

Progress in AI and the Future of Work

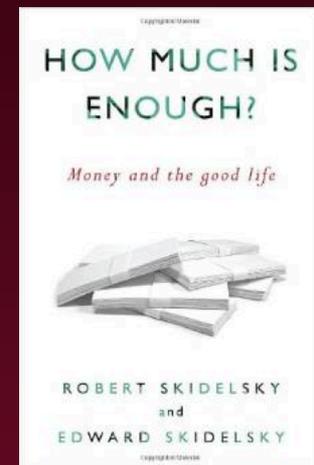
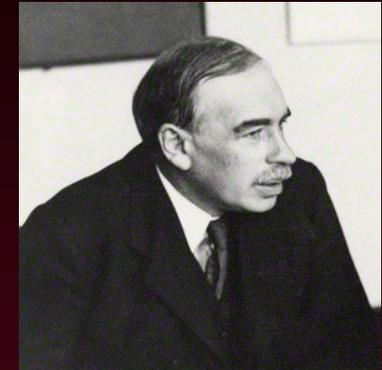
Miles Brundage
Arizona State University

Overview

- **AI and the future of work**
- **Existing theories of AI progress/job risks**
- **Limitations**
- **New approaches**
- **What to do about it?**

“Economic Possibilities for Our Grandchildren”

- Written 85 years ago
- Expected six-fold rise in US per capita GDP
- Expected decline in working hours, rise in leisure by 2030.
- Skidelsky and Skidelsky (2012) on what he got wrong:
 - Wants vs. needs
 - Pressure to work
 - Enjoyment of work



What's Changed Since 1930?

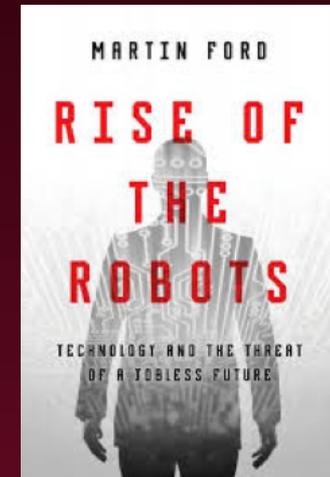
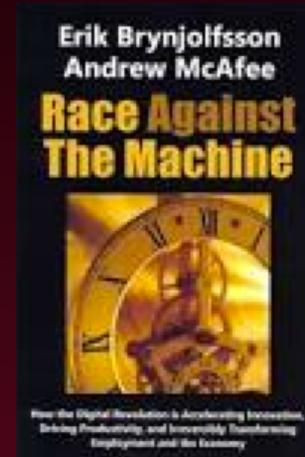
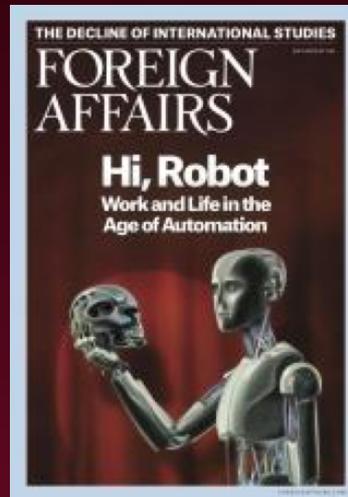
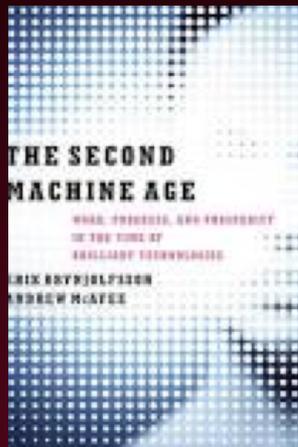
- **Progress in AI/robotics/tech in general**
- **Types of jobs available**
- **More/better education**
- **Wider workforce participation within developed countries**
- **More people involved in global economy**

What Hasn't Changed as Much?

- **Skidelsky and Skidelsky's three factors are still important:**
 - **Wants vs. needs**
 - **Pressure to work**
 - **Enjoyment from work**
- **The terms of the debate on jobs, technology, etc. (Bix 2001)**
- **Very few socio-technologically detailed/plausible scenarios for post-work futures are being seriously debated, let alone implemented**

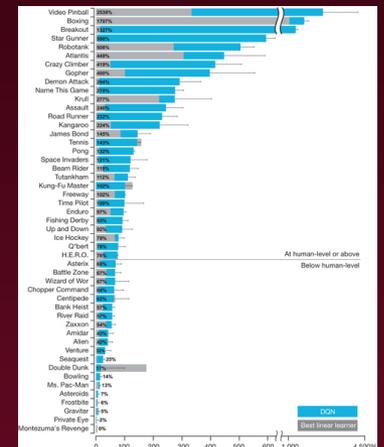
Is This Time Different?

- One key question: AI progress



Views on AI Progress

- **Media, Stuart Russell, etc.:**
 - Deep reinforcement learning (Mnih et al. 2015), other recent “breakthroughs” as significant
- **Alternative view:**
 - Progress = long, incremental slog – deep RL the result of decades of work (Stone 2015)
- **Common view:**
 - More data, computation have been key drivers of recent progress



Some Micro-theories

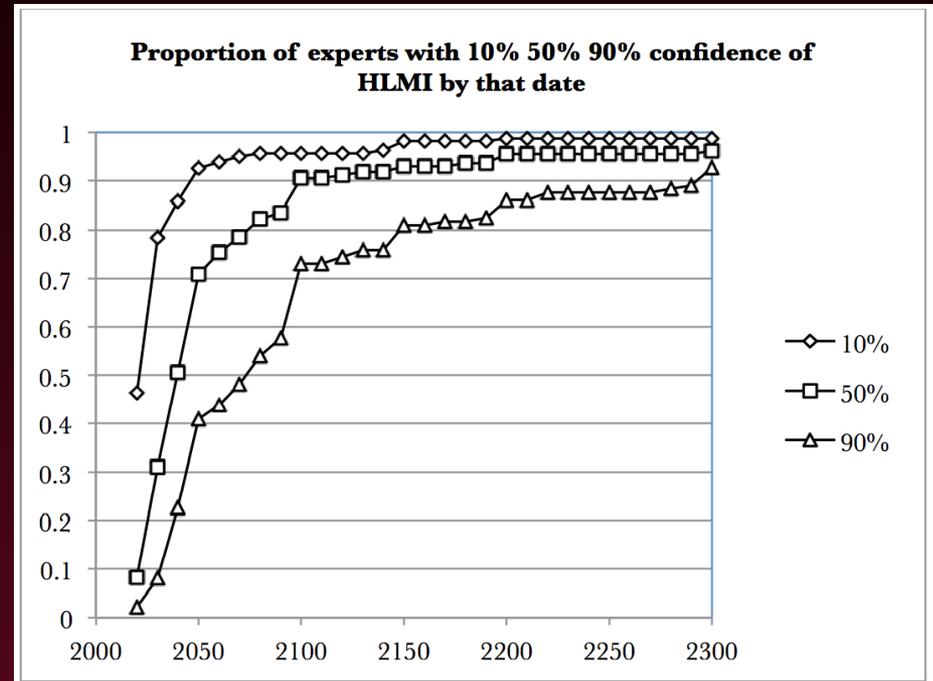
- **Theories of progress in particular sub-fields:**
 - NLP, machine learning, commonsense reasoning, cognitive architecture, planning, etc.
- **Evaluating (aspects of) AI progress:**
 - Visual Turing Test (Geman et al. 2014)
 - Winograd Schema Challenge
 - DARPA Robotics Challenge
 - College entrance exam (Japan)

Some Macro-theories

- **Moravec's paradox**
 - What's easy for us is hard for machines, and vice versa (for evolutionary reasons)
- **X (e.g. integration) as insufficiently worked on**
- **X (e.g. commonsense) as grand challenge**
 - Implies X unlikely to be solved soon
- **Need to learn from humans**

The Pacing Question

- **Human-level AI**
 - Timing?
 - Right framing?
- **“Decade or two”**
 - Frey and Osborne 2013:
 - ~47% of US jobs at “high risk”
 - Knowles-Cutler, Frey, and Osborne 2014:
 - ~35% of UK jobs (note: not directly comparable)



Bostrom and Mueller 2014

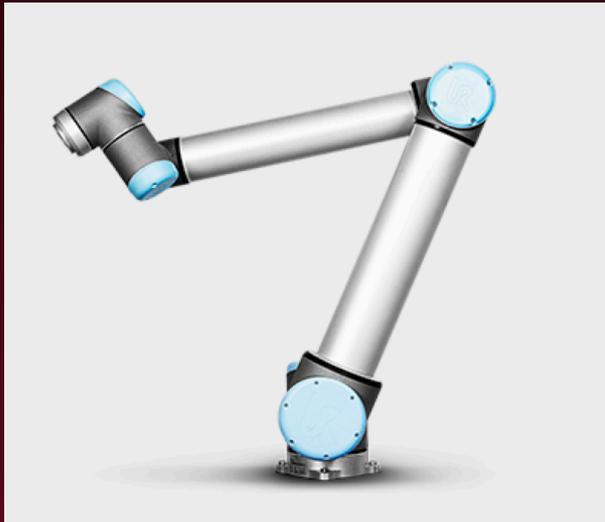
US Robotics Roadmap

- **Excerpt:**

- **“5 years: Low-complexity hands with small numbers of independent joints will be capable of robust whole-hand grasp acquisition.**
- **10years: Medium-complexity hands with ten or more independent joints and novel mechanisms and actuators will be capable of whole-hand grasp acquisition and limited dexterous manipulation.**
- **15years: High-complexity hands with tactile array densities approaching that of humans and with superior dynamic performance, will be capable of robust whole-hand grasp acquisition and dexterous manipulation of objects found in manufacturing environments used by human workers.”**

The Complementarity Question

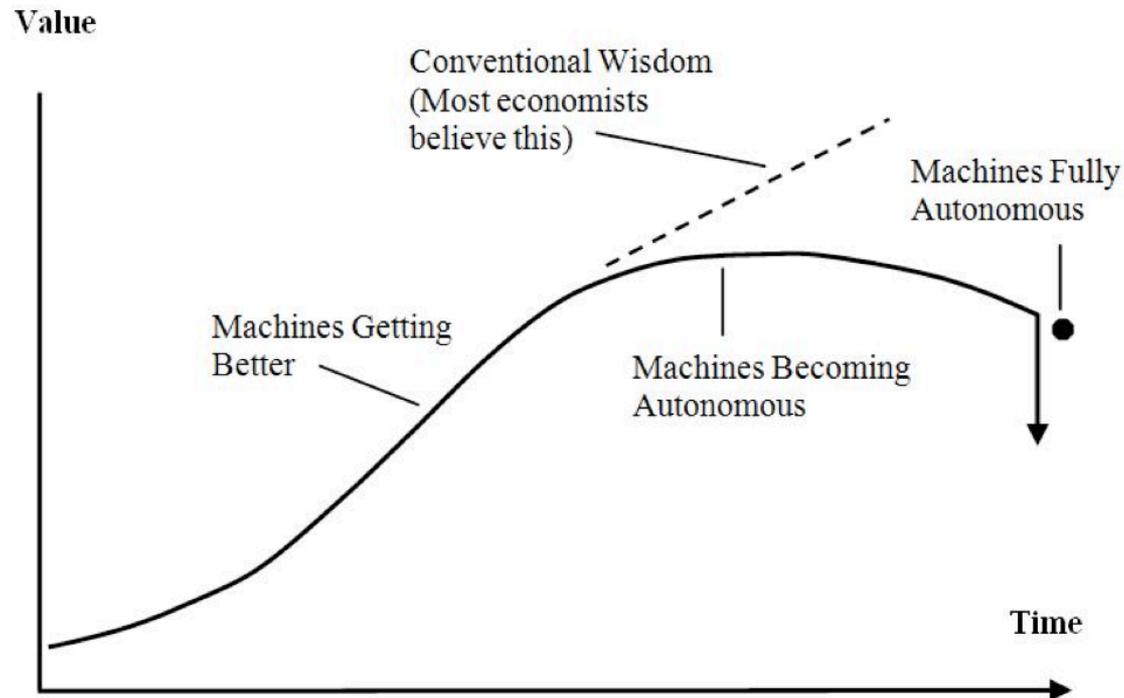
- **Substituting vs. complementing human labor**



Complementarity and Wages

Value Added (Wage) of Average Worker Operating Average Machine

Also: Overall Wealth of Society (GDP per capita will look similar)



Copyright © David J. Atkinson 2012

The Generality Question

- **Some systems touted as more general purpose than others (e.g. Baxter, Watson as teachable, deep RL agents able to learn many games...)**
- **Raises issues of training cost/time, brittleness, diffusion scale, order of impact...**

The Skills Question

- **Tasks vs. jobs (Autor 2013)**
 - Task bundling
- **Hollowing out of middle skill tasks**
 - Polarization: “lovely” and “lousy” jobs (Goos and Manning 2009)

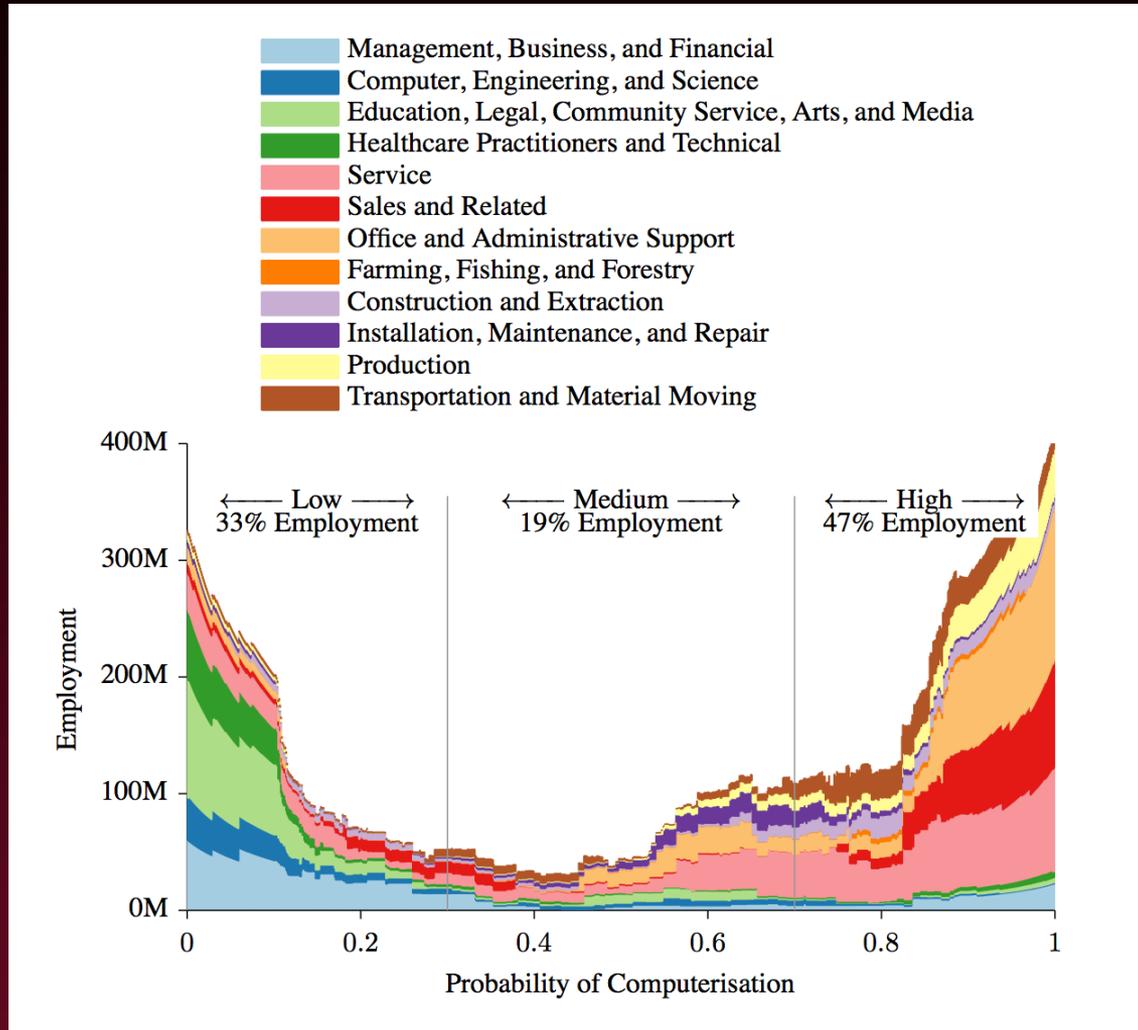
Supply and Demand

- **Human social interaction in product/service purchases – how highly valued?**
- **Cost of human labor alternative (e.g. Japanese vs. Indian car manufacturing)**

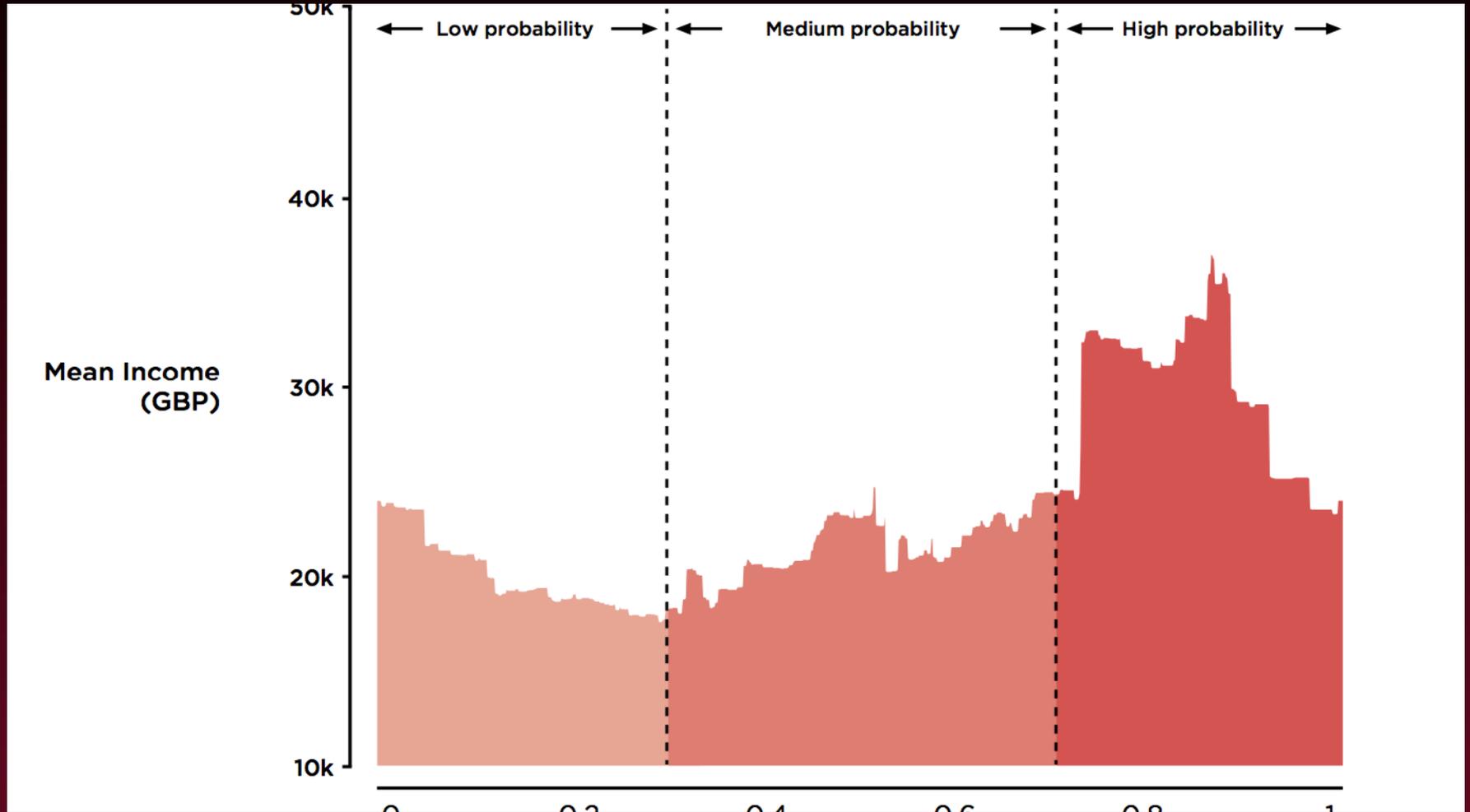
Theories of AI progress/jobs (partial list)

- **Murnane/Levy 2004:**
 - Routine vs. non-routine
- **Frey and Osborne 2013:**
 - Social intelligence
 - Creative intelligence
 - Perception and manipulation
- **Autor 2013:**
 - Novelty of tasks
- **Brynjolfsson/McAfee 2014:**
 - Creativity
- **Rus 2015:**
 - Perception/manipulation – it depends
 - Abstraction, creativity

Frey and Osborne 2013: US Employment Risk



Creativity requirement of jobs and income (Bakhshi, Frey, and Osborne 2015)



Problems with Existing Approaches

- **Static re: supply and demand**
- **Subjective in terms of:**
 - **AI state of the art/future**
 - **Job characteristics**
- **Ignoring diffusion issues**
 - **One state of the art**
- **Ignoring roadmaps**

Questions for the Future

- **Progress in:**
 - **social interaction**
 - **Perception and dexterity**
 - **Creativity**
 - **Generality**
 - **?**
- **Consumer demand**
- **Diffusion**
- **New job creation**

Towards a Theory of AI Progress

- **Philosophy of science**
 - Progress through incorporation and reduction, progress as problem-solving ability, etc.
- **Evolutionary models of science and technology**
- **Cognitive science, cognitive psych., developmental psych., technological forecasting, etc. potentially helpful**

Sketch of an AI progress theory

- **Multi-scale a la CHC theory (McGrew 2009)**
 - **General intelligence**
 - **Broad competencies**
 - **Tasks**
- **Multiple metrics**
- **Human contribution considered**
- **Exogenous factors (hardware, data) considered**
- **Environmental complexity considered**

Tracking and Anticipating AI Progress

- **One approach:**
 - **Leverage the AI literature**
 - Trends, metrics
 - Grants
- **More diverse, systematic testing**
 - **Narrow and integrated systems**
 - **Consider human dimension**
 - **State of the art --> states of the art**

Leveraging Models of AI Progress (and Diffusion)

- **Scenarios**
 - **Socio-technical possibilities/options in different AI progress scenarios**
- **Policy**
 - **Education**
 - **Education for complementarity?**
 - **Refocusing of education's aims?**
 - **Economic**
 - **Welfare**
 - **Guaranteed employment**
 - **R+D**
 - **Relative priorities**

The AI Community and the Future of Work

- **Division of labor**
 - AI community vs. policy-makers, funding agencies, consumers
- **Anticipation and public/policy-maker engagement – 15-20 year timeframe key?**
- **Applications**

Responsible AI Innovation and the Future of Work (cont'd)

- **Patent approaches**
 - Recent Google patents – dropout, deep RL, robot personalities
 - Ethical licensing (Cooper 2013)
- **Research choices**
 - Substitution/complementarity?
 - Pros/cons depend on broader social context, application domain

Conclusions

- **Many open research questions**
 - Supply
 - Demand
- **Untapped resources**
 - Roadmaps
 - AI literature
- **Big open social/ethical questions**

Thanks!

- Miles.brundage@asu.edu
- www.milesbrundage.com