# AI and Life in 2030

ONE HUNDRED YEAR STUDY ON ARTIFICIAL INTELLIGENCE | REPORT OF THE 2015 STUDY PANEL | SEPTEMBER 2016

# To Cite

- Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, AnnaLee Saxenian, Julie Shah, Milind Tambe, and Astro Teller. "Artificial Intelligence and Life in 2030." One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA, September 2016. Doc: http://ai100.stanford.edu/2016-report. Accessed: September 6, 2016.

**Study Panel: Peter Stone**, *Chair*, University of Texas at Austin, **Rodney Brooks**, Rethink Robotics, **Erik Brynjolfsson**, Massachussets Institute of Technology, **Ryan Calo**, University of Washington, **Oren Etzioni**, Allen Institute for AI, **Greg Hager**, Johns Hopkins University, **Julia Hirschberg**, Columbia University, **Shivaram Kalyanakrishnan**, Indian Institute of Technology Bombay, **Ece Kamar**, Microsoft Research, **Sarit Kraus**, Bar Ilan University. **Kevin Leyton-Brown**, University of British Columbia, **David Parkes**, Harvard University, **William Press**, University of Texas at Austin, **AnnaLee (Anno) Saxenian**, University of California, Berkeley, **Julie Shah**, Massachussets Institute of Technology, **Milind Tambe**, University of Southern California, **Astro Teller**, X

**Standing Committee of the One Hundred Year Study of Artificial Intelligence:** Barbara J. Grosz, *Chair*, Russ Altman, Eric Horvitz, Alan Mackworth, Tom Mitchell, Deidre Mulligan, Yoav Shoham

# TABLE OF CONTENTS

# A Short History of AI

- Based primarily on Nilsson's book and written from the prevalent current perspective, which focuses on data intensive methods and big data

# 1950s:  When AI was born …

- The field of Artificial Intelligence (AI) was officially born and christened at a 1956 workshop organized by John McCarthy at the Dartmouth Summer Research Project on Artificial Intelligence
- **The goal was to investigate ways in which machines could be made to simulate aspects of intelligence—the essential idea that has continued to drive the field forward**
- McCarthy is credited with the first use of the term "*artificial intelligence*" in the proposal he co-authored for the workshop with Marvin Minsky, Nathaniel Rochester, and Claude Shannon

# … before the birth of AI …

- 18th century: Thomas Bayes's framework for reasoning about the **probability** of events
- 19th century: George Boole showed that **logical reasoning**—dating back to Aristotle—could be performed *systematically* in the same manner as solving a system of equations
- 20th century:  progress in the experimental sciences had led to the emergence of the field of **statistics**, which enables inferences to be drawn rigorously from data.
- 1950s:
  - The idea of physically engineering a machine to execute sequences of instructions had matured by the 1950s, and resulted in the construction of the first **electronic computers**
  - Primitive **robots**, which could sense and act autonomously, had also been built

# Alan Turing & AI

- The most influential ideas underpinning computer science came from Alan Turing, who proposed a formal model of computing
- Turing's classic essay, *Computing Machinery and Intelligence*,
  - imagines the possibility of computers created for simulating intelligence
  - explores many of the ingredients now associated with AI
    - how intelligence might be tested
    - how machines might automatically *learn*
- Though these ideas inspired AI, Turing did not have access to the computing resources needed to translate his ideas into action

# 1950s – 1970s

- Newell and Simon pioneered the foray into **heuristic search**, an efficient procedure for finding solutions in large, combinatorial spaces
  - In particular, they applied this idea to construct proofs of mathematical theorems, first through their Logic Theorist program, and then through the General Problem Solver
- In the area of **computer vision**, early work in character recognition by Selfridge and colleagues laid the basis for more complex applications such as face recognition
- By the late 1960s, work had also begun on **natural language processing**
- “Shakey”, a wheeled robot built at SRI International, launched the field of **mobile robotics**

# 1950s – 1970s, Cont'd …

- Samuel's Checkers-playing program, which improved itself through self-play, was one of the first working instances of a **machine learning** system
- Rosenblatt's *Perceptron*, a computational model based on biological neurons, became the basis for the field of **artificial neural networks**
- Feigenbaum and others advocated the case for building **expert systems**—knowledge repositories tailored for specialized domains such as chemistry and medical diagnosis

# 1980s …. AI Winter

- But by the 1980s, the field still could boast no significant *practical* successes
  - This gap between theory and practice arose in part from an insufficient emphasis within the AI community on *grounding* systems physically, with direct access to environmental signals and data.
  - An overemphasis on Boolean (True/False) logic, overlooking the need to quantify uncertainty
- The field was forced to take cognizance of these shortcomings in the mid-1980s
  - interest in AI began to drop, and funding dried up
- Nilsson calls this period the “AI winter”

# 1990s …

- A much needed resurgence in the nineties built upon the idea that "Good Old-Fashioned AI" was **inadequate as an end-to-end approach to building intelligent systems**
- Rather, intelligent systems needed to be built from the ground up, at all times *solving* the task at hand, albeit with different degrees of proficiency
- **Technological progress had also made the task of building systems driven by real-world data more feasible**
  - from Soh: where Informatics now come into play

# … The Past Two Decades

- Cheaper and more reliable hardware for sensing and actuation made robots easier to build
- Further, the Internet's capacity for **gathering large amounts of data**, and the **availability of computing power and storage** to process that data, enabled statistical techniques that, by design, **derive solutions from data**
- These developments have allowed AI to emerge in the past two decades as a profound influence on our daily lives

# Traditional Sub-Areas of AI

- **Search and Planning** deal with reasoning about goal-directed behaviors
  - Search plays a key role, e.g., in chess-playing programs such as Deep Blue, in deciding which move (behavior) will ultimately lead to a win (goal)
- The area of **Knowledge Representation and Reasoning** involves processing information (typically when in large amounts) into a structured form that can be queried more reliably and efficiently
  - IBM's Watson program, which beat human contenders to win the Jeopardy challenge in 2011, was largely based on an efficient scheme for organizing, indexing, and retrieving large amounts of information gathered from various sources

# Traditional Sub-Areas of AI 2

- **Machine Learning** is a paradigm that enables systems to automatically improve their performance at a task by observing relevant data
  - key contributor to the AI surge in the past few decades
    - search and product recommendation engines
    - speech recognition
    - fraud detection
    - image understanding
    - countless other tasks that once relied on human skill and judgment
  - The automation of these tasks has enabled the scaling up of services such as e-commerce

# Traditional Sub-Areas of AI 3

- More and more intelligent systems get built → how such systems will interact with each other
  - The field of **Multi-Agent Systems** considers this question, which is becoming increasingly important in on-line marketplaces and transportation systems
- From its early days, AI has taken up the design and construction of systems that are embodied in the real world
  - The area of **Robotics** investigates fundamental aspects of sensing and acting—and especially their integration—that enable a robot to behave effectively
  - Robots and other computer systems share the living world with human beings → **Human Robot Interaction** has also become prominent in recent decades

# Traditional Sub-Areas of AI 4

- **Machine perception** has always played a central role in AI, partly in developing robotics, but also …
  - The most commonly studied perception modalities are **Computer Vision** and **Natural Language Processing**
- Several other focus areas within AI today are consequences of the growth of the Internet
  - **Social Network Analysis** investigates the effect of neighborhood relations in influencing the behavior of individuals and communities
  - **Crowdsourcing** relies on harnessing human intelligence (typically from thousands of humans) to solve hard computational problems

# Note on Integrating Ideas …

- Although the separation of AI into sub-fields has enabled deep technical progress along several different fronts, **synthesizing intelligence at any reasonable scale invariably requires many different ideas to be integrated**
- For example, the AlphaGo program that recently defeated the current human champion at the game of Go used multiple **machine learning algorithms for training itself**, and also **used a sophisticated search procedure while playing the game**

# Notes on Executive Summary

- AI is a science and a set of computational technologies that are inspired by—but typically operate quite differently from—the ways people use their nervous systems and bodies to sense, learn, reason, and take action
  - **Computer vision and AI planning** drive the video games that are now a bigger entertainment industry than Hollywood
  - **Deep learning** has made speech-understanding practical on our phones and in our kitchens, and its algorithms can be applied widely to an array of applications that rely on pattern recognition
  - **Natural Language Processing (NLP) and knowledge representation and reasoning** have enabled a machine to beat the Jeopardy champion and are bringing new power to Web searches

# Notes on Executive Summary 2

- **While impressive, these technologies are highly tailored to particular tasks**
  - Each application typically requires years of specialized research and careful, unique construction
- Substantial increases in the future uses of AI technologies:
  - self-driving cars
  - healthcare diagnostics and targeted treatments,
  - physical assistance for elder care can be expected
- AI and robotics will also be applied across the globe in industries struggling to attract younger workers, such as agriculture, food processing, fulfillment centers, and factories
  - facilitate delivery of online purchases through flying drones, self-driving trucks, or robots that can get up the stairs to the front door

# Notes on Executive Summary 3

- This report is the first in a series to be issued at regular intervals as a part of the **One Hundred Year Study on Artificial Intelligence (AI100)**
- Starting from a charge given by the AI100 Standing Committee to **consider the likely influences of AI in a typical North American city by the year 2030**, the 2015 Study Panel focused their attention on eight domains they considered most salient
  - **transportation; service robots; healthcare; education; low-resource communities; public safety and security; employment and workplace; and entertainment**

# Notes on Executive Summary 4

- Contrary to the more fantastic predictions for AI in the popular press, the Study Panel **found no cause for concern that AI is an imminent threat to humankind**
- **No machines with self-sustaining long-term goals and intent have been developed, nor are they likely to be developed in the near future**
- **Increasingly useful applications of AI, with potentially profound positive impacts on our society and economy are likely to emerge between now and 2030**

# Notes on Executive Summary 5

- Many of these developments will **spur disruptions in how human labor is augmented or replaced by AI, creating new challenges for the economy and society more broadly**
- Application design and policy decisions made in the near term are likely to have long-lasting influences on the nature and directions of such developments, **making it important for AI researchers, developers, social scientists, and policymakers to balance the imperative to innovate with mechanisms to ensure that AI's economic and social benefits are broadly shared across society**

# Notes on Executive Summary 6

- If society approaches these technologies primarily with fear and suspicion → missteps that slow AI's development or drive it underground → impeding important work on ensuring the safety and reliability of AI technologies
- If society approaches AI with a more open mind → AI technologies could profoundly transform society for the better in the coming decades

# Notes on Overview

- Beneficial AI applications in schools, homes, and hospitals are already growing at an accelerated pace
- Major research universities devote departments to AI studies, and technology companies (e.g., Apple, Facebook, Google, IBM, and Microsoft) spend heavily to explore AI applications they regard as critical to their futures
  - Even Hollywood uses AI technologies to bring its dystopian AI fantasies to the screen

# Notes on Overview 2

- The Study Panel further narrowed its inquiry to eight domains where AI is already having or is projected to have the greatest impact:
  - **Transportation**
  - **Healthcare**
  - **Education**
  - **Low-resource communities**
  - **Public safety and security**
  - **Employment and workplace**
  - **Home/service robots**
  - **Entertainment**

# Notes on Overview 3

- Each domain faces varied AI-related challenges, including
  - the difficulty of creating safe and reliable hardware for sensing and effecting (transportation and service robots)
  - the difficulty of smoothly interacting with human experts (healthcare and education),
  - the challenge of gaining public trust (low-resource communities and public safety and security),
  - the challenge of overcoming fears of marginalizing humans (employment and workplace)
  - the risk of diminishing interpersonal interaction (entertainment)

# Notes on Overview 4

- Some domains are primarily business sectors
  - E.g., transportation and healthcare
- Some are more oriented to consumers
  - E.g., entertainment and home service robots
- Some cut across sectors
  - E.g., employment/workplace and low-resource communities
- In each domain, AI also raises important ethical and social issues
  - Privacy concerns
  - Robots and other AI technologies have already begun to displace jobs in some sectors

# Notes on Overview 5

- As a society, we are now at a crucial juncture in determining **how to deploy AI-based technologies in ways that promote, not hinder, democratic values such as freedom, equality, and transparency**
- **For individuals, the quality of the lives we lead and how our contributions are valued are likely to shift gradually, but markedly**
- Over the next several years, **AI research, systems development, and social and regulatory frameworks** will shape how the benefits of AI are weighed against its costs and risks, and how broadly these benefits are spread

# Notes on Overview: Transportation

- Autonomous transportation will soon be commonplace and will strongly influence the public's perception of AI
- As cars become better drivers than people → city-dwellers will own fewer cars, live further from work, and spend time differently → leading to an entirely new urban organization
- **In the typical North American city in 2030, physically embodied AI applications will not be limited to cars, but are likely to include trucks, flying vehicles, and personal robots**

# Notes on Overview: Home/Service Robots

- Have already entered people's houses, primarily in the form of vacuum cleaners
- Better chips, low-cost 3D sensors, cloud-based machine learning, and advances in speech understanding will enhance future robots' services and their interactions with people
- Special purpose robots will deliver packages, clean offices, and enhance security.
- **But technical constraints and the high costs of reliable mechanical devices will continue to limit commercial opportunities to narrowly defined applications for the foreseeable future**

# Notes on Overview: Healthcare

- Collecting useful data from personal monitoring devices and mobile apps, from **electronic health records (EHR)** in clinical settings and
  - to a lesser extent, from surgical robots designed to assist with medical procedures and service robots supporting hospital operations
- Advances in healthcare can be promoted via the development of incentives and mechanisms for **sharing** data and for removing overbearing policy, regulatory, and commercial obstacles
- Though clinical applications have been slow to move from the computer science lab to the real-world, there are **hopeful signs that the pace of innovation will improve**
  - **AI systems will have to work closely with care providers and patients to gain their trust**
  - **Advances in how intelligent machines interact naturally with caregivers, patients, and patients' families are crucial**

# Notes on Overview: Education

- Though quality education will always require active engagement by human teachers, **AI promises to enhance education at all levels, especially by providing personalization at scale**
- **Interactive machine tutors** are now being matched to students for teaching science, math, language, and other disciplines
- **Natural Language Processing, machine learning, and crowdsourcing** have boosted **online learning**
  - higher education multiplying the size of their classrooms while addressing individual students' learning needs and styles
- **Over the next fifteen years in a typical North American city, the use of these technologies in the classroom and in the home is likely to expand significantly, provided they can be meaningfully integrated with face-to-face learning**

# Notes on Overview: Low-Resource Communities

- With targeted incentives and funding priorities, AI technologies could help address the **needs of low-resource communities**
- **Data mining and machine learning** have been used to create predictive models to help government agencies address issues such as prevention of lead poisoning in at-risk children and distribution of food efficiently
- **These *budding* efforts suggest more could be done, particularly if agencies and organizations can engage and build trust with these communities**

# Notes on Overview: Public Safety and Security

- North American cities and federal agencies have already begun to deploy AI technologies in border administration and law enforcement
- **By 2030, they will rely heavily upon them, including improved cameras and drones for surveillance, algorithms to detect financial fraud, and predictive policing**
- **Potential harm of predictive policing**: innocent people being unjustifiably monitored
  - Care must be taken to avoid systematizing human bias and to protect civil liberties
  - Well-deployed AI prediction tools **have the potential to provide new kinds of transparency about data and inferences**, and may be applied to detect, remove, or reduce human bias, rather than reinforcing it

# Notes on Overview: Employment and Workplace

- AI is poised to replace people in certain kinds of jobs, such as in the driving of taxis and trucks.
- **However, in many realms, AI will likely replace tasks rather than jobs in the near term, and will also create new kinds of jobs**
  - The new jobs that will emerge are harder to imagine in advance than the existing jobs that will likely be lost
- **AI will also lower the cost of many goods and services**, effectively making everyone better off
- **Longer term, AI may be thought of as a radically different mechanism for wealth creation in which everyone should be entitled to a portion of the world's AI-produced treasures**
  - **It is not too soon for social debate on how the economic fruits of AI technologies should be shared**

# Notes on Overview: Entertainment

- Some used AI to compose music, create stage performances, and even to generate 3D scenes from natural language text
  - Rely on **NLP, information retrieval, image processing, crowdsourcing, and machine learning**
  - Transformed by social networks and other platforms for sharing and browsing blogs, videos, and photos
- The enthusiasm with which people have already responded to AI-driven entertainment has been surprising
- Ongoing debate about the extent to which the technology replaces or enhances sociability
- **AI will increasingly enable entertainment that is more interactive, personalized, and engaging. Research should be directed toward understanding how to leverage these attributes for individuals' and society's benefit**

# Informatics' Roles in These 8 Domains?

# Notes on Overview: What's Next for AI?

- **Large-scale machine learning** concerns the design of learning algorithms, as well as scaling existing algorithms, to work with extremely large data sets
- **Deep learning**, a class of learning procedures, has facilitated object recognition in images, video labeling, and activity recognition, and is making significant inroads into other areas of perception, such as audio, speech, and natural language processing

# Notes on Overview: What's Next for AI? 2

- **Reinforcement learning** is a framework that shifts the focus of machine learning from pattern recognition to experience-driven sequential decision-making
  - It promises to carry AI applications forward toward taking actions in the real world
- **Robotics** is currently concerned with how to train a robot to interact with the world around it in generalizable and predictable ways, how to facilitate manipulation of objects in interactive environments, and how to interact with people
  - Advances in robotics will rely on commensurate advances to improve the reliability and generality of computer vision and other forms of machine perception

# Notes on Overview: What's Next for AI? 3

- **Computer vision** is currently the most prominent form of machine perception
  - For the first time, computers are able to perform some vision tasks better than people
  - Much current research is focused on automatic image and video captioning
- **Natural Language Processing**, often coupled with automatic speech recognition, is quickly becoming a commodity for widely spoken languages with large data sets
  - Research is now shifting to develop refined and capable systems that are able to interact with people through dialog, not just react to stylized requests
  - Great strides have also been made in machine translation among different languages, with more real-time person-to-person exchanges on the near horizon
- **Collaborative systems** research investigates models and algorithms to help develop autonomous systems that can work collaboratively with other systems and with humans

# Notes on Overview: What's Next for AI? 4

- **Collaborative systems** research investigates models and algorithms to help develop autonomous systems that can work collaboratively with other systems and with humans
- **Crowdsourcing and human computation** research investigates methods to augment computer systems by making automated calls to human expertise to solve problems that computers alone cannot solve well

# Notes on Overview: What's Next for AI? 5

- **Algorithmic game theory and computational social choice** draw attention to the economic and social computing dimensions of AI
  - e.g., how systems can handle potentially misaligned incentives, including self-interested human participants or firms and the automated AI-based agents representing them
- **Internet of Things (IoT)** research is devoted to the idea that a wide array of devices (e.g., appliances, vehicles, buildings, and cameras) can be interconnected to collect and share their abundant sensory information to use for intelligent purposes

# Notes on Overview: What's Next for AI? 6

- **Neuromorphic computing** is a set of technologies that seek to mimic biological neural networks to improve the hardware efficiency and robustness of computing systems
  - often replacing an older emphasis on separate modules for input/output, instruction-processing, and memory

# Informatics' Roles in These Trends?

# Notes on Overview: Policy

- **The measure of success for AI applications is the value they create for human lives**

# Notes on Overview: Policy 2

- Designed to enable people to understand AI systems successfully, participate in their use, and build their trust
- Public policies should help ease society's adaptation to AI applications, extend their benefits, and mitigate their inevitable errors and failures
- Debate about how AI is deployed, including concerns about how privacy is protected and AI's benefits fairly shared, should be encouraged
- The Study Panel **recommends that**
  - **All layers of government acquire technical expertise in AI**
  - **Research on the fairness, security, privacy, and societal implications of AI systems be encouraged by removing impediments and increasing private and public spending to support it**

# Notes on Overview: Policy 2

- Designed to enable people to understand AI systems successfully, participate in their use, and build their trust
- Public policies should help ease society's adaptation to AI applications, extend their benefits, and mitigate their inevitable errors and failures
- Debate about how AI is deployed, including concerns about how privacy is protected and AI's benefits fairly shared, should be encouraged
- The Study Panel **recommends that**
  - **All layers of government acquire technical expertise in AI**
  - **Research on the fairness, security, privacy, and societal implications of AI systems be encouraged by removing impediments and increasing private and public spending to support it**

# Notes on Overview: Policy 3

- Currently in the United States, at least 16 separate agencies govern sectors of the economy related to AI technologies
- Rapid advances in AI research and, especially, its applications require experts in these sectors to develop new concepts and metaphors for law and policy
- **As people integrate AI more broadly and deeply into industrial processes and consumer products, best practices need to be spread, and regulatory regimes adapted**

# Notes on Overview: Policy 4

- Who is responsible when a self-driven car crashes or an intelligent medical device fails?
- How can AI applications be prevented from promulgating racial discrimination or financial cheating?
- Who should reap the gains of efficiencies enabled by AI technologies and what protections should be afforded to people whose skills are rendered obsolete?

# Notes on Overview: Policy 5

- **While the Study Panel does not consider it likely that near-term AI systems will autonomously *choose* to inflict harm on people, it will be possible for people to *use* AI-based systems for harmful as well as helpful purposes**

# Notes on Overview: Policy 6

- A vigorous and informed debate about how to **best steer AI** in ways that enrich our lives and our society, while encouraging creativity in the field, **is an urgent and vital need**
- AI technologies could widen existing inequalities of opportunity if access to them—along with the high-powered computation and large-scale data that fuel many of them—is unfairly distributed across society
  - These technologies will improve the abilities and efficiency of people who have access to them
- Policies should be evaluated as to whether they foster democratic values and equitable sharing of AI's benefits, or concentrate power and benefits in the hands of a fortunate few

# Notes on Overview: Policy 7

- In the coming years, as the public encounters new AI applications in domains such as transportation and healthcare, they must be introduced in ways that build trust and understanding, and respect human and civil rights
- While encouraging innovation, policies and processes should address **ethical, privacy, and security implications**, and should work to ensure that the **benefits of AI technologies will be spread broadly and fairly**
- **Doing so will be critical if Artificial Intelligence research and its applications are to exert a positive influence on North American urban life in 2030 and beyond**