Experiences and Insights for Collaborative Industry–Academic Research in Artificial Intelligence

Lisa Amini, Ching-Hua Chen, David Cox, Aude Oliva, Antonio Torralba

■ *The factors that define and influence the success of industry–academic research in artificial intelligence have evolved significantly in the last decade. In this article, we consider what success means from both sides of a collaboration and offer our perspectives on how to approach the opportunities and challenges that come with achieving success. These perspectives are grounded on the recent and significant investments that have been made between IBM and several higher education institutions around the world, including IBM's Artificial Intelligence Horizons Network, the Massachusetts Institute of Technology–IBM Watson Artificial Intelligence Lab, and the Massachusetts Institute of Technology Quest for Intelligence.*

Imagine for a moment that you are part of a small leadership team for the Research division of a global, quarter-billion-plus employee company, whose existence depends critically on research to invent and innovate the future of science and technology. Suppose the business of that company is to create the computer software and hardware on which industries — from healthcare, to supply chain, to transportation, to financial services, to local, state, and national governments — depend, and in turn, their customers, businesses, and families depend. Now, imagine

  ISSN 0738-4602

a technology anticipated to change, disrupt, reimagine, or at least touch every aspect of how all of those businesses are conducted. That technology is artificial intelligence (AI). Analysts forecast it will generate an additional $3.5 to 5.8 trillion in value annually across these industries (Chui et al. 2018) and a staggering 37 percent compound annual growth rate in market size for the foreseeable future.[1] The race for talent, breakthroughs, and market share is on, and the stakes are high.

The pressures facing academe from this most recent AI renaissance are no less stark than those facing industry. According to the National Science Foundation (2018), the share of new US computer science PhD graduates taking industry jobs has risen from 38 percent to 57 percent over the last decade.[2] Many top research institutions are also raising concerns over industry luring experienced AI researchers from academe (Loizos 2016; Kahn 2018; Metz 2018). In 2015, Uber hired 40 researchers and technical engineers from Carnegie Mellon University's robotics laboratory to staff a self-driving car operation in Pittsburgh. In 2018, Facebook hired five top university professors from Carnegie Mellon University, University of Washington, University of Oxford, and the University of California, Berkeley, to set up or strengthen offices in Pittsburgh, London, Seattle, and Menlo Park. This is at a time when student demand for machine learning courses is seeing unprecedented highs. At the Massachusetts Institute of Technology (MIT), 700 students signed up for the Fall 2017 course 6.036, Introduction to Machine Learning (Kirsner 2017). Stanford University reported over a thousand enrollees for a similar course.

The AI resurgence in both academe and industry is also evidenced in conference submission and attendance statistics. As shown in figure 1, Neural Information Processing Systems (NeurIPS, formerly NIPS) conference attendance has dramatically increased in recent years. In 2017, NeurIPS registration sold out within 2 weeks, and in 2018, in under 12 minutes. NeurIPS also illustrates the exploding diversity of AI subtopics and the intensity of academic and corporate participation. In 2017,[3] there were: 156 subtopics, $1.76 million in corporate sponsorships, and three of the top 12 paper-producing institutions were corporate (Google, Microsoft, IBM). Overall, academic institutions still topped the NeurIPS paper leaderboard, with the top three being Carnegie Mellon University, MIT, and Stanford. Moreover, analysis of NeurIPS submissions (Allen 2017) shows that many of the papers have mixed academic–corporate authorship. The AI Index[4] illustrates the dramatically rising AI trends are pervasive, with a nine-time increase since 1996 in the number of computer science papers published and tagged with the keyword *artificial intelligence* in the Scopus database of academic papers.

With this tremendous uptick in investment of time, expertise, and money in AI research by academe and industry, it is important to ask the following questions about academic–industry collaborations: What aspects of the current industry–academic collaboration models are working well? What are the incentives for each party, and how well are they met? What are the specific challenges in meeting the growing demands for AI research? How might we sustain investment and progress in AI, knowing that achieving our loftiest goals for the field of AI could take decades?

In this article, we offer our recent thought processes and first-hand experiences in shaping academic–industry collaboration for our respective institutions in this new age of AI. The authors' current responsibilities include a mix of independent and joint leadership roles for major new models of academic–industry collaborations, including IBM's AI Horizons Network (AIHN),[5] the MIT–IBM Watson AI Lab,[6] and The MIT Quest for Intelligence.[7]

Answering the questions we posed for all of industry and academe is beyond the scope of this article, but we hope our reflections on these three initiatives will provide a context in which others in both academe and industry might better identify and evaluate their own objectives. We believe a rethinking of possible approaches to maximize and sustain the scientific and technological impact from the current and unprecedented investment in AI research could hold vast benefits for society at large. Our goal is not to promote a single model, but instead to provide the building blocks for designing models to suit a variety of circumstances. Additionally, because our perspective cannot possibly be all-encompassing, we hope that by sharing it, we can spawn a dialog that the entire community can participate in and benefit from.

Traditional Academic–Industry Research Collaborations

The field of AI has a rich and impactful tradition of academic–industry collaboration. Indeed, the now famous 1956 Dartmouth Conference that started the field of AI was proposed by researchers from Dartmouth, Harvard, IBM, and Bell Labs (McCarthy et al. 1955, 2006). In 1980, Carnegie Mellon University created an expert system called XCON for Digital Equipment Corporation. XCON was estimated to save Digital Equipment Corporation 40 million dollars annually by 1986 (Crevier 1993). The winner of the Defense Advanced Research Projects Agency 2005 Autonomous Driving Grand Challenge was a team led by Stanford University that also included experts from Volkswagen and Intel (Thrun et al. 2006). The Semantic Web, with inventors from MIT, University of Maryland, and Nokia, went well beyond a single industry–academic collaboration to create a framework that continues to enable widespread, direct and indirect, industry and academic collaboration to make data from the World Wide Web directly analyzable by machines (Berners-Lee, Hendler, and Lassila 2001). Many of the collaborative methods leveraged

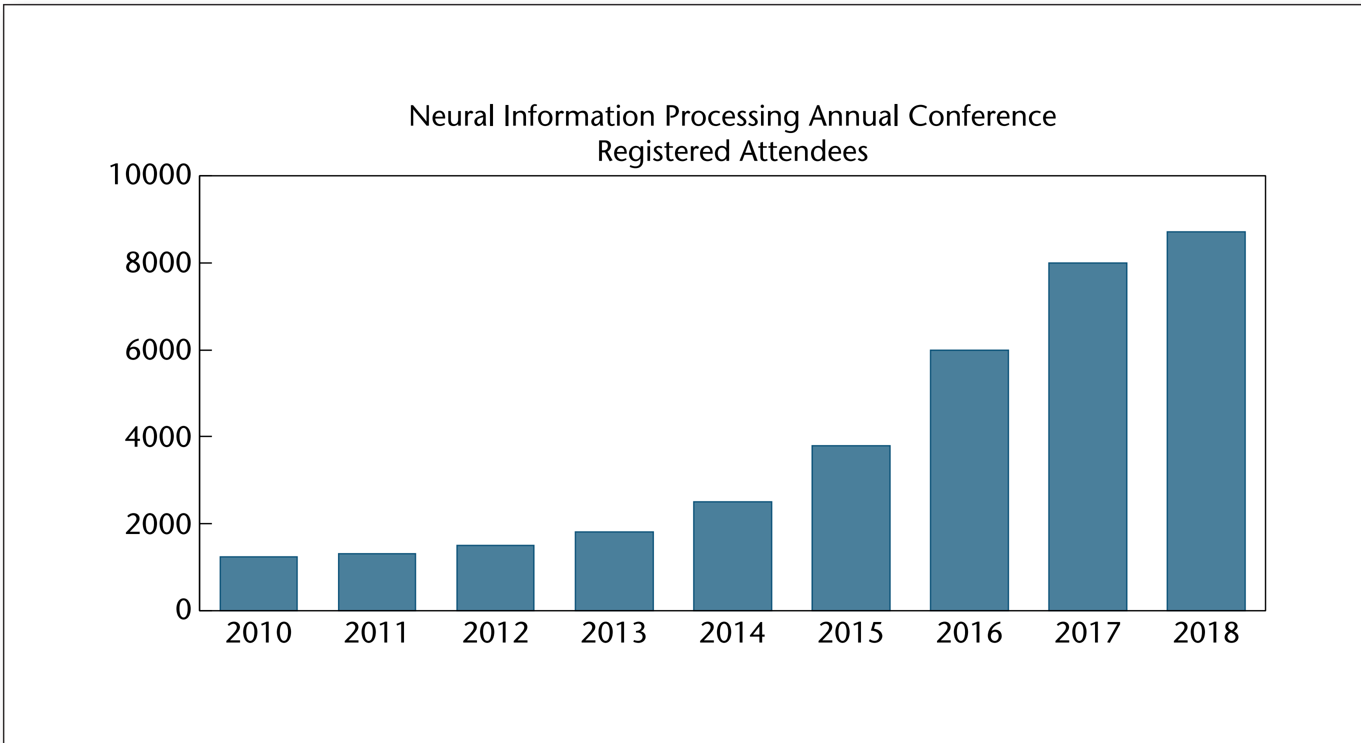

Figure 1. Attendance at the Neural Information Processing Annual Conference exemplifies the growth in academic and industry AI research.

The 2018 conference sold out in under 12 minutes.

by such projects have remained consistent through the decades, such as openly published and reproducible experiments, benchmarks, Grand Challenges, and in-market experimentation.

A key challenge in conducting research in any setting is determining how to measure success. In the industry setting, success is primarily measured by the impact resulting from efforts and investment. Nearly all research results, including publications, patents, and software, can be considered in terms of impact. For example, the impact of publications can be assessed in terms of the number and quality of publications, the quality and impact factor of the publishing venue, the citations from the scientific and technical community, and whether inflection points representative of a seminal work eventually result. This approach requires tracking publication impact over the course of months, years, and even decades. Patent impact can be measured in licensing revenue and freedom of action for products. In the field of Computer Science, results are often embodied in computer software, for which impact can be measured by the level of adoption internally, in for-profit products, openly accessible software, and the developer and user communities of all. To assess impact, commercial institutions may evaluate whether the research contribution, often small in terms of the total lines of code or investment in bringing a product or service to market, is the differentiating feature that drives a step change in adoption and revenue. Our perspective is shaped primarily from experiences in our respective institutions, and through formal collaborations and informal conversations with researchers from other institutions.

Commercial institutions leverage a variety of approaches for collaborating with and funding academic research, as summarized in table 1. Some programs focus on directly sponsoring basic research or engaging faculty expertise in industry projects, such as faculty awards and industry-directed research, respectively. Others, such as consortia, enable leveraged investment alongside other industry sponsors. Similarly, funding may sponsor the development of emerging research talent through PhD fellowships and internships. Visiting scientist roles facilitate cross-pollination through sabbaticals of faculty in industry laboratories and industry researchers in academic settings.

Industry research laboratories with significant AI programs such as Google, Facebook, Microsoft, Amazon, and IBM all have programs in most or all of the categories of table 1, but the specifics vary. For example, some faculty or fellowship awards are completely open, while others target particular geographical regions or researcher demographics. One clear trend is in the significant investment in internship programs in terms of the number of summer interns (in the thousands), competitive salaries, and

Program Type	Purpose of Funding Program
Faculty Award	Sponsor basic research and foster university collaborations in an area of interest
Industry-Directed Research	Engage faculty expertise in an industry-led research program
Visiting Scientist	Sponsor faculty sabbatical in a commercial research laboratory for increased in-person collaborations. Industry researcher sabbaticals may also be sponsored in academic laboratories
Collaborative Open Source	Support academe creating code or data for engaging scientific and developer communities of industry interest
Academic Consortium	Academe-led initiative in which multiple industries can cosponsor research in a given research area
PhD Fellowships	Recognize and invest in student research toward a PhD degree
Internships	Fund students to work in a commercial laboratory to gain industry research experience

Table 1. Traditional Funding Programs for Industry-Sponsored Academic Research.

perks such as flexible hours, access to top leaders, signature events, and outings.

The following seven topics highlight industry incentives for funding such academic collaborations, with complementary academic incentives interwoven.

Agile Access to Deep Expertise

Industry research institutions must often double-down in the most relevant areas and projects, and shifts may be abrupt. Working with academic institutions enables industry to augment research teams with experts in targeted fields. Such projects conversely provide academic researchers with experience and insights from solving challenges with industry transformative potential.

Interdisciplinary Innovation

Academic collaborations provide access to a substantially broader set of disciplines and perspectives, honed over decades of focus on that area.

Hiring Pipeline

Connections with top faculty provide improved visibility into the hiring pipeline of students with fresh ideas and expertise. By working with students before graduation through internships, fellowships, and co-advising, industry can help prepare students with expertise needed for industrial research. From a faculty perspective, placing a student in a laboratory with the potential for future collaborations, funding, and joint publications is highly attractive.

Broad Influence

Industry influence in the scientific community can be built by releasing open data and AI challenges, such as ImageNet (Deng et al. 2009), the Moments in Time dataset (Monfort et al. 2018), the DREAM Challenges,[8] and the AI2 Reasoning Challenge (Clark et al. 2018). In addition to helping to engage a larger community by enabling researchers to objectively evaluate competing approaches on common data, such benchmark data sets give students insights into the data and challenges of importance to industry. This industrial context can help accelerate the student's future integration into industrial research laboratories.

Noncompetitive Peer Research

University collaborations provide a means for industry researchers to collaborate and freely exchange ideas, without competitive issues. Such outside perspectives can be invaluable to both the industry and academic researchers.

Retention

Many industry researchers enjoy collaborating with and mentoring students. Many also enjoy teaching and hold adjunct positions with nearby universities. Supporting such interests helps industry researchers to grow their research profile in a variety of ways.

Risk Management

Although most industry–academic partnership funding goes to the areas industry is confident is strategic to their business, it is also a way for industry and academic partners to share the risk associated with exploring nascent ideas, whose business relevance is still uncertain.

These mechanisms and incentives as described remain relevant to industry–academic collaborations today. However, we also believe we are in the midst of profound changes to how AI research should be organized to better address current AI pressures, both within institutions and in collaborations with others.

New Challenges for AI Research

The field of AI is in a period of rapid and large-scale growth, innovation, and emerging industry adoption. In terms of the future of AI, we believe that narrow AI, which characterizes current machine learning-based deployments of AI, faces major limitations such as requiring large amounts of curated training data and careful, human-guided training to reliably perform a single or limited set of tasks. We similarly believe that broad AI, which moves beyond current limitations to tackle complex problems, to adapt robustly in changing environments, and to transfer learning to new domains, is critical to sustaining large-scale progress and industry investment in AI. Finally, we believe that artificial general intelligence, when a machine can successfully perform any intellectual task a human can, is still decades away.

Thus, a successful industry–academic collaborative model must include long-term sustainable industry investment in academic research capable of both advancing science and impacting technology adoption. The following subsections highlight pressures the current AI context is placing on industry–academic collaborations.

Industry Expectations of AI

Industry anticipation for productivity increases and new business opportunities from industrial applications of AI is high and gaining momentum. Industry analysts report that AI-based features will be pervasive in nearly all new applications by 2020 (Elliot and Andrews 2017). As industry applications and deployments rapidly emerge, the pressure for fundamental advances is also mounting. The briskly evolving frontier of AI is not always cleanly separable into basic, applied, and use-inspired research. Tackling an industry challenge with AI may start with applied research, but algorithmic limitations inspire basic and use-inspired research. Breakthroughs in basic and use-inspired research can open new doors for applied research.

For example, consider AI systems for question answering in domains such as medicine, manufacturing, finance, law, and logistics. The text targeted by machine learning algorithms may be in relatively small numbers of domain-specific documents or company-specific manuals. The vocabulary and semantics can be highly specialized and used only the domain- or company-specific documents. In some cases, vocabulary is defined specifically and precisely within the document itself. With today's narrow AI, examples of such content must often be painstakingly labeled by humans to train the machines. The labeling can be expensive because domain expertise is required to understand such documents and to reason about intent. Even within a single domain (such as law), there are many subdomains (such as regulatory compliance, criminal law, real estate law), with potentially little overlap in essential vocabulary. Vocabulary used in different contexts and settings imply different or nuanced meanings, for which background knowledge is required to understand, generalize, or make inferences. Finally, many industrial applications of AI have low error tolerance because business decisions are being made or safety critical operations are being undertaken.

Breakthroughs in basic and use-inspired research can help address these challenges by creating algorithms that reduce or eliminate the need for human labeling and other limitations in scaling to industry demands. For example, transfer learning focuses on algorithms and representations that could be learned in data-rich environments and then adapted to domains that are similar but have limited data and labels. Multitask learning targets algorithms and representations that learn from multiple tasks, while exploiting the commonalities and differences across tasks for improved accuracy. Algorithms for learning inferentially productive representations, such as learning causal structure from observational data, are needed. Improved representations are sought to better facilitate other objectives such as machine-generated explanations and transfer learning. New methods are needed to ensure AI is robust enough to undertake safety critical applications, even in the face of adversarial attacks, and to avoid shortcomings such as bias or unfairness, which can lurk in observational data.

Industry researchers are faced with broad and accelerating needs for AI research, technology, and expertise, from within our companies, across our client and collaborators ecosystem, and throughout the scientific and technical community at large. The academic mission of advancing science and educating students is fundamentally different than the corporate mission. This difference can sometimes make it difficult to align objectives between industry and academe. However, there are significant opportunities in working together to create a fluid flow of ideas and results across applied, use-inspired, and basic research in AI.

Academic and Industry Researcher Contexts

Experienced academic and industry researchers operate in different contexts. They share certain attributes, such as deep expertise in their chosen area that comes from years, potentially decades, of focus, failures, and breakthroughs; scientific rigor and principled approaches; and an understanding of the broader context in which technology can be brought into service. However, they also operate in different contexts, experience different career paths, and bring different perspectives.

The experience level of collaborators, their primary research horizon, their strategy, and pivot timing are meant to help contrast the complementary paths and perspectives of industrial and academic researchers, but certainly are not universal.

Experience Level of Collaborators

Academic researchers tend to collaborate more with students, and industry researchers tend to collaborate more with experienced industry researchers and potentially product teams. This can translate into differences in professional maturity of collaborators and units of work. Working with students requires a keener focus on breaking problems into semester- or degree-sized units; the ability to direct those units toward understanding; and in aggregate, as a community, better ability to investigate broader and deeper spaces. Collaborations within teams of experienced industrial researchers enables more easily building upon in-house progress and experience.

Primary Research Horizon

Over a career, an individual researcher may focus primarily on basic research, applied research, use-inspired basic research, or some combination. Traditionally, the path of academic researchers follows the quest for understanding — that is, basic research. In industry, even research directed toward a quest for understanding is often at least influenced by a desire to bring results to market at some point — that is, use-inspired basic research.

Strategy

Industry researchers are immersed in an environment influenced by business strategy, market growth, and competitive differentiation. This can be helpful in learning to understand which projects have the potential for market impact, which breakthroughs may be critical to transitioning a new technology from research to a product or services team and enterprise adoption, and how to inspire corporate sponsorship. Academic researchers often have experience with startups and learn how to pique industry and government interest and investment by helping them to see beyond current pressing problems to a potentially very different future.

Pivot Timing

Academic researchers learn to manage in the context of a very flexible workforce (students), tools (whatever works best for a project), and funding sources. Industry researchers often have the advantage of scale (larger, more experienced workforces, production level tools). Industry researchers attempt to invent at the forefront of technology or science relevant for their company, so changes are a balance of personal passion, long-term research agenda, and market or business forces.

These differences in context and incentives can make it difficult to create and maintain long-term industry–academic collaborations. A key goal in developing an effective alliance model for AI research is to leverage these differences as complementary strengths.

Accelerated Timeline for Disseminating Results and Success Metrics

The field of AI abounds with new and rapidly expanding subfields. The traditional means for disseminating research results through scientific journals and conferences cannot keep pace with today's AI field. Many AI researchers now publish their results on the open arxiv.org forum, in parallel with submitting to peer-reviewed venues. Accelerated timelines for disseminating research can help to drive the field more quickly, but can also present challenges in industrial settings, where articles may require vetting for confidential information prior to submission.

The accelerated timeline for disseminating research results also has an impact on traditional measures of success for research. For example, traditional research success metrics include the number, impact factor, and citations of publications, albeit with many shortcomings as a success metric (Shema 2013). The accelerated pace and sheer volume of publications in AI present new challenges for using publication attributes as a measure of success.

Moreover, the lack of sufficient details regarding the software and training conditions for key research results have made it difficult for AI researchers to reproduce prior results. A recent article characterized the reproducibility issue in AI as a crisis (Hutson 2018).

The field of AI does, however, have the advantage of targeting automation of intelligence tasks at or above human-proficiency. That is, targeted tasks can be precisely specified, with metrics of success that can be objectively evaluated, on data that can be open for the community to reproduce and build upon results. Well-defined tasks, with objective, verifiable measures of success for AI's near-term challenges and ambitious objectives can offer valuable metrics for tracking progress in AI, and can also provide a more open and consistent way to communicate research objectives across industry and academe.

Need for Diversity

AI is expected to transform many, if not all, industries and professions. While the authors do not anticipate AI taking on entire professions in coming years, we do anticipate AI will change the future of work, and life, by taking on increasingly complex tasks within a variety of professions (Brynjolfsson and Mitchell 2017). The need for a diversity of perspectives, cultures, disciplines, and expertise to participate in creating technology that will directly affect so many lives is a critical challenge for AI.

Despite the broad implications for AI, the field does not currently have broad representation. For example, recent years have actually seen a decline in the US doctoral graduation rates in computer science for women. According to the Computer Research Association's 2017 Taulbee Survey (Zweben and Bizot 2017), only 18 percent of US doctoral graduates in computer science are women, and the total population of doctoral graduates across all American Indian or Alaskan Native, Black or African American, Hispanic or Multiracial Non-Hispanic, Native Hawaiian, and Pacific Islander is less than 3 percent.

In top AI research conferences, such as NeurIPS 2017, the International Conference on Machine Learning 2017, or the International Conference on Learning

Representations 2018, only 12–17 percent of registered attendees were women and substantially fewer were underrepresented minorities, based on our informal queries into conference demographics. Several conferences with whom we checked do not yet track diversity demographics.

In addition to transforming industries and professions, AI has the potential to transform our built infrastructure, making it more accessible to individuals with health conditions or other physical, sensory, or mental impairments. However, recent reports have shown that smart infrastructures sometimes also hinder access, when potential impairments are not well considered (Hamraie 2018). The United Nations Convention on the Rights of People with Disabilities[9] defines disabilities not as an attribute of an individual, but instead as a concept resulting from a mismatch between the available infrastructure and the needs of an individual. Clearly, broad understanding and diverse perspectives are needed to ensure equal access to the benefits AI can bring.

AI Talent Shortage

Computer science had benefited tremendously over the years from the bidirectional learning that can happen between industry and academe. In the current age of AI and ML, industry can offer very large computer infrastructure needed for training and exploration of novel approaches, extensive data sources, insight from real-world challenges, and funding. Academe offers deep expertise in a diversity of disciplines, an agile and ambitious workforce, and a pipeline for future talent.

However, the enormous demand for AI expertise is also attracting faculty and pre-degree students away from academe. This reduces the capacity to conduct fundamental research, independent of what may be short-term commercial goals, and the capacity to produce future generations of expertise.

Improvements needed for academic–industry alliance models include better models for leveraging the assets each brings to the field of AI and for amplifying (instead of hollowing out) the pipeline of fundamental and curiosity-driven research.

In summary, the current state of AI research presents many new challenges and opportunities for improved academic–industry alliance models. Key challenges to traditional industry–academic approaches include industry's soaring demand for AI capabilities and talent; the need for effective, concurrent progress across applied, use-inspired, and basic research; sustaining research investments necessary for long-term AI goals; incompatible incentives for academic and industry researchers; a rate and pace of publications at odds with reproducible research results and traditional success metrics; and student and researcher demographics that are not commensurate with broad, transformative potential of AI.

In the following sections, we describe and explain the strategies we are using to address many of the challenges previously described. Although we cannot claim a single model that addresses all of the above challenges, we can highlight approaches we believe are most relevant for promoting short- and long-term AI research success.

Scaling Joint Industry–Academic AI Research in Large, Multiyear Collaborations

Addressing all of the limitations of traditional models of collaboration is out of scope for this article. However, we posit that an essential aspect of addressing them is to scale the volume of high-quality research and its positive influence on the larger industry and academic communities. In this section, we focus on strategies for achieving such scale in large, multiyear collaborations. At a high level, these strategies include creating a dynamic portfolio, measuring progress using AI challenges, collaborating as peers and across disciplines, and sharing intellectual property.

To provide some context for these strategies, the MIT–IBM Watson AI Lab includes 100 full-time equivalent researchers across MIT and IBM, and a portfolio of 70 active research projects (as of September 2019) involving 23 university departments and centers. These projects were selected from among 250 project proposals received over three competitive project proposal rounds.

Creating a Dynamic Portfolio

Large, multiyear programs require a portfolio management strategy that is responsive to the fast pace of today's AI research and achieves continued strong buy-in from all participants. This can be accomplished by instituting periodic (for example, annual) competitive calls for project proposals, and a process for peer review and project selection. A joint steering committee, composed of equal numbers of researchers from the industry and academic institutions, provides a means for peer decision-making on the strategic objectives of the laboratory, the focus areas for the competitive project proposal process, and decisions on which projects are funded to achieve the desired portfolio strategy. A joint steering committee can also provide a peer review process for monitoring the ongoing progress of projects, guidance to project teams and individual researchers, and representation of the overall activity to executive-level stakeholders.

Supporting multiple project durations and orientations can also help with crafting a portfolio that meets the needs of both entities. Exploratory projects may be small scale (for example, 2–3 researchers) and relatively short in duration (for example, 2–3 semesters). Larger, multiyear projects are also needed to support more well-defined research efforts that warrant larger teams (for example, 4–6 researchers) engaged over longer periods (for example, 3 years). Exploratory and larger projects can be independently executed or federated for some larger objective.

Over time, exploratory projects may spawn further exploratory projects or larger multiyear projects. Merging or reforming projects from successful threads of multiple projects can enable tackling larger-scale projects. Managing the vitality of the overall collaboration also requires sustaining a healthy flow of new exploratory projects, discontinuing less successful projects, and adapting project goals based on new discoveries.

Measuring Progress Using AI Challenges

Effectively measuring short-term progress against long-term research goals can be difficult, especially across industry and academic institutions. The AI research community is increasingly organizing challenges in which multiple teams compete to achieve the highest score against a well-specified AI challenge task. AI challenges aim to improve machine performance on a well-defined task to the point of approaching or surpassing human performance, as measured against a predetermined benchmark(s) and metric(s). Typically, the benchmark is anticipated to be surmountable with concentrated effort.

Three examples of projects that meet the criteria of being an AI Challenge include: Use knowledge acquisition and reasoning to solve questions on standardized tests with the goal of achieving state-of-the-art performance on the textbook question answering dataset; apply probabilistic program induction to learn graphics programs (for example, in a language like LaTeX) that can reproduce hand-drawn images; and develop an AI system that can automatically compete at the level of a Master in a Kaggle data science competition, without human intervention or parameter tuning.

The AI challenge approach enables measuring incremental progress by decomposing a long-term goal into a progression of shorter-term challenges, or progressively formulating a longer-term research agenda by incrementally identifying and measuring success against shorter challenges. Thus, well-specified AI tasks with objective performance metrics also provide a natural means to scope projects, measure shorter-term progress against longer-term research goals, and impose meaningful decision points.

Defining an AI challenge task with agreed-upon, quantitative performance metrics enables multiple approaches to be tested and objectively compared by collaborating or independent teams. Metrics of interest include accuracy against the task goals, the resource efficiency of the approach, the scalability in deployment (both human and machine requirements), and reliability.

The AI challenge approach is gaining momentum and is useful for organizing a variety of industry–academic collaborations. Examples include collaborative projects to tackle one or more challenges, time-constrained competitions such as leaderboards or hackathons, and publishing open challenges to the broader community. As part of a larger portfolio, the challenge format is effective for driving focused efforts for addressing already well-defined problems.

Collaborating as Peers and across Disciplines

Industry and academic researchers working as peers — shoulder-to-shoulder on conceiving, formulating, and executing research projects — can also play a major role in effectively scaling AI research. This construct enables jointly setting goals and evaluating results, as opposed to the industry-directed or academic-led consortia, where goals are set by the industry or academic leader of the program. Academic principal investigators gain access to experienced industry researchers, working as part of their team, on problems they view as moving the field forward, potentially co-mentoring students, with the potential to substantially improve the results and publications of their team. Conversely, the industry principal investigators get the mirror effect (academic principal investigators and students working on the industry team's projects).

The value of the peer research approach is not only in the number of researchers assigned to the effort. An even greater value comes from the diversity of perspectives, expertise, experience levels, and even potentially conflicting incentives. Students bring fresh ideas, eagerness to immerse and learn quickly, experiment and take risks, deeply focus on a novel problem or solution, and earn a scientific reputation (and a degree) for themselves. Experienced academic and industry researchers bring deep expertise, experience in driving a project to completion, and a broader understanding of related work and potential applications.

A collaboration model open to disciplines and departments, well beyond those specifically pursuing core AI and machine learning research, can also aid in scaling AI research. An open and inclusive approach is especially critical to bringing novel, cross-disciplinary perspectives to the difficult challenges AI brings. For example, AI has an urgent and critical need for learning causal models from observational data, an area requiring a sound grasp of statistical analysis, principles of identification, and other mainstays of data science. Conversely, automated techniques for learning causal structure could bring powerful new tools to many scientific fields. In another example, information theoretical approaches to understanding information flow in deep neural networks could enable provable bounds on learning efficiency or robustness. AI for ethical decision-making and ethics for AI are additional areas with a deep need for an open and inclusive collaboration model.

Sharing Intellectual Property

The authors are frequently asked: What is the industry value in funding research with an intellectual property model that promotes openly and expeditiously publishing research results and enabling spinouts that may eventually compete with the industry sponsor? While the following position and rationale is heavily influenced by the current context of the authors, we believe it offers insights on why others may also gain advantages from this model.

Specifically, as an enterprise that creates technologies for other enterprises, IBM relies not only on being at the forefront of technology, but also on ensuring there is an ecosystem at this envisioned future. Such an ecosystem requires startups with whom larger institutions can work. It requires many more trained experts than a given institution will hire, who will instead help staff partner and client institutions. It benefits from the technical staff in those client and partner institutions also having access to publications, code, and benchmark data so they can experiment and translate cutting-edge technology for their own business problems, make informed decisions on which technologies are best for their needs, and provide feedback from their experiences.

Additionally, for researchers to be inventing and innovating at the forefront, there must be other talented researchers from across industry and academe also publishing in the same venues, peer-reviewing and providing feedback, and moving the entire field forward together. This is not a new model for industry, as traditional industry–academic alliances have also supported participation in open source, open data, and open publication for decades.

Scaling AI research through academic–industry alliances benefits from an intellectual property model that ensures researchers are not only able, but also are strongly encouraged, to publish their results openly in top quality scientific conferences and journals and more expeditious venues such as arxiv.org. Industry collaborators may require vetting publications for confidential information, but by having industry and academic researchers working side-by-side, the industry researchers can ensure early, incremental vetting that avoids delays in publishing.

The open collaboration model also makes it feasible to provision shared computing infrastructure for joint research, co-development, test, evaluation, and experimentation. The shared infrastructure and research collaboration tools enable code and data sharing within and across teams, and also facilitates asset capture and reproducibility, by enabling teams to leverage one another's code and data repositories.

In summary, we believe advances in AI require not only scientific rigor and disciplined approaches, but also agility, incubation, relentless entrepreneurial spirit, and an innovative hack culture.

Practical Successes and Challenges

The goal of this section is to provide concrete examples of how the challenges and approaches described in previous sections are working in practice for IBM's AIHN and the MIT–IBM Watson AI Lab. AIHN and the MIT–IBM Watson AI Lab have been instrumental in forming and refining the model described in this article. The collaborations were established over the course of a few years, while we were also learning from and adapting our model. Thus, not all collaborations operate under the same model. The following paragraphs describe some of the realities of implementing the strategies previously presented.

First, current metrics for the program overall are aligned with the impact-based metrics described earlier and teams are experiencing successes against those metrics. All of the in-progress research teams are yielding coauthored publications to top-tier AI journals and conferences. Several of these industry–academic teams have jointly released open datasets and toolkits that are publicly posing AI challenges to the broader AI research community (Monfort et al. 2018; Gunasekara et al. 2019; Li et al. 2019). Joint teams have also produced top-ranked submissions to other open AI challenge competitions (Shi et al. 2018; Wei et al. 2018) and garnered best paper awards (Boratko et al. 2018; Zhang et al. 2018; Frankle and Carbin 2019; Pearson et al. 2019; Wright et al. 2019).

Second, we have also experienced some challenges in executing our strategies. To begin with, the appropriate use and adoption of the AI challenge format takes practice. Not all principal investigators submit proposals that comprise one or a small number of well-defined AI challenge tasks with objective metrics for success, against specific datasets. This is at least partially because, in many cases, the formulation of a well-specified AI challenge task itself often requires research to lay the groundwork. Awarding exploratory and seedling projects helps to mitigate this challenge. In initial phases, well over half the grants are awarded to exploratory and seedling projects. We hope to see exploratory projects follow up with improved formulations, backed by evidence of the suitability as an ambitious objective to drive forward the field of AI. This will facilitate growing the successful exploratory projects into larger projects with even more ambitious objectives.

Another challenge we experience is in reconciling differences in research philosophies, expertise, incentives, and working styles. All project proposals are to have at least one university principal investigator and at least one industry principal investigator. Every proposal is meant to be jointly conceived and refined, with a plan where each researcher brings their individual expertise and effort to jointly conduct the research to the team. However, it is not uncommon for university principal investigators to prefer unfettered funding, which typically does not include engaging industry collaborators as peer researchers. Likewise, it is not uncommon for industry principal investigators to prefer industry-directed research for specific business challenges, or for their priorities to shift over a multiyear collaboration. Balancing the need to fully explore the potential of an idea, while being responsive to changes in the business landscape requires a flexible, yet strategic mindset from both sides.

Finally, it can also be challenging to form project teams that optimally leverage the expertise and bandwidth of interested researchers. In some cases, individual researchers are already aware of one another's research and simply reach out to discuss potential collaborative projects. This is not as straightforward for interdisciplinary teams, where researchers may not be aware of one another's work. There is also the potential challenge of well-known researchers getting too many collaboration requests, and less-known researchers getting too few. Open houses, poster receptions, working sessions, and openly accessible team messaging channels help researchers identify potential collaborators.

Concluding Remarks

We believe the recent advances in AI and machine learning, alongside the opportunities they present, are worthy of a reconsideration of how industry and academe can best work together. As researchers, we share a lot in common. We all revel in thinking big, taking chances, solving pressing problems with broad scientific, technological or societal impact, the opportunity to fail while trying, and the deeper understanding that comes from a healthy mix of failures and breakthroughs. Developing new approaches, and most importantly, starting new laboratories and initiatives with ambitions to be famous for our science and its impact, is both daunting and exhilarating. We hope some of the aspects we have taken into consideration and the approaches we are pursuing can be valuable to the broader community.

Acknowledgments

The authors are directly involved in the industry–academic collaboration of our respective institutions, playing roles ranging from primary investigator to overall program leadership. MIT embarked on a transformative AI and industry collaborative model with the MIT Quest for Intelligence initiative in 2018, and also announced a billion-dollar investment in the new MIT Stephen A. Schwarzman College of Computing. IBM began transforming its university collaboration model in 2015, when it began establishing multiyear, multimillion-dollar, AI research alliances with a network of world-class faculty and top graduate students. This AI Horizon Network currently includes projects in nine universities: Indian Institute of Technology, MIT, Rensselaer Polytechnic Institute, Mila–Quebec Artificial Intelligence Institute, University of California San Diego, University of Illinois Urbana-Champagne, University of Michigan, University of Maryland at Baltimore County, and University of Massachusetts at Amherst. The goal of this network of academic institutions is to accelerate the advancement of AI and to apply AI to some of the world's most enduring challenges, ranging from disease and the environment, to transportation and education.

Notes

1. www.marketwatch.com/press-release/automotive-artificial-intelligence-market-is-expected-to-be-worth-us-1050-billion-by-2024-2018-05-14.

2. www.nsf.gov/statistics/2018/nsf18304/static/report/nsf18304-report.pdf.

3. medium.com/syncedreview/a-statistical-tour-of-nips-2017-438201fb6c8a.

4. aiindex.org/2017.

5. www.research.ibm.com/artificial-intelligence/horizons-network.

6. mitibmwatsonailab.mit.edu.

7. quest.mit.edu.

8. dreamchallenges.org.

9. www.un.org/development/desa/disabilities/resources/general-assembly/convention-on-the-rights-of-persons-with-disabilities-ares61106.html.

References

Allen, R. 2017. NIPS Accepted Paper Statistics. Medium December 5. medium.com/machine-learning-in-practice/nips-accepted-papers-stats-26f124843aa0.

Berners-Lee, T.; Hendler, J.; and Lassila, O. 2001. The Semantic Web. A New Form of Web Content That Is Meaningful to Computers Will Unleash a Revolution of New Possibilities. *Scientific American* 284: 29–37.

Boratko, M.; Padigela, H.; Mikkilineni, D.; Yuvraj, P.; Das, R.; McCallum, A.; Chang, M.; Fokoue-Nkoutche, A.; Kapanipathi, P.; Mattei, N., et al. 2018. A Systematic Classification of Knowledge, Reasoning, and Context within the ARC Dataset. Paper presented at the Workshop on Machine Reading for Question Answering, April 9. doi.org/10.18653/v1/W18-2607.

Brynjolfsson, E., and Mitchell, T. 2017. What Can Machine Learning Do? Workforce Implications. *Science* 358(6370): 1530–4. doi.org/10.1126/science.aap8062.

Chui, M.; Manyika, J.; Miremadi, M.; Henke, N.; Chung, R.; Nel, P.; and Malhotra, S. 2018. *Notes from the AI Frontier: Applications and Value of Deep Learning. Technical Report, April 2018*. New York: McKinsey Global Institute. www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning.

Clark, P.; Cowhey, I.; Etzioni, O.; Khot, T.; Sabharwal, A.; Schoenick, C.; and Tafjord, O. 2018. *Think you have Solved Question Answering? Try ARC, the AI2 Reasoning Challenge. arXiv preprint. arXiv:1803.05457v1*. [cs.AI]. Ithaca, NY: Cornell University Library.

Crevier, D. 1993. *AI: The Tumultuous Search for Artificial Intelligence*. New York: BasicBooks.

Deng, J.; Dong, W.; Socher, R.; Li, L.; Li, K.; and Li, F. 2009. ImageNet: A Large-Scale Hierarchical Image Database. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*. 248–255. Piscataway, NJ: Institute for Electronic and Electrical Engineers. doi.org/10.1109/CVPR.2009.5206848.

Elliot, B., and Andrews, W. 2017. *A Framework for Applying AI in the Enterprise. Report ID: G00336031, 2 May*. Stamford, CT: Gartner, Inc. www.gartner.com/en/documents/3700317/a-framework-for-applying-ai-in-the-enterprise.

Frankle, J., and Carbin, M. 2019. The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks. Paper presented at the 7th International Conference on Learning Representations, New Orleans, LA, May 6–9. arxiv.org/abs/1803.03635.

Gunasekara, C.; Kummerfeld, J.; Polymenako, L.; and Lasecki, W. 2019. DSTC7 Track 1: Noetic End-to-End Response

Selection. Paper presented at the Seventh Dialog System Technology Challenges at the Association for Advancement of Artificial Intelligence Conference on Artificial Intelligence, Honolulu, HI, January 27. workshop.colips.org/dstc7/papers/dstc7_task1_final_report.pdf.

Hamraie, A. 2018. A Smart City Is an Accessible City. *The Atlantic* 6(November). www.theatlantic.com/technology/archive/2018/11/city-apps-help-and-hinder-disability/574963.

Hutson, M. 2018. Artificial Intelligence Faces Reproducibility Crisis. *Science* 359(6377): 725–6. doi.org/10.1126/science.359.6377.725.

Kahn, J. 2018. Samsung Will Open AI Research Center in Cambridge, England. Bloomberg News May 22. www.bloomberg.com/news/articles/2018-05-22/samsung-will-open-ai-research-center-in-cambridge-england.

Kirsner, S. 2017. What Makes This the Hottest Class at MIT? The Boston Globe April 9. www.boston.com/culture/education/2017/04/09/what-makes-this-the-hottest-class-at-mit.

Li, C.; Dakkak, A.; Xiong, J.; and Hwu, W. 2019. *Challenges and Pitfalls of Machine Learning Evaluations and Benchmarks. arXiv preprint. arXiv:1904.12437v2*. [cs.LG]. Ithaca, NY: Cornell University Library.

Loizos, C. 2016.It Isn't Just Uber: Carnegie Mellon's Computer Science Dean on Its Poaching Problem. TechCrunch April 26. techcrunch.com/2016/04/26/it-isnt-just-uber-carnegie-mellons-computer-science-dean-on-its-poaching-problem.

McCarthy, J.; Minsky, M. L.; Rochester, N.; and Shannon, C. E. 1955, 2006. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *AI Magazine* 27(4): 12.

Metz, C. 2018. Facebook Adds AI Lab in Seattle and Pittsburg, Pressuring Local Universities. The New York Times May 5. www.nytimes.com/2018/05/04/technology/facebook-artificial-intelligence-researchers.html.

Monfort, M.; Zhou, B.; Bargal, S.; Andonian, A.; Yan, T.; Ramakrishnan, K.; Brown, L.; Fan, Q.; Gutfreund, D.; Vondrick, C.; Oliva, A. 2018. Moments in Time Dataset: One Million Videos for Event Understanding. *IEEE Pattern Analysis and Machine Intelligence*. http://moments.csail.mit.edu/TPAMI.2019.2901464.pdf.

National Science Foundation. 2018. *2016 Doctorate Recipients from U.S. Universities: Special Report NSF 18-304*. Alexandria, VA: National Science Foundation. www.nsf.gov/statistics/2018/nsf18304/static/report/nsf18304-report.pdf.

Pearson, C.; Dakkak, A.; Hashash, S.; Li, C.; Chung, I.; Xiong, J.; and Hwu, W. 2019. Evaluating Characteristics of CUDA Communication Primitives on High-Bandwidth Interconnects. In *Proceedings of the 2019 Association for Computing Machinery/Standard Performance Evaluation Corporation International Conference on Performance Engineering*. New York: Association for Computing Machinery. doi.org/10.1145/3297663.3310299.

Shema, H. 2013. What's Wrong with Citation Analysis? Scientific American Information Culture Blog January 1. blogs.scientificamerican.com/information-culture/whats-wrong-with-citation-analysis.

Shi, H.; Wang, Z.; Zhang, Y.; Wang, X.; and Huang, T. 2018. Geometry-Aware Traffic Flow Analysis by Detection and Tracking. Paper presented at the AI City Challenge Workshop at Computer Vision and Pattern Recognition Conference, Salt Lake City, UT, June 18. doi.org/10.1109/CVPRW.2018.00023.

Thrun, S.; Montemerlo, M.; Dahlkamp, H.; Stavens, D.; Aron, A.; Diebel, J.; Fong, P.; Gale, J.; Halpenny, M.; Hoffmann, G., et al. 2006. Stanley: The Robot that Won the DARPA Grand Challenge. *Journal of Robotic Systems* 23(9): 661–92. robots.stanford.edu/papers/thrun.stanley05.pdf.

Wei, Y.; Liu, T.; Ruan, R.; Zhao, J.; Wei, S.; Zhao, Y.; Shi, H.; Xiong, J.; and Huang, T. 2018. Look into Person Challenge Submission. Paper presented at Workshop on Visual Understanding of Human Crowd Scenes at Computer Vision and Pattern Recognition Conference, Salt Lake City, UT, June 18–22. vuhcs.github.io/vuhcs-2018/index.html.

Wright, D.; Katsis, Y.; Mehta, R.; and Hsu, C. 2019. NormCo: Deep Disease Normalization for Biomedical KBC. Paper presented at the Automated Knowledge Base Construction Conference, Amherst, MA, May 20–21. openreview.net/forum?id=BJerQWcp6Q.

Zhang, X.; Wang, J.; Zhu, C.; Lin, Y.; Xiong, J.; Hwu, W.; and Chen, D. 2018. DNNBuilder: An Automated Tool for Building High-Performance DNN Hardware Accelerators for FPGAs. In *Proceedings of the 2018 International Conference on Computer-Aided Design*. New York: Association for Computing Machinery. doi.org/10.1145/3240765.3240801.

Zweben, S., and Bizot, B. 2017. *Computer Research Associates 2017 Taulbee Survey*. Washington, DC: Computing Research Association. cra.org/wp-content/uploads/2018/05/2017-Taulbee-Survey-Report.pdf.

Lisa Amini is the director of IBM Research Cambridge, which is home to the MIT–IBM Watson AI Lab and IBM Research's AIHN. She was previously director of knowledge and reasoning research in the Cognitive Computing group at IBM's TJ Watson Research Center in New York, and she is also an IBM Distinguished Engineer. Amini was the founding director of IBM Research Ireland, and the first woman laboratory director of any non-US IBM Research Global Lab (2010–2013). As senior manager of the Exploratory Stream Processing Research Group, Amini was the founding chief architect for IBM's InfoSphere Streams product and research predecessor. She earned her PhD in computer science from Columbia University in New York.

Ching-Hua Chen is a research staff member and manager at IBM Research. Her current research involves developing computational methods to enable technologies for health behavior modification. She is coprincipal investigator of the Center for Health Empowerment, Analytics, Learning and Semantics, which is a collaboration between IBM and Rensselaer Polytechnic Institute within the AIHN. Previously, she was the research director of the IBM Research Collaboratory in Singapore, where she collaborated with the Singapore government on solutions for traffic management. She earned her PhD in operations research and business administration from the Pennsylvania State University.

David Cox is the IBM director of the MIT–IBM Watson AI Lab, a first of its kind industry–academic collaboration between IBM and MIT, focused on fundamental research in AI. Prior to joining IBM, Cox was the John L. Loeb Associate Professor of the Natural Sciences and of Engineering and Applied Sciences at Harvard University, where he held appointments in computer science, the Department of Molecular and Cellular Biology and the Center for Brain Science. Cox's ongoing research is primarily focused on bringing insights from neuroscience into machine learning and computer vision research. He is a faculty associate at the Berkman–Klein Center for Internet and Society at Harvard Law School and is an agenda contributor at the World

Economic Forum. He has received a variety of honors, including the Richard and Susan Smith Foundation award for excellence in biomedical research, the Google Faculty Research award in computer science, and the Roslyn Abramson award for excellence in undergraduate teaching. His academic laboratory has spawned several startups across a range of industries, ranging from AI for healthcare to autonomous vehicles.

Aude Oliva is the executive director of the MIT–IBM Watson AI Lab and The MIT Quest for Intelligence. She is also a principal research scientist at the Computer Science and Artificial Intelligence Laboratory. She formerly served as an expert to the National Science Foundation, Directorate of Computer and Information Science and Engineering. Her research interests span computer vision, cognitive science, and human neuroscience. She was honored with the National Science Foundation CAREER Award, a Guggenheim Fellowship, and the Vannevar Bush Faculty Fellowship. She earned a MS and PhD in cognitive science from the Institut National Polytechnique de Grenoble, France.

Antonio Torralba is a professor of electrical engineering and computer science at MIT and a principal investigator at the Computer Science and Artificial Intelligence Laboratory. He is also the MIT director of the MIT–IBM Watson AI Lab and the inaugural director of The MIT Quest for Intelligence. Torralba researches computer vision, machine learning, and human visual perception, with an interest in building systems that can perceive the world the way humans do. He was honored with a National Science Foundation CAREER award, the J.K. Aggarwal Prize from the International Association for Pattern Recognition, the Frank Quick Faculty Research Innovation Fellowship, and the Louis D. Smullin ('39) Award for Teaching Excellence. Torralba earned a BS from TelecomBCN, and a PhD from the Institut National Polytechnique de Grenoble, France. He did his postdoctoral work at MIT.