

One Hundred Year Study on Artificial Intelligence

Eric Horvitz

December 2016

<https://ai100.stanford.edu>



Stanford | One Hundred Year Study on Artificial Intelligence

[Home](#)

[About](#)

[People](#)

[Framing](#)

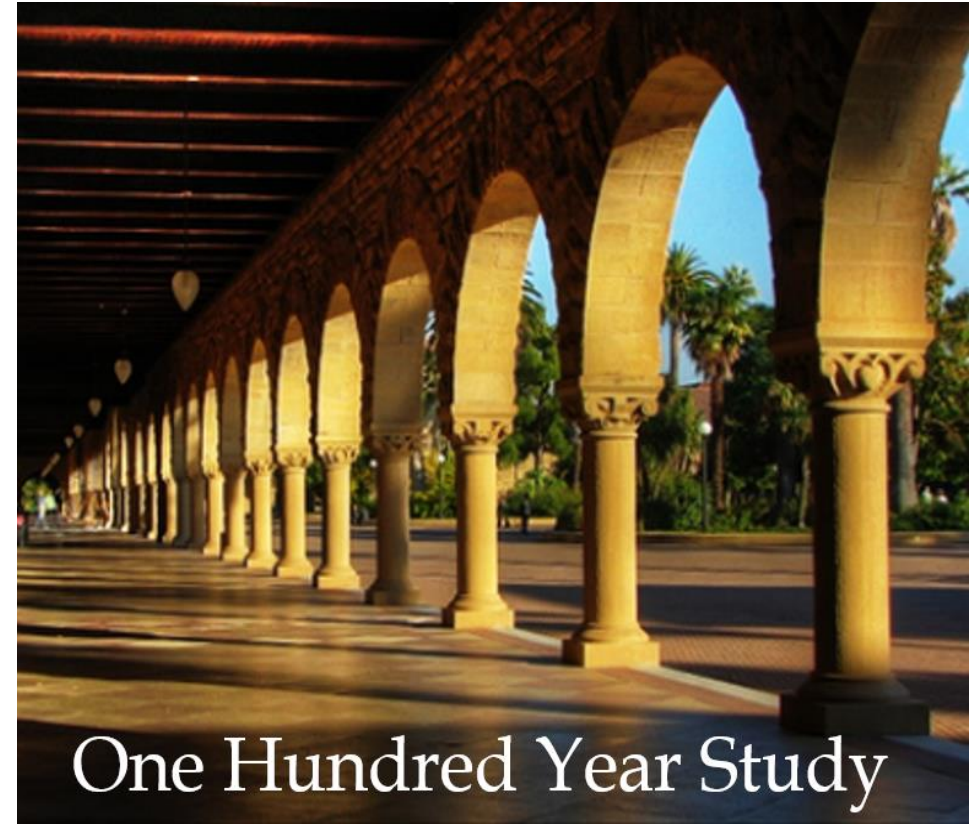


Goal: Create an Enduring Process

Motivation: Difficult to anticipate opportunities & issues ahead with advancement of AI.

Goal: Create ongoing study with long gaze & extended memory aimed at providing insights & guidance

Endowed to continue for as long as Stanford exists.

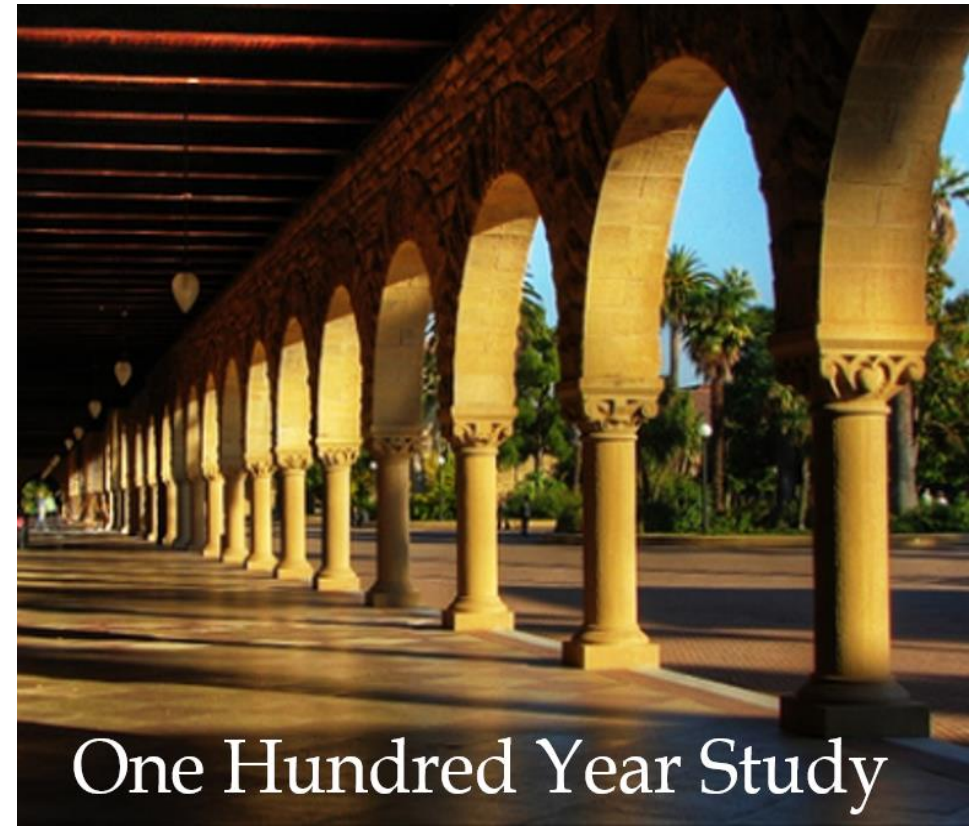


One Hundred Year Study

Intention

“To support a longitudinal study of influences of AI advances on people and society, centering on periodic studies of developments, trends, futures, and potential disruptions associated with the developments in machine intelligence, and on formulating assessments, recommendations, and guidance on proactive efforts.”

-July 2014

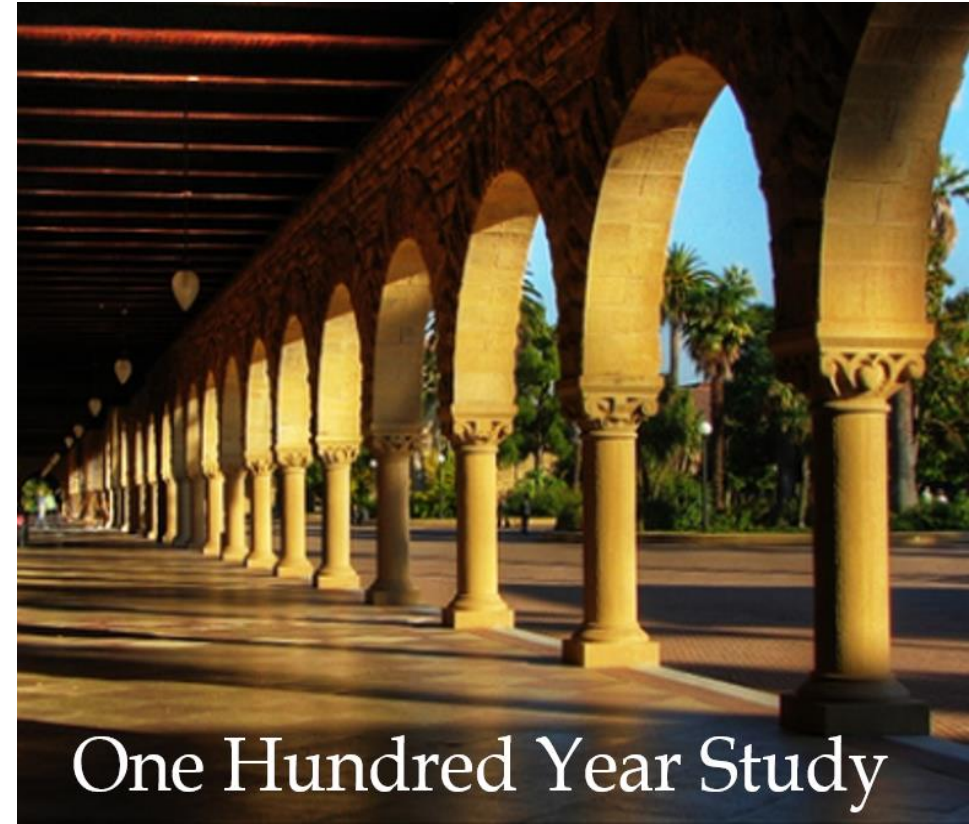


One Hundred Year Study

Intention

“To support a longitudinal study of influences of AI advances on people and society, centering on periodic studies of developments, trends, futures, and potential disruptions associated with the developments in machine intelligence, and on formulating assessments, recommendations, and guidance on proactive efforts.”

-July 2014



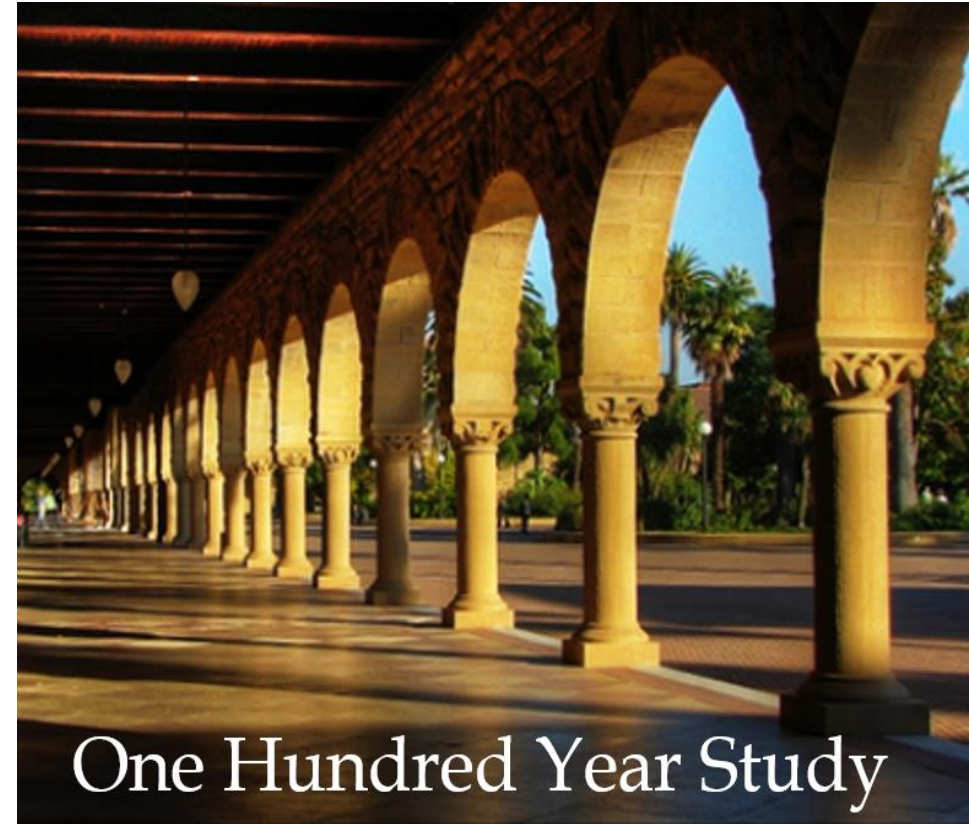
One Hundred Year Study

One Hundred Year Study

“Artificial intelligence is one of the most profound undertakings in science, and one that will affect every aspect of human life.”

“...we feel obliged and qualified to host a conversation about how artificial intelligence will affect our children and our children’s children.”

Stanford President John Hennessy



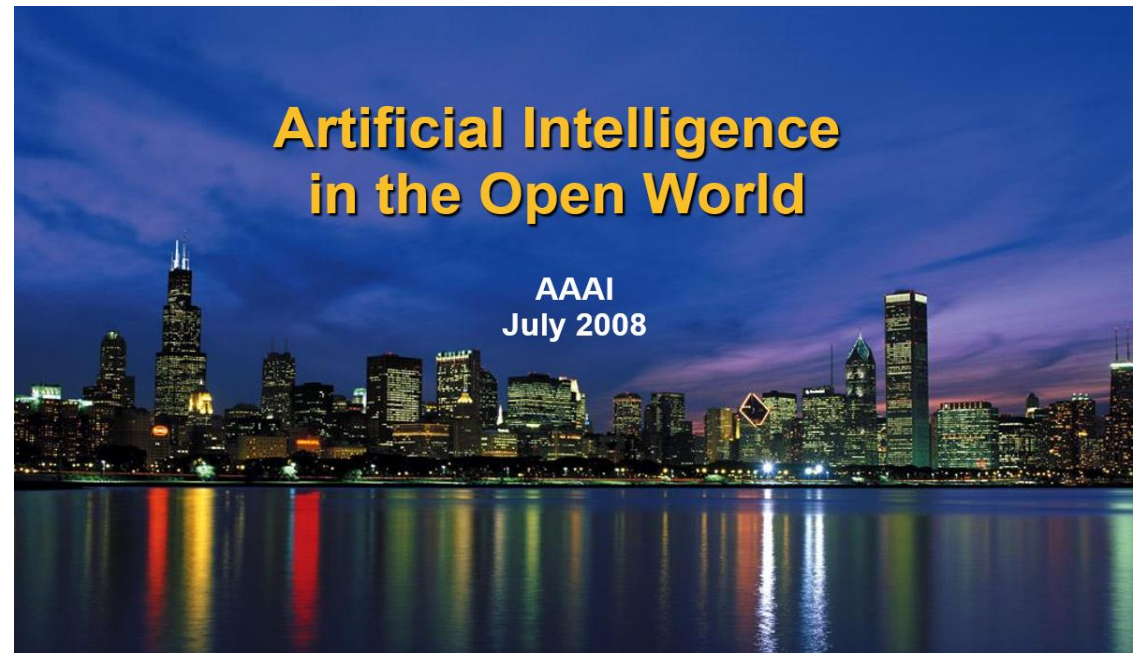
One Hundred Year Study

Background

AAAI Presidency, “AI in the Open World” 2007-2009

Technical challenges with AI systems in open world

Societal influences & challenges



Presidential Panel on Long-Term AI Futures (2008-2009)

“AAAI Asilomar Study”



<http://www.aaai.org/Organization/presidential-panel.php>

Panel on Long-Term AI Futures

Explore potential long-term directions & influences of AI advances, including safety, ethical, and legal issues

Subgroups:

- **Potential Disruptive Advances Over the Short-term**
- **Longer-term Pace, Concerns, Control**
- **Ethical and Legal Challenges**

<http://www.aaai.org/Organization/presidential-panel.php>

Findings

Shorter-term

- Surprises ahead? → Be ready for jump in ML competency
- “Criminal AI”, new attack surfaces, privacy
- Human-machine collaboration critical

Longer-term

- Differences on concerns on rise of superintelligence
- Need to study specification, robustness, control of behavior
- Research guidelines, relevance of containment

Legal & ethical

- Liability for autonomous & semi-autonomous systems
- Learn & represent human values
- Psychosocial issues with human-like experiences

Value in repeating exercise

Stanford | One Hundred Year Study on Artificial Intelligence (AI100)

Search this site...



[Home](#)

[About](#)

[People](#)

[Framing](#)



Framing Memo

<https://ai100.stanford.edu/reflections-and-framing>

Technical trends and surprises

Key opportunities for AI

Delays with moving AI into world

Privacy & machine intelligence

Democracy & freedom

AI advances & law

AI advances & ethics

AI & economics

AI & warfare

Collaborations with machines

AI and human cognition

Criminal uses of AI

Safety & robustness

Loss of control of AI systems

Psychology of people & smart machines

Communication, understanding, outreach

Neuroscience & AI

AI and philosophy of mind

Standing Committee



Barbara Grosz, Chair



Russ Altman



Alan Mackworth



Tom Mitchell



Deirdre Mulligan



Yoav Shoham



Eric Horvitz

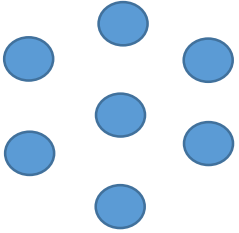
Syncopated Timeline

AAAI Asilomar study

2015



Standing committee



Syncopated Timeline

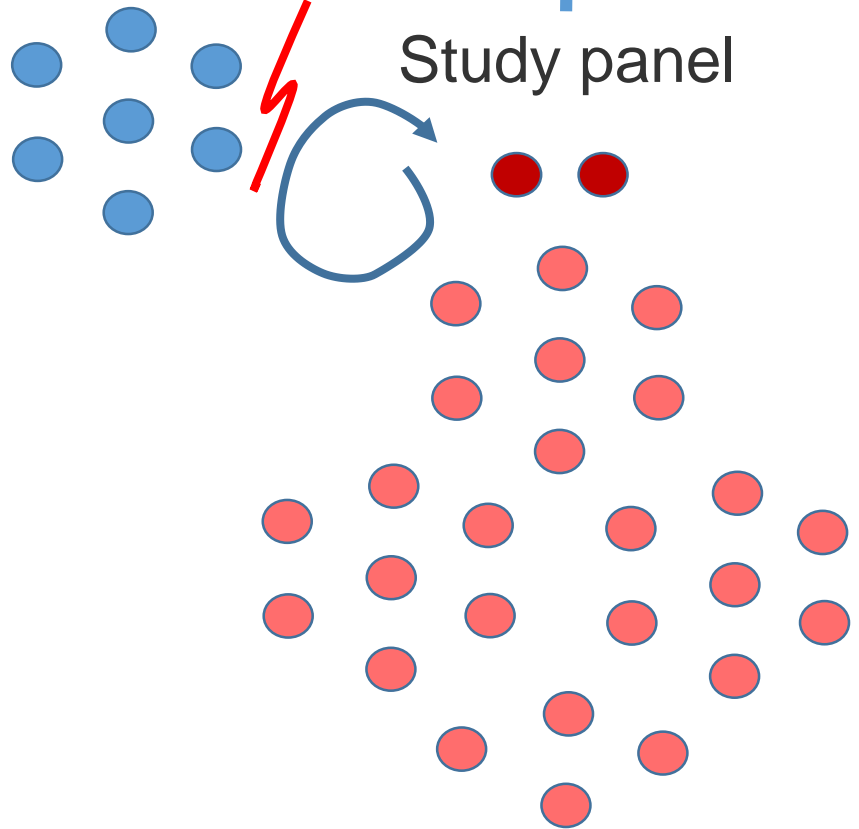
AAAI Asilomar study

2015

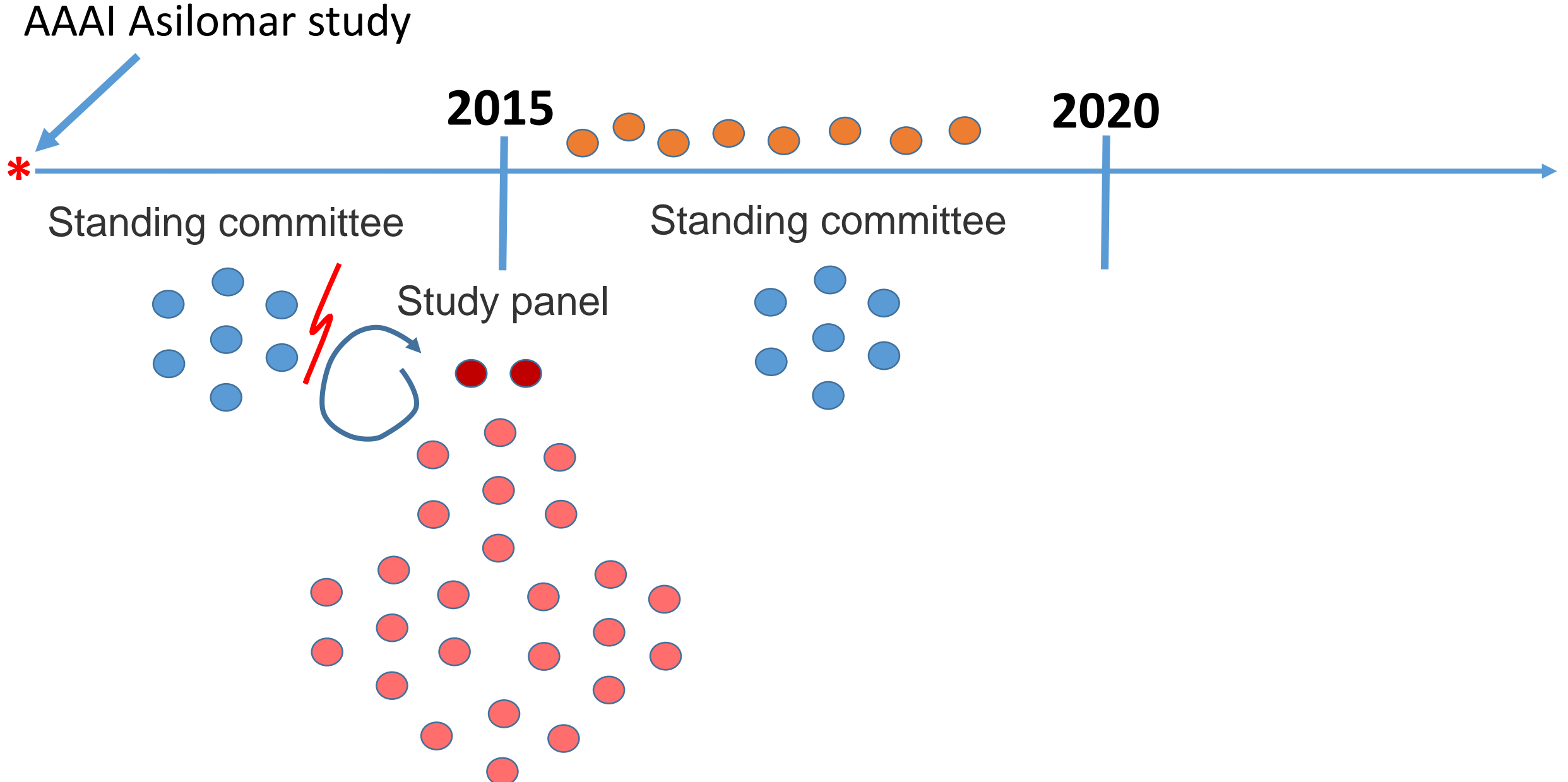


Standing committee

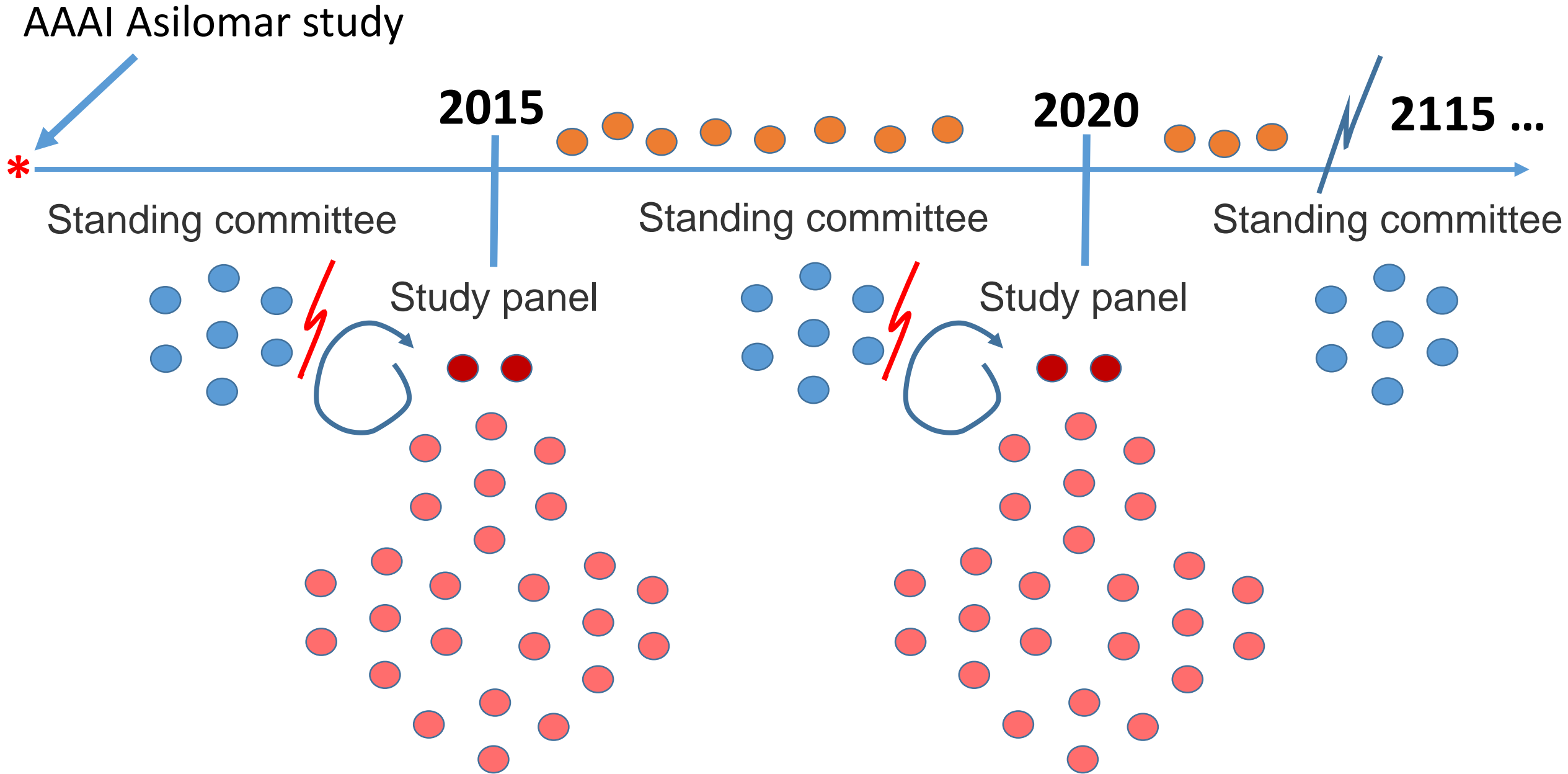
Study panel



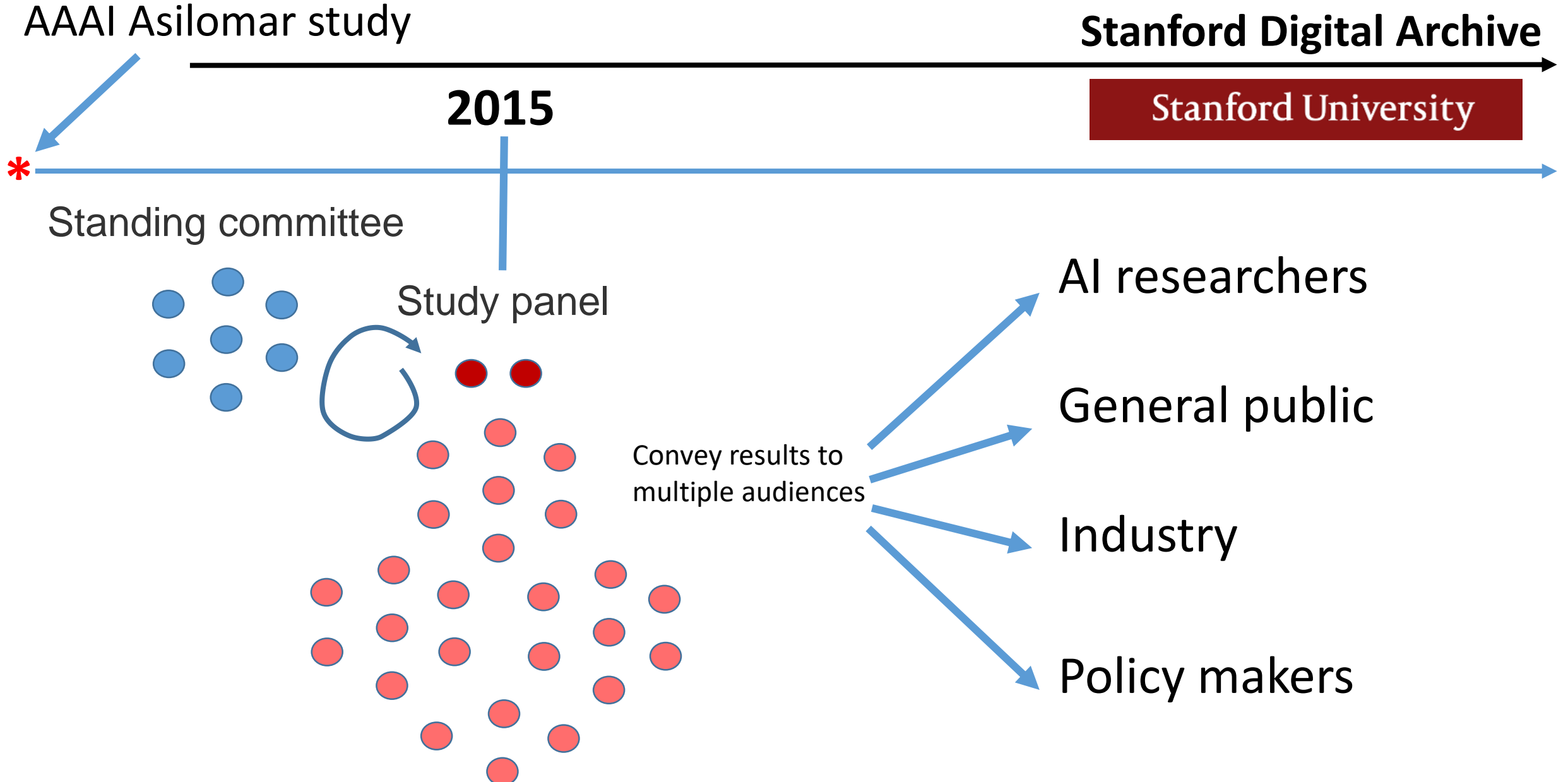
Syncopated Timeline



Syncopated Timeline



Syncopated Timeline



Charge to Study Panel

Artificial Intelligence and Life in 2030

<http://bit.ly/2idS3JC>

Consider AI advances & influences over next 15 years

- Potential influences on daily life
- Proactive efforts on technology, design, policy

Focus: Typical North American city

- Central role of cities in the human experience
- Influences, interdependencies of multiple AI services

Charge t

Artificial Intelligence and

<http://bit.ly/2idS3JC>

Consider AI advances & int

- Potential influences on
- Proactive efforts on te

Focus: Typical North Amer

- Central role of cities in
- Influences, interdepe



One Hundred Year Study on Artificial Intelligence

Charge from the Standing Committee to the 2015 Study Panel

“Artificial Intelligence and Life in 2030”

August 2015

By 2030, advances in the science and practice of artificial intelligence (AI) will likely have significant effect on such diverse aspects of daily life as the nature of work, transportation, healthcare, the environment, urban planning and development, public safety, community engagement, approaches to governance, interpersonal relationships, and self-perception. While many such influences may already be predicted from expected trends in AI research and applications, surprises are also likely—unanticipated new AI competencies as well as unexpected disruptive influences from predicted advances. The goals of the study on *AI and Life in 2030* are to identify possible AI advances over the next 15 years and their potential influences on daily life and to help shape future outcomes for the better proactively through deliberations about the design of systems incorporating these advances and policies.

which have played a central role in the human history and to do so in the 21st century. This study will focus on framing pro

2015-16 Study Panel



Peter Stone, Chair
UT-Austin

Rodney Brooks, Rethink Robotics

Erik Brynjolfsson, MIT

Ryan Calo, University of Washington

Oren Etzioni, Allen Institute for AI

Greg Hager, Johns Hopkins University

Julia Hirschberg, Columbia University

Shivaram Kalyanakrishnan, IIT Bombay

Ece Kamar, Microsoft

Sarit Kraus, Bar Ilan University

Kevin Leyton-Brown, UBC

David Parkes, Harvard

Bill Press, UT Austin

Julie Shah, MIT

Astro Teller, Google[X]

Milind Tambe, USC

AnnaLee Saxenian, Berkeley

AI experts & AI-savvy scholars in law, political science, policy, economics.

ARTIFICIAL INTELLIGENCE AND LIFE IN 2030

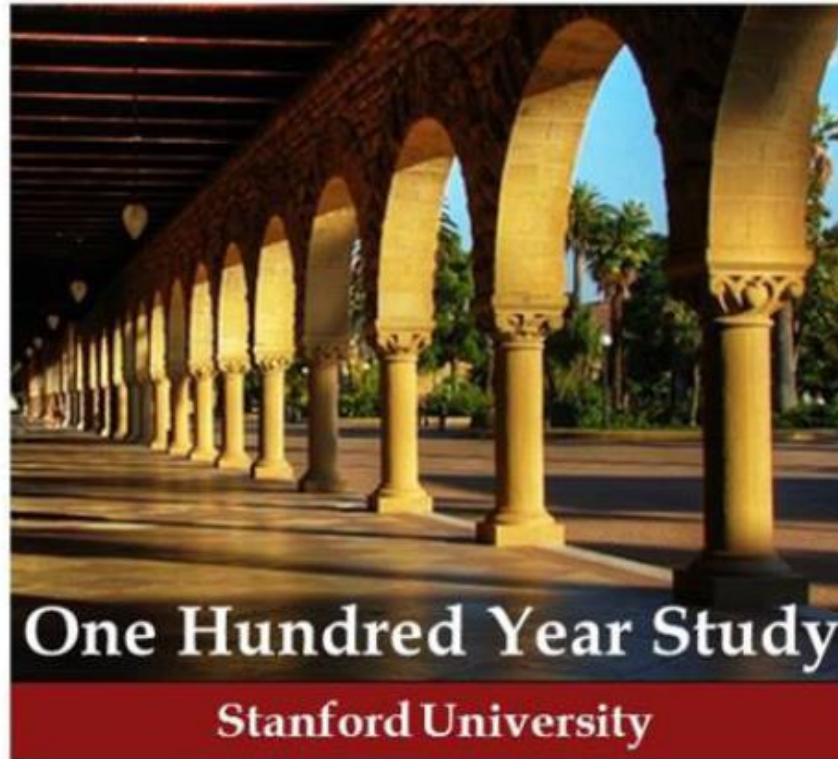
<https://ai100.stanford.edu/2016-report>

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

PREFACE

The One Hundred Year Study on Artificial Intelligence, launched in the fall of 2014, is a long-term investigation of the field of Artificial Intelligence (AI) and its influences on people, their communities, and society. It considers the science, engineering, and deployment of AI-enabled computing systems. As its core activity, the Standing Committee that oversees the One Hundred Year Study forms a Study Panel every five years to assess the current state of AI. The Study Panel reviews AI's progress in the years following the immediately prior report, envisions the potential advances that lie ahead, and describes the technical and societal challenges and opportunities these advances raise, including in such arenas as ethics, economics, and the design of systems compatible with human cognition. The overarching purpose of the One Hundred Year Study's periodic expert review is to provide a collected and connected set of reflections about AI and its influences as the field advances. The studies are expected to develop syntheses and assessments that provide expert-informed guidance for directions in AI research, development, and systems design, as well as programs and policies to help ensure that these systems broadly benefit individuals and society.¹

The One Hundred Year Study is modeled on an earlier effort informally known as



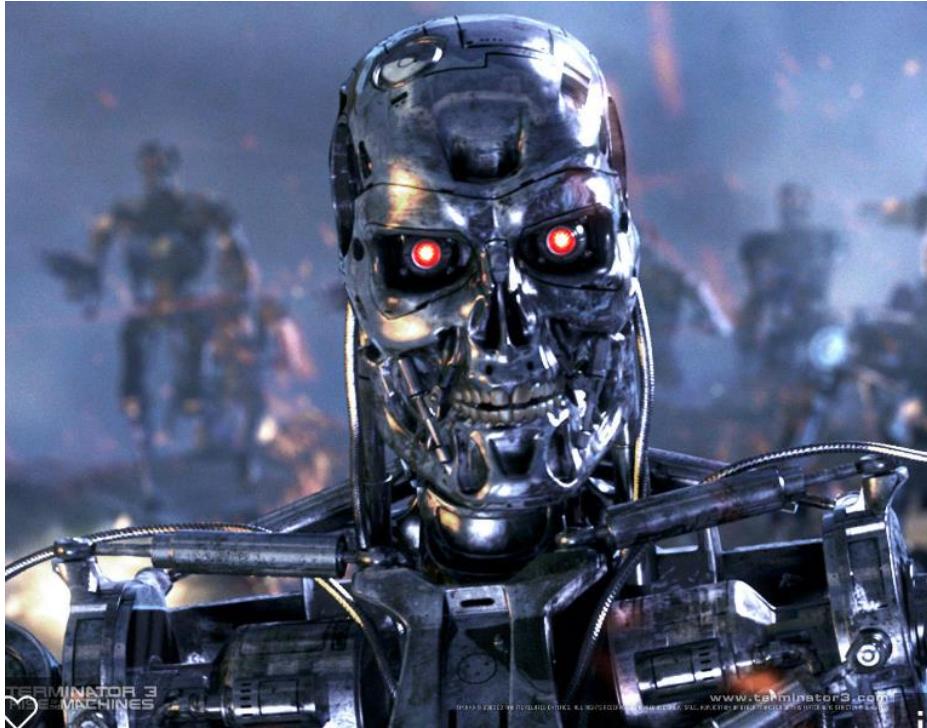
**The overarching purpose
of the One Hundred Year
Study's periodic expert**



Eight areas of focus

- Transportation
- Service robots
- Healthcare
- Education
- Low-resource communities
- Public safety and security
- Employment and workplace
- Entertainment

On Concerns with “Rise of the Machines”



BBC Sign in News Sport Weather Shop Earth

NEWS

Out of control AI will not kill us, believes Microsoft Research chief

© 28 January 2015 | Technology



Alex Garland's film *Ex Machina* is one of several movies to consider the threat posed by AI

On Concerns with “Rise of the Machines”

“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.”

Need to Engage

Emerging technologies have potential to profoundly transform society & economy for the better by 2030.

Near-term design & policy decisions likely to have long-lasting influences

AI researchers, social scientists, policymakers need to work together to balance technical innovations with mechanisms that ensure that AI's economic & social benefits are broadly shared across society.

Focus of Attention

We are underinvesting resources in studying the societal implications & uses of AI



Inadequate funds for AI research that lacks commercial application

Targeted incentives & funding could help address needs of low-resource communities

e.g., lead poisoning in at-risk children (Flint, MI),
pregnant women at risk for adverse birth outcomes (Illinois IDHS)
HIV reduction among homeless (LA)

Private & public dollars should support interdisciplinary teams

Human—AI Collaboration

Need to increase focus on building systems that can collaborate effectively with people



AI systems become more central, shift to building intelligent systems that are *human-aware* and *trustworthy*

Future engagements with machines will become ever more nuanced, fluid, and personalized.

Directions include scalable ways for people to teach intelligent systems & robots

Transportation

Autonomous transportation will soon be commonplace: cars, trucks, aircraft, drones



Strong influence on public's perception: First experiences with *physically embodied AI*

People will own fewer cars, live further from work, spend time differently, leading to new urban organization

Public transportation could become personal rapid transit using small capacity vehicles to transport people on demand

Healthcare

AI methods promise to change cognitive tasks of clinicians by 2030 if sufficient data and well-targeted systems.

Healthcare field structurally ill-suited to absorb and deploy advances, and held back by regulatory, professional, and commercial obstacles.

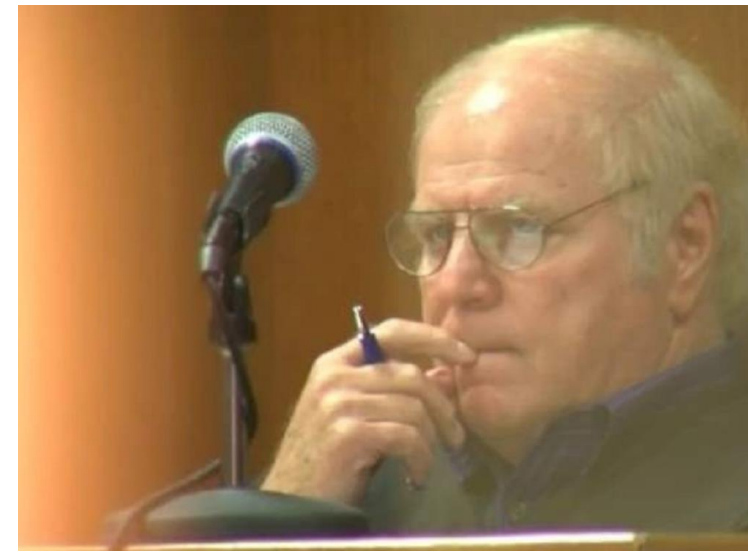
Opportunities to learn from millions of clinical records & scientific literature, to personalize diagnosis & treatment and to create true cognitive assistants.

“Hands-on” experience of physician will be critical, but challenge with integration of human care & automated reasoning

Key area for innovation is elder care: US elderly to grow by 50% over 15 years, home health aides by 40% over 10 years

Governance & Criminal Justice

Cities, federal agencies already deploying AI methods in criminal justice & law enforcement. By 2030, they will rely heavily on them.



Concern: Innocent people may be unjustifiably monitored & targeted

Care must be taken to avoid systematizing human bias, protect civil liberties

AI tools can provide new kinds of transparency to detect, remove, reduce human bias, rather than reinforce it.

Society at crucial juncture per how to deploy AI-based technologies so as to promote rather than hinder democratic values such as freedom, equality, and transparency.

Legal & Ethical

“As AI applications engage in behavior that:

“...were it done by a human, would constitute a crime”

...courts and other legal actors will have to puzzle through whom to hold accountable and on what theory.”

Ethical challenges rising where human injury or death is likely, and split-second choices are made about whom to put at risk.

Legal & Ethical

Attempts to regulate “AI” in general would be misguided as AI isn’t one thing.

Risks and considerations are very different in different domains

Various industries may need distinct, appropriate, regulations

Gov’t will need AI expertise to scrutinize standards & technology developed by the private & public sector, and to craft regulations where necessary

.

Jobs & Economy

AI advances will spur disruptions in how human labor is augmented or replaced by AI, creating challenges.

Near-term: AI will likely replace *tasks* rather than *jobs*

Emerging new jobs harder to imagine than existing jobs that will likely be lost

AI advances will lower cost of goods & services; can make everyone better off

Jobs & Economy

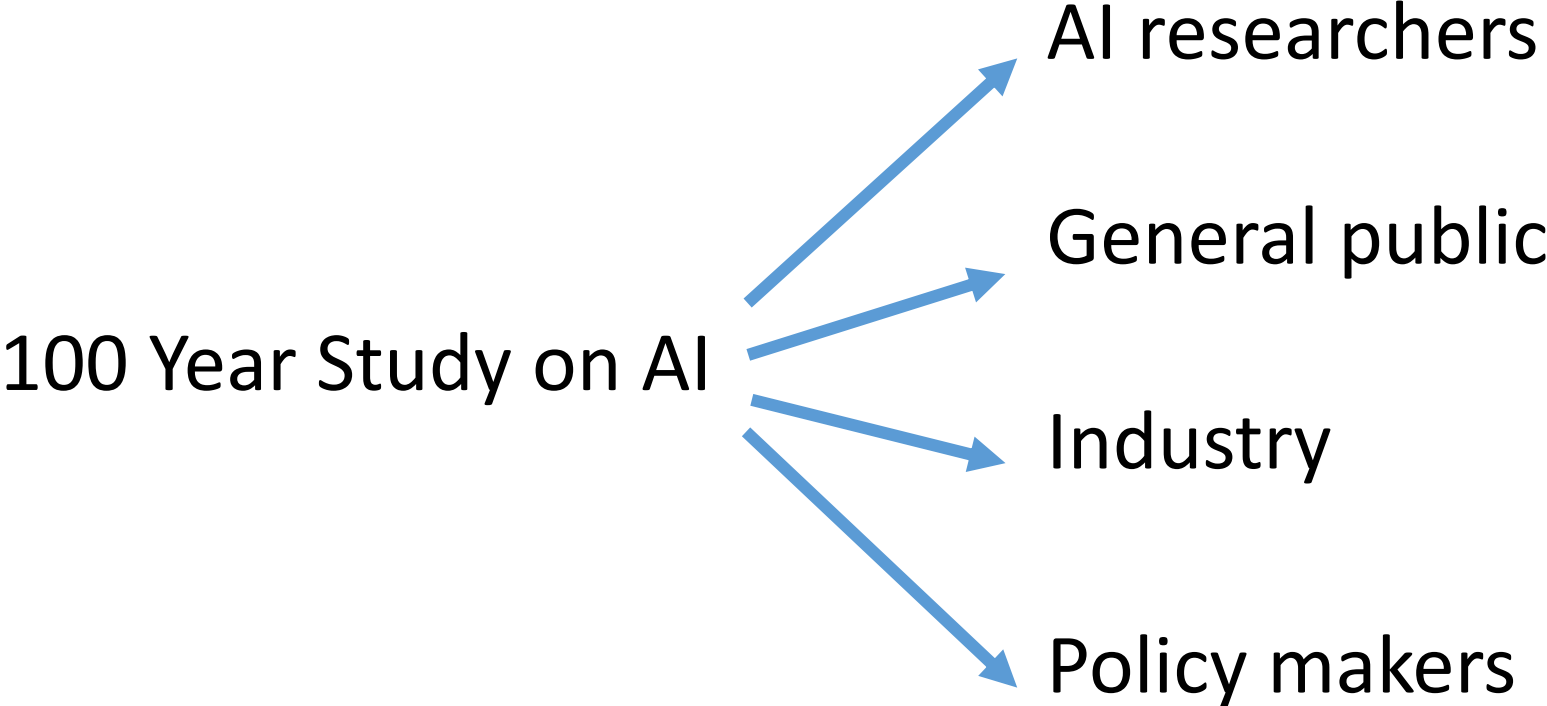
Labor will become less important for production vs. owning intellectual capital

For many, labor may not support desired standard of living

Long-term: AI as radically different mechanism for wealth creation
“where everyone entitled to a portion of AI-produced treasures”

“It is not too soon for social debate on how the economic fruits of AI technologies should be shared.”

Timing of Study & Report



Timing of Study & Report

