The Impact of Artificial Intelligence on Innovation

Iain M. Cockburn
Boston University and NBER

Rebecca Henderson
Harvard Business School and NBER

Scott Stern
MIT and NBER

Paper prepared for the NBER Conference on Research Issues in Artificial Intelligence
Toronto, September 2017

Preliminary and Incomplete!
Please do not cite or quote!

This draft: September 11, 2017

We thank Michael Kearney for extraordinary research assistance.

The Impact of Artificial Intelligence on Innovation:

An Exploratory Analysis

ABSTRACT

Artificial intelligence promises to improve existing goods and services, and, by enabling automation of many tasks, to greatly increase the efficiency with which they are produced. But it may have an even larger impact on the economy by serving as a new general-purpose "new method of invention" that can reshape the nature of the innovation process and the organization of R&D. This exploratory essay considers this possibility in three interrelated ways. First, we review the history of artificial intelligence, focusing in particular on the distinction between automation-oriented applications such as robotics and the potential for recent developments in "deep learning" to serve as a general-purpose method of invention. We then assess preliminary evidence of this differential impact in changing nature of measurable innovation outputs in artificial intelligence, including papers and patents. We find strong evidence of a "shift" in the importance of application-oriented learning research since 2009 (relative to developments in robotics and symbolic systems research), and that a significant fraction of this upswing in application-oriented learning research was initially led by researchers outside the United States. Finally, we consider some of the implications of our findings, with a focus on both likely changes in the organization of the innovation process as well as the appropriate policy and institutional response that might be required if deep learning represents a meaningful general-purpose method of invention. From an organizational perspective, there is likely to be significant substitution away from more routinized labor-intensive research effort (often directed towards testing specific hypotheses in relatively small purpose-built datasets) towards research that takes advantage of the interplay between passively generated large datasets and enhanced prediction algorithms for phenomena that result from complex interdependencies. At the same time, the potential commercial reward is likely to usher in a period of racing, driven by powerful incentives for individual companies to acquire and control critical large datasets and application-specific algorithms. We suggest that policies which encourage transparency and sharing of core datasets across both public and private actors can stimulate a higher level of innovation-oriented competition, and also allow for a higher level of research productivity going forward.

I. Introduction

Rapid advance in the field of "artificial intelligence" has profound implications for the economy as well as society at large. Artificial intelligence has the potential to directly influence products and services (and the tasks required to create these goods), with important implications for productivity, employment, and competition. But, as important as these effects are likely to be, artificial intelligence also has the potential to change the innovation process itself, with consequences that may be equally profound, and may, over time, come to dominate the direct effect.

Consider the case of Atomwise, a startup firm which developing novel technology for identifying potential drug candidates (and insecticides) by using neural networks to evaluate the 3D structure of candidate molecules relative to that of target proteins and thus predict their bioactivity. The company reports that its deep convolutional neural networks "far surpass" the performance of conventional "docking" algorithms. After appropriate training on vast quantities of data, the company's AtomNet product appears to be able to recognize foundational building blocks of organic chemistry and is capable of generating highly accurate predictions of the outcomes of real-world physical experiments (Wallach et al., 2015). Such breakthroughs hold out the prospect of substantial improvements in the productivity of early stage drug screening, both by reducing the number of unnecessary tests, and identifying candidate molecules with greater chances of success. Beyond the domain of drug discovery, Atomwise is now deploying this approach to discovery and development of new pesticides and agents for controlling crop diseases. Of course, the commercial and medical productivity promise of Atomwise (and other companies leveraging artificial intelligence to advance drug discovery or medical diagnosis) is still at an early stage: though their initial results seem to be promising, no new drugs have actually come to market using these new approaches. But, whether or not Atomwise delivers fully on its promise, it represents a meaningful attempt to develop a new innovation "playbook" for drug discovery, one that leverages large datasets and learning algorithms to engage in precise prediction of biological phenomena in order to guide drug candidate selection.

This example illustrates two key features of how advances in artificial intelligence have the potential to impact innovation. First, though the origins of artificial intelligence are broadly in the field of computer science, and the principal commercial applications of artificial intelligence so far such as industrial robots have been in relatively narrow domains of applications, the types of learning algorithms used by Atomwise and similar companies illustrate that artificial intelligence applications may span a very wide range. From the perspective of the economics of innovation (among others, Bresnahan and Trajtenberg (1995)), there is an important distinction between the problem of providing innovation incentives for technologies that impact a relatively narrow domain of application, such as the field of traditional "robotics" which are purpose-built for narrow tasks (as exemplified by products such as the Roomba), versus technologies with a wide—advocates might say almost limitless—domain of application, as may be true of advances in neural networks and machine learning. As such, a first question to be asked is the degree to which artificial intelligence (and which parts of artificial intelligence) are not simply examples of new technologies, but rather "general purpose technologies" (hereafter GPT) that have been so influential in long-term technological progress.

Second, while many applications of artificial intelligence will surely be to provide a lower-cost or higher-quality input into existing production processes (spurring concerns about the potential for large job displacements), the qualitative shift in the nature of prediction enabled by machine learning may not simply be an opportunity for direct application across a wide variety of domains but may itself allow for a further change in the nature of the innovation process within those domains. As articulated famously by Griliches (1957), the "invention of a method of invention" (hereafter an IMI) has the potential for a more influential impact than a single invention, but is also likely to be associated with wide variation in the ability to adapt the new tool to particular settings, resulting in a heterogeneous pattern of diffusion and use over time. Here again it is useful to draw the distinction between the impact of artificial intelligence tools in areas such as traditional robotics versus advances in learning-oriented approaches such as neural networks. Whereas innovation in areas such as robotics has likely had important direct effects through the introduction of new tools and techniques specific to the problems of that domain, machine learning seems more likely to satisfy the properties of an IMI, making the conditions under which different potential innovators are able to gain access to these tools and use them in a pro-competitive way a central concern for policy.

This essay begins to unpack, in a preliminary way, the potential impact of advances in artificial intelligence on innovation, and identify the role that policy and institutions might play in promoting effective incentives for innovation, diffusion, and competitive structure. To do so, we begin in Section 2 by highlighting the distinctive economics of research tools such as artificial intelligence, focusing on the interplay between the degree of generality of application of a new research tool and the role of research tools not simply in enhancing the efficiency of research activity but in creating a new "playbook" for innovation itself. We then turn in Section 3 to briefly contrasting three key technological trajectories within the domain of AI – robotics, neural networks and machine learning, and symbolic systems. We propose that these often conflated aspects of artificial intelligence likely hold very different roles for the future of innovation itself; most notably, we suggest that whereas areas such as robotics have been relatively narrow in application and have low potential to themselves change the nature of invention, neural networks seem to be an area of research that is highly general-purpose and has the potential to change the innovation process itself.

We then explore these contrasts more systematically by drawing out the empirical implications of this framing through an exploration of the evolution of different areas artificial intelligence in terms of scientific and technical outputs as measured (imperfectly) by the publication of papers and patents from 1990 through 2015. In particular, we develop what we believe is the first systematic database that classifies the corpus of scientific paper and patenting activity in the broad areas of artificial intelligence and divides these outputs into those associated with robotics, learning, and symbolic systems. Though preliminary in nature (and inherently imperfect given that key elements of artificial intelligence may not be observable using these traditional innovation metrics), we find striking evidence for a rapid and meaningful shift in the application orientation of learning-oriented publications, particularly after 2009. The timing of this shift is informative, since it accords with qualitative evidence about the surprisingly strong performance of so-called "deep learning" multi-layered neural networks in a range of tasks including computer vision and other prediction tasks. As well, though not as systematic as our work on the corpus of artificial intelligence, supplementary evidence based on the citation patterns to particular authors (such as Geoffrey Hinton) suggest a striking acceleration of work building on a small number of algorithmic breakthroughs related to multi-layered neural networks in just the last few years.

Though not a central aspect of the analysis for this paper, we further find that, whereas learning-oriented algorithms have had a slow and steady upward swing outside of the United States, US researchers have had a less sustained commitment to learning-oriented research prior to 2009, and have been in a "catch up" mode ever since.

Finally, we begin to explore some of the organizational, institutional and policy consequences of our analysis. We see machine learning as an IMI that depends, in each application, on having access to both the underlying algorithms but also large and granular datasets on physical and social behavior that allow for high levels of prediction of (usually rare) events that had previously defied systematic empirical analysis. As one of the most distinctive aspects of artificial intelligence, the nature of this IMI raises the question of, even if the underlying scientific approaches are open (i.e., the basic multi-layered neural networks algorithms), whether the data to achieve the most consequential advances in this area are likely to be public or private and the consequences of that difference. Specifically, if there are increasing returns to scale from data acquisition (there is more learning to be had from the "larger" dataset), it is possible that early or aggressive entrants into a particular application area may be able to engender a significant and long-lasting advantage over potential rivals merely through the control over data rather than through formal intellectual property or demand-side network effects. Strong incentives to maintain data privately has the additional potential downside that data is not being shared across researchers, thus reducing the ability of all researchers to access an even larger set of data that would arise from public aggregation. As the competitive advantage of incumbents is reinforced, the power of new entrants to drive technological change may be weakened. Though this is an important possibility, it is also the case that, at least so far, there seems to be a significant amount of entry and experimentation across most key application sectors.

II. The Economics of New Research Tools: The Interplay between New Methods of Invention and the Generality of Innovation

At least since Arrow (1962) and Nelson (1959), economists have appreciated the potential for significant underinvestment in research, particularly basic research or domains of invention with low appropriability for the inventor. And, considerable insight has been gained into the conditions under which the incentives for innovation may be more or less distorted, both in terms of their overall level and in terms of the direction of that research.

As we consider the potential impact of different aspects of "artificial intelligence" on innovation, two dimensions seem particularly important – the potential for contracting problems in the face of a new "invention for the method of invention" and the potential for coordination problems arising from a new "general purpose technology." As we develop further in the rest of this paper, it is possible that, relative to the relatively narrow domains of application of traditional automation and industrial robots, those areas of artificial intelligence evolving most rapidly -- "machine learning" – have the potential for both of these features. To understand the likely economic consequences of this potential change in the nature of artificial intelligence innovation, we therefore consider each of these elements in turn, and then consider the impact of their interplay.

First, as initially highlighted by Griliches in his classic studies of hybrid corn, there are certain types of inventions – particularly those that might be considered "research tools" – which not only have a direct impact on a particular area of productive economic use, but may be more constructively thought of as an "invention of a method of invention." (Griliches, 1957). An invention such as hybrid corn not only yielded a new variety of a particular corn strain, but suggested a new approach for breeding novel strains of agricultural products, including different types of corn, but also various grains, rice, and even fruits and vegetables. Research tools such as hybrid corn are not simply a mechanism for reducing the costs of subsequent innovation, but perhaps more consequentially enable a new approach to innovation itself, by altering the "playbook" for innovation in the domains impacted by the new tool. For example, prior to the systematic understanding of the power of "hybrid vigor," a primary focus in agriculture had been improved techniques for self-fertilization (i.e., allowing for more and more specialized natural varietals over time). Once the rules governing hybridization (i.e., heterosis) were systematized, and the performance advantages of hybrid vigor demonstrated, the techniques and conceptual approach for agricultural innovation was shifted, ushering in a long period of systematic innovation using these new tools and knowledge.

Not simply a descriptive fact about a certain type of innovation, inventions that are themselves methods for invention may be particularly subject to a lack of appropriability. As emphasized by Scotchmer (1990), providing appropriate incentives for an upstream innovator that develops only the first "stage" of an innovation can be particularly problematic when contracting is imperfect and the ultimate application of a particular tool is uncertain. Scotchmer and her co-authors emphasized a key point about a multi-stage research process: when the ultimate innovation that creates value requires multiple steps, appropriate innovation incentives are not about whether to provide property rights in general, but how to distribute property rights and incentives across the multiple stages of the innovation process. Lack of incentives for early-stage innovation can mean that the tools required for innovation do not even get invented; strong early-stage property rights without adequate contracting opportunities may result in "hold-up" for later-stage innovators and so reduce the ultimate impact of the tool in terms of commercial application.

Not simply a problem of appropriate intellectual property policy for industrial organization, the vertical research spillovers engendered by new research tools are exemplars of the core innovation externality highlighted by endogenous growth theory (Romer, 1990; Aghion and Howitt, 1992); a central source of underinvestment in innovation in the aggregate results from the intertemporal spillovers from innovators today to innovators tomorrow who do not need to pay the full cost of earlier findings as they "stand on the shoulders of giants." Not simply a theoretical consideration, an increasing body of evidence has accumulated over the last several years highlighting the central importance of research tools and institutions supporting intertemporal spillovers (among others, Furman and Stern, 2011; Williams, 2014). A central insight of this work is that control – both in the form of physical exclusivity as well as formal intellectual property rights -- over tools and data can shape both the level and direction of innovative activity, and that rules and institutions governing control over these areas has a powerful influence on the realized amount and nature of innovation.

A second and distinct potential challenge in providing appropriate innovation incentives is when an innovation has potential across a wide number of distinct applications. These "general purpose technologies" (David, 1990; Bresnahan and Trajtenberg, 1995) are most often in the form of core inventions that have the potential to significantly enhance productivity or quality across a wide number of fields or sectors. As famously argued by David (1990), the electric motor ushered in subsequent technological and organizational change (over a long period of time) across a wide number of economic sectors, from established areas such as parts manufacturing or agricultural processing to the creation of entirely new sectors including X and Y. As emphasized by Bresnahan and Trajtenberg (1995), the presence of a general-purpose technology gives rise to both vertical and horizontal externalities in the innovation process that can both lead to underinvestment as well as distortions in the direction of investment, depending on the private versus social returns to innovation across different application sectors. Most notably, if there are "innovation complementarities" between the general purpose technology and each of the application sectors, lack of incentives in one sector can also result in an indirect externality resulting in a systemwide reduction in innovative investment itself. In particular, the private incentives for innovative investment in each application sector depend on the market structure and appropriability regime of that sector; however, innovation in each sector enhances innovation in the GPT itself, which then induces subsequent demand (and further innovation) in other downstream application sectors. Lack of coordination between the GPT and application sectors, as well as across application sectors, can significantly reduce realized innovative investment and incentives. Despite these challenges, this reinforcing cycle of innovation between the GPT and a myriad of application sectors can result not simply in the improvement in a particular sector, but a more systemic economywide transformation; a rich empirical literature examining the productivity impacts of information technology point to the role of the microprocessor as a GPT as a way of understanding the impact of IT on the economy as a whole (among many others, Bresnahan and Greenstein (1995); Brynjolfsson and Hitt (1999); and Bresnahan, Brynjolfsson, and Hitt (2001)).

It is useful to consider the interplay between these distinct but interrelated aspects of research tools. On the one hand, most research tools are relatively narrow in scope (they are not a GPT) and their primary impact is to reduce the cost or enhance the quality of an existing innovation process. For example, in the pharmaceutical industry, there are particularized materials that promise enhancements to the efficiency of narrow research processes. At the same time, it is possible that a research tool, for example a "faster" workstation, is general in nature and applies across a wide domain but does not necessarily change the nature of the research itself in an important fashion. As well, there are important IMIs that are nonetheless relatively narrow in application. For example, the introduction of genetically engineered research mice (such as the Oncomouse) have had a profound impact on the conduct and "playbook" of research in fields such as cancer and even in many areas of medicine, but are by construction limited to a relatively narrow influence in terms of their domain (i.e., genetically modified research mice have had essentially zero impact on areas such as information technology, energy, or aerospace). And, of course, there are research tools that are simultaneously general in nature and also have the potential to change the method of innovation itself. These types of technologies – from the telescope to the broader introduction of the computer in general – have the potential for profound and unanticipated implications across the economy and society in general.

This framework covers of course only a subset of the key informational and competitive distortions that might arise when considering whether and how to provide appropriate innovation incentives both in terms of the level of research as well as its direction. However, we highlight these two areas in particular as they are likely to be important in the consideration of the role of artificial intelligence, and in particular understanding the rapidly changing landscape of artificial intelligence arising from dramatic improvements in machine learning over the past few years. We therefore turn in the next section to a brief outline of these changes, with an eye towards bringing the framework here to bear on how we might consider the innovation challenges arising from AI moving forward.

III. The Evolution of Artificial Intelligence: Robotics, Neural Networks, and Symbolic Systems

In his omnibus historical account of artificial intelligence (AI) research, Nils J. Nilsson (2010) defines AI as "that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment." His account details the contributions of multiple fields to achievements in AI, including but not limited to biology, linguistics, psychology and cognitive sciences, neuroscience, mathematics, philosophy and logic, engineering and computer science. And, of course, regardless of their particular approach, artificial intelligence research has been united by its earliest motivations such as Turing (1950), and his discussion of the possibility of mechanizing intelligence.

But, though often grouped together, the intellectual history of artificial intelligence as a scientific and technical field is usefully informed by some of the key distinctions and differences among three interrelated but separate areas: robotics, neural networks, and symbolic systems.

On the one hand, perhaps the most successful line of research in the early years of artificial intelligence – dating back to the 1960s – falls under the broad heading of symbolic systems. Though early pioneers such as Turing had emphasized the importance of teaching a machine as one might a child (i.e., emphasizing artificial intelligence as a learning process), the "symbol processing hypothesis." (Newell, Shaw, and Simon, 1958; Newell and Simon, 1976) was premised on the attempt to replicate the logical flow of human decision making through processing symbols. Early attempts to instantiate the logical approach to artificial intelligence yielded striking demonstration projects, such as the ability of a computer to be able to navigate elements of a chess game (or other board games) or engage in relatively simple conversations by following specific heuristics and rules embedded into a program. However, while research based on the concept of a "general problem solver" has continued to be an area of significant academic interest, and there have been explosions of interest over time in the ability to use such approaches to assist human decision-making (e.g., in the context of early-stage expert systems to guide medical diagnosis), the symbolic systems approach has been subject to fundamental critiques about the ability to develop a level of performance that can meaningfully impact real-world processes in a scalable way. While it is of course possible that there will be breakthroughs in the future using this approach, it is fair to say that, while symbolic systems continues to be an area of academic research, this research stream, grounded as it is on a systematic logical system that is pre-determined, has not been central to the commercial application of artificial intelligence nor is it at the heart of the recent reported advances in artificial intelligence that are associated with the area of machine learning and prediction.

A second influential trajectory in artificial intelligence has been broadly around the area of robotics. While the concepts of "robots" as machines that can perform human tasks dates back at least to the 1940s, the field of robotics began to meaningfully flourish from the 1980s onwards through a combination of the advance of relatively mechanistic numerically controlled machine tools and through the development of more adaptive but still rules-based robotics that relied on active sensing of a known environment. On the one hand, perhaps the most consequential application of the broad science of artificial intelligence has been in the area of industrial robotics. These machines are precisely programmed to undertake a given task in an highly controlled environment. Often located in "cages" within highly specialized industrial processes (most notably automobile manufacturing), these purpose-built tools are perhaps more aptly described as highly sophisticated numerically controlled machines rather than as robots with significant artificial intelligence content. However, over the past twenty years, the scientific field of robotics has had an important impact on manufacturing and automation, most notably through the introduction of more responsive robots that are provided with precisely programmed response algorithms as they encounter certain types of stimuli. This approach, famously pioneered by Rod Brooks (1990), focused the commercial and innovation orientation of artificial intelligence away from the modeling of human-like intelligence and instead focused on providing feedback mechanisms that would allow for practical and effective robotics for specified applications (an insight which led to among other applications the Roomba and also adaptable industrial robots that could interact with humans such as Rethink Robotics' Baxter). While these advances are indeed important, and are indeed exemplars of many of the most advanced robots that have captured public imagination with artificial intelligence, it is likely that these innovations are not themselves inventions for the method of invention itself. The principal ways in which these robots are used are in terms of specialized end-use applications, and are not centrally connected to the underlying ways in which researchers themselves might develop approaches to undertake innovation itself (across many domains).

Finally, a third stream of research that has been a central element of artificial intelligence since its founding can be broadly characterized as a "learning" approach. Rather than being focused on symbolic logic, or precise sense-and-react systems, the learning approach was based on the idea of creating reliable and accurate methods for prediction of particular events (either physical or logical) in the presence of particular inputs. Though the field of learning is quite large, of particular importance was the development of the concept of a neural network. A neural network is a program that translates a set of inputs through a combination of weights and thresholds into outputs, measures the "closeness" of the outputs to reality, and then adjusts weights to narrow the distance between outputs and reality by following the gradient of the weights. In this way, the neural network could learn as it is fed more inputs (Rosenblatt, 1958; 1963). Over the course of the 1980s, Hinton and co-authors further advanced the conceptual framework for neural networks through the development of back-propagating neural networks that enhance the potential for supervised learning.

Though initially heralded as having significant promise, the field of neural networks has been subject to considerable ebb and flow, particularly within leading institutions within the United States. From the 1980s through the mid-2000s, the challenge of neural networks was that there seemed to be significant limitations on the ability to enhance prediction in the face of larger datasets or through the introduction of additional "layers" for the learning process (i.e., additional layers that allow for reinforcement through back propagation) (Edwards, REF). However, in the mid-2000s, several a small number of new algorithmic approaches held the potential to develop multi-layer networks that had the ability to enhance prediction through the propagation of increasing layers, could increase their predictive power with larger and larger datasets, and were able to scale to an arbitrary level (among others, a key reference here is Hinton and Salakhutdinov (2006)). These conceptual advances were then shown to exhibit a "surprising" level of practical performance improvement through several research projects from 2009 onwards, notably the dramatic improvements in performance that were able to be achieved using "deep learning" in the context of the ImageNet visual recognition project competition pioneered by Fei-Fei Li at Stanford (Krizhevsky, Sutskever and Hinton, 2012).

How Might Different Fields within Artificial Intelligence Impact Innovation?

The purpose of this discussion is not simply to provide an intellectual history of artificial intelligence. Instead, this discussion allows us to form a more specific set of hypotheses about how different types of AI – each of which is associated with different types of underlying functionality – are likely to influence the innovation process going forward.

First, though a significant amount of public discussion of AI focuses on the potential for AI to achieve super-human cognitive capabilities over a wide range of human capabilities, it is useful to note that, at least so far, the significant advances in AI have not been in the form of the "general problem solver" approaches that were at the core of early work in symbolic systems (and were the motivation for considerations of human reasoning such as the Turing test). Instead, both the mechanical advances in robotics and the algorithmic advances in machine learning are by and large methods that require a significant level of human planning and focus a machine, in a given instance, on a relatively narrow domain of problems-solving (e.g., learning chess, picking up a particular object, etc.) While it is of course possible that an independent stream of invention will arise that meaningfully mimics the nature of human subjective intelligence and emotion, the precise advances that have attracted scientific and commercial attention are well removed from these domains.

Second, though most economic and policy analysis of artificial intelligence draws out consequences from the last two decades of automation to consider the impact of artificial intelligence going forward (e.g., in the domain of job displacement for an ever-increasing number of tasks), it is important to emphasize that there is a sharp difference between the advances in robotics that were a primary focus of application during the 2000s and the potential applications of learning which have come to the fore over the last few years.

To consider this more systematically, we build on our discussion from Section 2 and contrast a number of different underlying new “tools” that have been enabled by artificial intelligence (see Figure A). It is useful to note that robotics by and large has been associated with applications that were highly specialized and also focused on end-user applications rather than the innovation process itself. As well, though there has been significant advance in robotics in enhancing their generality (pick-and-place robots such as Baxter clearly are destined for a much wider range of application than traditional industrial robots), these advances do not seem as of yet to have translated to a more general invention in the method of invention. Robotics is an area where we might focus on the nature of the innovations themselves in terms of job displacement versus enhancement, but it is unlikely to be a direct source for a change in the innovation process. At the same time, many of the most important algorithmic advances in the last twenty years (which many have used to consider the potential impact of deep learning) have represented important IMIs, but have lacked generality. For example, powerful algorithms to scan brain images (so-called functional MRI imaging) have transformed our understanding of the human brain (even though those algorithms themselves are imperfect) not simply through their direct research findings but by establishing an entirely new paradigm and protocol for brain research. However, despite its role as a powerful IMI, fMRI lacks the type of general-purpose applicability that has been associated with the most profound GPTs. In contrast, the specific form of advance that has been achieved by techniques through deep learning have the potential to not simply be general-purpose nor simply be an IMI but to be a general-purpose IMI for the field of innovation.

It is useful to articulate more precisely how the promise of deep learning as a general-purposes IMI might be realized. Specifically, the core technical advance promised by deep learning is to provide an enormously powerful new tool that allows for unstructured "prediction" for physical or logical events that had previously resisted systematic empirical characterization. The development of this new approach to prediction suggests a new approach to undertaking scientific and technical research. Specifically, rather than focusing on small well-characterized datasets or testing settings, it is possible to instead focus instead on identifying large pools of unstructured data and leverage these new tools in order to construct and exploit the ability to predict both technical and consumer phenomena. To return to our initial example of Atomwise, it is possible to consider an unstructured approach to predictive drug candidate selection that brings together a vast array of previously disparate clinical and biophysical data in a way that fundamentally reshapes the "ideas production function" for drug candidates.

FIGURE 1

Invention in the Method of Invention \ General Purpose Technology	NO	YES
NO	**INDUSTRIAL ROBOTS**	**PICK AND PLACE ROBOTS**
YES	**fMRI ALGORYTHM**	**DEEP LEARNING**

To be clear, if true (and it is of course just a hypothesis), the arrival of a general-purpose IMI would have enormous economic, social, and technological consequence (over the long run). First, such a shift would imply that the near-term impact of AI on jobs, organizations, and productivity are likely to be only a very small fraction of the impact that will be felt over a longer time frame as this new IMI diffuses across the widest number of application sectors. A more subtle implication of this point is that "past is not prologue": even if automation over the recent past has resulted in job displacement (e.g., Acemoglu and Restrepo, 2017a), the job design and employment consequences of artificial intelligence will be equally shaped by its ability to enhance the potential for "new tasks" (as in Acemoglu and Restrepo, 2017b!).

Second, the arrival of a general-purposes IMI is a sufficiently uncommon occurrence that its impact could be profound for economic growth and its broader impact on society. There are only a handful of previous general-purposes IMIs –think of the impact of optical lenses in the 17th century – and each of these has had an enormous impact not primarily through their direct effects (e.g., glasses) but through their ability to reshape the ideas production function itself (telescopes and microscopes). It would be useful to understand whether it is possible that deep learning might allow researchers to significantly reshift their approach in order to enhance research productivity (in the spirit of Jones (2009)).

Finally, if deep learning has the potential to serve as a general-purpose IMI, it will be important to consider and develop institutions and a policy environment conductive to enhancing innovation through this approach and doing so in a way that promotes competition and social welfare. A central concern here may be the interplay between the key input required for deep learning – large unstructured databases that provide information about physical or logical events – and competitive structure. While the underlying algorithms for deep learning are primarily open (and can and are being improved on rapidly), the data "pools" that can be used to form predictions may be public or private, and will depend on organization, policy and institutions. As a general-purposes IMI, it may be possible, in a particular application area, for a specific company (either an incumbent or start-up) to invest aggressively in developing a learning algorithm that is significantly advanced relative to alternatives, and thus allows that company to have a significant and persistent innovation advantage due to their control over data that is independent of traditional economies of scale or demand-side network effects. This "deep learning competition for the market" is likely to come with several consequences. First, it will lead to a duplicative race of establish a data advantage in particular application sectors (say, search, autonomous driving, or navigation) followed by the establishment of durable barriers to entry that may be of significant concern for competition policy. And, perhaps even more importantly, this strategic behavior will result itself in a balkanization of data within that application area, reducing innovative productivity in that sector, which itself reduces the spillovers back up to the deep learning GPT sector, and then to other application sectors. Accordingly, both from the perspective of competition policy as well as innovation policy, proactive development of institutions and policies that encourage competition, data sharing, and openness are likely to be important.

Of course, our discussion so far has been highly exploratory, and it would be useful to consider whether the key claims underlying our hypothesis – that deep learning may be a general-purposes IMI distinct from earlier generations of technology associated with AI – have empirical content. To do so, we turn in the next section to a preliminary examination of the evolution of artificial intelligence, with an eye towards identifying the role of learning network as a GPT and IMI, and then drawing out consequences in terms of competitive and industrial structure.

IV. Data

This analysis draws upon two distinct datasets, one that captures a set of Artificial Intelligence (AI) publications from Thompson Reuters Web of Science, and another that identifies a set of Artificial Intelligence patents from the U.S. Patent and Trademark Office. In this section, we provide detail on the assembly of those datasets and summary statistics for variables in the sample.

The primary challenge of this work is to identify and sort AI publications and patents into a cohesive set of subgroups. As previously discussed, peer-reviewed and public-domain literature on artificial intelligence points to three different fields of AI: robotics, learning systems and symbol systems, each comprised of numerous AI subfields (Appendix 1 lists the classifiers we used to group AI-oriented research into these three fields). In short, the robotics field includes AI approaches in which a system engages with and responds to environmental conditions; the symbolic systems field which attempts to represent complex concepts through symbols, e.g. consider a digital computer as a collection of zeros and ones, and the learning systems field which processes data through analytical programs modeled on neurologic systems.

IV.A. Publication Sample and Summary Statistics

Our analysis focuses on journal articles and book publications through the Web of Science from 1955 to 2015. We conducted a keyword search utilizing the keywords described in

Appendix A (we tried several variants of these keywords and alternative algorithmic approaches but this did not result in a meaningful difference in the publication set). We are able to gather detailed information about each publication, including publication year, journal information, topical information, as well as author and institutional affiliations.

This search yields 98,124 publications. We then code each publication into one of the three main fields of AI, as described above. Overall, relative to an initial dataset of 98,124, we are able to uniquely classify 95,840 as symbolic systems, learning systems, robotics, or "general" artificial intelligence (we drop papers that involve combinations of these three fields). Table 1 reports the summary statistics for this sample.

Of the 95,840 publication in the sample, 11,938 (12.5 percent) are classified as symbolic systems, 58,853 (61.4 percent) as learning and 20,655 (21.6 percent) as robotics, with the remainder being in the general field of "artificial intelligence." To derive a better understanding of the factors that have shaped the evolution of artificial intelligence, we create indicators for variables of interest including organization type (private versus academic), location type (US domestic versus International), and application type (computer science versus other application area, in addition to individual subject spaces, e.g. biology, materials science, medicine, physics, economics, etc.).

We identify organization type as academic if the organization of one of the authors on the publication is an academic institution. 81,998 publications (85.5 percent) and 13,842 (14.4 percent) are produced by academic and private sector authors, respectively. We identify publication location as US domestic if one of the authors on the publication lists the United States as his or her primary location. 22,436 publications (25 percent of the sample) are produced domestically.

We also differentiate between subject matter. 44 percent of the publications are classified as Computer Science, with 56 percent classified as other applications. Summary statistics on the other applications are provided in Table 3. The other subjects with the largest number of publications in the sample include Telecommunications (5.5 percent), Mathematics (4.2), Neurology (3.8), Chemistry (3.7), Physics (3.4), Biology (3.4), and Medicine (3.1).

Finally, we create indicator variables to document publication quality, including journal quality (top 10, top 25 and top 50 journals by impact factor[1]) and a count variable for cumulative citation counts. Less than one percent of publications are in a top 10 journal with two percent and 10 percent in top 25 and top 50 journals. The average citation count for a publication in the sample is 4.9.

IV.B. Patent Sample and Summary Statistics

We also undertake a similar approach for gathering a dataset of patents that relate to the broad areas of artificial intelligence. We start with the public-use file of USPTO patents (Marco, Carley et al., 2015; Marco et al., 2015,), and filter the data in two ways. First, we assemble a subset of data by filtering the USPTO Historical Masterfile on the U.S. Patent Classification System (USPC) number. [2] Specifically, USPC numbers 706 and 901 represent “Artificial Intelligence” and “Robots,” respectively. Within USPC 706, there are numerous subclasses including “fuzzy logic hardware,” “plural processing systems,” “machine learning,” and “knowledge processing systems,” to name a few. We then use the USPC subclass to identify patents in AI fields of symbolic systems, learning systems and robotics. We drop patents prior to 1990, providing a sample of 7,347 patents through 2014.

Second, we assemble another subset of AI patents by conducting a title search on patents, with the search terms being the same keywords presented in Table 1.[3] This provides an additional 8,640 artificial intelligence patents. We then allocate each patent into an AI field by associating the relevant search term with one of the overarching fields. For example, a patent that is found through the search term “neural network,” is then classified as a “learning” patent. Some patents found through this search method will be duplicative of those identified by USPC search, i.e. the USPC number will be 706 or 901. We drop those duplicates. Together these two subsets create a sample of 13,615 unique AI patents. Summary statistics are provided in Table 4.

[1] The rankings are collected from Guide2Research, found here: http://www.guide2research.com/journals/

[2] We utilized data from the Historical Patent Data Files. The complete (un-filtered) data sets from which we derived our data set are available here: https://www.uspto.gov/learning-and-resources/electronic-data-products/historical-patent-data-files

[3] We utilized data from the Document ID Dataset that is complementary to Patent Assignment Data available on the USPTO website. The complete (un-filtered) data sets from which we derived our data set are available here: https://www.uspto.gov/learning-and-resources/electronic-data-products/patent-assignment-dataset

Unlike the distribution of learning systems, symbolic systems and robotics in the publication data, the fields are more evenly distributed in the patent data: 3,832 (28 percent) learning system patents, 3,930 (29 percent) symbolic system patents, and 5,524 (40 percent) robotics patents. The remaining patents are broadly classified only as artificial intelligence.

Using ancillary datasets to the USPTO Historical Masterfile, we are able to integrate variables of interest related to organization type, location, and application space. For example, Patent Assignment Data tracks ownership of patents across time. Our interest in this analysis relates to upstream innovative work, and for this reason, we capture the initial patent assignee by organization for each patent in our sample. This data enables the creation of indicator variables for organization type and location. We create an indicator for academic organization type by searching the name of the assignee for words relating to academic institutions, e.g. “University”, “College” or “Institution.” We do the same for private sector organizations, searching for “corp”, “business”, “inc”, or “co”, to name a few. We also search for the same words or abbreviations utilized in other languages, e.g. “S.p.A.” Only seven percent of the sample is awarded to academic organizations, while 91 percent is awarded to private entities. The remaining patents are assigned to public organizations, e.g. U.S. Department of Defense.

Similarly, we create indicator variables for patents assigned to U.S. firms and international firms. The international firm data can also be more narrowly identified by specific country (e.g. Canada) or region (e.g. European Union). 59 percent of our patent sample is assigned to U.S. domestic firms, while 41 percent is assigned to international firms. Next to the United States, firms from non-Chinese, Asian nations account for 28 percent of patents in the sample. Firms from Canada are assigned 1.2 percent of the patents, and firms from China, 0.4 percent.

Additionally, the USPTO data includes NBER classification and sub-classification for each patent (Marco, Carley, et al., 2015). The USPTO utilize the same classification methods as Hall, Jaffe and Trajtenberg (2001). These sub-classifications provide some granular detail about the application sector for which the patent is intended. We create indicator variables for NBER sub-classifications related to chemicals (NBER sub-class 11, 12, 13, 14, 15, 19), communications (21), computer hardware and software (22), computer science peripherals (23), data and storage (24), business software (25), medical fields (31, 32, 33, and 39), electronics fields (41, 42, 43, 44, 45, 46, and 49), automotive fields (53, 54, 55), mechanical fields (51, 52, 59), and other fields (remaining). The vast majority of patents (71 percent) reside in NBER subclass 22, Computer Hardware and Software. Summary Statistics of the distribution of patents across application sectors are provided in Table 5.

Using the organization information, we are able to consider the degree of concentration in each AI field over time. To accomplish this, we create an indicator variable for the top 30 and top 5 patent assignees within each field by year, which we will use to evaluate concentration ratios within each application-sector year.

V. Deep Learning as a GPT: An Exploratory Empirical Analysis

As described in Section 3, the history of artificial intelligence suggests that a specific trajectory, focused on learning, experienced a meaningful shift over the past decade, and may represent the nucleus of a general-purpose invention for the method of invention. The purpose of this section is to begin examining this claim in more detail by considering the evolution of artificial intelligence as a field using the data described in Section IV.

We begin in Figure 2 with a simple description of the evolution over time of the three main fields we identified in the corpus of patents and papers. The first key insight is that the overall field of artificial intelligence has experienced sharp growth since 1990. While there are only a small handful of papers (counted at less than a hundred per year), each field of artificial intelligence now record more than a thousand papers per year. At the same time, there is a striking divergence in activity across fields: though starting from a similar base, there is a steady relative increase in the learning publications relative to robotics and symbolic systems, particularly after 2009. Interestingly, at least through the end of 2014, there is more similarity in the patterns for all three fields in terms of patenting, with robotics patenting continuing to hold an advantage over learning and symbolic systems. However, there does seem to be an acceleration of learning-oriented patents in the last few years of the sample, and so there may be a relative shift towards learning over the last few years which will manifest itself as one looks at more recent trends in patenting.

In terms of the publication data, there is a striking pattern over time in terms of the geographic origin of these publications. Figure 3A shows the overall growth in learning publications by US versus rest-of-world, and Figure 3B captures the fraction of learning publications by US versus rest-of-world (i.e., what fraction of publications within each geography are learning publications?). For the US researchers, there is both a lower level of research on learning, and, in an environment where the overall size of the field is increasing, the relative focus of the US on learning is far more variable. This is consistent with the claim that learning research has had a "faddish" quality in qualitative histories of artificial intelligence that have focused primarily on the United States, with the additional insight that the rest of the world (notably Canada) seems to have taken advantage of this inconsistent focus in the United States to develop capabilities and comparative advantage in this field.

With these broad patterns in mind, we turn to our key empirical exercise: whether there has been a shift in the late 2000s of learning relative to other areas of artificial intelligence towards more "application-oriented" research and away from academic computer science. We begin in Figure 4 with a simple graph that examines over time the number of publications (across all three fields) in computer science journals versus application-oriented outlets. While there has actually been a stagnation (even a small decline) in the overall number of artificial intelligence publications in computer science journals, there has been a dramatic increase in the number of artificial intelligence-related publications in application-oriented outlets. By the end of 2015, we estimate that nearly 2/3 of all publications in artificial intelligence were in fields beyond computer science.

We then focus more closely on the areas of computer science from which this growth emanates in Figure 5. We divide all publications into the three areas of computer science and by whether the publication is classified as belonging to the field of computer science versus an application sector. Several patterns are worthy to note. First, as earlier, we can see the relative growth through 2009 in learning versus other fields in both computer science and application publications. Also, consistent with more qualitative accounts of the fields, we see the relative stagnation of symbolic systems research relative to robotics and learning. But, after 2009, we see that there is a significant increase in application publications in both robotics and learning, but that the learning boost is both steeper and more long-lived. Over the course of just seven years, learning-oriented application publications more than double in number, and now represent just under 50% of *all* artificial intelligence publications.[4]

These patterns are if anything even more striking if one disaggregates by the geographic origin of the publication. In Figure 6, we examine learning publications by computer science versus applications, disaggregated by the US versus rest-of-world. The striking upward swing beginning in 2009 is actually led by international publications (which are at a much higher overall level than the United States), though US researchers begin a period of catch-up at an accelerating pace towards the final few years of the sample.

Finally, we consider how these applications areas themselves have changed over time. In Table 3, , we examine the number of publications by applications field in each of the three areas of computer science across two three-year cohorts (2004-2006 and 2013-2015). There are a number of patterns of interest. First, and most importantly, in a range of application fields including medicine, radiology and economics, there is a large relative increase in learning-oriented publications relative to robotics and symbolic systems. As well, there are some sectors that realize a large increase in both learning-oriented research as well as other AI fields, such as neuroscience or biology. And, there are some more basic fields such as mathematics that have experienced a relative decline in publications (indeed, learning-oriented publications in mathematics experienced a small absolute decline, a striking different relative to most other fields in the sample). Overall, though it would be useful to dig much more deeply into precisely the type of research that is being conducted and what is happening at the level of particular subfields, these results are consistent with our broader hypothesis that, alongside the overall growth of artificial intelligence, learning-oriented deep learning research may represent a general-purpose technology that is now beginning to be exploited far more systematically across a wide range of application sectors.

Together, these exploratory findings seem to provide direct empirical evidence for at least one hypothesis about the economics of artificial intelligence: rather than simply being an extrapolation of past efforts at artificial intelligence that have had relatively little broad-scale

[4] The precise number of publications for 2015 are estimated from the experience of the first nine months (the Web of Science data run through September 30, 2015). We apply a linear multiplier for the remaining three months (i.e., estimating each category by 4/3).

economic impact (but have of course been highly important in particular application areas), learning-oriented artificial intelligence seems to have some of the signature hallmarks of a general-purpose technology, and one that is already seeing diffusion across a wide number of fields and sectors that had previously been isolated from developments in artificial intelligence.

VI. Deep Learning as a General-Purpose Invention for the Method of Invention: Considerations for Organizations, Institutions and Policy

With these results in mind, it is useful to consider the potential implications for innovation if deep learning is indeed a general-purpose technology. As discussed briefly in Section 3, it is useful to distinguish the case between the potential for deep learning to serve as a GPT and the perhaps even more consequential contention that it also represents an IMI. In the former, deep learning will allow for enhanced research productivity across a range of applications (with potential for spillovers both back to the learning GPT and also to other application sectors) but would not itself change the nature of the innovation production function itself. However, though systematic empirical evidence has not yet been developed in this area, cases such as Atomwise and other areas that deep learning is a specific approach that may have the potential to reorient the research process itself.

As discussed in Section III, this would involve a shift towards research that uses large datasets to create meaningful predictions for physical and logical events that had previously resisted systematic empirical scrutiny. To be clear, the underlying data would often be passively created as the result of events both online (e.g., search or online purchasing behavior) and in the physical world (i.e., from various types of sensor data) or from prior knowledge (as in the case of "learning" prior literatures as in Watson or Atomwise). Suppose for a moment that this potential shift was in fact real. What would be the likely consequences in terms of appropriate organization of innovation, the institutions we have for training and conducting research over time, and in terms of policy, particularly as we think about private incentives to maintain proprietary datasets and application-specific algorithms?

The Management and Organization of Innovation

Perhaps most immediately, the rise of general-purpose predictive analytics using large datasets to substitute towards capital and away from labor in the research production process.

Many types of R&D and innovation more generally are effectively problems of labor-intensive search with high marginal cost per search (Evenson and Kislev, 1975, among others). and the promise of deep learning is to substitute away from labor in these search processes towards undertaking fixed costs investments in the development of application-specific AIs that can then reduce marginal search costs dramatically. On the one hand, this opens up the possibility for researchers to investigate a much wider range of social, physical, and natural phenomena than has previously been considered feasible, and should open up significant new opportunities to expand the range of subjects and phenomena that are covered under the domain of systematic scientific and empirical research.

Three interrelated related implications arise as one considers the likely impact of deep learning as a new IMI in the research sector. First, it is possible that the ability to substitute away from specialized labor and towards capital (that in principle could be rented or shared) may lower the "barriers to entry" in certain scientific or research fields (where data and algorithms are made available) while erecting new barriers to entry in other areas (e.g, by restricting access to data and algorithms). As of yet, there are few if any organized "markets" for AI capital for research services, and few standards to evaluate alternatives. The development of markets for shared AI services is likely a precursor to broad adoption and dissemination of deep learning into a variety of scientific, technological and research fields.

At the same time, the arrival of this new research paradigm is likely to require a significant shift in the management of innovation itself. For example, it is possible that the democratization of innovation will also be accompanied by the lack of investment by individual researchers in developing deep research skills and expertise in a given area, reducing the level of underlying theoretical or technical depth. This shift away from career-oriented research trajectories and towards the ability to find new findings based on deep learning may undermine long-term incentives for breakthrough research that can only be conducted over a long period of time.

Finally, it is possible that deep learning will itself change the nature of scientific and technical advance. To a first approximation, classical science and engineering fields are dominated by relatively sophisticated disciplines that nonetheless focus on identifying a relatively small number of causal drivers of underlying phenomena (a parsimony famously attributed to Einstein that theory should be "as simple as possible but no simpler.") However, deep learning offers an alternative paradigm based on the ability to "predict" complex multi-causal phenomena in a "black box" approach that abstracts away from underlying causes but does allow for a singular prediction index that can yield sharp insight. But, this substitution away from causal mechanisms may come at a deep expense: it may be possible to train a deep learning algorithm to discover the double helix structure of DNA but it would likely require a level of human judgment to notice that the proposed structure suggested a direct mechanism for heredity.

Innovation and Competition Policy and Institutions

A second and related area of impact will not simply be on the organization of individual research projects or the nature of what counts as "science" in a particular field, but on the appropriate design and governance of institutions governing the innovation process. Three implications stand out.

First, as discussed in Section 2, research over the past two decades has emphasized the important role for institutions that encourage cumulative knowledge production through low-cost independent access to research tools, materials and data (Furman and Stern, 2012; Murray, et al, 2015). At least so far, there has only been a modest level of attention to ensure transparency and replicability within the deep learning community. Particular initiatives such as those through github and reddit are grassroots efforts to encourage openness. But, it is useful to emphasize that there is likely to be a significant gap between the private and social incentives to share and aggregate data outside of formal institutional mechanisms (even among academic researchers or private sector research communities). Specifically, to the extent that any individual research result will depend on the aggregation of data from many sources, it will be important to develop rules concerning the appropriate credit and attribution, and also the ability to replicate results to detect false inferences (particularly important given the potential for p-hacking in the case of deep learning).

A particular area of concern will be in the design and enforcement of formal intellectual property rights. Though the level of patenting in the area of deep learning has so far been relatively low (and we have yet to hear of the "patent on intelligence"), the history of the economics of research tools has been that the emergence of new breakthrough approaches – combined with relatively few capabilities in the relevant technical area at the patent offices – results in long periods of uncertainty, hold-up, and potential lowered research productivity and competition due to ineffective intellectual property policy.

In addition to these traditional innovation policy institutions, the prospect for deep learning raises a wide variety of other issues, including issues relating to privacy, the potential for bias (deep learning has been found to reinforce stereotypes already present in society), and consumer protection (related to areas such as search, advertising, and consumer targeting and monitoring). The key is that, to the extent that deep learning is general-purpose, the issues that arise across each of these domains (and more) will play out across a wide variety of sectors and contexts and at a global rather than local level. Little analysis has been conducted that can help design institutions that will be responsive at the level of application sectors that also internalize the potential issues that may arise with the GPT.

Finally, as we briefly outlined in Section 3, the recognition that this new general-purpose IMI may yield significant prediction insights across a range of sectors is likely to engender a race, within each sector, to establish a proprietary advantage leveraging these approaches. As such, the arrival of deep learning raises issues for competition policy. In each application sector, there is the possibility that firms that are able to establish an advantage at an early stage, and so be able to generate more passive data (about their technology, about customer behavior, about their organizational processes) will be able to erect a deep-learning-driven barrier to entry that will ensure market dominance over at least the medium term. As such, rules ensuring data accessibility are not simply a matter of research productivity or aggregation, but instead impinge on the potential to guard against lock-in and anticompeitive conduct. At least so far, there seems to be a large number of individual companies attempting to take advantage of artificial intelligence across a wide variety of domains (e.g., there are probably more than 20 firms engaging in significant levels of research in autonomous vehicles, and no firm has yet to show a decisive advantage), but this high level of activity today likely reflects an expectation for the prospects for significant market power in the future. Ensuring that deep learning does not enhance monopolization and increase barriers to entry across a range of sectors will be a key topic going forward.

VI. Concluding Thoughts

The purpose of this exploratory essay has not been to provide a systematic account or prediction of the likely impact of artificial intelligence on innovation, nor clear guidance for policy or management of innovation. Instead, our goal has been to raise a specific possibility --- that deep learning represents a new general-purpose method for the method of invention -- and draw out some preliminary implications of that hypothesis for management, institutions, and policy.

Our preliminary analysis highlights a few key ideas that have not been central to the economics and policy discussion so far. First, at least from the perspective of innovation, it is useful to distinguish between the significant and important advances in fields such as robotics (that nonetheless are unlikely to change the nature of innovation itself) from the potential arrival of a general-purpose method of invention from multi-layered neural networks. Our preliminary empirical evidence documents a striking shift since 2009 towards application-oriented research that focuses on learning, consistent with qualitative evidence from the field about these underlying breakthroughs. Second, and relatedly, the prospect of a change in innovation raises key issues for a range of policy and management areas, ranging from how to evaluate this new type of science to the potential for prediction methods to induce new barriers to entry across a wide range of industries. Proactive analysis of alternative policies and institutions from these breakthrough seems like an extremely promising area for research going forward.

REFERENCES

Aghion, P. and P. Howitt (1992) "A Model of Growth Through Creative Destruction," *Econometrica*, 60(2), 323-251.

Bresnahan, T., E. Brynjolfsson, and L. Hitt (2002) "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *The Quarterly Journal of Economics*, 117(1), 339-376.

Bresnahan, T. and S. Greenstein (1999) "Technological Competition and the Structure of the Computer Industry," *Journal of Industrial Economics*, 47(1), 1-40.

Bresnahan, T. and M. Trajtenberg (1995) "General Purpose Technologies 'Engines of Growth'?" *Journal of Econometrics*, 65 (1995) 83-108.

Brooks, R. (1990) "Elephants Don't Play Chess," *Robotics and Autonomous Systems*, 6, 3-15.

Brooks, R. (1991) "Intelligence Without Representation," *Artificial Intelligence*, 47, 139-159.

Brynjolfsson, E. and K. McElheran (2017) "The Rapid Adoption of Data-Driven Decision-Making," *American Economic Review*, 106(5), 133-139

Griliches, Z. (1957) "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 25(4), 501-522.

Henderson, R. and K. Clark (1990) "Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms," *Administrative Science Quarterly*, 35(1), 9-30.

Krizhevsky, A., I. Sutskever, G. Hinton (2012) "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing*, 25, *MIT Press*.

Marco, A., A. Myers, S. Graham, P. D'Agostino, and K. Apple (2015) "The USPTO Patent Assignment Dataset: Descriptions and Analysis," *USPTO Working Paper No. 2015-02*, 1-53.

Marco, A., M. Carley, S. Jackson and A. Myers (2015) "The USPTO Historical Patent Data Files," *USPTO Working Paper No. 2015-01*, 1-57.

Minsky, M. (1961) "Steps Toward Artificial Intelligence," *Proceedings of the IRE*, 8-30.

Nilsson, N. (2010), *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge University Press.

Romer, P. (1990) "Endogenous Technological Change," *Journal of Political Economy*, 98(5), S71-S102.

Rosenberg, N. and M. Trajtenberg (2004) "A General Purpose Technology at Work: The Corliss Steam Engine in the Late-Nineteenth-Century United States," *Journal of Economic History*, 61(1), 61-99.

Rumelhart, D., G. Hinton, and R. Williams (1986) "Learning Internal Representations by Error Propagation," in J. McClelland and D. Rumelhart (editors), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 2: Psychological and Biological Models*, MIT Press, 7-57.

Scotchmer, S. (1991) "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law," *Journal of Economic Perspectives*, 5(1), 29-41.

Turing, A. (1950) "Computing Machinery and Intelligence," *Mind*, 59, 433-460.

Wallach, I. Dzamba, M. and Heifels, A. "AtomNet: A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery." arXiv:1510.02855 [cs.LG]

Table 1A: Publication Data Summary Statistics

	Mean	Std. Dev.	Min	Max
Publication Year	2007	6.15	1990	2015
Symbolic Systems	.12	.33	0	1
Learning Systems	.61	.48	0	1
Robotics	.21	.41	0	1
Artificial Intelligence	.06	.23	0	1
Computer Science	.44	.50	0	1
Other Applications	.56	.50	0	1
US Domestic	.25	.43	0	1
International	.75	.43	0	1
Observations	95840			

Table 1B: Patent Data Summary Statistics

	Mean	Std. Dev.	Min	Max
Application Year	2003	6.68	1982	2014
Patent Year	2007	6.98	1990	2014
Symbolic Systems	.29	.45	0	1
Learning Systems	.28	.45	0	1
Robotics	.41	.49	0	1
Artificial Intelligence	.04	.19	0	1
Computer Science	.77	.42	0	1
Other Applications	.23	.42	0	1
US Domestic Firms	.59	.49	0	1
International Firms	.41	.49	0	1
Org Type Academic	.07	.26	0	1
Org Type Private	.91	.29	0	1
Observations	13615			

Table 2A: Distribution of Publications across Subjects

	Mean	Std. Dev.
Biology	.034	.18
Economics	.028	.16
Physics	.034	.18
Medicine	.032	.18
Chemistry	.038	.19
Mathematics	.042	.20
Materials Science	.029	.17
Neurology	.038	.19
Energy	.015	.12
Radiology	.015	.12
Telecommunications	.055	.23
Computer Science	.44	.50
Observations	95840	

Table 2B: Distribution of Patents across Application Sectors

	Mean	Std. Dev.
Chemicals	.007	.08
Communications	.044	.20
Computer Hardware and Software	.710	.45
Computer Peripherals	.004	.06
Data and Storage	.008	.09
Business software	.007	.09
All Computer Science	.773	.42
Medical	.020	.14
Electronics	.073	.26
Automotive	.023	.15
Mechanical	.075	.26
Other	.029	.16
Observations	13615	

Table 3: Publications Across Sectors, by AI Field, 2004-2006 versus 2013-2015

		Biology	***Economics***	***Physics***	***Medicine***	***Chemistry***	***Math***	***Materials***	***Neuro.***	***Energy***	***Radiology***	***Telecom.***	***CompSci***
Learning Systems	2004-2006	258	292	343	231	325	417	209	271	172	94	291	3889
	2013-2015	600	423	388	516	490	414	429	970	272	186	404	4582
	% growth	133%	45%	13%	123%	51%	-1%	105%	258%	58%	98%	39%	18%
Robotics	2004-2006	33	10	52	69	24	45	36	31	6	47	653	1431
	2013-2015	65	12	122	83	92	80	225	139	18	25	401	1322
	% growth	97%	20%	135%	20%	283%	78%	525%	348%	200%	-47%	-39%	-8%
Symbol Systems	2004-2006	93	8	68	96	139	54	32	35	15	82	51	827
	2013-2015	105	10	125	84	149	60	101	73	22	56	88	1125
	% growth	13%	25%	84%	-13%	7%	11%	216%	109%	47%	-32%	73%	36%

Table 4: Herfindahl-Hirschman Index for Application Sectors

Application	**H=** $\sum \boldsymbol{PatShare}^2$
Chemical Applications	153.09
Communications	140.87
Hardware and Software	86.99
Computer Science Peripherals	296
Data and Storage	366.71
Computer Science Business Models	222
Medical Applications	290.51
Electronic Applications	114.64
Automotive Applications	197.03
Mechanical Applications	77.51
Other	129.20

Figure 2A: Publications by AI field over Time

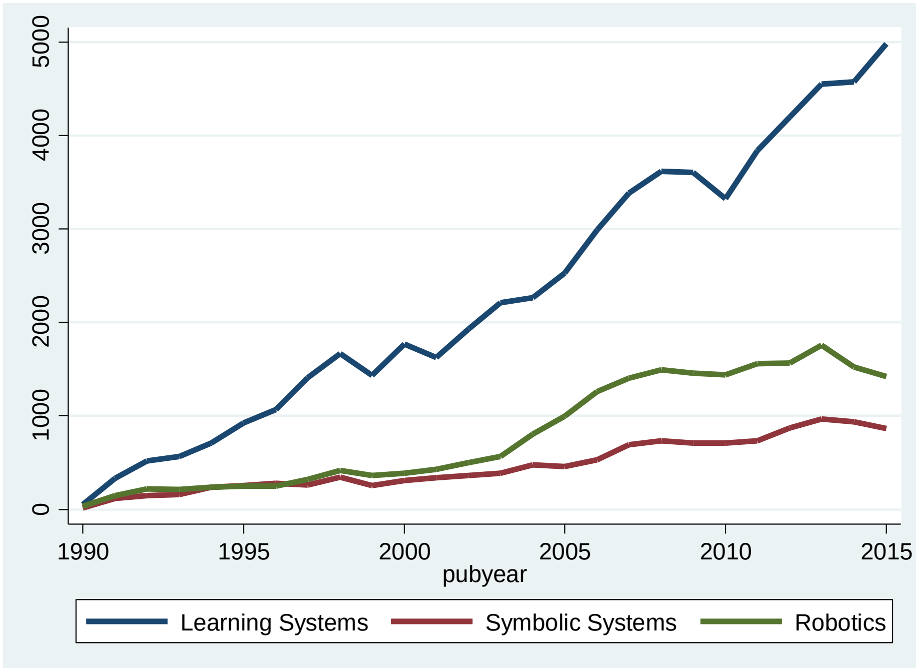


Figure 2B: Patents by AI field over Time

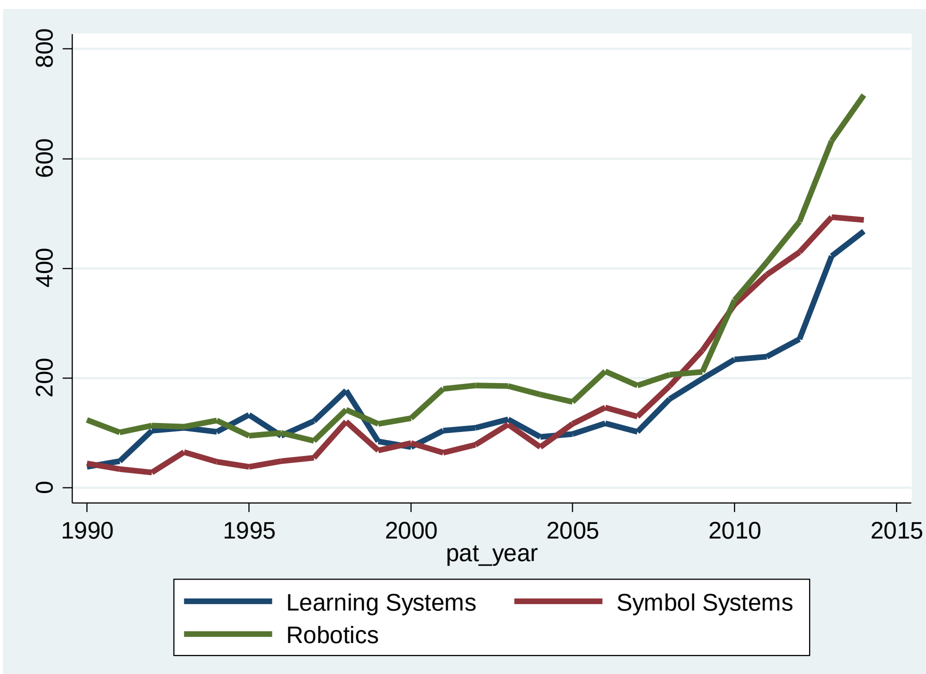


Figure 3A: Academic Institution Publication Fraction by AI Field

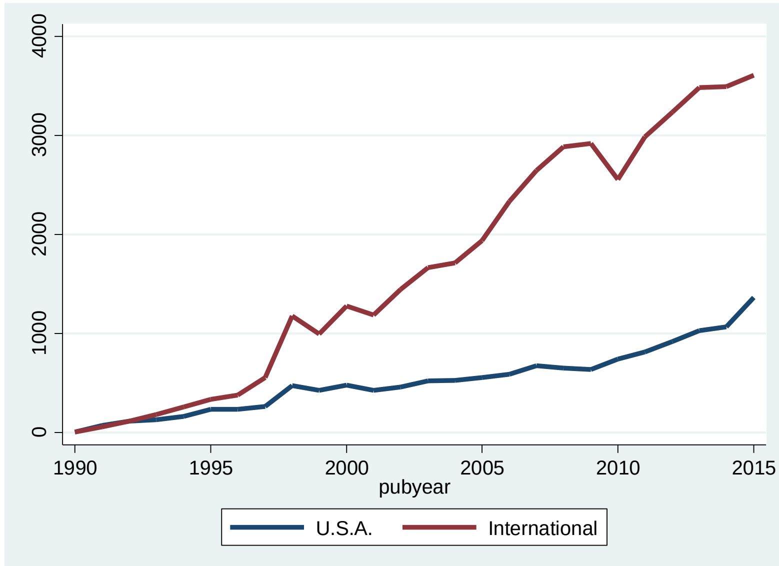


Figure 3B: Fraction of Learning Publications by US versus World

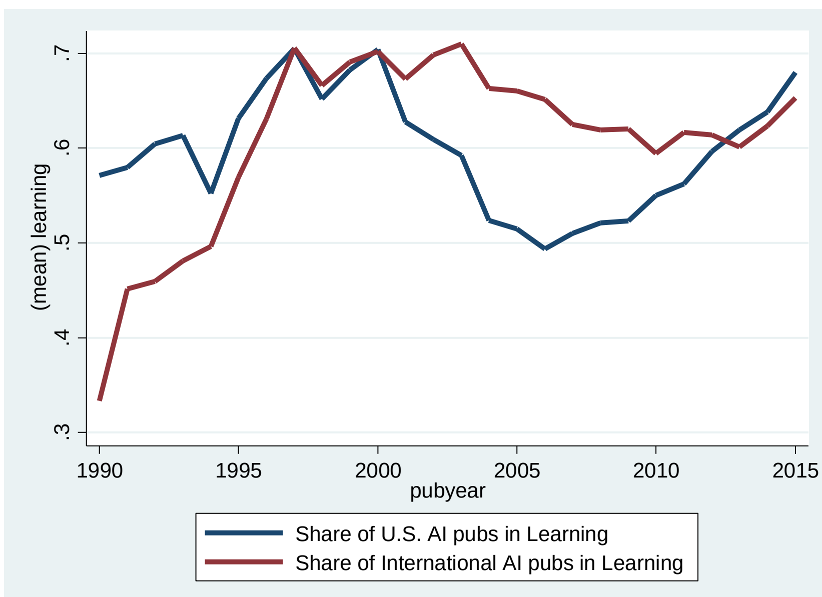


Figure 4: Publications in Computer Science versus Application Journals

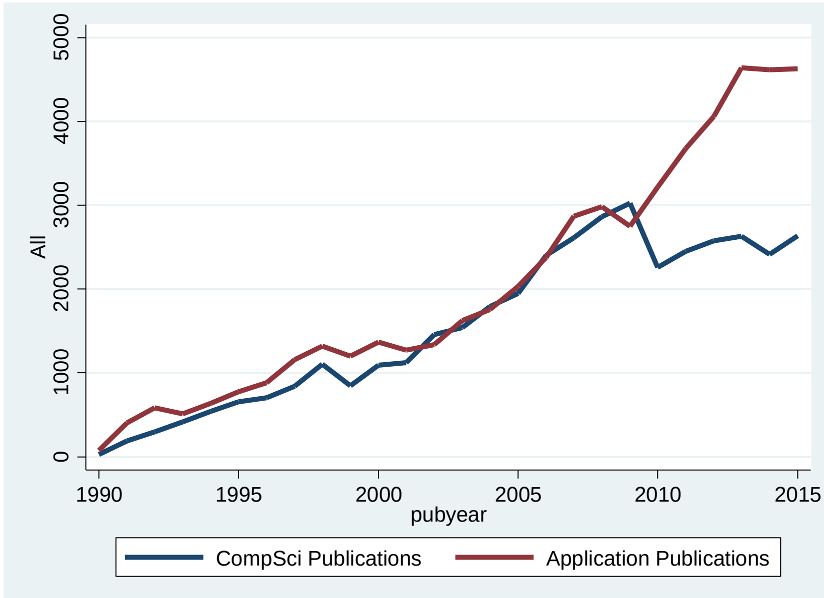


Figure 5: Publications in Computer Science versus Application Journals, by AI Field

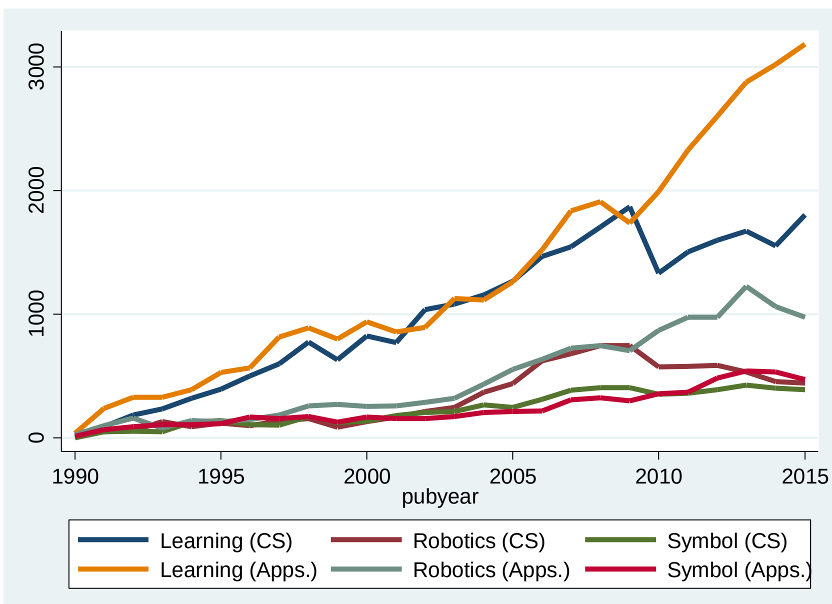
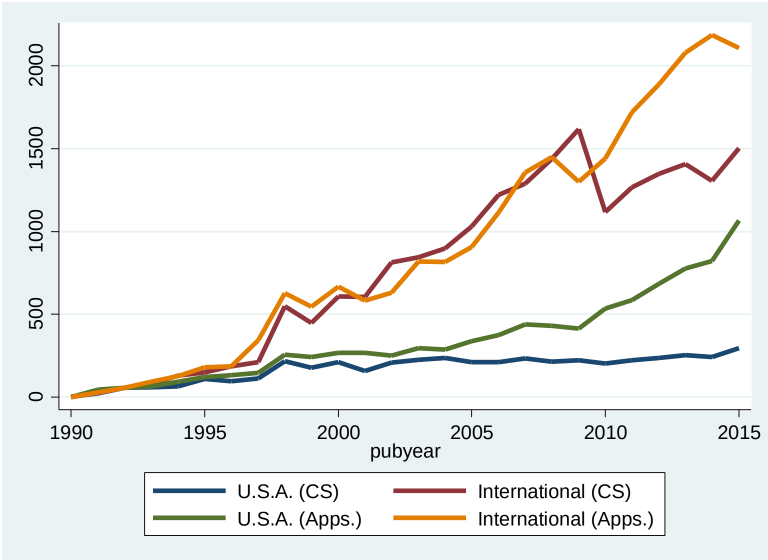


Figure 6: Learning Publications in Computer Science versus Applications, By US versus ROW



Appendix A

Appendix Table 1: Artificial Intelligence Keyword Allocation

Symbols	Learning	Robotics
natural language processing	machine learning	computer vision
image grammars	neural networks	robot
pattern recognition	reinforcement learning	robots
image matching	logic theorist	robot systems
symbolic reasoning	bayesian belief networks	robotics
symbolic error analysis	unsupervised learning	robotic
pattern analysis	deep learning	collaborative systems
symbol processing	knowledge representation and reasoning	humanoid robotics
physical symbol system	crowdsourcing and human computation	sensor network
natural languages	neuromorphic computing	sensor networks
pattern analysis	decision making	sensor data fusion
image alignment	machine intelligence	systems and control theory
optimal search	neural network	layered control systems
symbolic reasoning		
symbolic error analysis		