

Artificial Intelligence #1

Lecture 10

Introduction on artificial intelligence

Reveal and discuss the social and ethical issues existed in artificial intelligence

Lecturer YANG Lu

Agenda

- What is Artificial Intelligence
- The history of Artificial Intelligence
- AI applications
- Existing challenges and Issues

Artificial Intelligence

Intelligence

- The ability to learn from experience
- The power of thought
- The ability to reason
- The ability to perceive relations
- The power of insight
- The ability to use tools
- Intuition
- etc

Many Definitions of AI:

- Making computers **perceive, reason and act** in ways that **resemble human beings**.
- Enables computer to **do things that human beings are better at, at least presently**.
- Enables computers to do things that **make human beings seem intelligent**.
- Etc.

Artificial Intelligence

Thinking Humanly

“The exciting new effort to make computers think . . . machines with minds, in the full and literal sense.” (Haugeland, 1985)

“[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . .” (Bellman, 1978)

Acting Humanly

“The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil, 1990)

“The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)

Thinking Rationally

“The study of mental faculties through the use of computational models.” (Charniak and McDermott, 1985)

“The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)

Acting Rationally

“Computational Intelligence is the study of the design of intelligent agents.” (Poole et al., 1998)

“AI . . . is concerned with intelligent behaviour in artifacts.” (Nilsson, 1998)

The History of Artificial Intelligence

“Can Machine Think?”

Alan Turing (1912-1954):

British mathematician who designed world's first electronic digital computer.

Asked a famous question “Can machines think?”

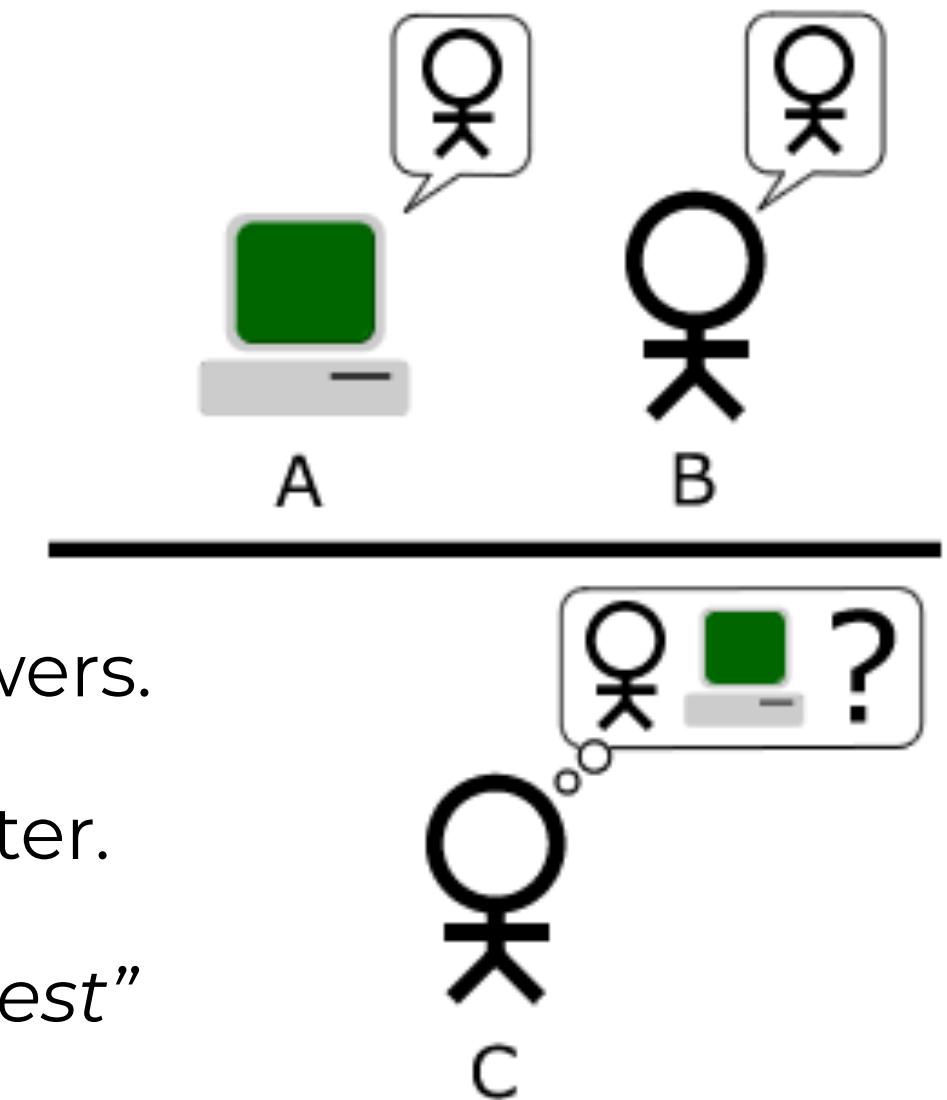
Turing Test (1950) - “Imitation Game”

Involves 2 humans and 1 computer.

Human interrogator asked questions and human responder and computer provide answers.

Human interrogator guesses whether answers provided by human responder or computer.

“A machine may be deemed intelligent when it can pass for a human being in a blind test”



Loebner Prize:

An annual competition since 1991 in AI that awards prizes to the chatterbot considered by the judge to be the most human-like.

The History of Artificial Intelligence

Alan Turing

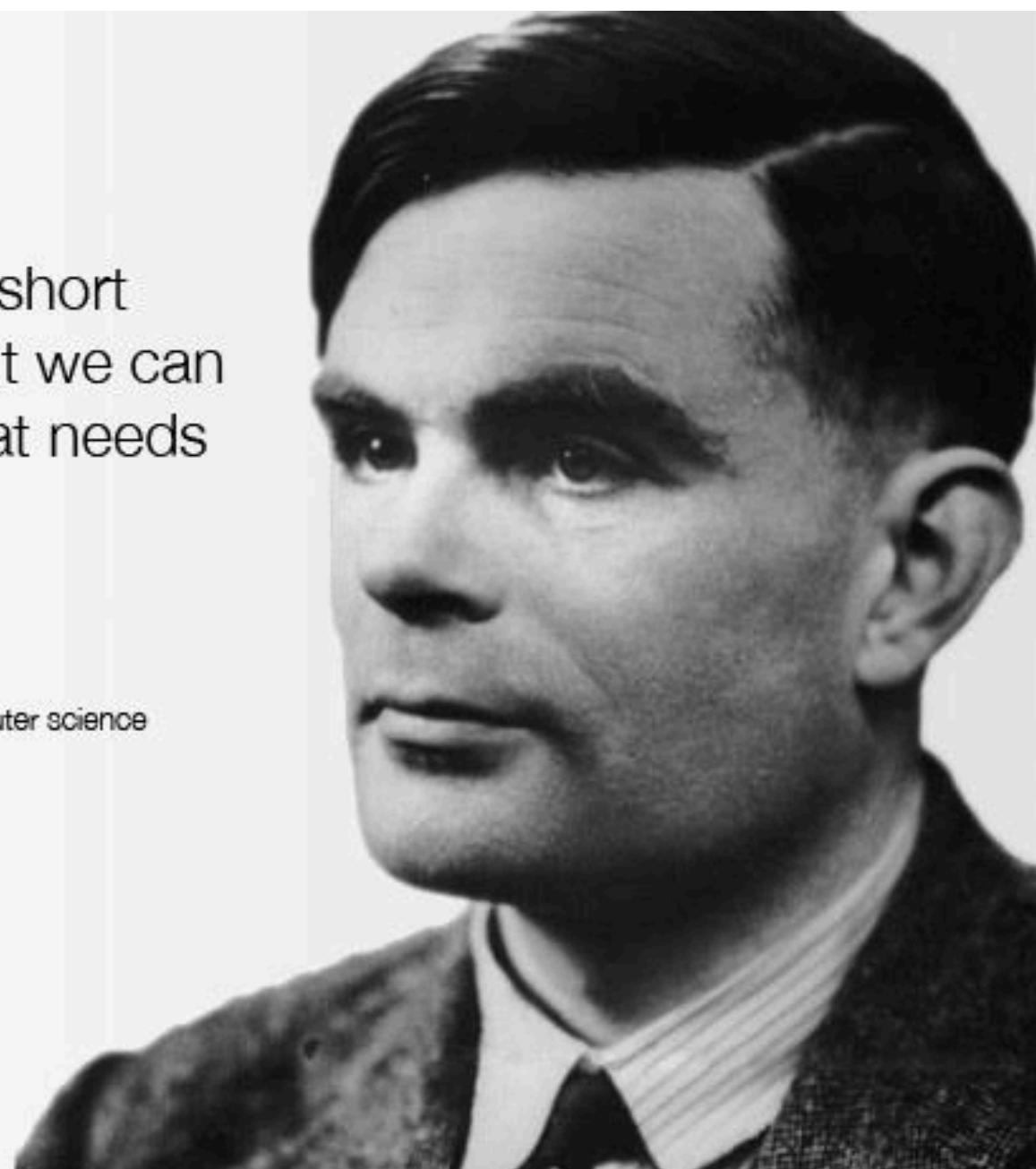
VOL. LIX. No. 236.]

[October, 1950]

"We can only see a short distance ahead, but we can see plenty there that needs to be done."

~ Alan Turing

the father of modern computer science



"I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9 , to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after 5 minutes of questioning"

-Alan Turing (1950)

'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is

The History of Artificial Intelligence

- The birth of AI (1956)

- A two month workshop at Dartmouth in the summer of 1956.
 - “Dartmouth Summer Research Project on Artificial Intelligence”
- Participants:
 - John McCarthy (who coined the term “Artificial Intelligence”)
 - Claude Shannon (father of information theory)
 - Marvin Lee Minsky (cognitive scientist, co-founder of MIT AI lab)
 - Nathaniel Rochester (AI scientist at IBM)
 - Arthur Samuel (who coined the term “machine learning”)
 - Allen Newell (Turing Award)
 - Herbert Alexander Simon (Nobel Prize in Economics, Turing Award)
 - etc

The propose states:

“We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is 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. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.”

The History of Artificial Intelligence

- Early enthusiasm, great expectations (1952–1969)
 - Newell and Simon's General Problem Solver (GPS)
 - Aim to simulate human problem-solving protocols
 - Limited puzzles used (e.g., The Towers of Hanoi)
 - Probably the first program to embody the “thinking humanly” approach
 - Herbert Gelernter (1959) constructed the Geometry Theorem Prover
 - Prove theorem
 - John McCarthy's work on AI programming language Lisp
 - Microworld problems such as blocks world
 - Home to projects like computer vision, robotics, natural language processing, learning theories

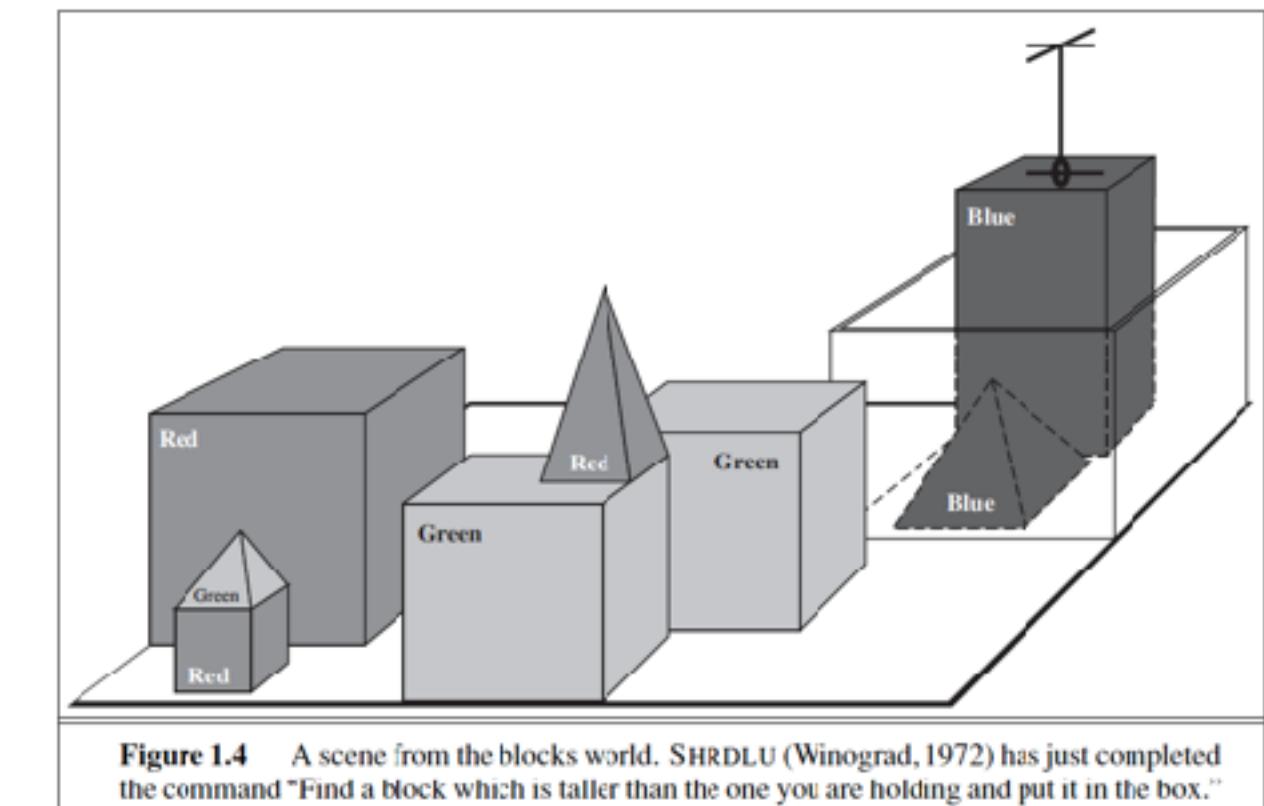


Figure 1.4 A scene from the blocks world. SHRDLU (Winograd, 1972) has just completed the command “Find a block which is taller than the one you are holding and put it in the box.”

The History of Artificial Intelligence

- Early enthusiasm, great expectations (1952-1969) - cont.
 - Arthur Samuel's work on checkers programs
 - Easy to represent in computer's digital memory
 - Clearly defined rules
 - The goals are unmistakable
 - Coined the term "machine learning"
 - "the field of study that gives computers the ability to learn without being explicitly programmed."
- Tom Michael Mitchell's definition (1998)
 - "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."
 - Example: playing checkers.
 - E = the experience of playing many games of checkers
 - T = the task of playing checkers.
 - P = the probability that the program will win the next game.



The History of Artificial Intelligence

- A does of reality (1966-1973)
 - Difficulty
 - 1) early programs know nothing about their subject matter
 - e.g., machine translation programs
 - Initial thoughts: grammar + word replacement is suffice
 - The accurate translation requires background knowledge in order to resolve ambiguity and establish the content of the sentence
 - 2) intractability of many problems
 - 3) some fundamental limitation on the basic structures being used to generate intelligent behaviour

Challenges and Difficulties

Sarcasm/Irony

Sarcasm uses words of a polarity to represent another polarity.

Example: The perfume is so amazing that I suggest you wear it with your windows shut.

The History of Artificial Intelligence

- Knowledge-based systems (1969-1979)
 - Use more powerful, domain-specific knowledge that allows larger reasoning steps and can more easily handle typically occurring cases in narrow areas of expertise.
 - Capture tacit knowledge in very specific and limited domain of human expertise
 - Capture knowledge of skilled employees as **a set of rules** in software system that can be used by others in organization
 - Typically perform **limited tasks** that may take a few minutes or hours, for example
 - e.g., medical diagnosis
 - MYCIN program to diagnose blood infection (and 450 rules)

The History of Artificial Intelligence

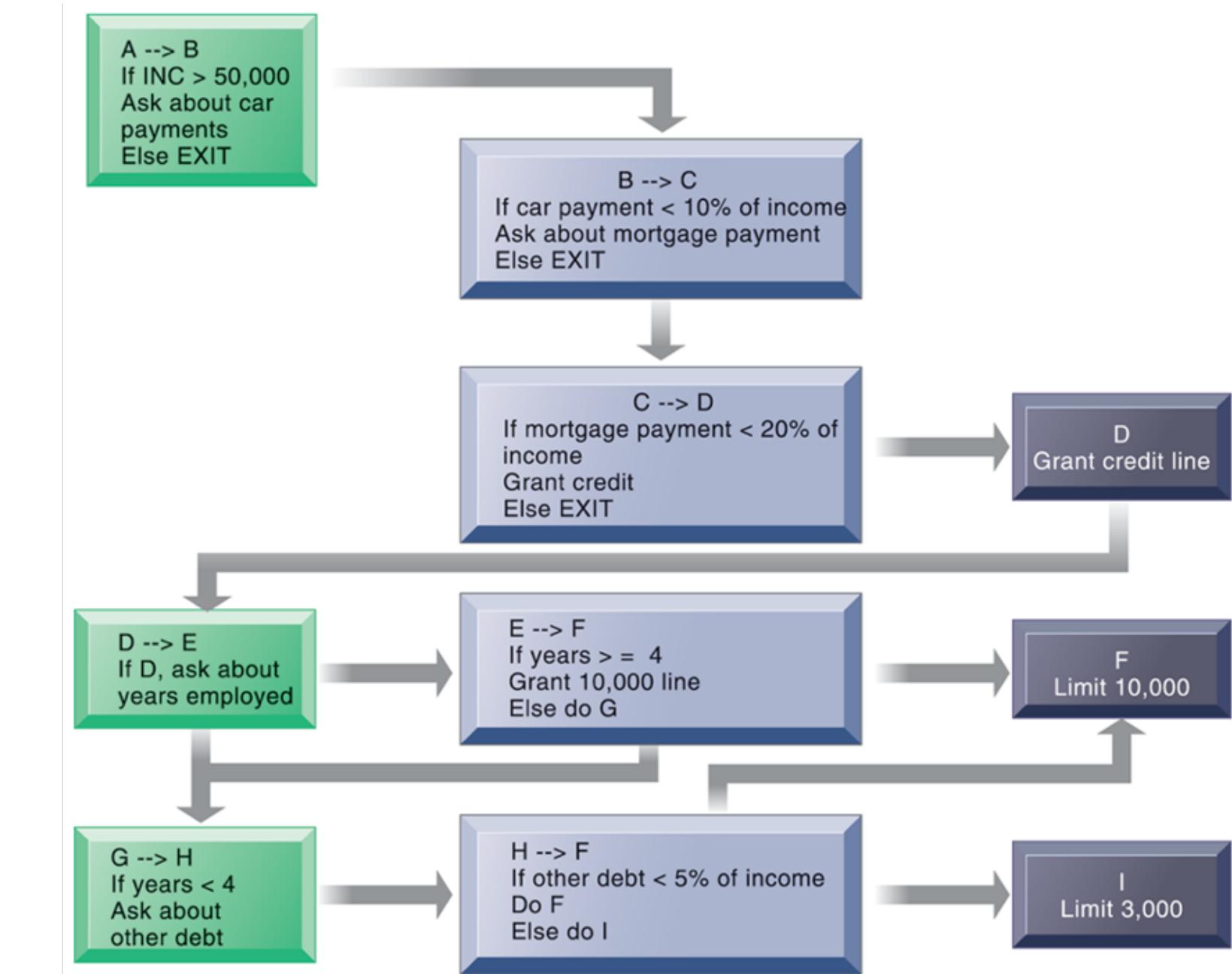
- Knowledge-based systems (1969-1979) - cont.

- How expert systems work

- Knowledge base:
 - Set of hundreds or thousands of rules
 - IF-THEN-ELSE rules
 - Inference engine: Strategy used to search knowledge base
 - *Forward chaining*: Inference engine begins with information entered by user and searches knowledge base to arrive at conclusion
 - *Backward chaining*: Begins with hypothesis and asks user questions until hypothesis is confirmed or disproved

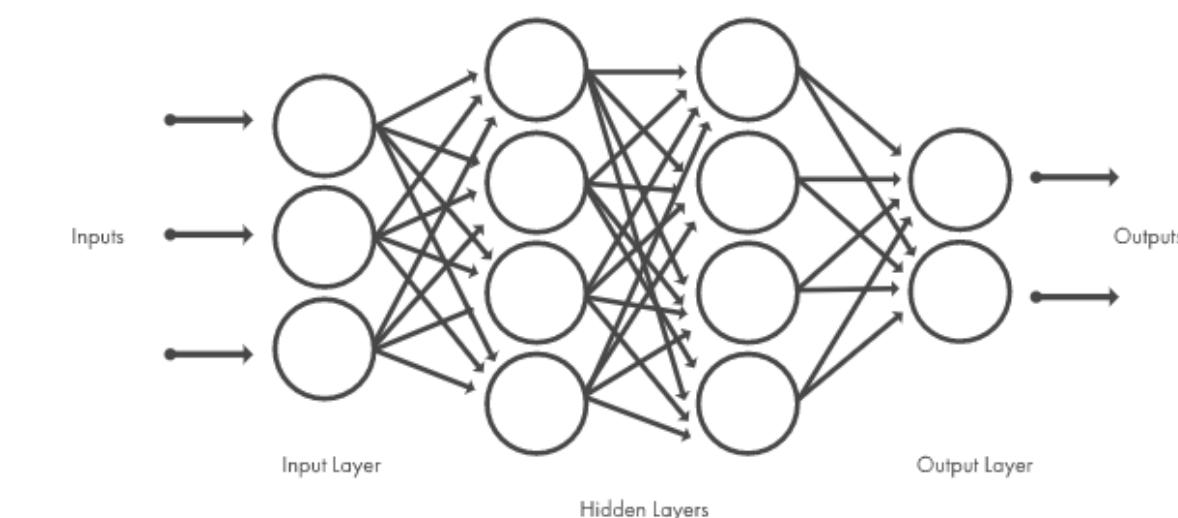
An expert system contains a number of rules to be followed. The rules are interconnected; the number of outcomes is known in advance and is limited; there are multiple paths to the same outcome; and the system can consider multiple rules at a single time. The rules illustrated are for simple credit-granting expert systems.

FIGURE 11-5



The History of Artificial Intelligence

- AI becomes an industry (1980–present)
 - the AI industry boomed from a few million dollars in 1980 to billions of dollars in 1988, including hundreds of companies building expert systems, vision systems, robots, and software and hardware specialised for these purposes.
 - But “AI winter” soon came
- The return of neural network (1986-present)
 - creating effective network architectures and algorithms and understanding their mathematical properties
 - careful modelling of the empirical properties of actual neurons and ensembles of neurons.
- AI adopts the scientific method (1987–present)
 - Rigorous mathematical theory
 - Trained on very large data sets



The History of Artificial Intelligence

- The emergence of intelligent agents (1995-present)
 - A complete intelligent agent, Human-Level AI (HLAI)
 - Machines that think, that learn, and that create
 - Requires huge knowledge bases (e.g., cognitive thinking, bionics)
 - Artificial general intelligence (AGI)
- The availability of very large data sets (2001–present)
 - Billions of images from the web
 - Trillions of words of English from the web



Artificial Intelligence Domains

Disciplines:

- **Natural language processing:** enable it to communicate successfully
- **Knowledge representation:** to store the information
- **Automated reasoning:** to use the stored the information to answer questions and to draw conclusions
- **Machine learning:** to adapt to new circumstances and to detect and extrapolate patterns
- **Computing vision:** to perceive objects
- **Pattern recognition:** Identifying recurring patterns in input data with the goal of understanding or categorising that input
- **Robotics:** to manipulate objects and move about

Artificial Intelligence Applications

The current state of art:

- Robotic Vehicles: e.g., driverless car
- Speech recognition
- Face recognition
- Autonomous planning and scheduling
- Game playing: e.g., AlphaGo, AlphaGo Zero, OpenAI Five in Dota 2
- Spam fighting
- Machine translation
- Robotics: e.g., robotic vacuum cleaner, PackBot in military application, Atlas
- Etc.

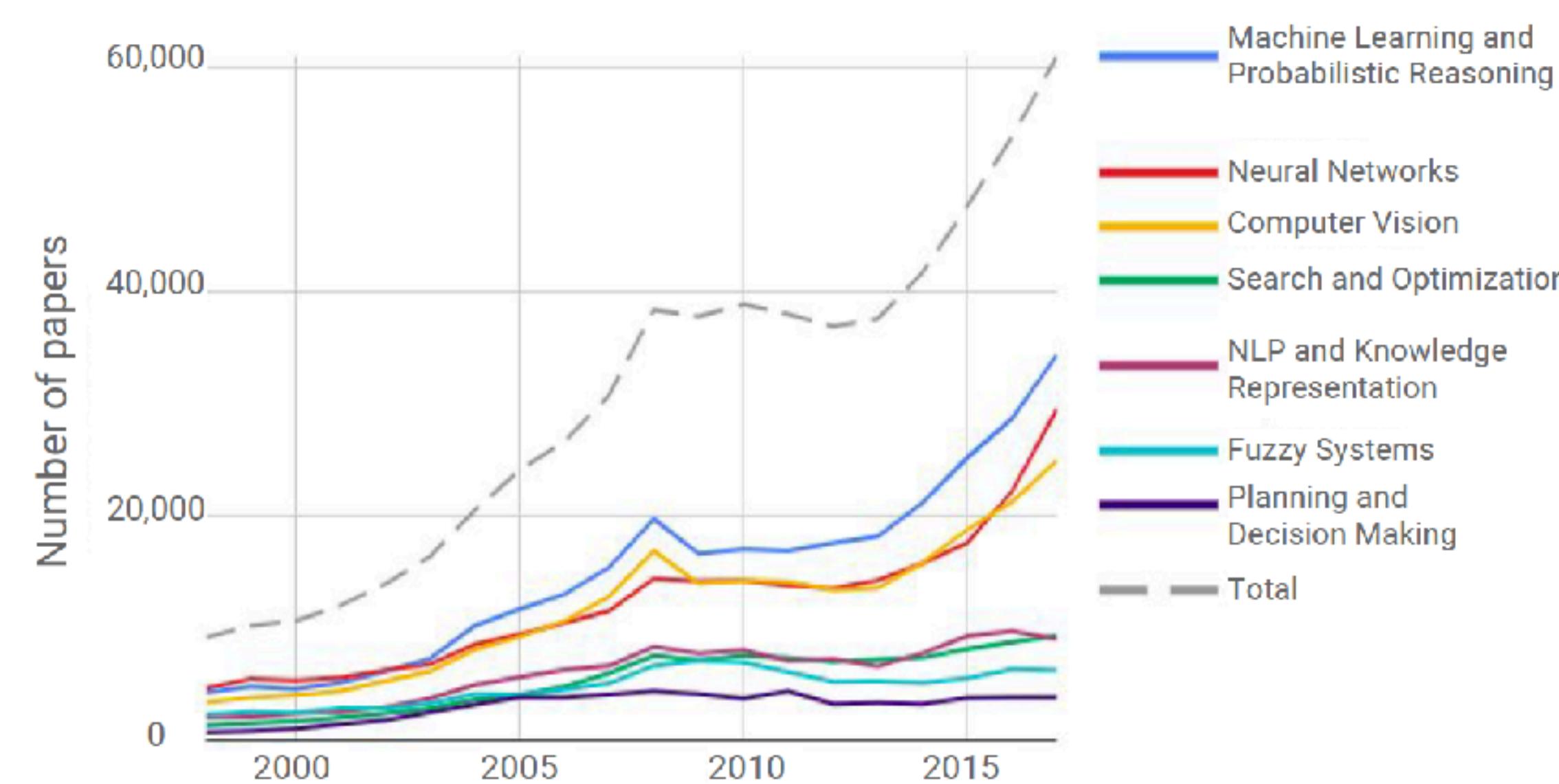
Bill Gates (@BillGates) posted a tweet on June 27, 2016, at 7:25 AM. The tweet reads: "#AI bots just beat humans at the video game Dota 2. That's a big deal, because their victory required teamwork and collaboration – a huge milestone in advancing artificial intelligence. b-gat.es/2KqAlzU". The tweet has received 14.6K likes and 6,301 people are talking about it. The URL in the tweet is https://b-gat.es/2KqAlzU.

Artificial Intelligence Index

[AI Index 2018 Annual Report](#)

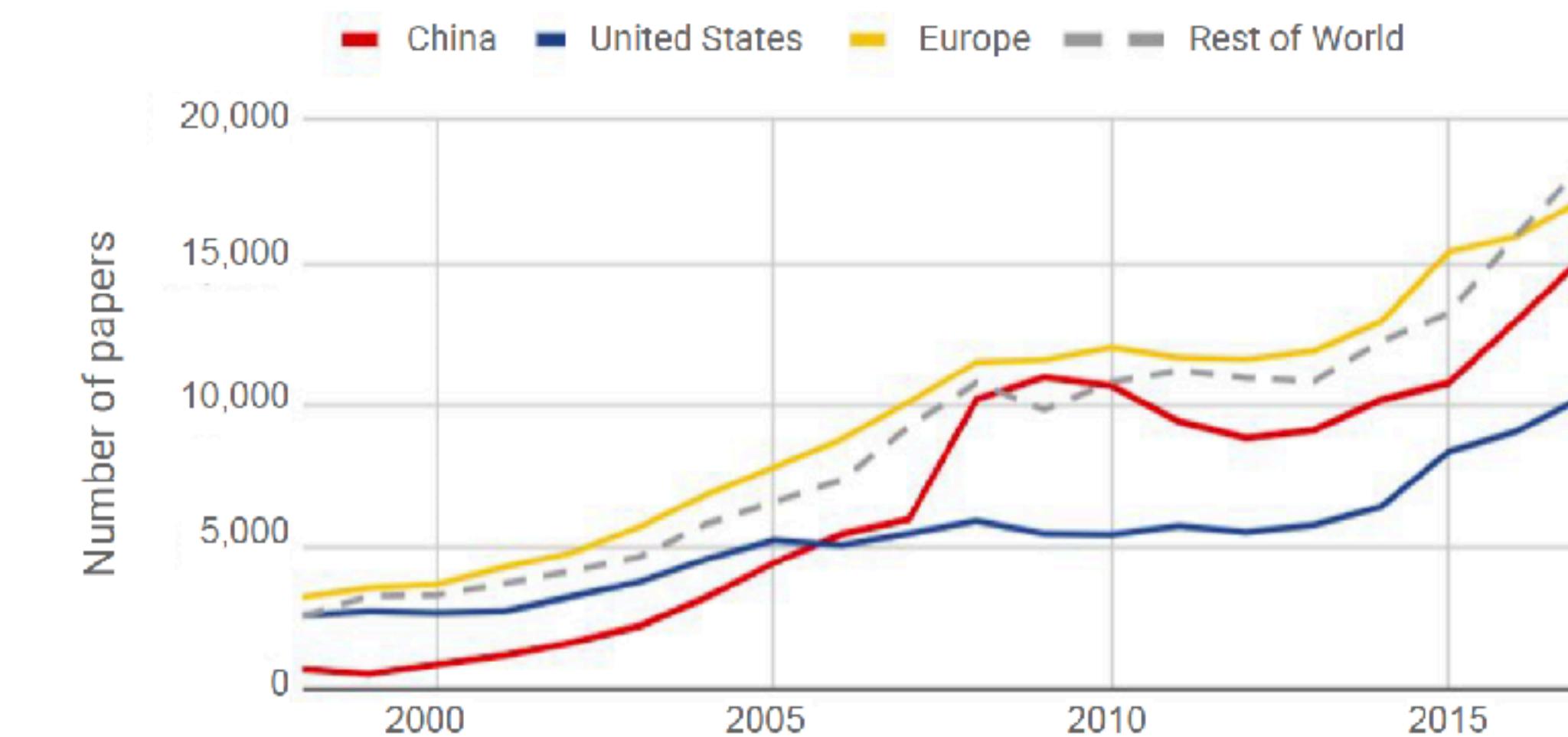
1. AI paper by subcategories

Number of AI papers on Scopus by subcategory (1998–2017)
Source: Elsevier



2. AI paper by region

Annually published AI papers on Scopus by region (1998–2017)
Source: Elsevier

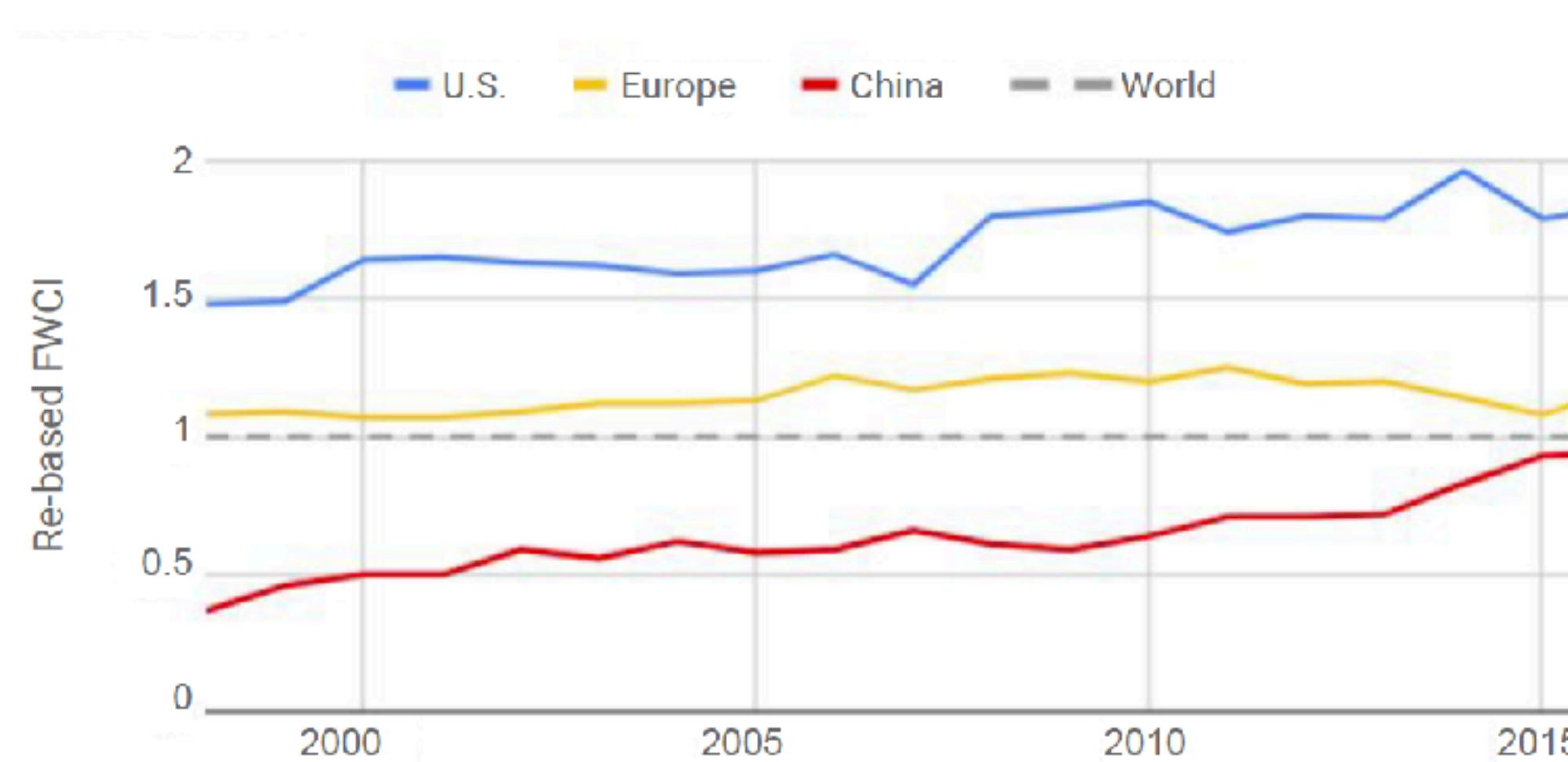


Artificial Intelligence Index

[AI Index 2018 Annual Report](#)

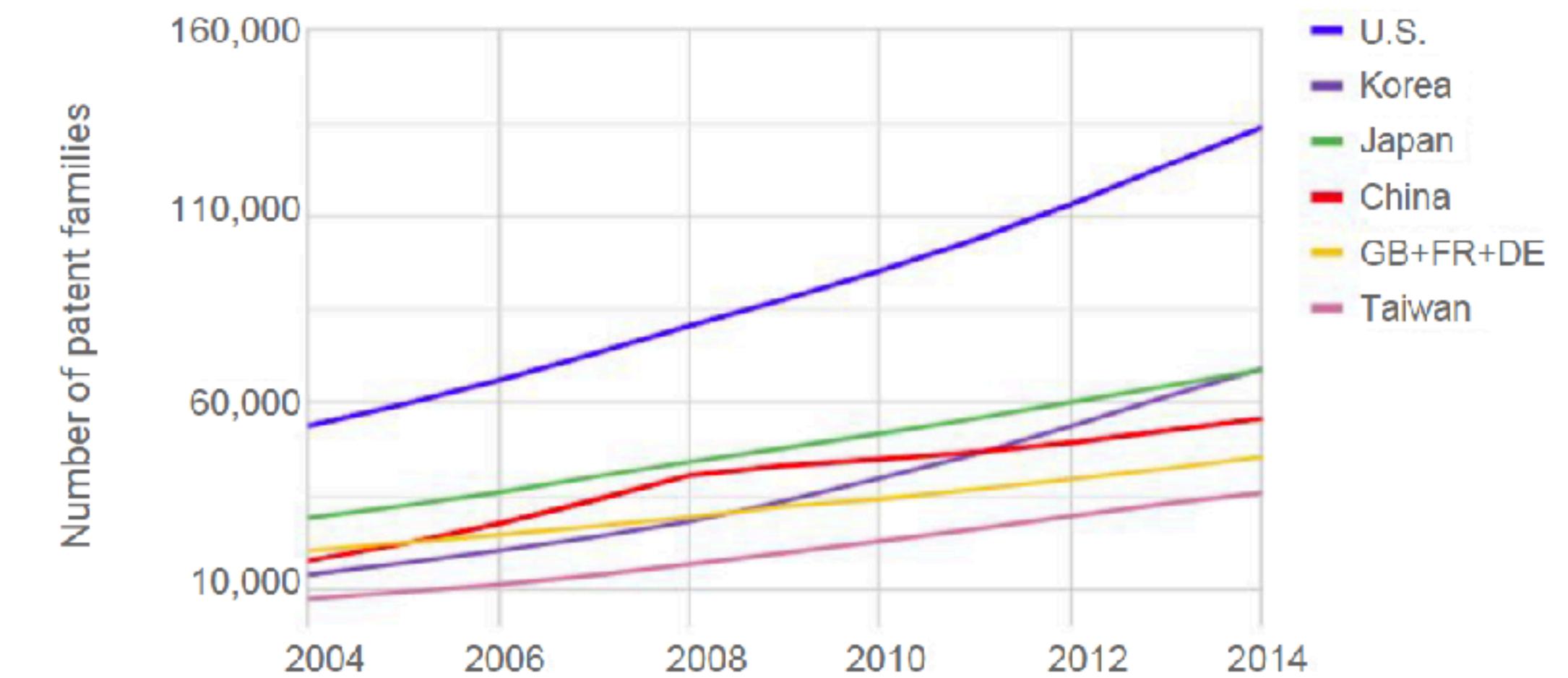
3. AI paper citation impact by region

Field-Weighted Citation Impact of AI authors by region (1998–2016)
Source: Elsevier



4. AI patent by inventor region

AI patents by inventor region (2004–2014)
Source: amplified



Artificial Intelligence Race

AI Index 2018 Annual Report: Recent Government Initiatives

USA	EUROPE (EU, U.K., and France only — does not include several initiatives by other European countries)	CHINA
<p>October 2016: The Obama administration proposes a national AI R&D strategy with the goal of investing in research, developing methods for human-AI collaboration, addressing the safety, ethical, legal and, societal implications of AI, creating public data sets for AI training, and evaluating AI technologies through standards and benchmarks.</p> <p>The U.S. government also published the first policy report on AI: Preparing for the Future of Artificial Intelligence.</p> <p>May 2018: The Trump administration's Summit on AI announces its goals to (1) maintain American leadership in AI (2) support the American worker (3) promote public R&D, and (4) remove barriers to innovation. The Select Committee on Artificial Intelligence is created to advise the White House on AI R&D priorities and to consider the creation of partnerships with academia and industry.</p>	<p>EU</p> <p>April 2018: EU member states sign the Declaration of Cooperation on AI, where they agree to work together. EC also issues a Communication on AI with the objectives of (1) boosting the EU's research investments (2) preparing for socioeconomic changes (3) building an ethical and legal framework. The communication also dedicates 1.5B euros (1.7B USD) to support AI research during 2018—2020. This is in addition to the 2.6B euros from the Horizon 2020 program. The EC's goal is to invest 20B euros in AI research over the next 10 years.</p> <p>June 2018: the EC proposes a Digital Europe program with a budget of 9.2B euros (10.4B USD) for 2021—2027. The program focuses on advancing AI technology and ensuring the use of AI across the economy and society. The EC also proposes to develop common 'European libraries' of algorithms that are accessible to all.</p>	<p>Since 2014, the Chinese government has launched a series of key national AI initiatives with the goal of creating a \$14.7B AI market in China by 2018 and ensure that China leads the world in AI by 2030.</p> <p>July 2015: Internet+ initiative focuses on intelligent manufacturing in China. AI-specific objectives include: increasing public support for development of AI; promoting the popularization of AI in areas like the smart home, smart car, and robots; building a large training database including voice, image, video, maps, research; and developing and industrializing key AI technologies such as computer vision, language processing, and human computer interaction.</p> <p>2016—2020: The robot industry development plan makes way for the development of intelligent industrial and service robots in China.</p>

Artificial Intelligence Race

AI Index 2018 Annual Report: Recent Government Initiatives-cont.

USA	EUROPE (EU, U.K., and France only — does	CHINA
<p>September 2018: DARPA announces \$2B+ investment plan to overcome limitations on AI technology. The AI Next program begins. The Subcommittee on Information Technology of the U.S. House Committee on Oversight and Government Reform publishes a white paper on AI and its impact on policy: Rise of the Machines: Artificial Intelligence and its Growing Impact on U.S. Policy.</p> <p>Private companies play the central role in AI development / investment in the U.S. In 2017, private technology companies like Amazon and Alphabet invested \$16.1B and \$13.9B, respectively, in R&D. To put this in perspective, the total budget for the NSF, together with DARPA and DOT's investment in autonomous and unmanned systems totals \$5.3 billion in the 2019 budget.</p>	<p>France</p> <p>2018: France publishes three recent reports: (1) The French AI plan, which proposes a strategy on research, education, innovation, (2) A report produced by the Office of the Parliament for Science and Technology that focuses on social and regulatory issues, and (3) The Mission Villani report, which focuses on policy and reducing the brain drain, and also addresses the importance of diversity in AI research.</p> <p>United Kingdom</p> <p>2016: The U.K. Parliament House of Lords Select Committee on Artificial Intelligence began releasing a series of AI policy reports.</p> <p>April 2018: The government publishes its AI Sector Deal which invests 950M pounds (1.2B USD) to support research / education, and enhance the UK's data infrastructure.</p>	<p>July 2017: New Generation AI Development Plan strategizes AI development by 2030. It plans for China's AI technology to be on par with most advanced levels worldwide by 2020. By 2025 it states that China will have major breakthroughs in AI theory and AI will become the driving force for industrial upgrading and economic restructuring. And by 2030, China will become the world's major AI innovation center.</p>

Artificial Intelligence Race

Singapore Government Initiatives

AI Singapore (AISG)



A national AI programme launched by the National Research Foundation (NRF) to anchor deep national capabilities in Artificial Intelligence (AI) in 2017. NRF will invest up to \$150 million over five years in AI Singapore.

1. AI for everyone: free introductory AI program for 10,000 people ranging from secondary school students to working adults, over 3 years

2. AI for industry (AI4I): a three-month foundational programme for 2,000 working professionals and students who are technically inclined and eager to learn programming

3. AI Ethics Council: AI ethical conundrum - introduced by *having machines that make decisions on our behalf* – are a growing question which needs to be tackled and answered.

The formation of an Advisory Council on the Ethical Use of AI and Data

4. AI Apprenticeship Programme (AIAP): intensive 9-month programme

5. 100 Experiments (100E) Programme: aim to solve industrial problem by using AI technics

Artificial Intelligence Race

Singapore Government Initiatives

The full composition of Singapore's Advisory Council on the Ethical Use of AI and Data (Advisory Council) was announced by Minister for Communications and Information Mr S Iswaran at AI Singapore's first year anniversary.

SINGAPORE – 30 August, 2018: The full composition of Singapore's Advisory Council on the Ethical Use of AI and Data (Advisory Council) was today announced by Minister for Communications and Information Mr S Iswaran at AI Singapore's first year anniversary. This follows the earlier announcement in June on the establishment of the Advisory Council, to be chaired by former Attorney-General V.K. Rajah SC.

The eleven Advisory Council members come from diverse backgrounds and have been selected for their ability to contribute to the Advisory Council's objectives. Members comprise international leaders in AI such as Google, Microsoft and Alibaba; advocates of social and consumer interests; and leaders of local companies who are keen to make use of AI. The full list of Advisory Council members can be found in [Annex A](#).

The formation of the Advisory Council is one of three structured, interlinked initiatives¹ to support the engagement of stakeholders to collaboratively develop a trusted and vibrant AI ecosystem and position Singapore as a leading hub for AI.

The Advisory Council will advise and work with the Infocomm Media Development Authority (IMDA) on the responsible development and deployment of AI. Amongst other things, the Advisory Council will assist IMDA in engaging stakeholders on issues that support the development of AI governance capabilities and frameworks. These include engaging ethics boards of commercial enterprises on ethical and related issues arising from private sector use of AI and data; consumer representatives on consumer expectations and acceptance of the use of AI; as well as members of the private capital community on the need to incorporate ethical considerations in their investment decisions into businesses which develop or adopt AI.

The Advisory Council will also assist the Government in developing ethics standards and reference governance frameworks, and publish advisory guidelines, practical guides, and codes of practice for the voluntary adoption by the industry.

Supported by a secretariat, the Advisory Council will be able to tap on technical, legal, ethics and international experts from different fields and representatives for consumers and civil society to guide its work.

Issues existed in Artificial Intelligence

Name a few:

- AI and job displacement
- Do humanlike machines deserve human rights?
- People might lost their sense of being unique
 - The boundary of human
- Human-robot relationship
- AI systems might be used towards undesirable ends
 - Massive surveillance program
- The use of AI systems might result into a loss of accountability, explainability, and transparency
- System bias in AI systems
- The boundary of AI
- etc.

Automation and Job Displacement

Examples

Self-Driving Trucks

Tractor-trailers without a human at the wheel will soon barrel onto highways near you. What will this mean for the nation's 1.7 million truck drivers?



Drone Deliveries Are No Longer Pie In The Sky



Deutsche Post/DHL is testing deliveries of medicine from a pharmacy in Bonn. (Photo by Ulrich Baumgarten/Getty Images)

Related - Technology and Job Displacement

South Korea cab drivers protest Uber-like ride share app

Drivers say the service being introduced by KakaoTalk -- the country's largest mobile chat app -- will threaten their livelihoods

Last Published: Thu, Dec 20 2018, 07:32 PM IST

Bloomberg

Enter email for newsletter

Sign Up



Traffic jams were reported on some highways leading to the capital as taxi drivers drove their vehicles at a snail's pace across the river Han, raising concerns of evening traffic chaos on the giant Mapo Bridge. Photo: AFP

Seoul: Tens of thousands of taxi drivers in South Korea went on a nationwide strike Thursday, snarling up traffic in Seoul, in the latest protest at a planned Uber-like ride-sharing service.

MORE FROM TECHNOLOGY »

- What were some of the new features WhatsApp introduced in 2018? Here's a list
- Picture-in-Picture comes to WhatsApp Web
- Keep your Aadhaar, other docs safe. Shift to DigiLocker app
- Itching for a quick doodle? Chrome Canvas lets you do just that
- Realme U1 available with Rs 1,500 discount till January 2

BUSINESS

Spain restricts ride-hailing services like Uber

After years of protests, Spanish taxi drivers claimed a major victory over app-based ride sharing services such as Uber. In a close vote, the Spanish parliament resolved a measure to regulate providers, but leave them with a 4-year frame to adjust to the changes.



▶ Watch video

01:35

Automation and Job Displacement

Solution?

Singapore SkillsFuture Program

SkillsFuture is a national movement to provide Singaporeans with the opportunities to develop their fullest potential throughout life, regardless of their starting points.

SkillsFuture Credit

All Singaporeans aged 25 and above will receive an opening credit of S\$500 from January 2016. Your credit will not expire and the government will provide periodic top-ups, so you may accumulate your credit.

Singapore

More than 285,000 Singaporeans benefited from SkillsFuture Credit since launch



Office workers in the central business district of Singapore. (File photo: AFP/Simin Wang)

Robots Identity

On October 25, 2017, Sophia, a delicate looking woman with doe-brown eyes and long fluttery eyelashes made international headlines. She'd just become a full citizen of Saudi Arabia -- the first robot in the world to achieve such a status.

What does it mean to be a citizen?

What rights does Sophia hold?

Everything You Need To Know About Sophia, The World's First Robot Citizen



Human-Robots Relationship

China

Chinese man 'marries' robot he built himself

Zheng Jiajia had grown tired of pressure to get married so he turned to Yingying, a robot spouse he constructed last year



0

Benjamin Haas in Hong Kong

@haasbenjamin

Tuesday 4 April 2017
10.58 BST



Zheng Jiajia, 31, decided to 'marry' Yingying after failing to find a suitable human alternative. Photograph: Qiangjing Evening News

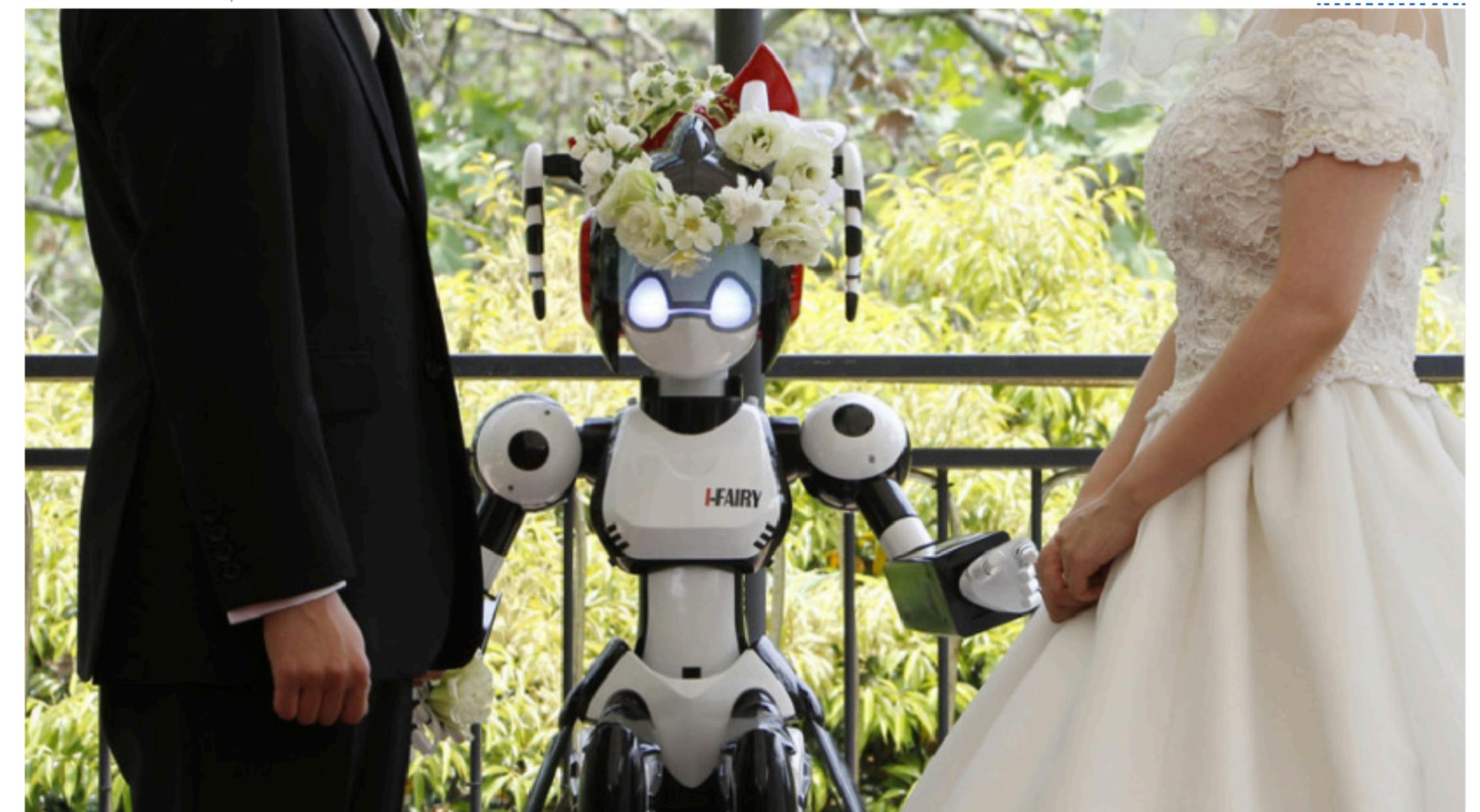
Source: <https://www.theguardian.com/world/2017/apr/04/chinese-man-marries-robot-built-himself>

25% of millennials think human-robot relationships will soon become the norm - study

Published time: 9 Dec, 2017 18:08

Edited time: 10 Dec, 2017 10:28

[Get short URL](#)



Stock images of a humanoid robot in Japan. © Reuters

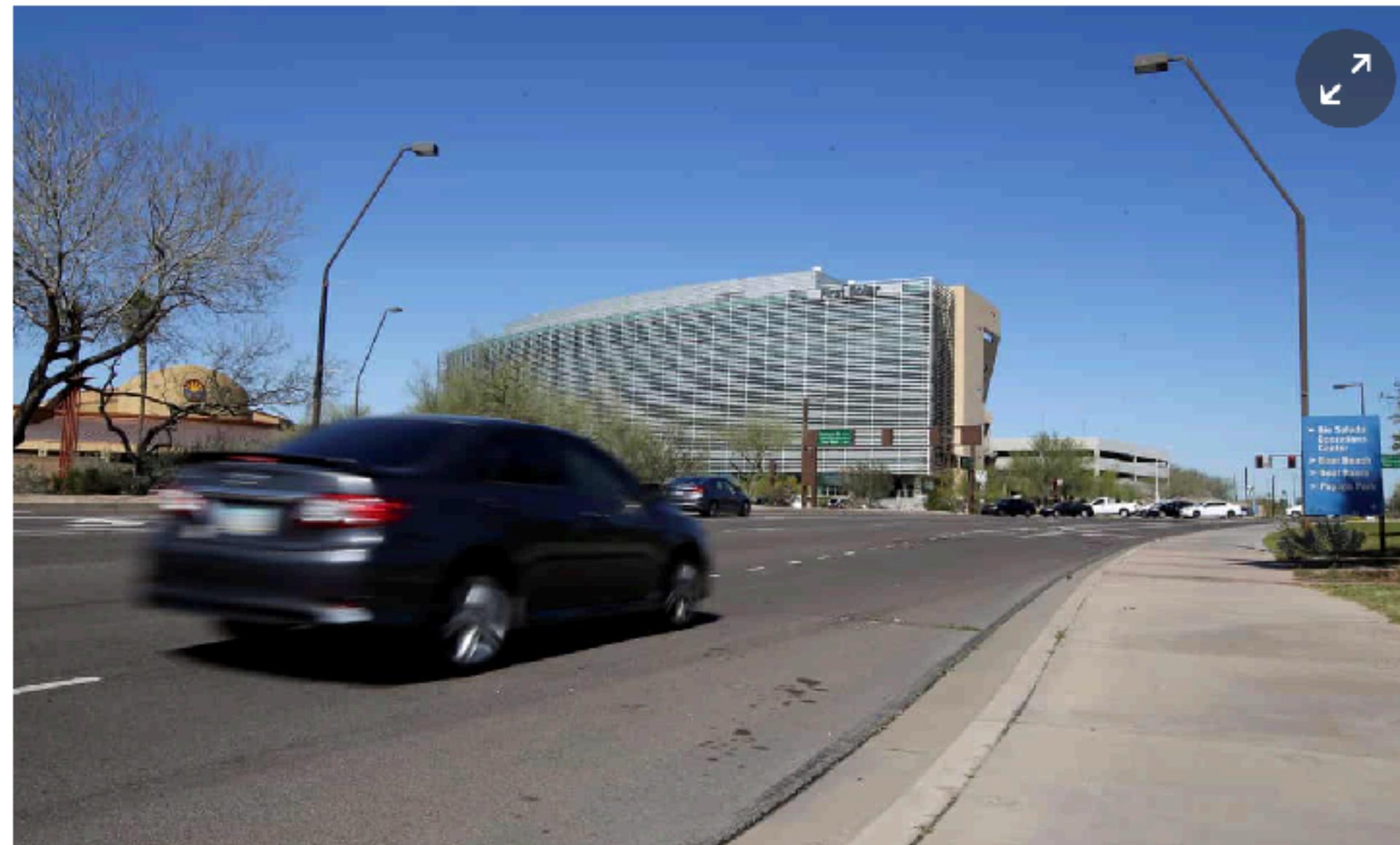
A study of 12,000 people by French advertising firm Havas has revealed that human-robot relations could become ever more cozy in the years to come, with some 25 percent of people aged between 18-34 believing that it'll be normal for humans and robots to develop deep friendships, or even fall in love.

Self-driving Car Accidents

Recent Cases

Self-driving Uber kills Arizona woman in first fatal crash involving pedestrian

Tempe police said car was in autonomous mode at the time of the crash and that the vehicle hit a woman who later died at a hospital



▲ A car passes the location where a woman pedestrian was struck and killed by an Uber self-driving sport utility vehicle in Tempe, Arizona, on Monday. Photograph: Rick Scuteri/Reuters

Tesla in fatal California crash was on Autopilot

⌚ 31 March 2018

f Share



The driver of the Tesla Model X died shortly after the crash

REUTERS

System Accountability

Whose fault?

Preliminary report on Uber's Arizona fatal case from NTSB:

"According to data obtained from the self-driving system, the system first registered radar and LIDAR observations of the pedestrian about 6 seconds before impact, when the vehicle was traveling at 43 mph. As the vehicle and pedestrian paths converged, **the self-driving system software classified the pedestrian as an unknown object, as a vehicle, and then as a bicycle with varying expectations of future travel path.** At 1.3 seconds before impact, the self-driving system determined that an emergency braking maneuver was needed to mitigate a collision (see figure 2).² According to Uber, **emergency braking maneuvers are not enabled while the vehicle is under computer control, to reduce the potential for erratic vehicle behavior. The vehicle operator is relied on to intervene and take action. The system is not designed to alert the operator.**"

<https://www.ntsb.gov/investigations/AccidentReports/Reports/HWY18MH010-prelim.pdf>

System Accountability

Whose fault?

Preliminary report on Tesla's California fatal case from NTSB:

- The Autopilot system was engaged on four separate occasions during the 32-minute trip, including a continuous operation for the last 18 minutes 55 seconds prior to the crash.
- During the 18-minute 55-second segment, the vehicle provided **two visual alerts and one auditory alert for the driver to place his hands on the steering wheel**. These alerts were made more than 15 minutes prior to the crash.
- During the 60 seconds prior to the crash, the driver's hands were detected on the steering wheel on three separate occasions, for a total of 34 seconds; for the last 6 seconds prior to the crash, the vehicle did not detect the driver's hands on the steering wheel.
- At 8 seconds prior to the crash, the Tesla was following a lead vehicle and was traveling about 65 mph.
- At 7 seconds prior to the crash, the Tesla began **a left steering movement while following a lead vehicle**.
- At 4 seconds prior to the crash, the Tesla was no longer following a lead vehicle.
- At 3 seconds prior to the crash and up to the time of impact with the crash attenuator, the Tesla's speed increased from 62 to 70.8 mph, **with no precrash braking or evasive steering movement detected**.
- <https://www.ntsb.gov/investigations/AccidentReports/Reports/HWY18FH011-preliminary.pdf>

Dilemma in Autonomous Vehicles Design

“If you delegate transportation from point A to Point B entirely to a machine, are you responsible if it hits someone?”

In a self-driving car, the control of the vehicle is shared between the driver and the car’s software. How the software behaves is in turn controlled—designed—by the software engineers.

In Aviation domain:

“while the plane is being run by software, the pilots in the cockpit are legally responsible for its operation. US Federal Aviation Administration (FAA) regulations specify this directly, and courts have consistently upheld it. So when something goes wrong, we observe pilots becoming “**moral crumple zones**”—largely totemic humans whose central role becomes soaking up fault, even if they had only partial control of the system.”

Case: Air France flight 447 in 2009

Automation Paradox:

the more efficient the automated system, the more crucial the human contribution of the operators. Humans are less involved, but their involvement becomes more critical.

Issues to be Solved

Just list a few:

- How can we ensure that AI/AS (Autonomous Systems) do not infringe human rights (fair)?
- How can we assure that AI/AS are accountable?
- How can we ensure that AI/AS are transparent?
- How can we ensure that AI/AS are explainable?
- How can we extend the benefits and minimise the risks of AI/AS technology being misused?
- Lack of value-based ethical culture and practices for industry
- Lack of values-aware leadership
- Lack of ownership or responsibility from tech community
- Lack of empowerment to raise ethical concerns
- Lack of an independent review organization
- Lack of law regulation

Recommended Materials

A. M. TURING; I.—Computing Machinery and Intelligence, Mind, Volume LIX, Issue 236, 1 October 1950, Pages 433–460 by Azim Shariff et al. <https://www.csee.umbc.edu/courses/471/papers/turing.pdf>

Artificial Intelligence, A Modern Approach, 3rd Edition, Stuart Russell and Peter Norvig

Artificial Intelligence Index 2018 Annual Report, <http://cdn.aiindex.org/2018/AI%20Index%202018%20Annual%20Report.pdf>

AI Singapore, <https://www.aisingapore.org>

2001: A Space Odyssey, [https://en.wikipedia.org/wiki/2001:_A_Space_Odyssey_\(film\)](https://en.wikipedia.org/wiki/2001:_A_Space_Odyssey_(film))

Her, [https://en.wikipedia.org/wiki/Her_\(film\)](https://en.wikipedia.org/wiki/Her_(film))