Draft March 29, 2017

Tom Kalil

AI for Good: Maximizing the economic and societal benefits of AI

Summary: Advances in AI will have a broad and profound impact on our economy and our society. An overarching strategy for the responsible and beneficial development of AI has at least six components: (1) advancing the state-of-the-art of the technology; (2) understanding and pro-actively addressing risks and ethical dimensions (e.g. AI safety, AI and cybersecurity, transparency and fairness of algorithmic decision-making); (3) expanding the workforce needed to both develop and use AI; (4) supporting activities such as scenario planning, given the inherent uncertainty of future impacts of AI, especially over the medium and long-term; (5) engaging the public and other key stakeholders in meaningful discussions and collaborations; and (6) identifying mechanisms to advance the economic and societal benefits of AI.

This paper discusses the building blocks of a strategy to maximize the economic and societal benefits of AI, with a focus on identifying and pursuing compelling, ambitious and achievable goals. This approach is inspired by other examples of government and private sector participation in the pursuit of ambitious goals (some of which have not yet been achieved) including:

- Vice President Biden's Cancer Moonshot. Make 10 years of progress in the fight against cancer in the next 5 years.

- DOE's SunShot. Make solar energy as affordable as coal by the end of the decade.

- BRAIN Initiative: Determine how the brain encodes and processes information by developing the tools needed to study the brain (and large-scale neural circuits) in action.

- Qualcomm. Develop a portable wireless device that can diagnose 13 health conditions as accurately as a board-certified physician.

- Elon Musk/SpaceX, "I want to die on Mars but not on impact."

- IBM. Beat Gary Kasparov at chess and Ken Jennings at Jeopardy.

- Microsoft. Eliminate language as a barrier to human collaboration and communication.

- Google: Dramatic reduction in the 1.4 million deaths from traffic accidents with self-driving cars.

- UCLA. Develop the technologies, policies and strategies to make LA County 100 percent sustainable by 2050, as measured by generation and use of clean energy, use of local water, and enhanced ecosystem health.

An actionable, challenge-based strategy for advancing AI would provide answers to the following questions:

1) What is the challenge and how should success be measured?

2) What are actions that could be taken to make these challenges achievable?

3) Who are the actors who can enable progress?

The remainder of this paper addresses each of these questions in turn.

Examples of Possible AI Grand Challenges

Below are some examples of potential goals and projects that are relevant to these goals. This list is illustrative, as opposed to comprehensive. Additional work is required to determine whether the specific goals strike the right balance between ambition and achievability, and what the appropriate metrics are.

AI and Workforce: Increase the wages of non-college educated workers (or unemployed/under-employed veteran's) by $10,000 in 6 months or less, by enabling them to master a skill that is a ticket to a middle class job.[1]

- DARPA-funded "Education Dominance" program created a digital tutor that is allowing new Navy recruits with a high school degree to outperform Navy IT technicians with 7-10 years of experience, using both written exams and ability to solve real-world trouble tickets.

- The impact of AI for accelerated training would be increased with other types of innovation. For example, firms could collaborate to identify critical skills, sponsor the development of competency-based assessments that are accurate predictors of on-the-job-performance, share these assessments with training providers, and embrace hiring on the basis of skills as opposed to credentials. Training providers could offer a "pay for success" model, where they are paid on the basis of an increase in future earnings of a worker.

- **AI and Workforce:** Allow teams of American workers and machines to beat the world price for a broad range of manufactured products, enabling the "onshoring" of manufacturing.

[1] Of course, if advanced training technologies dramatically increased the supply of workers with a given skill, it might reduce the wages for workers with this skill.

- **AI and Infrastructure:** Increase the capacity of America's roadways by 50% while reducing the number of accidents and injuries due to traffic accidents.

- **AI and Education:** Develop intelligent toys that increase the school readiness (e.g. executive function, vocabulary size) of children from low-income families. Longer-term, provide the equivalent of Neal Stephenson's "Young Lady's Illustrated Primer" to every child.[2]

- **AI and Education:** Increase the performance of students in a core academic subject such as middle school math and science by the equivalent of 1-2 letter grades.

- **AI and Healthcare:** Reduce error rates in medical diagnostics by 80 percent.

    - Harvard researchers have met this metric for metastatic breast cancer in the 2016 Camelyon Grand Challenge. See https://camelyon16.grand-challenge.org/results/ and https://blogs.nvidia.com/blog/2016/09/19/deep-learning-breast-cancer-diagnosis/

    - There is a near team opportunity to digitize and annotate millions of pathology slides by collaborating with the Defense Department's Joint Pathology Center, which has 55 million pathology glass slides and 32 million tissue samples. See https://obamawhitehouse.archives.gov/the-press-office/2016/10/17/fact-sheet-vice-president-biden-delivers-cancer-moonshot-report for a commitment from DOD to explore this opportunity, including "public-private partnerships."

- **AI and Healthcare:** Identify X medical conditions where it is possible to both improve health outcomes and lower costs by at least $5-$10 billion.

    - Aetna Innovation Labs and GNS Healthcare have teamed up to use predictive analytics to identify patients with the largest risk of getting type II diabetes, and to effectively target programs to prevent diabetes. See http://www.gnshealthcare.com/ajmc-publishes-results-showing-big-data-analytics-can-predict-risk-of-metabolic-syndrome/

- **AI and Energy:** Significantly increase the energy efficiency of residential and commercial buildings and industrial systems.

    - Deep Mind has reduced the energy used for the cooling of data centers by up to 40 percent. See https://deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-40/

[2] One important caveat is that interaction between adults and young children is important for early language acquisition. See http://thirtymillionwords.org/

- **AI and Energy:** Double the capacity of the grid to use intermittent sources of energy such as wind and solar with no reduction in grid reliability

    - Colorado's Xcel Energy has installed more wind than any other utility and has changed its policy positions on renewable energy because of a collaboration with the National Center for Atmospheric Research, which uses data, AI and high-performance computing to increase the accuracy of wind power forecasts. See https://www.technologyreview.com/s/526541/smart-wind-and-solar-power/

- **AI and Public Good:** Create an infrastructure for continuous improvement in health, education and other critical services using low-cost experimentation and learning systems.

- **AI and Science**: Dramatically increase the productivity of scientists and engineers in both basic and applied R&D – e.g. reduce by 50 percent the time and cost needed to discover a new drug or a new material through machine reading of the scientific literature, robotic experimentation, automated or semi-automated knowledge extraction from large, disparate sources of data, etc.

    - The DARPA "Big Mechanism" project seeks to "read research abstracts and papers to extract pieces of causal mechanisms, assemble these pieces into more complete causal models, and reason over these models to produce explanations." See http://www.darpa.mil/program/big-mechanism

- **AI for People with Disabilities:** Increase the quality of life and independence of people with disabilities with automatic translation of digital content and services, allowing blind/low-vision individuals to navigate the physical world, etc.

    - A variety of organizations (Department of Education, companies, etc.) are supporting the development of a Global Public Inclusive Infrastructure, with the goal of using the cloud and platform services to making online information and services available for people facing accessibility barriers. See https://ischool.umd.edu/news/ischool-receives-18-million-grant-dept-education-disability-innovation-fund

- **AI and Aging:** Increase the capacity of senior citizens to live independently in "smart" homes by at least 5 years.

- **AI and Agriculture:** Increase the "crop per drop" or reduce the use of fertilizer and pesticides in agriculture by using AI and data from sensors and satellites.

How can these goals be achieved?

Identifying concrete, mutually reinforcing and high-impact public and private actions that can be taken will increase the chances that these goals and other goals can be met. For many AI applications (e.g. AI for the financial sector, AI for advertising), market forces will be sufficient to drive the development, validation and deployment of the technology. However, in other cases, a more intentional effort may be required to (a) identify the public and private actions that will increase the chances that these goals will be met; (b) secure credible commitments from relevant organizations; (c) monitor and evaluate progress over time; and (d) build momentum by celebrating success and expanding the coalition of individuals and organizations that are contributing to the goal.

There is a taxonomy of the public and private actions that could be taken to promote the definition, development, evaluation, and widespread use of "AI for Good" applications. Some examples include:

1. *Increase the availability of data that can be used to train and test AI systems*

   Data availability is a defining aspect of many AI applications. In the same way that the ImageNet data set accelerated progress in image recognition, organizations could either make available training data that already exists, or support the collection of new data sets.

   For example, the National Geospatial-Intelligence Agency, Digital Globe, In-Q-Tel, Amazon and Nvidia are collaborating on SpaceNet. The project's data sets will eventually include photos of 0.5 million square kilometers of Earth to enable automated interpretation of high-resolution satellite imagery. See https://aws.amazon.com/public-datasets/spacenet/ and https://medium.com/the-downlinq

2. *Support the use of "market-shaping approaches" for AI-enabled applications*

   Given that some applications of AI (as with most digital goods) have high fixed costs and low marginal costs, there should be more experimentation with models that allow companies to recover or defray their costs for investments that have high social returns and low or uncertain private returns. This could have a large impact, since the marginal cost of allowing more people to benefit from an AI-enabled application could be very low. This economics (high fixed costs, low marginal costs) has allowed people to download free courses from institutions such as MIT and Khan Academy hundreds of millions of times.

   The global health community has developed a set of "market shaping" tools that combine the best of public purpose (health technologies for the poor) and private sector resources and expertise (R&D, product development, marketing and distribution.) See https://www.usaid.gov/cii/market-shaping-primer

For example, five countries and the Gates Foundations told Pfizer and GSK that if they could develop a vaccine for diseases of the poor, they would be guaranteed a market. This one intervention (a so-called Advance Market Commitment) will save the lives of 7 million poor children over the next 20 years from diseases such as meningitis and bacterial pneumonia. Other market shaping approaches include incentive prizes, milestone payments, volume guarantees, and pooled procurement.

Some of these approaches are being used for IT applications. An example is the $7 million Adult Literacy X Prize, which seeks to increase the literacy level of low-literate adults within 12 months using IT.

Philanthropists could play an important role by supporting one or more university-based centers of excellence, which could suggest how market-shaping approaches might be used to increase private sector investment in "AI for Good" applications.

3. *Support rigorous, independent evaluations*

   Organizations could sponsor rigorous evaluation of AI-enabled solutions to determine whether they are effective, and if so, whether they are cost-effective relative to other interventions as measured by outcome per dollar. It is also important to define metrics for evaluating progress, like the Message Understanding Conferences (MUC) for evaluating progress on message extraction or the Text Retrieval Conference (TREC) for evaluating progress on text retrieval

4. *Support multidisciplinary R&D that brings together AI researchers, domain experts and users*

   Most of the challenges listed above require transdisciplinary teams engaged over long periods of time, and willing to adopt a multi-disciplinary mind-set. Some of this work could be done at universities, which would also have the advantage of exposing the "next generation" to new ways of thinking and create a workforce that has the capacity to work at the intersection of AI and a particular application of AI.

5. *Invest in infrastructure or test sites that can enable real-world experimentation with AI technologies in a controlled environment*.

   This is being done to allow experimentation with the safe integration of drones in the National Air Space.

6. *Create development methodologies and regulatory or approval pathways that facilitate rapid translation of promising approaches from prototype to deployment at scale.*

   Many of the challenges listed above are limited by institutional, organizational, or regulatory barriers. For example, promising healthcare tools may take years to put into

practice due to barriers in access to data, barriers to deployment, barriers to testing and revision, and market barriers to wider dissemination.

7. *Support "regulatory science" or the development of performance-based standards to support regulatory decision-making*.

   FDA is partnering with the private sector and university researchers to develop and validate biomarkers that can serve as "surrogate endpoints" for purposes of drug approval. This can reduce the time and cost associated with clinical trials. FDA has called the research that is needed to support science-based regulatory decision-making "regulatory science." This approach could be useful applications of AI, such as self-driving cars. See http://www.fda.gov/ScienceResearch/SpecialTopics/RegulatoryScience/default.htm?utm_campaign=Goo

Who can help advance this agenda?

Some of the challenges listed above have a ready market, and will be shaped in part by market forces. However, many highly beneficial AI applications (e.g. the development of AI-based systems that that are specifically designed to meet the needs of low-income families and communities, applications of AI that address market failures or require investments in what economists call "public goods") will require more collaboration between different sectors and actors (e.g. companies, foundations and philanthropists, non-profits, impact investors, government agencies).

Companies may need to view their participation in the development of these applications through the lens of their corporate citizenship efforts, and as activities that improve their reputation and employee engagement. For example, leading AI companies have recently formed the *Partnership on AI to benefit people and society*. A potential next step for this partnership would be for each of the member companies to take the lead on the development of at least one "AI for Good" application, such as those described above. Each firm could tap the creativity of their employees and external partners to identify an ambitious economic or societal goal that AI could help achieve. Firms could play a leadership role in achieving a given goal by:

- Increasing funding for internal R&D, pilot projects, or university-based research;
- Investing in startups that are pursuing this goal through corporate venture arms; and
- Making available training data (companies like Microsoft and Facebook are doing this through organizations such as the Common Visual Data Foundation).

Other areas of AI will require different types of support or investment, and may require more intentional collaboration between government, industry, academia, and civil society. The private sector and the research community can play an important role in maximizing the societal "upside" of AI by setting ambitious goals, and identifying the public and private actions needed

to achieve these goals. Federal agencies and philanthropic organizations can play a leadership role in catalyzing participation around specific objectives, using their resources as leverage to maximize impact of research and development. For example, companies or foundations may offer “enhancement funding” to federal grants, allowing a research project to gain additional resources to develop a specific application using new AI techniques.

Finally, public outreach and solicitation could play a large role in advancing AI for Good. As noted at the outset, public education and information can be used to highlight the potential for improvements in quality of life using AI. There are already many successful examples of crowd-sourced funding for research ventures and commercial ventures. By creating concrete, visible and valued AI applications, it may be possible to augment existing efforts with small-donor support for important initiatives.