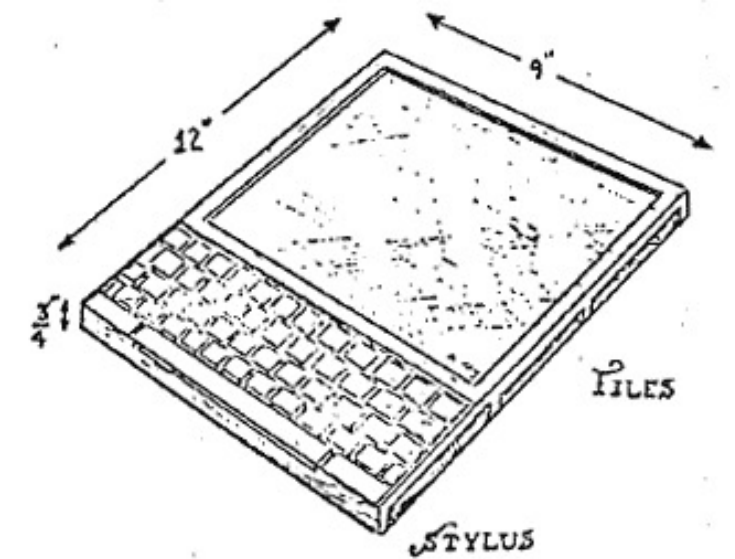
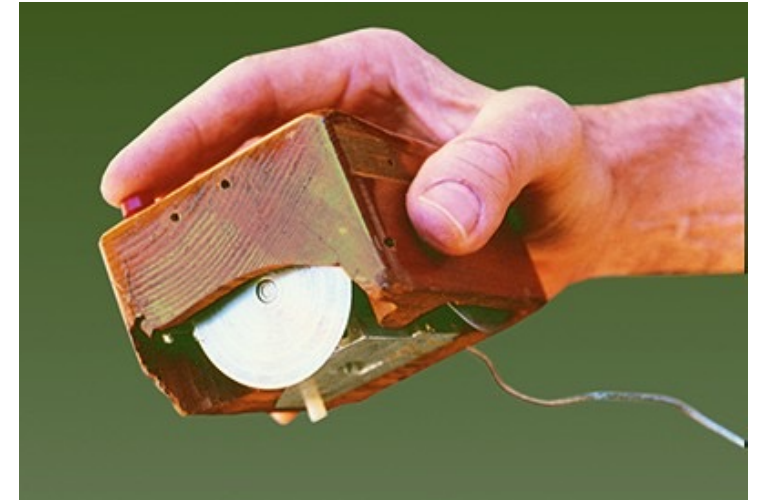
AI+HCI

- "From the earliest times in the development of computers, activities in the field of Human-Computer Interaction (HCI), and Artificial Intelligence (AI) have been intertwined. But as subfields of Computer Science, AI and HCI have always had a love-hate relationship."
- "Together, the community can make user interfaces a little less stupid and frustrating than they are today."

[source: Henry Lieberman, "User Interface Goals, AI Opportunities", AI Magazine, 2009, <https://doi.org/10.1609/aimag.v30i4.2266>]

Human-Computer Interaction (HCI)

- A multi-disciplinary field
- Concerned with the **design, evaluation, and implementation** of interactive computing systems for human use
 - and with the study of **major phenomena** surrounding them
- Involves two *entities* (the human and the computer) that determine each other behavior over time
 - framed in terms of humans' **goals** and related **tasks/pursuits**



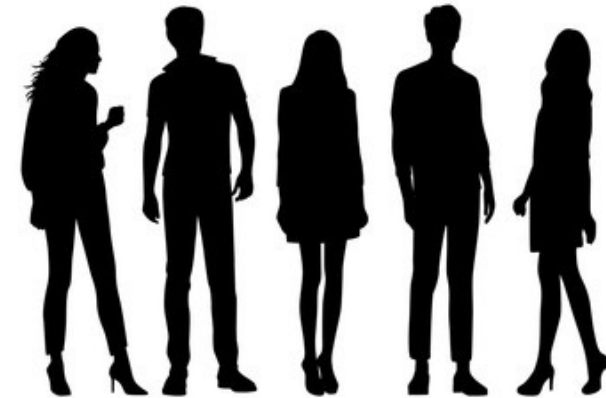
A Matter Of Perspective

Technology as an end



Technology as a mean to an end

- The end: satisfying users' needs, attitudes, and expectations



People first and at the center!

Main Concepts

▪ Key Attributes

- Usefulness: to accomplish what is required/expected
- Usability: "do it easily... don't let me think"
- Performance and robustness
- Attractiveness and engagement

▪ User Studies

- Understanding people and their needs
- Analyzing their behavior (with technology)
- Understanding how the designed solution affect people' perceptions, attitudes, and judgements

A Human-Centered Approach

- **Early focus on people and tasks:** observing humans doing their tasks and then involving them in the design process
- **User-based evaluation:** users' reactions and performance to scenarios, simulations, and later to prototypes are observed, recorded, and analyzed
- **Iterative design:** when problems are identified through user testing, fix them and carry out more tests

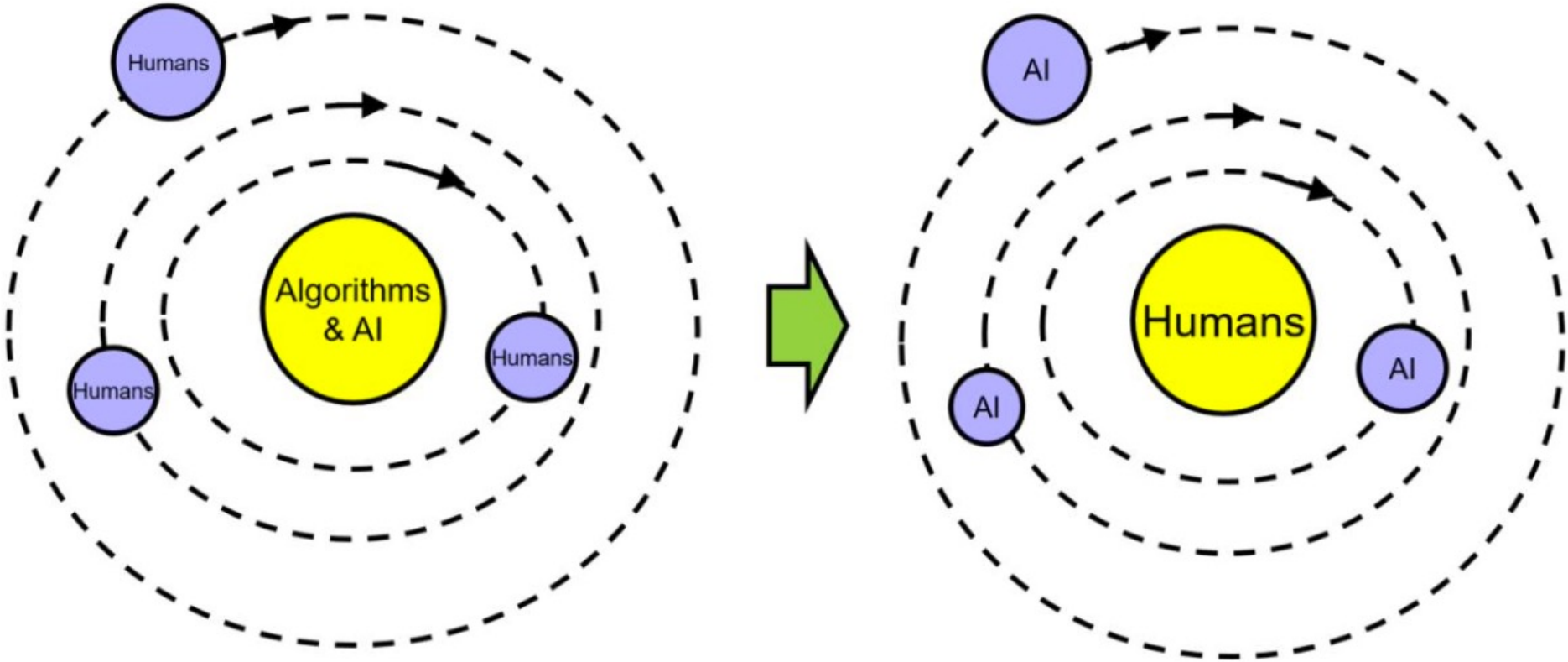
Benefits: ROI, Safety, Ethics, Innovation, ...

Human-Centered AI

- Human-centered AI focuses on **amplifying, augmenting, and enhancing** human performance in ways that make AI systems **reliable, safe, and trustworthy**
- Shift from measuring **only** algorithm performance to evaluating human performance and satisfaction, with **human-centered** and participatory approaches (for evaluation, too)

Ben Shneiderman, *Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-centered AI Systems*. ACM Transactions on Interactive Intelligent Systems, Vol. 10, No. 4, Article 26, 2020, <https://doi.org/10.1145/3419764>

A Paradigmatic Change



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/>

