

Lecture 1

AI & HCI: A historical perspective on why now is the right time

Who are we?

Ranjay Krishna - PhD from Stanford

- I worked with Fei-Fei Li (**AI**)
- And with Michael Bernstein (**HCI**)

I conduct two types of research inquiries:

- I study emergent **human behaviors** when they interact with AI systems
- I develop better AI (specifically **computer vision**) systems with these insights

Past courses:

Stanford CS 231N: Convolutional neural networks for computer vision

Stanford CS 131: Computer vision fundamentals and applications



Who are we?

Jiafei Duan - robotics PhD student at UW

- I work with Dieter Fox (**Robotics**)
- And with Ranjay Krishna (**Vision**)

I work on scaling up robotics manipulation data collection.

My research interests lie in embodied AI, robot learning, computational cognitive science, and computer vision.



Class activity: Who are you?

Lecture 1

A new hope for AI with HCI

A historical perspective on why now is the right time

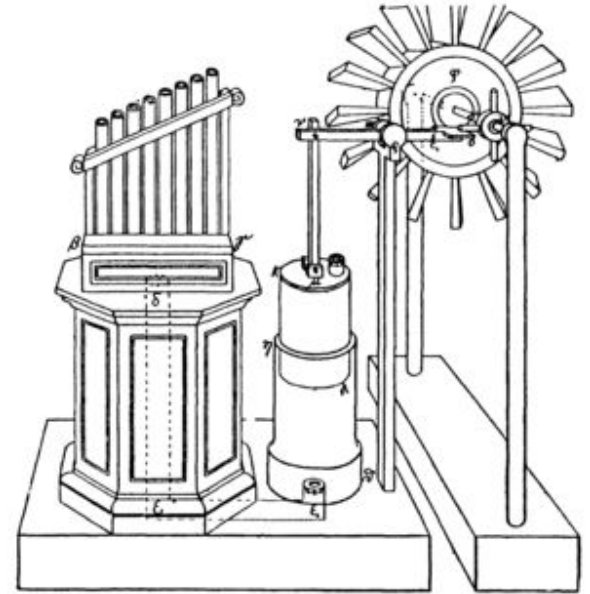
Science stands on the shoulder of giants

We can only design for the future if we understand the past

Automation: Prehistory of AI (2000 years ago)

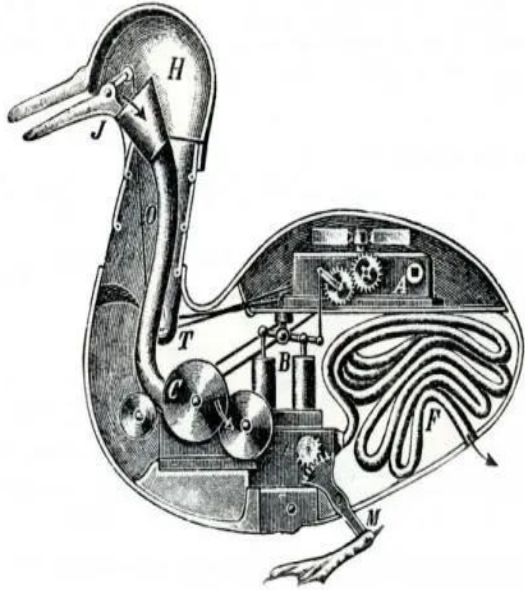


The Antikythera mechanism predicts astronomical positions and eclipses (circa 100 BC)



Hero's contraption to harness wind power (circa 60 AD)

Prehistory of AI (18th & 19th centuries)



Vaucanson's Duck (1738-1739)

Ranjay Krishna | ranjay@cs.washington.edu



Möllinger's dulcimer-playing automaton (Circa 1770)



Manzetti's flute-playing automaton (1840)

Depictions of AI: Tricks

Mechanical Turk: a fraudulent chess-playing machine constructed in the late 18th century. (A human grandmaster hid inside.)



Depictions of AI: Tricks

The screenshot shows the Amazon Mechanical Turk website. At the top, there's a navigation bar with 'Your Account', 'HITS', and 'Qualifications' tabs. Below that, a blue banner contains the text: 'Mechanical Turk is a marketplace for work. We give businesses and developers access to an on-demand, scalable workforce. Workers select from thousands of tasks and work whenever it's convenient. 264,053 HITS available. View them now.' The main content area is split into two columns. The left column is titled 'Make Money by working on HITS' and lists benefits for workers, such as working from home and flexible hours. It includes a flow diagram: 'Find an interesting task' (represented by a gear icon) -> 'Work' (represented by a gear icon) -> 'Earn money' (represented by a dollar sign icon). The right column is titled 'Get Results from Mechanical Turk Workers' and lists benefits for requesters, such as access to a global workforce and fast completion times. It includes a flow diagram: 'Fund your account' (represented by a plus sign icon) -> 'Load your tasks' (represented by a document icon) -> 'Get results' (represented by a star icon). The footer contains links for 'FAQ', 'Contact Us', 'Careers at Amazon', 'Developers', 'Press', 'Policies', and 'Blog', along with the copyright notice '©2005-2012 Amazon.com, Inc. or its Affiliates' and the text 'An amazon.com company'.

Mechanical Turk: a fraudulent chess-playing machine constructed in the late 18th century. (A human grandmaster hid inside)

Make Money by working on HITS

HITS - Human Intelligence Tasks - are individual tasks that you work on. [Find HITS now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work



or [learn more about being a Worker](#)

Get Results from Mechanical Turk Workers

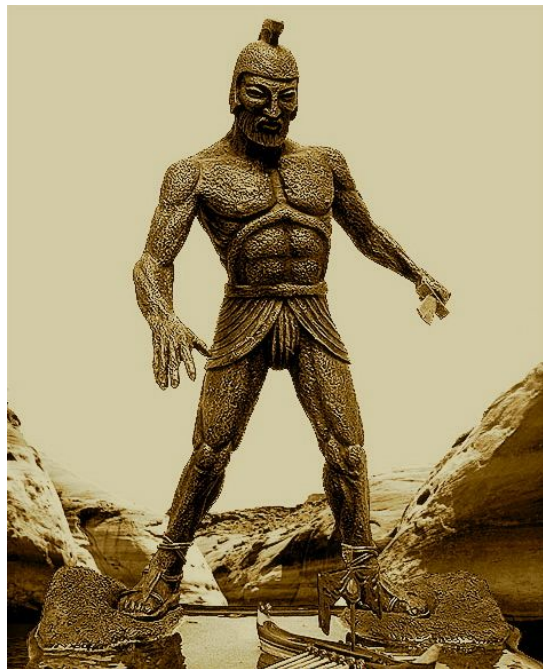
Ask workers to complete HITS - Human Intelligence Tasks - and get results using Mechanical Turk. [Register Now](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITS completed in minutes
- Pay only when you're satisfied with the results

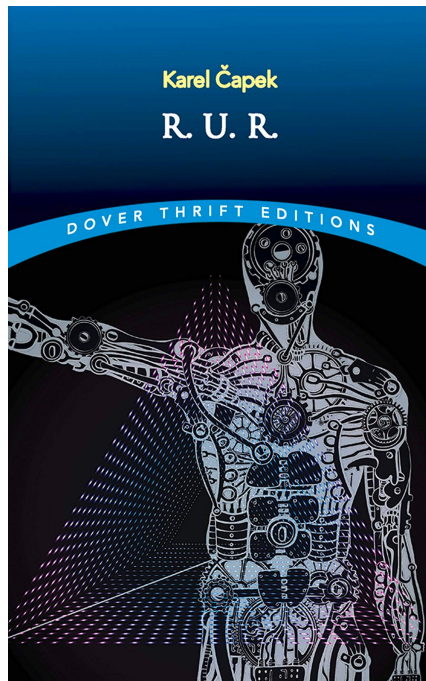


Depictions of AI: Myths and Stories



Legend of Talos

Adrienne Mayor, *Gods and Robots*



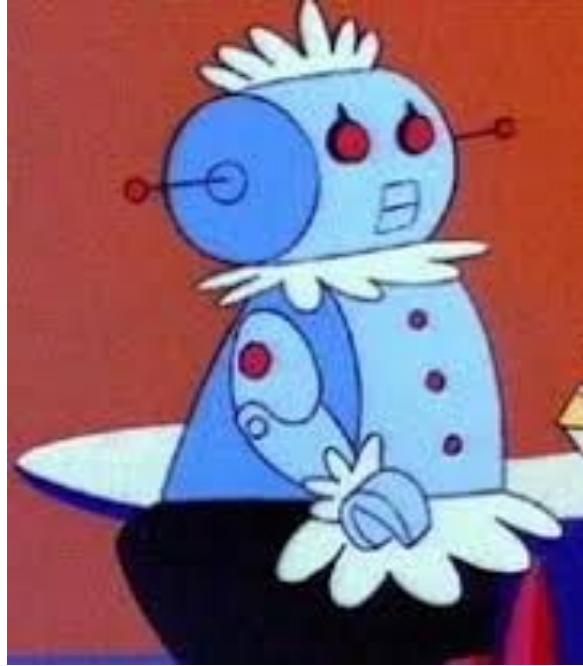
R. U. R. (1920)



Talking Computer in *Star Trek*
(1966)



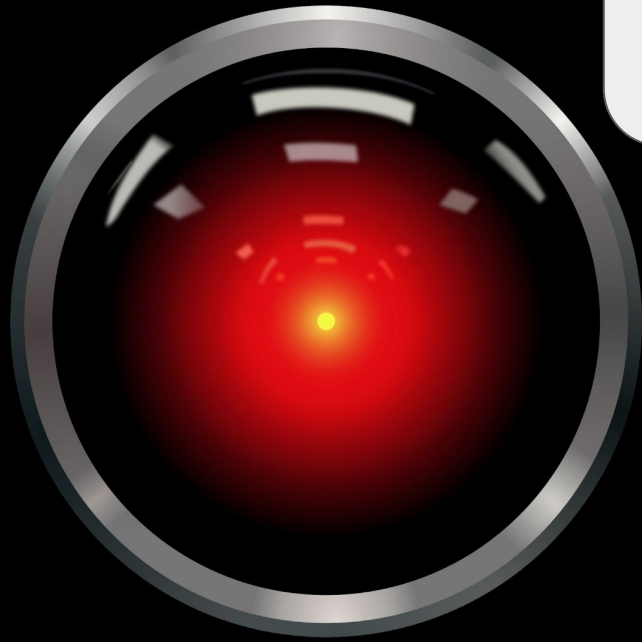
1956



1962

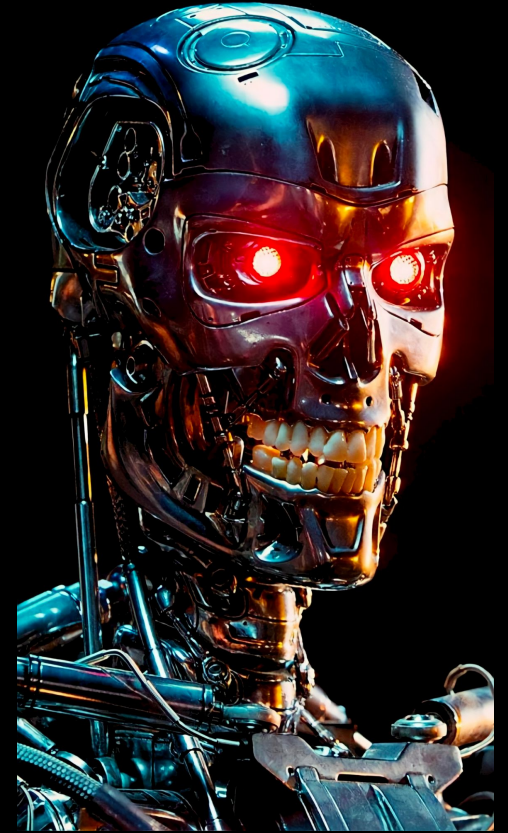


1965



1968

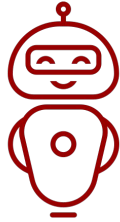
I'm sorry, Dave.
I'm afraid
I can't do that.



1984

But let's look at this history through an
academic lens

The academic start of AI and HCI



Artificial Intelligence

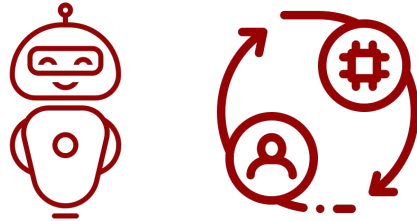
Originated in mathematics and engineering



Human-Computer Interaction

Originated in psychology

The academic start of AI and HCI



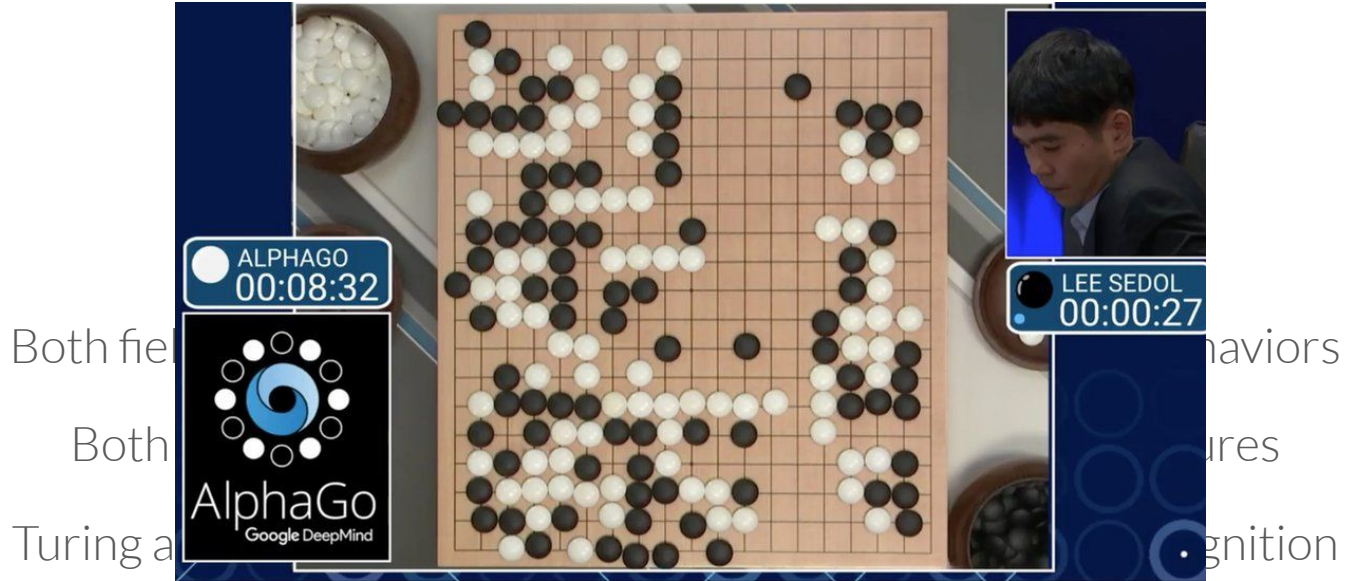
Both fields are at the nexus of computing and intelligent behaviors

Both claim **Allen Newell** and **Herb Simon** as founding figures

Turing award in 1975 for contributions to AI and human cognition

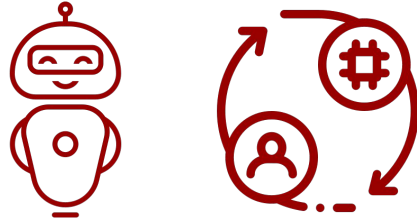
First AI program: [Logic Theorist](#) to solve mathematical theorems

The same challenge of traversing large search spaces



Wrote the [Logic Theorist](#) - program to solve mathematical theorems

So much in common: AI and HCI

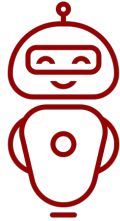


Both claim **Allen Newell** and **Herb Simon** as founding figures

General problem solver in 1959

Found CMU school of computer science

Two fields divided: AI and HCI



Artificial Intelligence

Goal: to create an artificial rival to human intelligence

Artifact: models of human intelligence

Long time horizon



Human-Computer Interaction

Goal: To improve applications as they approach widespread use

Artifact: designs for mass market products

Short time horizon

Entering the 1950's

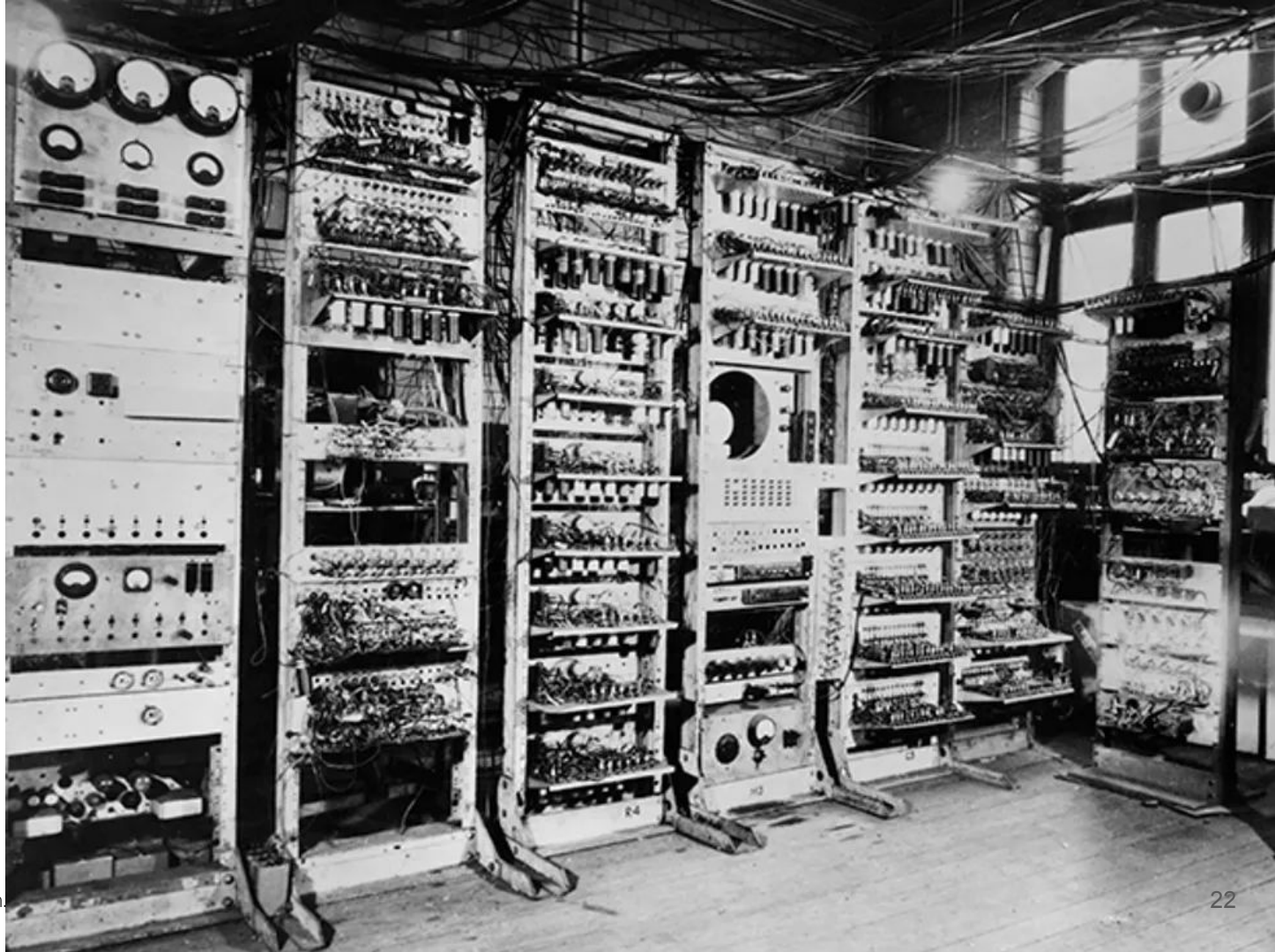
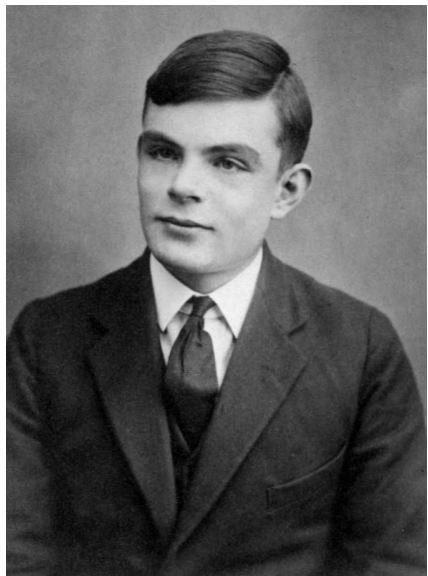


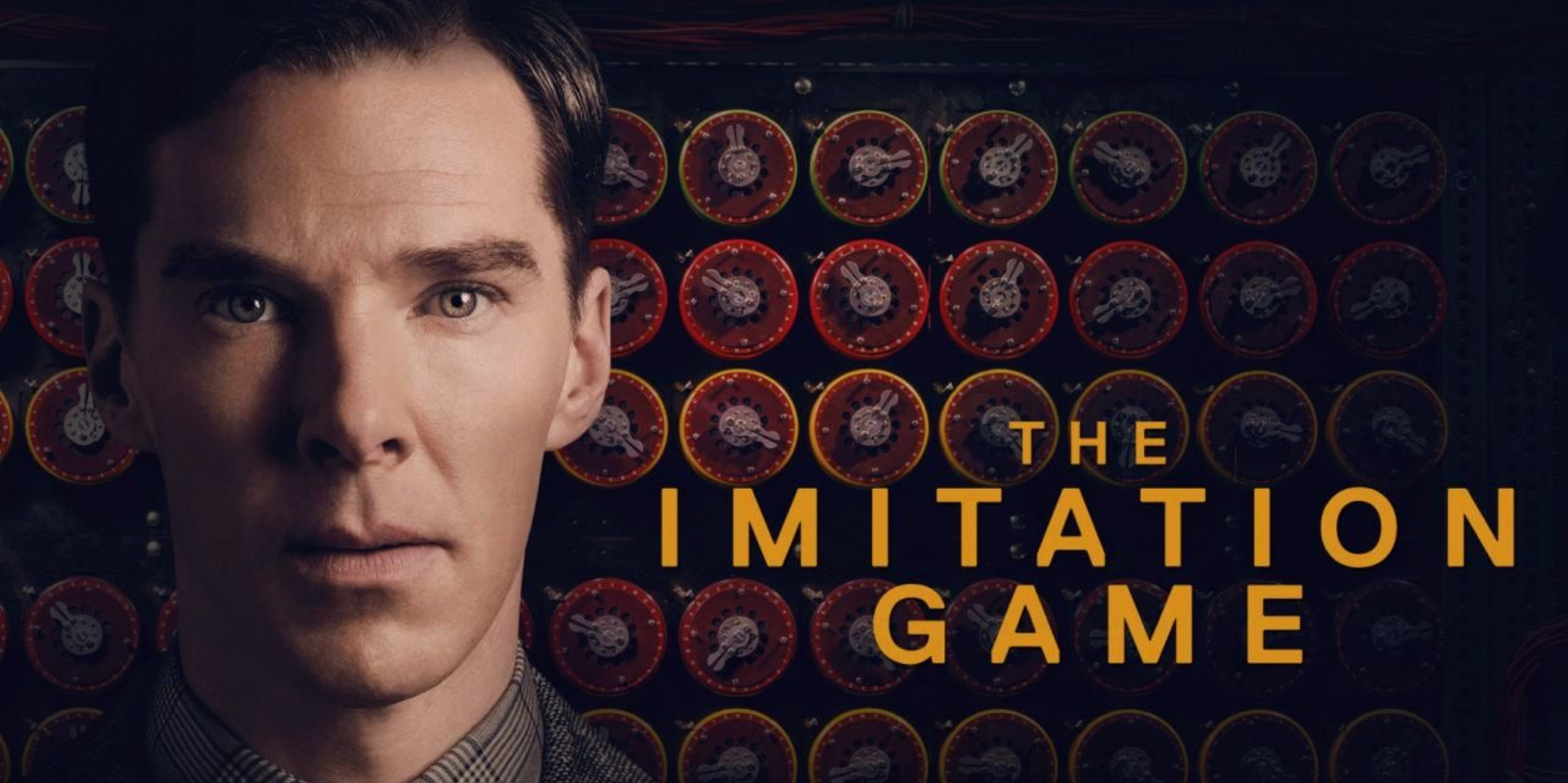
Does anyone know why the Government allocated a large budget for this?

Governments received large amounts of funding

-> large expensive computers at a few universities

Alan Turing







The Turing Test:

Computing Machinery and Intelligence - 1949

“I do not see why [the computer] should not enter any one of the fields normally covered by the human intellect, and eventually compete on equal terms. I do not think you can even draw the line about sonnets, though the comparison is perhaps a little bit unfair because a sonnet written by a machine will be better appreciated by another machine”

- Alan Turing



The first program for playing chess

Programming a Computer for Playing Chess - Claude Shannon

[\[link to chapter\]](#)

Prediction: Chess will be solved in 4 years

Reality: 40 more years





Popular culture began hyping AI

Isaac Asimov's laws of robotics:

First Law: A robot may **not injure** a human being or, through inaction, allow a human being to come to harm.

Second Law: A robot **must obey** the orders given it by human beings except where such orders would conflict with the First Law.

Third Law: A robot **must protect** its own existence as long as such protection does not conflict with the First or Second Law.





The term Artificial Intelligence

Coined by John
McCarthy in 1956

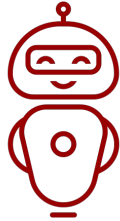
Everyone agreed
that AI





The first AI winter, HCI summer arrives

Two fields divided: AI and HCI



Artificial Intelligence

Goal: to create an artificial rival to human intelligence

Artifact: models of human intelligence

Long time horizon



Human-Computer Interaction

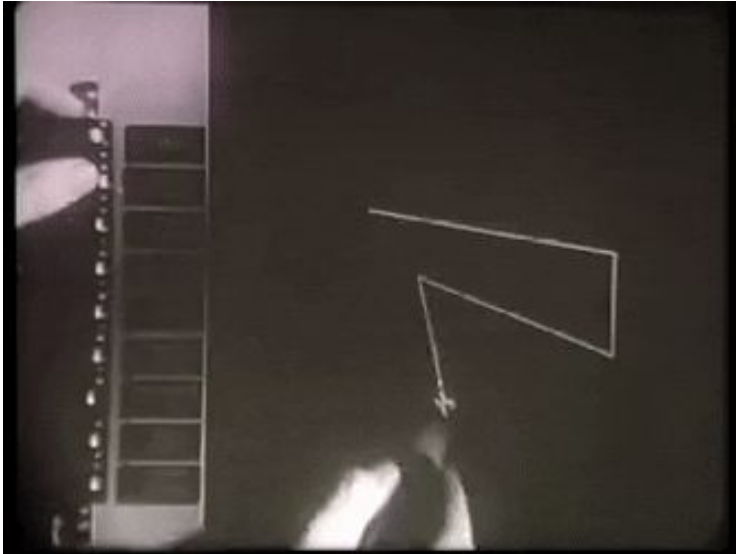
Goal: To improve applications as they approach widespread use

Artifact: designs for mass market products

Short time horizon



HCI innovations in the 60s



Ivan Sutherland (1963) introduced the *Sketchpad*

- Ancestor of modern computer-aided design (CAD) programs
- Ancestor of object-oriented programming
- Ancestor of computer graphics

Turing award in 1988



HCI innovations in the 60s



Grace Hopper (1960) started developed Fortran.

Programming languages should look like the language we speak.



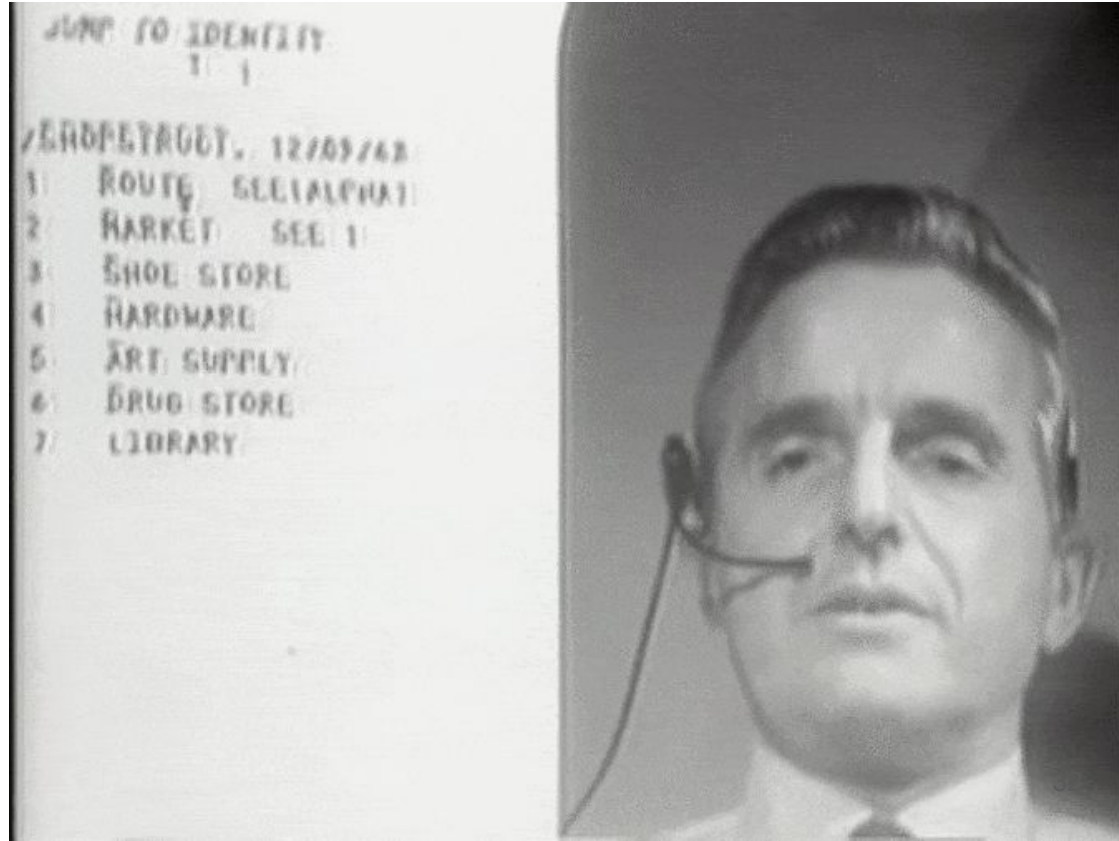
Douglas Engelbart - “father of HCI”

Demo included:

- the computer mouse,
- hypertext,
- windowing,
- videoconferencing,
- real time text editing

Not just a tool but a whole system around the tool

[Augmenting Human Intellect: A Conceptual Framework](#)





The new AI awakening in the mid 60s

HCI dwindles

Man-Computer Symbiosis

J.C.R. Licklider

- Believed that machines would surpass humans. The first AGI claim.
- Machine evolution is a basic stage of development in human evolution
- 15 year time frame





Overconfidence bloomed in AI

“Machines will be capable, within **twenty years**, of doing any work that a man can do.”

- Herb Simon (1960)



Overconfidence bloomed in AI

“In from **three to eight years** we will have a machine with the general intelligence of an average human being. I mean a machine that will be able to read Shakespeare, grease a car, play office politics, tell a joke, have a fight.”

- Marvin Minsky (MIT 1970)

MIT itself received nearly \$20 million annually.



The inception of speech and language understanding

Researchers believed that a machine that can not understand speech or language would be **unreliable** and **dangerous**



Andrew Ng ✓
@AndrewYNg

ChatGPT is sometimes amazing, and sometimes hilariously wrong. Here, it confidently explains to me why an abacus is faster than a GPU. 😊



Why is CPU computation faster than GPU computation for deep learning?



CPU computation is generally faster than GPU computation for deep learning because CPUs are designed to handle a wide range of tasks and can switch between different tasks quickly and efficiently. This allows them to handle complex operations, such as data preprocessing and model training, more efficiently than GPUs, which are specifically designed for parallel processing of large amounts of data. Additionally, CPUs typically have larger cache memory and more advanced memory management techniques, which can improve the speed and efficiency of deep learning algorithms.



Why is DNA computing faster than CPU computation for deep learning?



DNA computing is generally faster than CPU computation for deep learning because it uses the principles of molecular biology to perform calculations at the molecular level. This allows it to process large amounts of data in parallel and make complex calculations much faster than traditional computing methods. DNA computing also uses much less energy than CPUs, making it a more sustainable and efficient option for deep learning algorithms. Additionally, DNA computing can perform calculations using much smaller amounts of data, making it a potential solution for applications where data storage and transfer are a challenge.



Why is abacus computing faster than DNA computing for deep learning?



Contextualizing the overconfidence

The looming presence of thermonuclear war

AI was seen as the answer

- Just as it proved useful during the second world war.

TUESDAY, OCTOBER 23, 1962 ... 3d.
© 1962, by the Daily Sketch

BLOCKADE!



Ultimatum to Khrushchev 'Move those missiles'

Pictured in Washington—Kennedy as he said last night: We will not prematurely or unnecessarily risk the costs of world-wide nuclear war in which even the fruits of victory would be ashes in our mouth—but neither will we shrink from that risk at any time it must be faced.

PRESIDENT KENNEDY last night announced a full-scale blockade of Cuba to stop a build-up of Russian missiles there. He said that Cuba had been turned into an offensive base able to rocket destruction into the heart of America. Early this morning a defence spokesman said the United States is prepared to sink Soviet ships if necessary to prevent offensive weapons from reaching Cuba.

The President's ultimatum to Cuba and Russia was given in a radio and TV broadcast to the American people. He said: Within the past week, unmistakable evidence has established the fact that a series of offensive missile sites is now in preparation on the imprisoned island of Cuba.

Two types

The characteristics of these new missile sites include two distinct types of installations. Several of them include medium range ballistic missile capable of carrying a nuclear warhead more than 1,000 statute miles. Each of these missiles is capable of striking Washington, D.C., the Panama Canal, Cape Canaveral, Mexico City or any other city in the South-Eastern part of the United States, in Central America, or in the Caribbean area. Additional sites not yet completed appear to be designed for intermediate

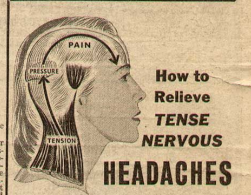


range ballistic missiles—capable of travelling more than twice as far—and thus capable of striking most of the major cities in the Western Hemisphere—as far North as Hudson Bay, Canada, and as far south as Lima, Peru. In addition, jet bombers, both piloted and privately owned, that the arm build-up in Cuba would retain its original defensive character, and that the Soviet Union had no need or desire to station strategic missiles on the island of Cuba into an important strategic base—by the presence of these large

offensive weapons of sudden mass destruction—constitutes an explicit threat to the peace and security of all the Americas. This action also contradicts the repeated assurances of Soviet spokesmen both public and private that the arm build-up in Cuba would retain its original defensive character, and that the Soviet Union had no need or desire to station strategic missiles on the island of Cuba into an important strategic base—by the presence of these large

offensive weapons of sudden mass destruction—constitutes an explicit threat to the peace and security of all the Americas. This action also contradicts the repeated assurances of Soviet spokesmen both public and private that the arm build-up in Cuba would retain its original defensive character, and that the Soviet Union had no need or desire to station strategic missiles on the island of Cuba into an important strategic base—by the presence of these large

offensive weapons of sudden mass destruction—constitutes an explicit threat to the peace and security of all the Americas. This action also contradicts the repeated assurances of Soviet spokesmen both public and private that the arm build-up in Cuba would retain its original defensive character, and that the Soviet Union had no need or desire to station strategic missiles on the island of Cuba into an important strategic base—by the presence of these large



How to Relieve TENSE NERVOUS HEADACHES

ANADIN Relieves Pain fast * Releases Pressure * Eases Tension

HOW THESE HEADACHES START Many headaches start somewhere you'd never suspect—in the muscles of the back of the neck and scalp. When you get "nervy" these muscles tighten up, causing pressure on nerves and tiny blood vessels and resulting in pain. In turn the pain builds up more tightness, more pressure and makes your headache worse.

HOW ANADIN RELIEVES NERVOUS HEADACHES The special Anadin formula gets straight to the root of the trouble by soothing nerves and relaxing the muscles as well as relieving pain. Anadin helps you to throw off depression, soothes you calm, cheerfully relaxed. Always insist on 'Anadin' tablets.

Nothing acts faster than ANADIN

© John Frost Newspapers



AI moved further away from HCI

“As suggested by the term ‘artificial intelligence, ... [The goal] was to **get away from studying human behavior** and consider the computer as a tool for solving certain classes of problems. Thus AI was created as a branch of computer science and **not as a branch of psychology.**”

- John McCarthy [1988](#)



Human-Computer Interaction

Originated in psychology

AI moved further away from HCI



Artificial Intelligence

Intelligence is the application of well defined rules to a small set of symbolic entities.



Human-Computer Interaction

Only through user testing can design principles emerge.

AI took on an engineering focus

The [Shaky robotics](#) system

- Easily broken
- Rarely did what it was supposed to





The second AI winter, HCI blossoms



The lighthill report and the usability studies

[[the whole lighthill report](#)]

- Identified multiple usability issues

User studies

- Terry Winograd, Don Norman
- Found that users found it hard to use or interact with AI

Thompson, H.; Winograd, T.; Bobrow, D.; Norman, D.; Kay, M.; and Kaplan, R. 1977. GUS, A Frame-Driven Dialog System. Artificial Intelligence

Norman, D. A. 1978. Analysis and Design of Intelligent Systems. Human and Artificial Intelligence



HCI blooms

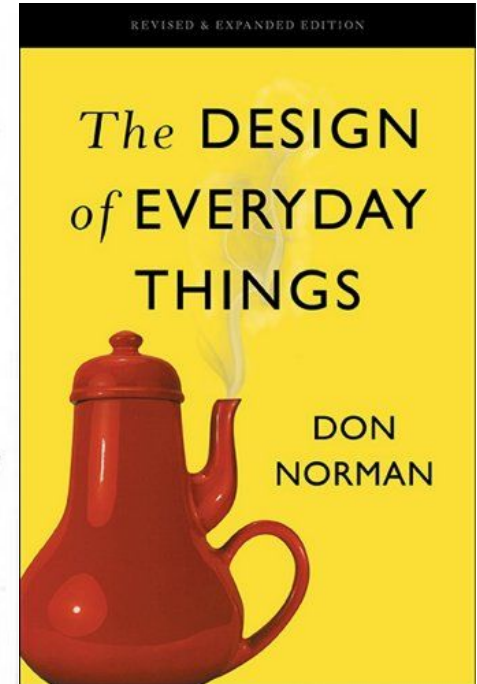
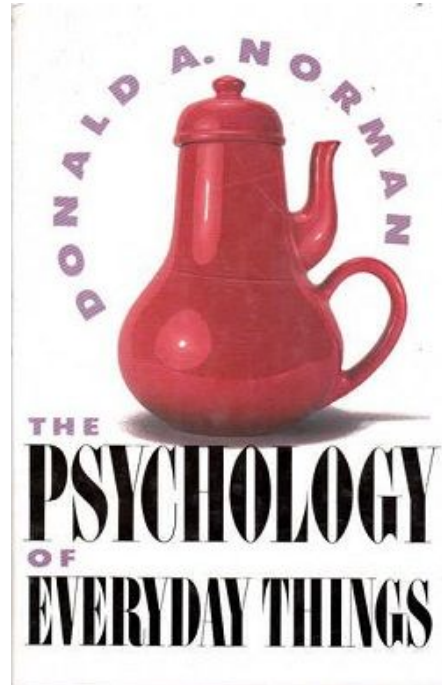




User-centered design gains attention

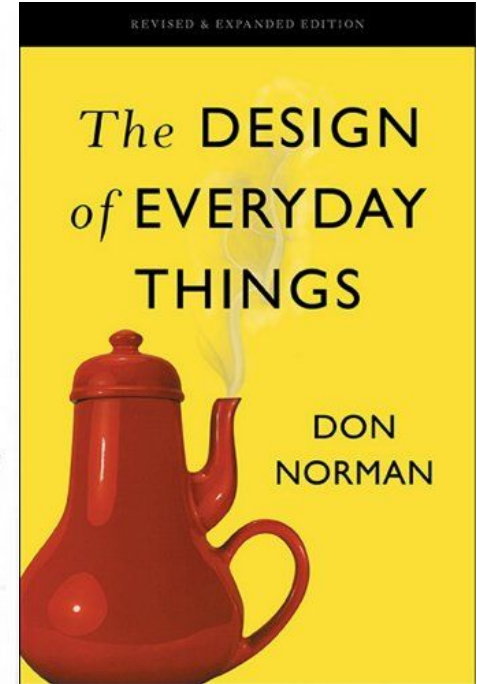
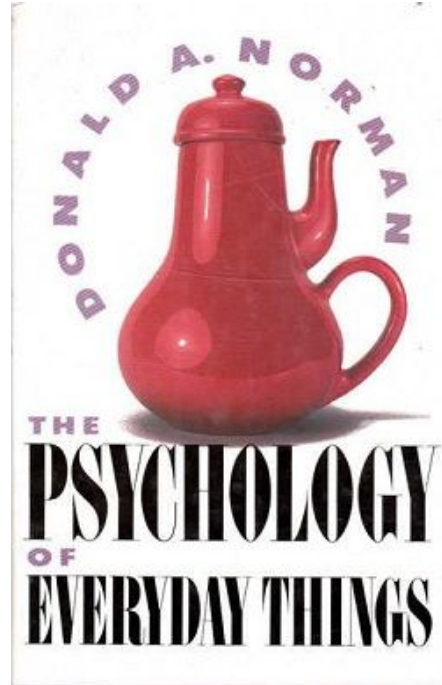
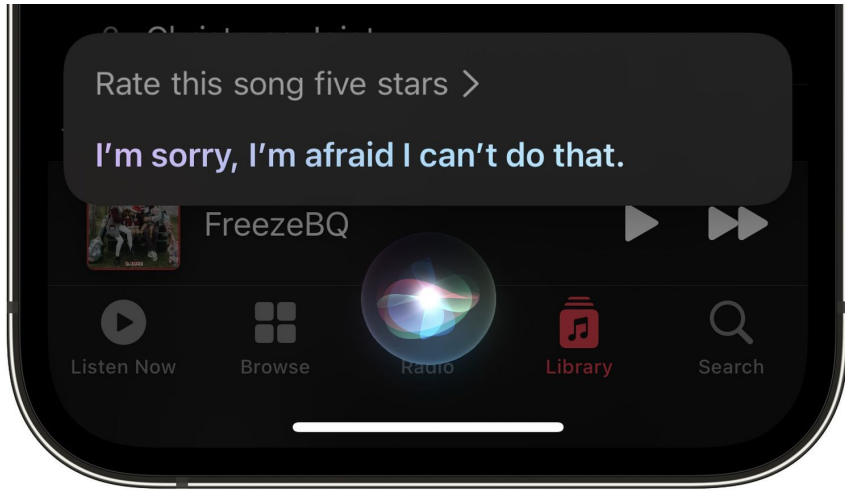
Don Norman

- Coins user-centered design
- Criticizes unix terminal as unfriendly
- Writes his now canonical book





Reframes interaction failures as system failures not user fault





A commercial GUI succeeds in the commercial market

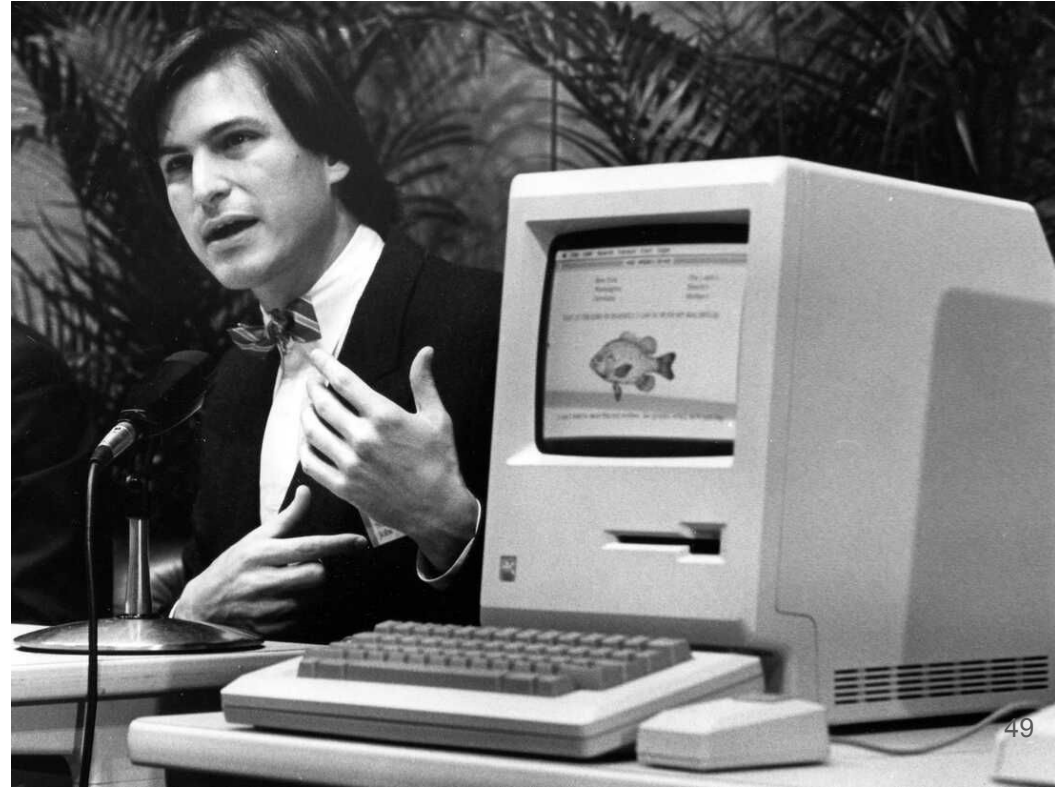
In 1985!!

Does anyone know what it is?



A commercial GUI succeeds in the commercial market

The Macintosh

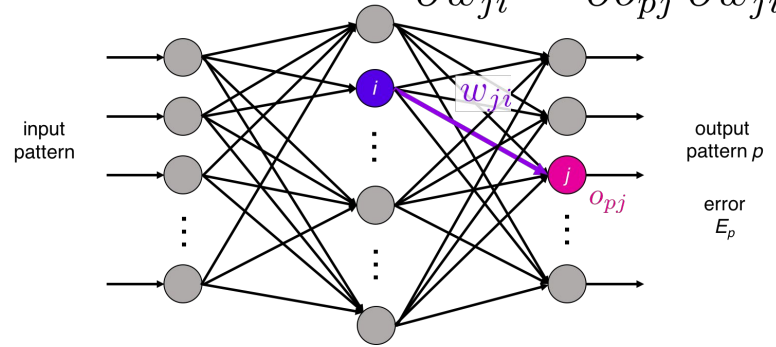




The bodacious 80s of AIHCI continues with less

Reasons of the new AI boom

$$\frac{\partial E_p}{\partial w_{ji}} = \frac{\partial E_p}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial w_{ji}}$$



'Fifth Generation' Became Japan's Lost Generation

By ANDREW POLLACK

Special to The New York Times

TOKYO, June 4 — A bold 10-year effort by Japan to seize the lead in computer technology is fizzling to a close, having failed to meet many of its ambitious goals or to produce technology that Japan's computer industry wanted.

After spending more than \$400 million on its widely heralded Fifth Generation computer project, the Japanese Government said this week that it was willing to give away the software developed by the project to anyone who wanted it, even foreigners.

Machines That Would Think

That attitude is a sharp contrast to the project's inception, when it spread fear in the United States that the Japanese were going to leapfrog the American computer industry. In response, a group of American companies formed the Microelectronics and Computer Technology Corporation, a consortium in Austin, Tex., to cooperate on research. And the Defense Department, in part to meet the Japanese challenge, began a huge long-term program to develop intelligent systems, including tanks that could navigate on their own.

Now, with a debate in the United States about whether the Government should help American companies compete, the Fifth Generation venture is a reminder that even Japan's highly regarded Ministry of International Trade and Industry can make mistakes in predicting which technologies will be important in the future.

The problem for Japan is that the computer shifted so rapidly that the technological path the Fifth Generation took — which seemed a wise choice in 1982 — turned out to be at odds with the computer industry's



"Ten years ago we faced criticism of being too reckless. Now we see criticism from inside and outside the country."

direction by 1992.

In a sense, Japan's ability to stay the course in pursuit of a long-term payoff — usually considered one of the country's strongest assets — turned into a liability. A similar challenge for Japan may now be arising

in high-definition

HDTV system development for coming to market just as some engineers believe that a major shift to digital television technology will make the Japanese analog approach

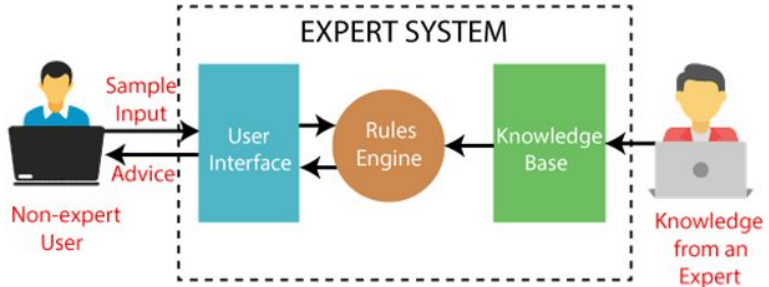
After spending more than \$400 million on its widely heralded Fifth Generation computer project, the Japanese Government said this week that it was willing to give away the software developed by the project to anyone who wanted it, even foreigners.

Continued on Page D3





Expert systems became the focus during this time



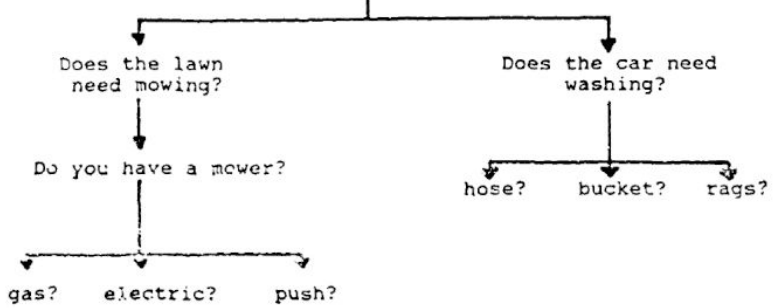
Led by Edward Feigenbaum

Originally called heuristic programming project.

BACKWARD CHAINING

GOAL: Make \$20.00

RULE: If the lawn is shaggy and the car is dirty and you mow the lawn and wash the car, then Dad will give you \$20.00



*** The inference engine will test each rule or ask the user for additional information.



Expert systems had some fundamental challenges:

Knowledge acquisition: Obtaining the time of domain experts is expensive and hard.

Scaling: more rules make the system more accurate but slower.

The result: **Brittle systems** that fail unexpectedly

- Impossible to account for all possible inputs
- By 1992, experts systems were dead



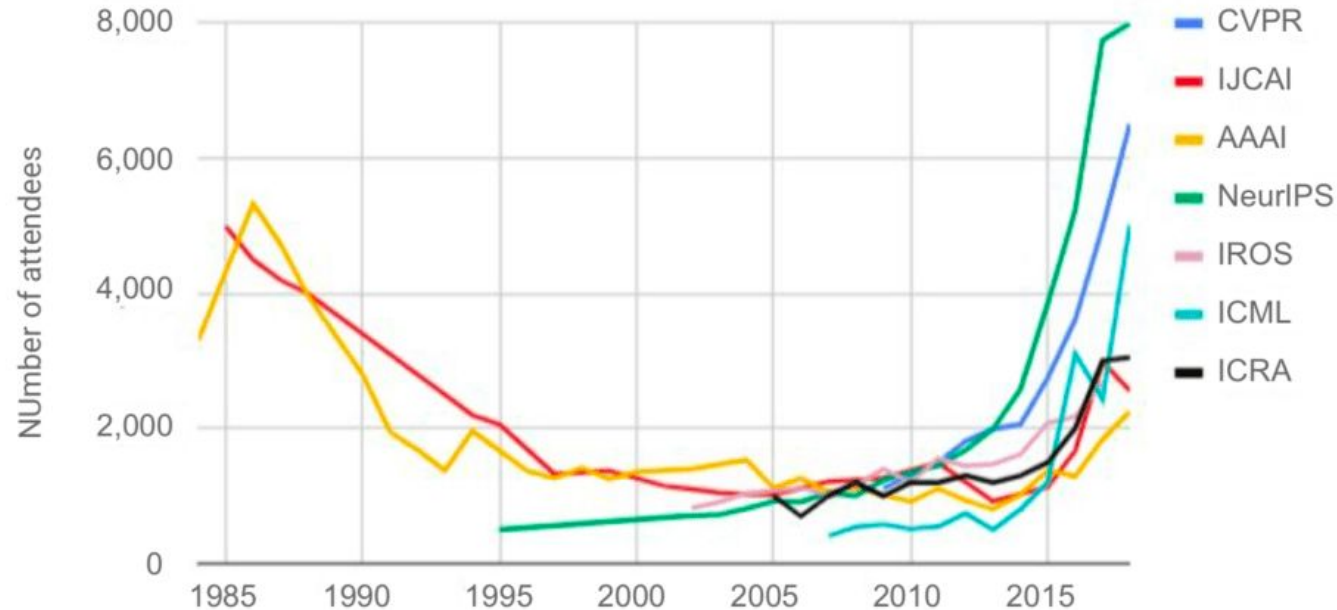
The third AI winter, HCI contributions ++



AAAI conference attendance between 85 and 95

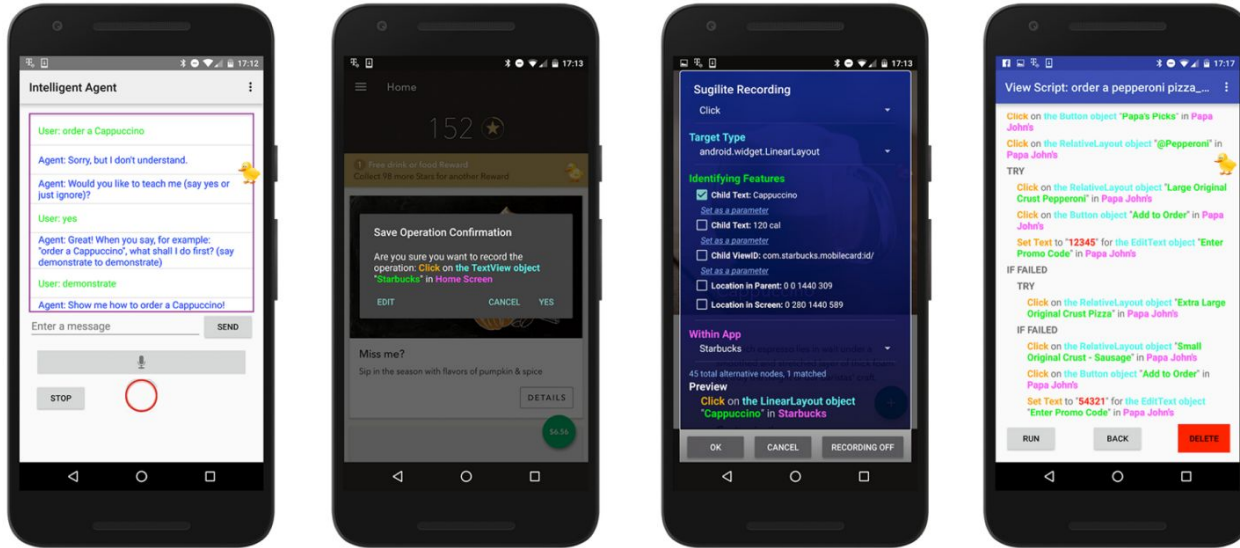
Attendance at large conferences (1984–2018)

Source: Conference provided data





Programming-by-example - Henry Lieberman



[[source](#)]

Collaborative filtering - Pattie Maes

Social networks

Interface agents

So so much more....



PROFILE

Firefly's Pattie Maes is still an agent of change

BY MARK WILLIAMS

THE PROFESSOR heading the software agents group at MIT's Media Laboratory is unfailingly courteous. No small feat, since—besides the demands of the holidays and of directing a startup slated to make its public debut in February—Pattie Maes is recovering from the flu. On the day of this December interview, she pauses in her answers for brief coughing spells: she currently has strep throat. In addition to everything else, she is a mother, and part of motherhood is sharing your two young children's infections.

LITTLE HELPERS

Best known as the brainy, photogenic cofounder of Firefly Networks, which grew out of the Media Lab, Ms. Maes moved from her native Belgium to Cambridge, Massachusetts, in 1989. Her MIT career began at the Artificial Intelligence Laboratory under Rodney Brooks (readers who've seen Errol Morris's documentary *Fast, Cheap & Out of Control* will recall Mr. Brooks' insectlike robots). Two years later, she switched her focus to the Media Lab and founded the

machine learning would go," she says. Software agents are autonomous, educable programs usually designed to migrate within computer networks. They can assist with tasks as elementary as email filtering or as elaborate as air-traffic control. "The goal is helping people with tasks and making their cognitive load lighter," Ms. Maes explains. "All these little programs

keep track of things so you yourself don't have to go out and search so much."

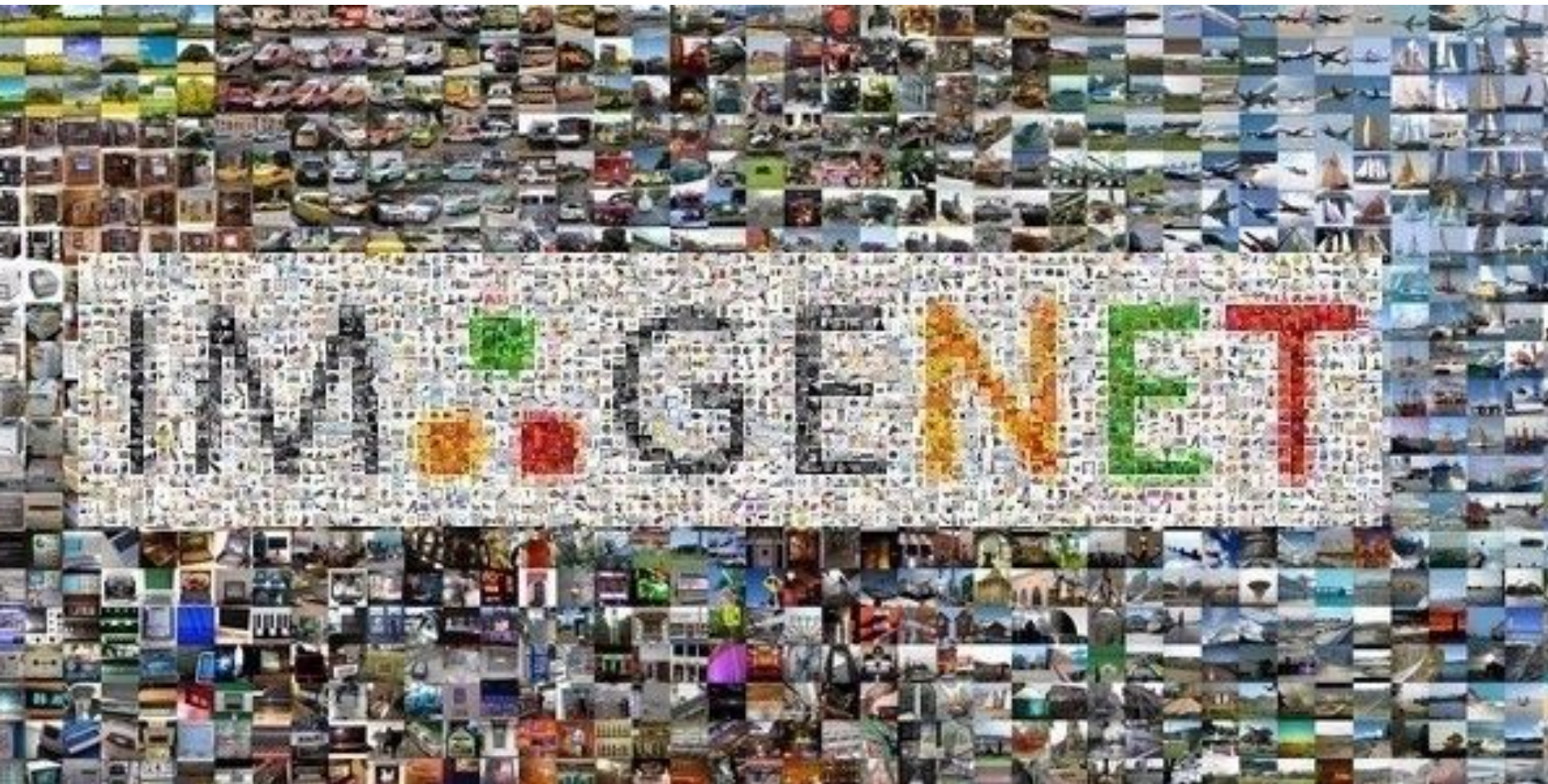
This technology could be seen as playing into our cultural tendency to render complex ideas through simplifying lenses, like graphical user interfaces or mass-media entertainment. The risk, critics contend, is that agents could escalate this dumbing-down to a point at which our attitude toward our technological infrastructure resembles a primitive, magical worldview—a point at which we no longer care to know how things work, we just assume things will.

"We're very aware of that danger," says Ms. Maes. "In my talks I often quote Marshall McLuhan, that 'every automation is amputation.' Every time you automate something for people it's like amputating—they're no longer as good at whatever's been automated or augmented." She points to the pocket calculator's effect on math abilities as an example. Her work since the mid-'90s has thus deemphasized smart agents, which might, by performing more and more tasks for people, shrink their skills. Instead, she's concentrated on technologies





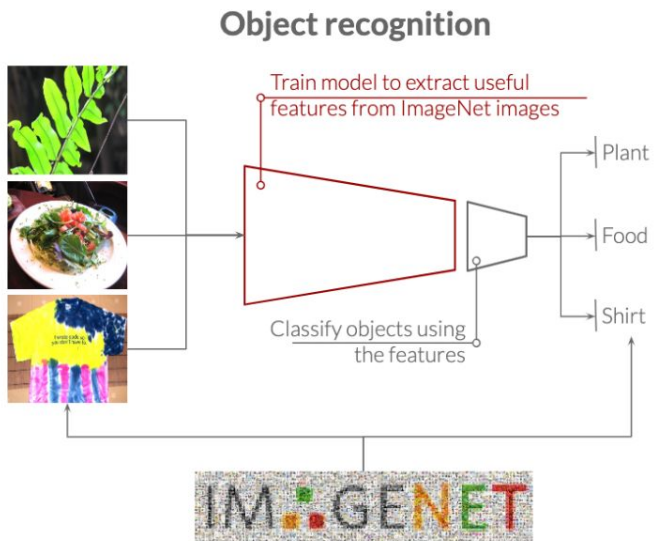
The data revolution of 2012 for AI



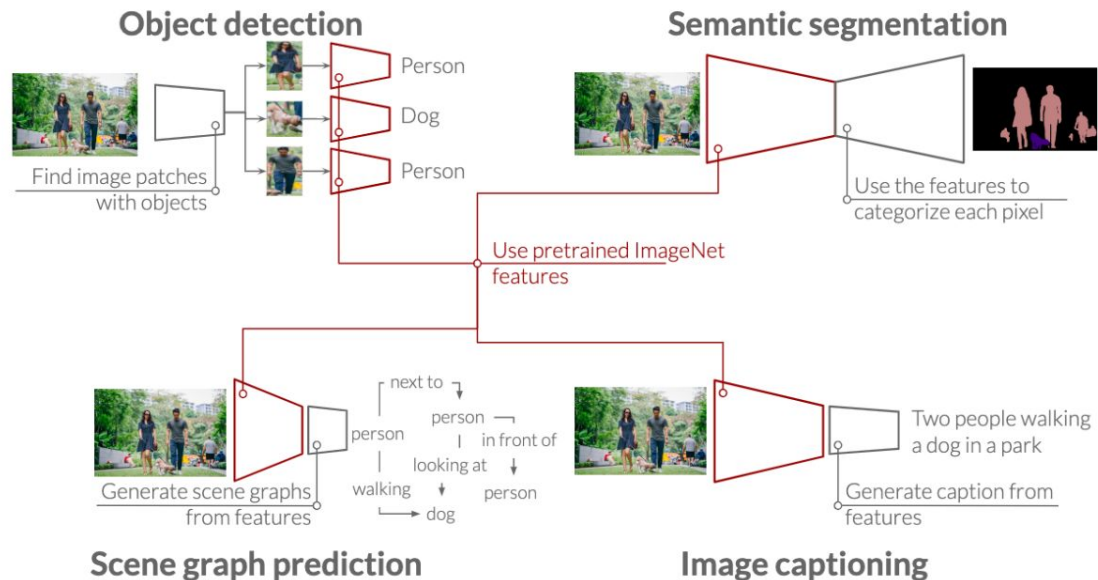


The rise of transfer learning and later the rise of pretraining

Pretraining on ImageNet for object classification



Transfer ImageNet features for many other tasks:





Pretraining LLMs

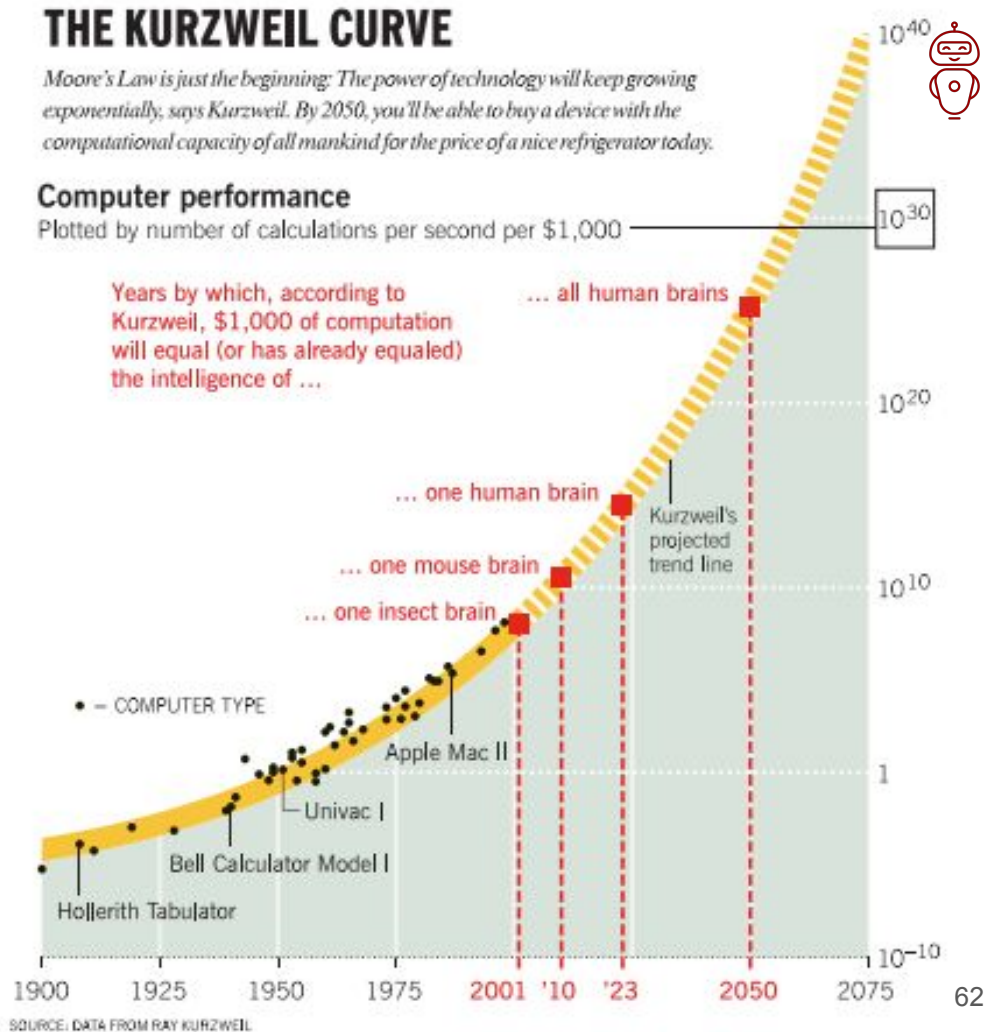
Shifting the paradigm from collecting annotated data to prompt design and engineering



The overconfidence is back

Ray Kurzweil's "The Age of Spiritual Machines" predicts human level intelligence in 30 years.

Are you ready? 2023 will be the year!





The promise of self-driving cars **by 2018**

Waymo, Alphabet's self-driving car company, was testing fully driverless cars and preparing for a commercial launch in the Phoenix area. [[source](#)]

Uber was racking up millions of test miles for its own cars. [[source](#)]

Tesla was selling a "full self-driving" upgrade for its vehicles—though the feature wasn't actually shipping yet. [[source](#)]

Several times in 2016, Elon Musk predicted that the technology would be ready in about two years. [[source](#)]





ChatGPT passes Turing test but is confidently wrong

NG

Why is abacus computing faster than DNA computing for deep learning?



Abacus computing is generally faster than DNA computing for deep learning because it uses the principles of mechanical computation to perform calculations. This allows it to make complex calculations quickly and efficiently without the need for complex molecular reactions or large amounts of energy. Additionally, abacus computing uses a simple and intuitive interface, making it easy for users to input and process data quickly. Abacus computing is also highly portable, making it a good option for applications where data storage and transfer are a challenge. Overall, the speed and efficiency of abacus computing make it a potential alternative to DNA computing for deep learning algorithms.



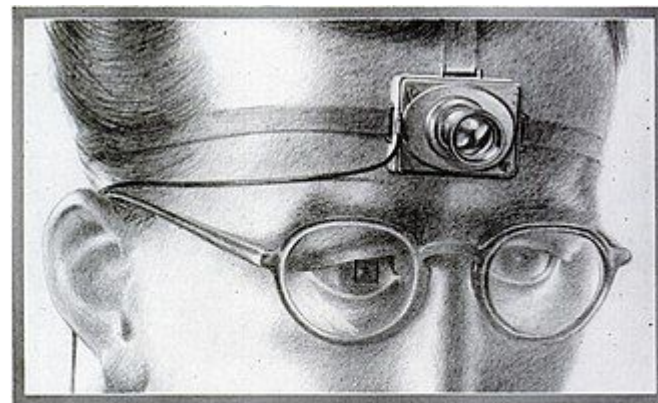
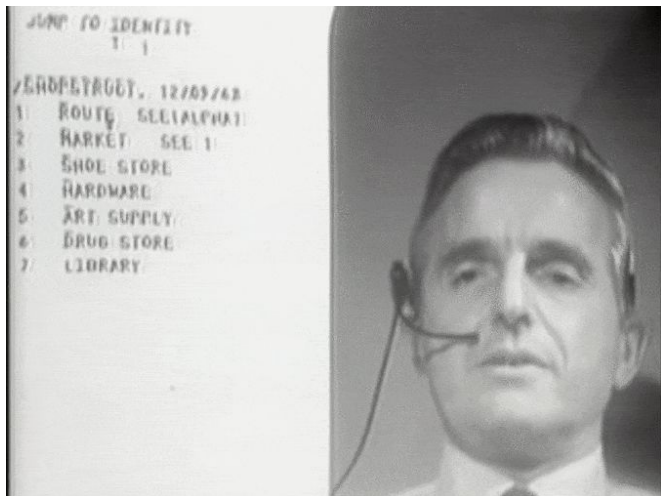


So why is this historical perspective important?



The visions for the future have already been laid out

Douglas Engelbart (“Mother of all demos”) inspired by an article from the 1945

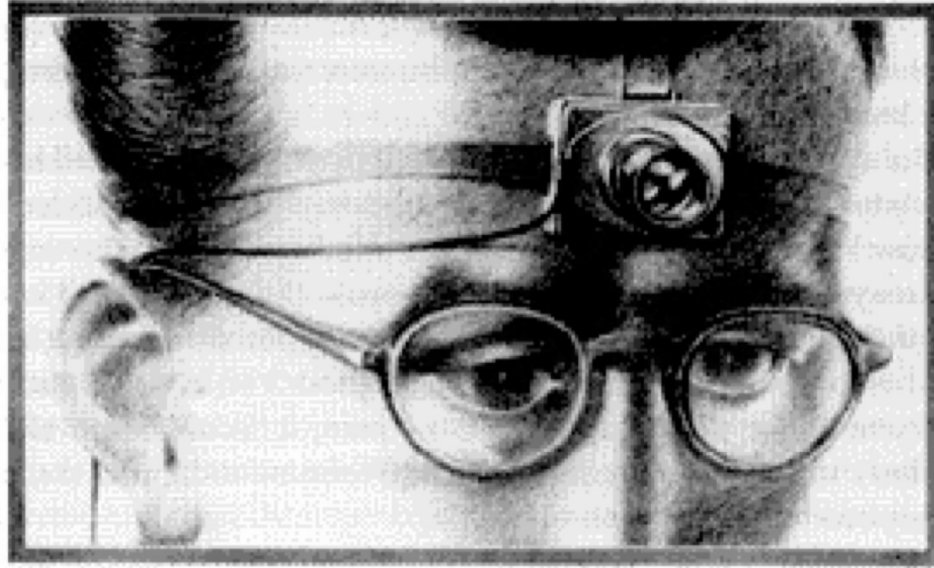


A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE VIEWFINDER AT THE LEFT SHOWS THE OBJECT

AS WE MAY THINK

A TOP U. S. SCIENTIST FORESEES A POSSIBLE FUTURE WORLD
IN WHICH MAN-MADE MACHINES WILL START TO THINK

Vannevar Bush (1945)



A scientist of the future records experiments with a tiny camera fitted with universal-focus lens. The small square in the eyeglass at the left sights the object (*LIFE* 19(11), p. 112).

Visions are alive today - not just scientists wearing them



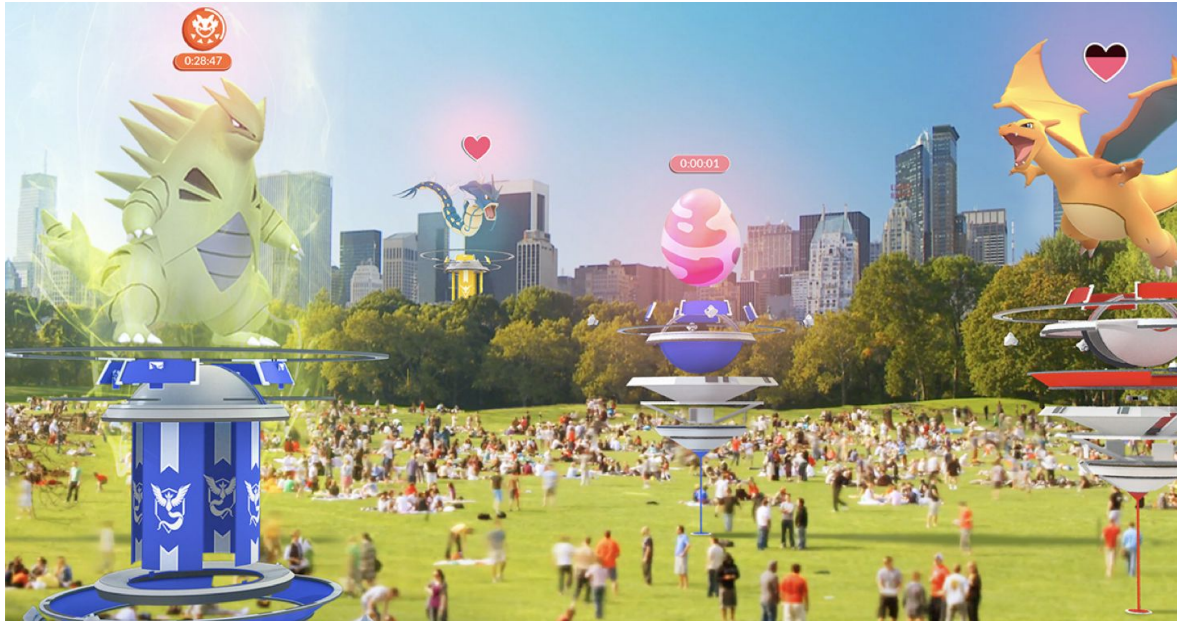
A scientist of the future records experiments with a tiny camera fitted with universal-focus lens. The small square in the eyeglass at the left sights the object (*LIFE* 19(11), p. 112).

Visions are alive today - taking us into the virtual



A scientist of the future records experiments with a tiny camera fitted with universal-focus lens. The small square in the eyeglass at the left sights the object (*LIFE* 19(11), p. 112).

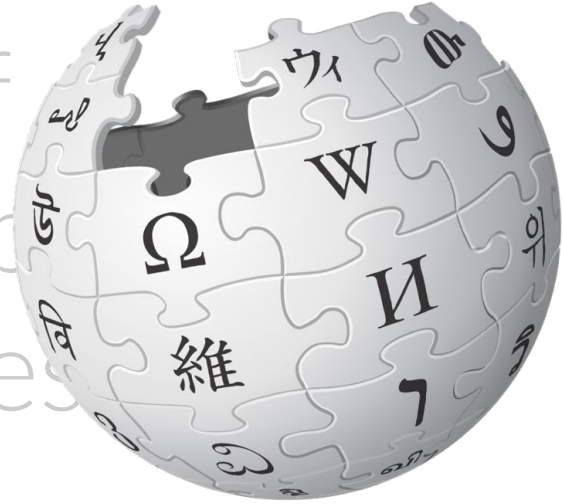
Visions are alive today - augmenting our reality



A scientist of the future records experiments with a tiny camera fitted with universal-focus lens. The small square in the eyeglass at the left sights the object (*LIFE* 19(11), p. 112).

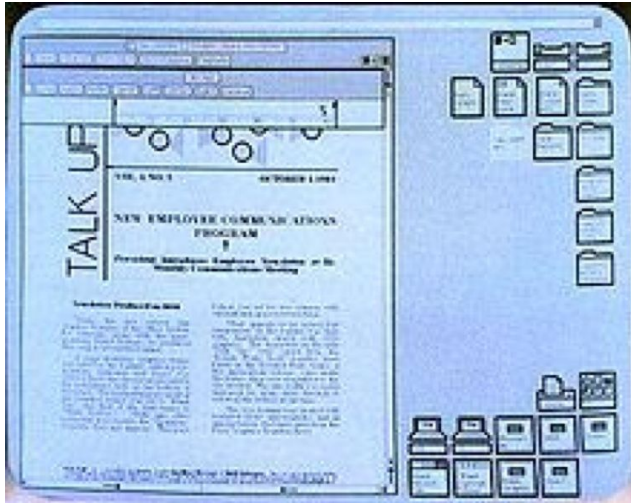
“Wholly new forms of encyclopedias will appear, ready-made with a mesh of associative trails running through them.”

“Wholly new forms of encyclopedias will appear ready-made with a mesh of associative trails running through them.”



WIKIPEDIA
The Free Encyclopedia

Continue to look back to design forward

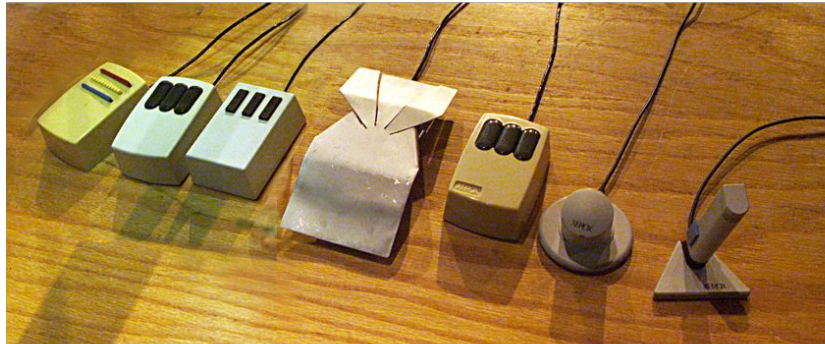


Xerox PARC. 1973.
The Xerox Alto.



Modern MacOS desktop

Continue to look back to design forward



Card, English and Burr. 1978. Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT.



Modern mouse

Continue to look back to design forward



Zhai and Kristensson. 2003. Shorthand writing on a stylus keyboard.

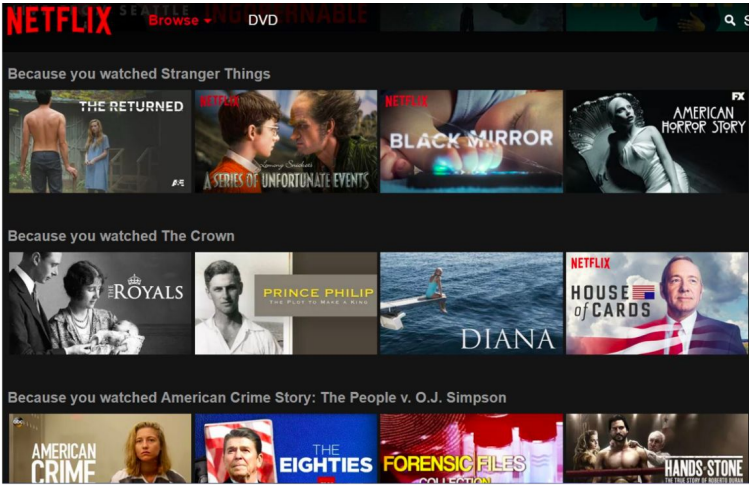


Swipe keyboards
(iOS, Android)

Continue to look back to design forward



Resnick et al. 1994. GroupLens: an open architecture for collaborative filtering of netnews.



Brought to you by your Netflix binges

Continue to look back to design forward



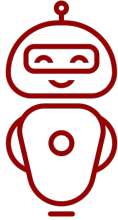
Consolvo et al. 2008. Activity sensing in the wild: a field trial of UbiFit Garden.



Modern fitness trackers

The past contextualizes
and hints at which
research directions to
pursue.

Remember the differences?



Artificial Intelligence

Goal: to create an artificial rival to human intelligence

Artifact: models of human intelligence

Long time horizon



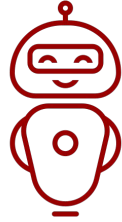
Human-Computer Interaction

Goal: To improve applications as they approach widespread use

Artifact: designs for mass market products

Short time horizon

AI is finally ready for mass market



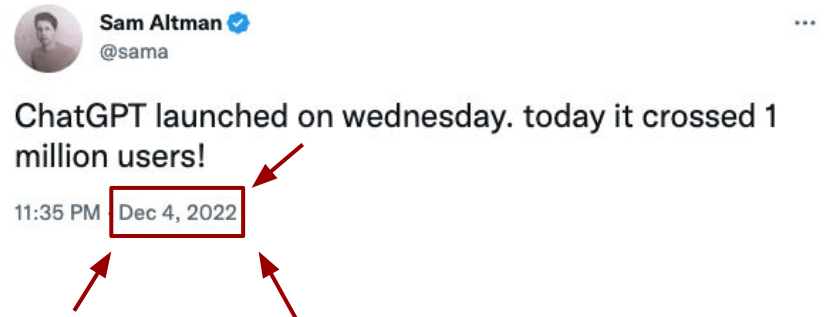
Artificial Intelligence



Goal: to align AI with human intelligence

Artifact: models for mass market use

No longer for long time horizon



HOT TAKE

The past hints at which
research directions
might disappear.

Language is a terrible modality for interacting with machines

Why?? Wait for the next lecture!!!

Lecture 2: The humans strike back,
The gaps that remain between AI and IA

Course logistics

Grading

15%: In-class active participation in discussion, feedback, and presentations throughout the quarter.

25%: Reading commentaries.

15%: Assignments (7.5 points for each assignment).

50%: Research project:

- 5%: project proposal.
- 10%: mid-point milestone report.
- 10%: final presentation.
- 20%: final report.

Participation expectations

- You are expected to attend and actively participate in every class session.
- Come ask questions, make guesses, be wrong.
- I want to see you thinking and proposing ideas. I don't care if those ideas are wrong.
- Be respectful of others' comments and let one another speak

Disability, Religious, and Family Accommodations: If you have any questions about disability or religious accommodations, please refer to university policies regarding [disability accommodations](#) or [religious accommodations](#). Feel free to also contact me for any reason.

Reading commentaries

- Most wednesdays have a discussion session.
- There are usually 2 (sometimes a 3rd short paper) that you have to read
 - Papers are available in the [schedule section](#) of the course website.
- You will be asked to **write a commentary for each** of the papers.
 - Worth 5% each. We will **drop the lowest grade** out of the 6 discussion sections
- Commentaries are due 48 hours before the discussion session.
 - So for a wednesday discussion at 1:30pm, you will be asked to submit your commentaries at **1:30pm on the monday** prior.
- Commentaries will be graded using the following **rubric**:
 - Check-minus: surface-level engagement, summary of the paper, critiques that engage obliquely.
 - Check: effective engagement, indicate that you reflected over the main ideas and provided nontrivial observations worth discussing.
 - Check-plus: excellent engagement (rare), unique insightful point of view worth sharing.

Example check-plus commentary

As We May Think

This paper was fascinating because it forces us to **consider technologies that nowadays we take for granted**. In some ways Bush was **overly optimistic**; for example walnut-sized wearable cameras are uncommon (even though they are possible), likely because optical and physical constraints favor handheld sizes. In other ways he **underestimated**, such as the explosion of data. For example, some modern cameras can store ten thousand photos rather than a hundred.

Underestimating the data explosion is also apparent in the disconnect between the initial problem description ("publication has been extended far beyond our present ability to make real use of the record") and the first two-thirds of the paper, which describe technologies that would (and did!) exacerbate the issue by further proliferating data. Yet, he recognizes this issue later in the paper, and then goes on to predict search engines.

It is remarkable **how many technologies are predicted** in this paper: digital photography, speech recognition, search engines, centralized record-keeping for businesses, hypertext (even Wikipedia?). At the same time, many of the predicted implementations are distorted by technologies and practices common at the time, like "dry photography" or "a roomful of girls armed with simple keyboard punches". While these presumably served to make the hypotheses more accessible to readers of the time, is it even possible to hypothesize technology without such artifacts.

Aside from predictions, this paper is important for the way Bush frames **science in the support of the human race**, by augmenting the power of the human mind. It is likely that many of the scientists (and physicists in particular) that were **his audience felt guilt and despair from the destruction wrought by advances in nuclear, and even conventional, weaponry in the war**. In that social context, seeing **science described as a powerful constructive tool** for good must have been inspiring.

Discussions

Inspired by [Alec Jacobson](#) and [Colin Raffel](#), we will be discussing papers in the [many-to-many role-playing format](#).

Roles and their descriptions can be found [here](#), though we will make the following modifications:

- Added role: **discussion leader**. This role will be able to see all the submitted commentaries submitted 48 hours before the discussion. They will be responsible for providing a short 8 minute summary of the papers, students' responses, and spur in-class discussion. They will be the first role to present.
- Modified role: The scientific peer reviewer has been split into two roles, a positive peer reviewer (**advocate**) and a negative peer reviewer (**skeptic**).

Logistics will work as follows:

- At least 1 week before each discussion session, please start a group message with your team on Slack (create a new channel if you need to).
- At least 2 hours before each discussion session (**11:30am on Wednesday**), upload your slides to the shared google slides and decide who will be presenting.
- All roles will have **4** minutes to present which will be followed by a 4 minute discussion, except the discussion leader who will have **8**.

Light assignments

Assignment 1: Reflections on personal AI use (7 min presentation due by Wed, Jan 11th)

Assignment 2: Hack something useful for someone else (7 min presentation due by Wed, Mar 1st)

Project

Project proposal (due Jan 24th at 11:59pm)

- 1 page

Project milestone (due Feb 16 at 11:59pm)

- 3-4 pages

Final project Presentations (due Mar 13 in class)

- 10 minute presentations

Final project report (due Mar 14th at 11:59pm)

- 5-6 pages
- ([overleaf template](#), [downloadable latex/word template](#))

Types of research projects (can be ongoing research)

- Design and implement a **new machine learning, computer vision, NLP, or AI model/system** from which people can derive value.
- Conduct an **empirical evaluation of a new or existing AI system** using real human subjects. Subjects can be participants you recruit around campus or hire from platforms like Amazon Mechanical Turk or Prolific.
- Perform a **qualitative or quantitative analysis** to understand people's problems or impressions of an existing AI model and produce implications for the design of new technology.
- Identify human social or cognitive capabilities that are missing in existing models. Either **evaluate how existing models** have such capabilities or propose a new training paradigm that can result in the emergence of such machine capabilities.
- Conduct a critical analysis of existing AI models or academic literature in HCI, to produce **design recommendations** for new AI models/systems.
- **New data collection or active learning methods** that study the implications of the data being collected.
- Create new or study the effects of existing **explanations or interpretable** machine learning methods in new tasks.

How you should expect this course to be...

- This is not a clear, tightly scripted Hemingway-esque story of ideas, this is more of a **survey** (that we'll try to stitch together into a whole by the end of the class)
- It's a collection of topics on how people and AI systems interact; what works and what doesn't work...
- We will read predominant machine learning methods in use today, learn about them and critique them
- Desiderata and guidelines on how to build them ... for **humans**

Learning goals

- **Contextualizing** artificial intelligence and human-computer interaction as two twin sub-fields of computer science.
- **Awareness** of the major researchers, terminology, historical debates, and long-standing problems that have shaped the two fields.
- **Knowledge** of the current state of the two fields, including key recent literature, outstanding research questions, and newer areas of interest.
- **Understanding** of different research and design methods within each, with a focus on the design of interactive systems, and which methods can practically best tackle which research questions.

Syllabus

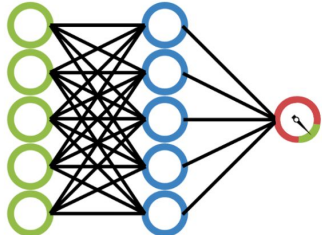
- Lecture 1, 2 and discussion 1: a historical perspective and what the challenges that exist
- Lectures 3 and discussion 2: evaluating models
- Lecture 4: Understanding models
- Discussion 3: Cognitive science and emergent machine behaviors
- Lecture 5 and discussion 6: Interacting with and teaching machines
- Lecture 6 and discussion 5: The role of data
- Lecture 7: Creativity and generative models
- Guest lectures

Questions

- Operational questions?
- Class content questions?
- General questions?

Related student group at UW

AI could change the world let's make it for the better



discuss solutions to AI
misalignment in the
AI technical
alignment program

AI alignment asks questions like:

How do we make systems like ChatGPT do what we want?

How can we understand what's going on inside a neural network?

How do we make AI that can generalize good behavior to situations where they haven't previously been given feedback, and could go catastrophically wrong?

learn more
& apply
bit.ly/uwalignment23wi

atmallen@cs.washington.edu

