RIETI Policy Discussion Paper Series 18-P-017

Understanding AI Driven Innovation by Linked Database of Scientific Articles and Patents

MOTOHASHI Kazuyuki
RIETI

RIETI
Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Policy Discussion Paper Series 18-P-017

November 2018

Understanding AI Driven Innovation by Linked Database of Scientific Articles and Patents[1]

MOTOHASHI Kazuyuki

(University of Tokyo, NISTEP and RIETI, Japan)

Abstract

The linked dataset of AI research articles and patents reveals that substantial contributions by the public sector are found in AI development. In addition, the role of researchers who are involved both in publication and patent activities, particularly in the private sector, increased over time. That is, open science that is publicly available in research articles, and propriety technology that is protected by patents, are intertwined in AI development. In addition, the impact of AI, combined with big data and IoT, which are defined as “new” IT innovations, is discussed by comparing it with traditional IT, which only consists of the technological progress of computer hardware and software developments. Both new and traditional IT can be understood by using the framework of GPT (general purpose technology), while the organization of new IT innovation can be characterized as an emergence of multiple ecosystems, instead of being organized in the pattern of platform leadership, found in traditional IT.

Keywords: Patent data, AI, General Purpose Technology

[1]. This study is conducted as a part of the Project “Empirical Analysis of Innovation Ecosystems in Advancement of the Internet of Things (IoT)” undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The conceptual part in section 2 is drawn from the author’s work at FFJ-EHESS as a Michelin Fellow from April 2017 to March 2018, and the statistical analysis in section 3 is drawn from the NISTEP Project on innovation process database. The author thanks for helpful comments from the participants at various workshops, including the EHESS-OECD workshop on Innovation policy in AI/IoT era: Investigating university industry collaboration 2.0 (December 2017), the NISTEP-Tokyo University Workshop on Science and Technology Policy in the Digital Innovation Era (March, 2018) and the MSI seminar at KUL (May 2018). The author also offers thanks for all comments made by RIETI discussion paper workshop as well as Prof. Nagaoka’s written comments. All views expressed in this article are those of authors’ and not of his affiliated organizations.

1. Introduction

Undoubtedly, scientific knowledge has become increasingly important in the industrial innovation process. Genome science has significantly changed the pharmaceutical industry's R&D process, while an understanding of materials' nanoscale physical properties has become essential to the miniaturization of the large-scale integration (LSI) circuit-fabrication process. Information technology has had substantial social and economic impacts, and Big Data analyses have brought deeper insights to business and management activities. Specifically, the advance of data science and artificial intelligence (AI) has contributed to a scientific understanding of business processes, which can be applied to the entire economy. Therefore, while scientific findings were previously applied to only one specific sector, such as the pharmaceutical industry, the "science economy" has emerged, with recent trends involving the economic impacts of scientific innovation, instead of science-based industries (Motohashi, 2014).

As the distance decreases between science and innovation, the relationship between these two activities has shifted, from a linear model—in which scientific knowledge is first gained at research institutions, such as universities, then used by companies to develop new products—to a co-occurrence model—in which scientific and innovation activities simultaneously occur through their interactions with each other. In this regard, we have developed a new indicator of science representing the co-occurrence of science and innovation, based on linked datasets of patents and research articles by the same author-inventor (Ikeuchi et al., 2016). This provides complementary information regarding this scientific indicator based on research papers' patent citations, a traditional science linkage indicator based on the linear model.

This paper analyzes the nature of AI-driven innovation based on the author-inventor linkage data from US patents and research articles. Attention to AI is increasing in the scientific sector, including in universities as well as in industry; further, AI is perceived as a key technology to fundamentally change the innovation landscape by making what is coming in the future less expensive and more certain (Agrawal et al., 2018). Therefore, private incentives to capture such potential value are substantial. Another characteristic of AI involves its potential application in a wide range of industries. In other words, AI is a killer application of IT as a general purpose technology. Further, AI can be called an invention as a method of inventing, in a sense that AI can be used in the process of inventing new applications, such as autonomous driving and condition-based maintenance, and in new drug discoveries (Cockburn et al., 2018).

The remainder of this paper is organized as follows: The next section describes the concept of AI-driven innovation, and discusses the interrelationship between AI and the complementary elements to innovation, Big Data, and the Internet of Things (IoT) is discussed. An analytical section then provides a linked dataset from research articles and patents. This is followed by the discussion section, which focuses on the differences in AI-driven innovation, or "new" IT innovation. Finally, this paper concludes with both managerial and policy implications.

2. Conceptualization of AI-driven innovation

The AI field has experienced significant progress, such as the construction of new machine-learning methods—including deep learning and generative adversarial networks—but the neural network as the foundation of such technologies is not new at all. The improved performance and capabilities in computing and Big Data availability have enabled progress in developing these methods. Specifically, the emergence of Big Data and developments in AI-based science innovation are inseparable. Additionally, the IoT concept is also important for AI technologies that result in industrial applications, such as AI speakers and automated driving systems. Thus, we will discuss the necessary elements for AI-based innovations—such as Big Data and the IoT—and discuss the mutually complementary relationships between these two items and AI.

(1) Big Data

Data analysis is being used now more than ever to solve companies' managerial issues. What is new in utilizing Big Data? Companies have conventionally used data to accompany business systems. On the one hand, corporate financial data accumulate in financial accounting systems, while other such data related to personal histories and employee salaries accumulate in human resource systems. On the other hand, a supply chain-management system manages materials and products' stock status and order records as data, which are respectively characterized as made for a specific reason. Since the 1990s, each business system has been increasingly integrated as an optimum resource management system for the entire company, or enterprise resource planning. Such systems emphasize the improving of corporate activities' operational efficiency.

Alternatively, Big Data is characterized by its construction for a non-specific use. For example, Amazon is frequently cited as a company that played a pioneering role in the use of Big Data in its maintaining of users' purchase records, as Amazon uses this information to make book recommendations that match its individual users' interests. Companies that engage in e-commerce automatically accumulate merchandise-related purchase data, such as book purchases, although this data is not specifically collected so companies can make recommendations to customers regarding books they would be highly likely to purchase. Additionally, a certain amount of data is required for highly accurate recommendations, including the number of users and the number of purchase histories. This is because the recommendation method is calculated with a stochastic model to estimate what books customers would be highly likely to purchase based on their purchase histories. As the number of samples increases, the recommendations' accuracy improves. In other words, the value can vary depending on the data size (Mayer-Schonberger and Cukier, 2013).

Another leading company in utilizing the Big Data innovation model is Google, which utilizes search histories as a source of Big Data. Google's business model involves using their search engine to find keyword-specific advertisements. Data that has been used in IT systems thus far has primarily been numerical, but various forms of data exist on the Internet, such as audio, visual, and text data. Big Data is also characterized by the difference from conventional IT use, and specifically, in the value added by converting data from conventional IT use, or "datafication."

Further, in addition to the development and use of the Internet, data is also currently collected using various kinds of sensor-based information to add economic value. For example, Komatsu has equipped its construction machinery with GPS systems and fuel gauges to collect location and operational status data. The company uses such data to provide various services, including an anti-theft functionality, and make recommendations for fuel cost savings, which will assist in differentiating their products from those of their competitors. These data-collection sensors gather data from various sources, such as travel data from cell phones and onboard GPS systems, to indicate the operational status of all types of industrial equipment. The use of these sensors is being actively developed (Kinukawa et al., 2014).

As previously described, various data types are available from e-commerce purchase records, online information, and sensor information, in addition to the data produced from companies' internal business systems. Big Data is characterized by the three "Vs": volume, or their large data size; variety, or the "datafication" of various information, including text, images, and audio; and velocity, or the continuous daily inflow of data from the Internet and sensors. Additionally, one data utilization method enables the wider, more comprehensive observation of various events related to business management, society, and human behavior, and allows these mechanisms to be analyzed at the micro-level.

(2) Artificial Intelligence

Artificial intelligence (AI) integrates technologies to exploit beneficial information in business management from Big Data. Applied AI technologies may be categorized as image or text data-recognition technologies, including the conversion to computer-readable data; human interfaces, including the visualization of data and interactive agents; and knowledge-discovering technologies related to the diagnosis, monitoring, and datamining of various types of equipment devices, among other categories. Additionally, machine-learning, fuzzy control, and such mathematical models as genetic models have been implemented as basic system technologies to realize those functions.

Specifically, the existence of Big Data has led to significant technological progress in the machine-learning methodology. Machine-learning model-estimation methods are generally grouped as either supervised or unsupervised learning, and the three “Vs” of Big Data are important in either method. In supervised learning, a large volume of the text and image data accumulated on the Internet can be used as training data. For example, Google’s translation services involve their translation system reading a large volume of documents written in two or more languages (training data) to construct translation models. Conventional machine-translation systems utilize rule-based models, which are based on sentences’ grammatical structure and a word dictionary. Alternatively, models that use machine learning incorporate computers to produce translation rules from a large volume of documents provided as input, such as corresponding documents between English and Japanese for a Japanese-English translation. In other words, computers automatically parse language as a basis for translation rules to replace those developed by linguists. Thus, we can consider this case as an example of AI, as this involves a computer replacing human thought.

In the image-recognition field, the news that “a computer has recognized a cat” quickly circulated worldwide based on a paper presented by Google in 2012. When machine-learning was performed by randomly extracting 10 million images from videos uploaded on YouTube, the computer could automatically select images containing the features of cats. Rather than a collection of cat images given to a computer as training data, this case is characterized by the fact that only 10 million images were given as input to conduct unsupervised learning. As it is necessary to use a human’s recognition ability to prepare training data, the supervised machine-learning model is developed through a cooperation between humans and computers. However, Google’s unsupervised image recognition model is revolutionary, as the model was developed without human input.

Machine-learning models for automatic translation and image recognition are constructed using the deep-learning method, which involves a multi-layered neural network, or a classical mathematical method with decades of history. Conventional deep learning ideas have involved creating a multilayered network layer, but this was problematic in that it was difficult to estimate its parameters, which increase as the multilayered network is created. Additionally, a computer’s abilities and performance are insufficient. However, deep learning has been reexamined in recent years, and AI studies are now popular as computer performance has improved and a large volume of information has been compiled on the Internet; this enables the utilization of Big Data when estimating models. Estimation methods have also been developed in recent years for respective types of data—such as image versus text data—and characters. Ultimately, these are implemented in various fields, including industrial applications, such as industrial robots and autonomous operation technologies; investment decision-making for financial institutions; financial advisory work; and in household appliances, such as cleaning robots and AI speakers.

(3) The Internet of Things

The sensor information from Komatsu's construction machinery as described in the Big Data example involves data originating from an object rather than people. Information on physical things—including electronic devices, automobiles, production equipment, household appliances, and industrial machinery—has been exchanged over the Internet, which has been called the Internet of Things (IoT). The IoT aims to spread opportunities for new business innovations once various equipment and devices are connected on the Internet.

The realization of IoT requires various elements—identification, or the labelling of each item using an IP address; sensing, or the measurement and "datafication" of the item; communication, which primarily involves data communication; computation, or an analysis of the item's data—and its implementation as a specific service, such as the maintenance and operation of industrial machinery or buildings' energy management systems (Al-Faqaha et al., 2015). The Big Data concept focuses on the data of things exchanged over the Internet, and AI can be considered as a key component of the technologies used in the aforementioned data analysis. Accordingly, the IoT concept can be considered more comprehensive, as it includes those technical elements as well as an implementation in society in the form of services.

It has been demonstrated that IoT can provide various solutions through Internet information by connecting various things respectively through a sensor network. It is believed that one trillion things, or 100 times the human population, could be connected by the year 2020. As networking expands from people to "things," the data volume will also dramatically increase. As it is unrealistic to exchange all information through the Internet, "edge" computing has also attracted attention, as this forms local networks and performs distributed processing. This can enable more expanded applications through the aggregation of a certain level of information from those local networks and the connection of "things" in wider areas through the Internet. Consequently, information of all kinds will become connected worldwide through the Internet.

The network is comprised of systems, including such wide-area networks as mobile nets, local area networks, various hardware devices as the cores of communication devices, and software. Further, a service layer exists for each device in a communications infrastructure, such as automobiles, household appliances, and industrial machinery. Innovations using AI are primarily advancing in terms of this service layer.

While standardization organizations—such as the ESTI and IEC—have advanced the technical standardization of communications infrastructures, the service layer has primarily developed through the activities of a consortium of private companies, or collective entities formed through companies' collaborations. For example, Apple Inc. has entered the field of household appliances by creating its standard HomeKit, and engages in collaborative activities with manufacturers related to home automation and network devices. Such activities are based on the idea of using their iPhone as a control hub to illuminate homes and manage home appliances and home security systems.

Regarding such industrial machinery as aircraft jet engines and power-generating gas turbines, manufacturers actively engage in such activities. Among these, General Electric launched its GE-Data as a cross-sectional organization for their business departments—including jet engine, energy and wind-power generating, healthcare, and railroad systems—and have developed an IoT system platform called "PREDIX." The company also started an Industrial Internet Consortium with various information communication companies, including IBM, Cisco, Intel, and AT&T, and has developed activities to expand their IoT platform to fields other than those pertaining to their business.

(4) Relationships among these New ITs

A mutually complimentary relationship exists between AI and IoT, as well as Big Data as discussed above. In addition to the substantial volume of image or text information, among other types, which has originated from humans and been compiled on the Internet, data has also been compiled that originates from things due to the advancement of IoT sensor network technology. This has markedly enhanced the three "Vs" of Big Data: volume, velocity, and variety. For example, the data volume generated daily during a factory's production process can be calculated as (the number of products in a production line) x (the number of production steps) x (data granularity). Here, "data granularity" refers to the data acquisition frequency per the unit of time (e.g., per minute). Accordingly, the IoT service produces data with volume and velocity, exceeding human's physical capabilities. Additionally, audio and image data can be generated from various sources, such as surveillance cameras for security services or self-driving cars.

As previously mentioned, the expansion of Big Data's potential uses has significantly contributed to the development of different AI technologies. Gathering various data types that are unsuitable for information processing as-is, such as audio, image, and text data, have improved perceptual and recognition technologies, including audio/image recognition and natural language processing. Additionally, using a large volume of data has resulted in remarkable advancements in various AI-related basic technologies for knowledge and discovery techniques and deep learning, such as information-searching and data-mining.

(Figure 1)

Finally, these AI-related technologies are key components in realizing every sort of IoT service, such as smart factories, smart appliances, and smart cities. A large volume of Internet information and sensor information has been made available, although such Big Data is not intentionally collected for a particular purpose. Specifically, IoT sensors naturally collect such data, which are perceived, interpreted, and realized as economically valuable systems, which is why these are labelled as "smart." In other words, as this differs from the human-led conventional form of data processing provided to computers, computers could provide services with much less human involvement. The innovations enabled by Big Data, AI, and the IoT are characterized as new services, which would not require service recipients, people, or companies to function, and which would further benefit their respective entities.

Specifically, this will not only make routine tasks more efficient, which is the strength of IT, but also enable the performance of non-routine applications. For example, the primary tasks of IT system applications in corporate finance/accounting and human resources typically involve financial/accounting computations and human resource databases. However, such finance or accounting computations may be performed in the future according to revised accounting standards and a personnel-allocation recommendation system according to the characteristics of the company's positions. Moreover, IT has been increasingly used in sales departments' customer information management. In this case, IT could support more strategic marketing activities with the development of new customers and determination of customer segments. Further, efforts made for IoT services—such as smart household appliances and smart cities—can be considered as the initial growth of service innovations beyond existing business models. Upon improving productivity, or the output per unit of input, conventional IT is often considered a tool to decrease input, or in making the existing work more efficient. However, IT should be recognized as a technology to enable expanding output in considering the technological advancements in Big Data, AI, and the IoT.

This will also enable us to grow beyond the business-by-business structure, and realize innovation involving a wide range of players who can work across industries. Although still in an empirical study stage, smart cities are a typical example of collaborative work among automobile and home appliance manufacturers as well as power companies, among others. For example, the Nest thermostat offers one practical application, as it learns residents' living patterns to automatically control household appliances' and equipment's ON/OFF functionality, as well as Google's Waymo self-driving car project. Accordingly, IoT services can be achieved by mutually connecting "things" that currently exist independently. Therefore, manufacturing companies should work with IT companies and venture companies that specialize in specific technical fields, including artificial intelligence technology. This will consequently bring new innovations, such as those described above, and may also eliminate conventional manufacturers' business models, in which profits are made by producing and selling things. As various household appliances—such as the Nest thermostat—become connected to a controller, their manufacturers will lose contact points with their end customers. In other words, those manufacturers will decrease in status to become merely suppliers of smart household appliance services, in providing only some parts to the entire household system. Automobile manufacturers will likely experience a similar decrease in status to become merely a transportation service provider with the increase in self-driving cars. It is highly likely that the IoT wave will significantly impact the business structures of manufacturers that handle “things.”

3. Measuring AI science and innovation

This section provides some statistical evidence of the interactions between science and innovation, with a specific focus on AI. Clearly, new ways of using Big Data, such as deep learning, have changed how business innovation is organized. Additionally, deep-learning techniques are used as tools to invent new products or innovate, as can be observed in the IBM Watson’s application in new drug development (Nayak et al., 2016). Thus, AI can not only serve new business applications but also act as a method of invention (Cockburn et al., 2018).

Another special feature of AI is found in its development style, as public science sectors are involved in such development. For example, the deep-learning concept (deep neural network) is not new, but academia has combined its actual implementation with a new methodology and powerful computing capabilities. Subsequently, computer science scholars have developed a series of deep-learning algorithms that are suited to various datasets, such as CNN for image data and RNN for text data. Additionally, the private sector has also substantially contributed to AI development. A typical example is Google Brain’s publication of “Alpha *Go*” (Silver et al., 2017); the Deep Mind team currently under the Google Brain AI department not only developed software to beat the world’s *go* (Chinese chess) champion, but also made this public as a research paper. Simultaneously, innovations are flourishing, such as the economic valuation of public technologies as can be observed in the previous section on IoT applications and the “smart” technology wave, as well as the mushrooming of startup firms that provide a specific AI technology with its associated fees. The tremendous speed of technological progress and business development has led both the public and private sectors to co-mingle in this process. Therefore, AI is a typical example of a science-based economy, in that scientific findings published for free and innovation in commercial activities co-exist as a result of the public sector’s cross-over style among universities and public research institutions, and the private sector as formed by firms.

This paper adopts a methodology that involves linking research articles and patents, or specifically, finding identical author-inventors in research article and patent databases (Ikeuchi et al., 2016), which approximates the idea of finding patent papers with similar content (Lissoni et al., 2013). Traditionally, an industry’s degree of scientific basis, or its “scientific intensity,” has been measured using non-patent literature, such as research articles with patent citations (Narin and Noma, 1985; Schmoch, 1997). Non-patent literature citations indicate the degree of disembodied scientific knowledge that flows into patents, while the patent-publication pair can capture the co-occurrence of scientific and invention activities among the same researchers, or the interplay of science and technology as embodied in human capital.

I used Elsevier's Scopus database for research articles, and data from the United States' Patent and Trademark Office (USPTO) for patents. In both datasets, I selected researchers working for organizations located in the United States, resulting in approximately eight million papers from Scopus and three million patents from the USPTO's data. These two datasets are linked by author-inventor names as well as their affiliates, and approximately 5% and 13% of all authors from Scopus and USPTO data can be linked, respectively (Motohashi, 2018).

This linked dataset of research articles and patents can then be used to identify AI papers and patents. Regarding AI papers, Elsevier's All Science Journal Classification (ASJC) is used; concretely, ASJC 1702 is an "artificial intelligence" label under a broader computer science category, which is used in this paper to identify AI papers.[2] In terms of patents, the G06N IPC code is used to parallel the JPO's technology trend survey publication (JPO, 2014); G06N signifies technologies that involve "computer systems based on specific computational models, such as neural networks, inference machinery, and fuzzy logic." This definition narrowly defines the AI concept, in a sense that this only includes basic methodologies for analytics and models.[3]

Figure 2 displays the trends regarding the AI papers and patents' proportion to all works. Both exhibit an upward trend until 2010, but stabilize afterwards. It should be noted that the USPTO only discloses information on granted patents; therefore, only data applied until 2011 can be used, with the datasets obtained from the USPTO's data download site (http://patentview.org) in 2016. Figure 1 also presents trends among AI patent applications, which have a relatively small truncation bias. A surge of AI patent applications can be observed after 2010, and it should also be noted that the shares of AI papers and patents are small—less than 1% of all papers and patents—as only core AI technologies are included in both definitions.

---

[2] The ASJC classification is made at the journal level, instead of the individual paper level, so that extracting AI papers by ASJC code will not capture the emerging trend of AI among non-AI-related journals. Therefore, the AI paper trend is checked for robustness using keyword matching in the titles of individual papers using keywords from work by Coburn et al. (2017). It was found that the share of AI paper does not substantially differ from that of the ASJC.

[3] Another approach involves taking a broad category, such as G07F, which includes general purpose computer software and applications (OECD, 2013).

(Figure 2)

Next, I examine the AI author-inventor's contribution to the aggregated trends among AI papers and patents to discover that such crossover researchers' shares of AI papers and patents are greater than those from pure authors and pure inventors, respectively (Figure 3). The difference in these two groups is particularly significant in AI patent shares; specifically, it is observed that AI scientists who also contributed to research article publication activities increasingly appropriated AI technology through patents.

(Figure 3)

I then decompose the author-inventor category by affiliation, or the whether their affiliation is in the private (for-profit firms) or public sector (for non-profit research organizations, such as universities and government laboratories). Figure 4 provides the share of AI patents/papers by researcher affiliation, and by whether an author-inventor exists, versus only an author or inventor. This figure reflects researchers' relative importance by type in the above AI patents and papers, normalized by the general trends across patents and papers during our observation period. Regarding the public sector, the author-inventor AI shares are higher than those for pure authors or inventors in both papers and patents, respectively. Therefore, relatively larger numbers of researchers involve both publication and patenting activities in this sector. Alternatively, the pure inventor's share is greater than that of the author-inventor for AI patents by private firm, while the author-inventor's contribution is greater in AI papers. This can indicate that substantial numbers of patents in the private sector are application-oriented, and invented by pure inventors, or those uninvolved in publication activities.

(Figure 4)

Finally, AI publication and patenting's impact on other fields is investigated. Figure 5 illustrates the share of patents, indexed by technology field then by AI inventors and authors. The "computer technology" category that includes G06N is excluded from this chart to determine how other technologies widely use AI.[4] It should be noted that the patents invented by AI authors have wider applications across the technology field than those from AI inventors. Differences between the two figures can be particularly observed in the "measurement," "medical technology," "transport," and "organic fine chemistry" categories. These findings reveal that AI publications play an important role in AI's wide applications across various technology fields.

[4] The share of computer technology is 53.2% for AI inventors and 41.6% for AI authors.

(Figure 5)

4. Discussion

The innovation caused by information technology can be observed across business sectors and in various business arenas. Its application also extends from a system that supports firm functions, such as personnel and financial accounting, to a wide range of operations, including production, customer, and supply chain management. Artificial intelligence is a core technology to support the application of these various information technologies. Economists have analyzed the general purpose technology (GPT) in such IT over the past 20 years (Bresnahan and Greenstein, 1996; Helpmann, 1998). For example, backbone systems related to factory operations in the manufacturing industry versus banks' financial transactions involve the information system performing completely different tasks. However, improved computer performance can realize more advanced tasks in both industries, and thus, information technology that geometrically progresses by Moore's Law will significantly benefit the entire economy. As a new methodology in computer modeling and analytics, AI clearly provides new possibilities for business applications, but what is new compared to information technology in general?

First, literature discusses the GPT feature in information technology. General purpose technology is characterized by its speed, pervasiveness, and downside applications (Helpmann, 1998). A driver of IT speed is Moore's Law, applicable to semiconductors, and developing both operational and application software enables the pervasive use of computers in various industries. However, IT investment does not automatically improve business performance and productivity; this requires organizational innovation, such as business process changes (Bresnahan et al., 2002). Therefore, IT is merely an "enabling" technology to create innovation that can be applied to businesses and impact management. For example, general-purpose technology can be used to innovate by manufacturing users' application to highly control factory production processes, or by a bank's biometric authentication system to improve ATM transaction security. Essentially, the "new IT" involves AI, Big Data, and the IoT, as discussed in Section 2, and extends general purpose information technology. In other words, these enable the realizing of an innovation, and not the innovation itself. Businesses must innovate and to realize these innovations by integrating them into activities with user-side business value.

However, some differences exist in its innovation organizations. The key technology in traditional IT (Figure 6, left-hand panel) to link computer technology to user innovation is the software bundle, and operational software plays a particularly important role in enabling various application software developments. This proprietary operational software is comprised of a platform and upstream platform leaders; for example, Intel and Microsoft in the personal computing industry drove entire systems of innovation (Gawer and Cusumano, 2013). Consequently, a division of innovative labor can be found between platform leaders and downstream user-innovation players (Gambardella and McGahan, 2010).

(Figure 6)

In contrast, such platform leaders cannot be found in the new IT environment (Figure 6, right-hand panel), or at least not at this moment. Two fundamental changes have occurred in traditional IT, as the left-hand panel in Figure 6 illustrates: First, the smartphone platform has emerged. While Intel has exhibited platform leadership in the personal computing sector, multiple platforms exist on the smartphone, such as Android and iOS. Second, Moore’s Law drives the entire GPT environment in IT, but this will end in the near future (Cross, 2017). Further, a substantial cloud computing Internet environment has emerged, which has replaced both the computer and OS environments in traditional IT, as the right-hand panel in Figure 6 demonstrates.

The development of AI technology—as well as Big Data, which has made such AI applications available—is currently positioned around the cloud computing environment and user innovation. The private sector’s role here has increased over time, in open scientific articles as well as proprietary patents, as the previous section indicated. We can also posit that the interactions between open science and proprietary technology have intensified, and no platform leader has clearly emerged in the new IT environment, where no single firm can play a leadership role, but rather, must intensify ecosystem-building activities among many players.

Naturally, some firms act like platformers, such as Google and Amazon. However, while Google dominates public Internet information, an expansion of the IoT will impede their role as a platform leader. The critical component in IoT business innovation is domain-specific proprietary data, such as autonomous driving records, and this domain has expanded to include a variety of fields: consumer electronics, medical treatments, smart factories, and financial services, among others. In this sense, the value of user innovation comes from the inter-connections among open and proprietary technologies, data, and business knowledge. As an innovation becomes more structurally complex, various eco-systems are created instead of platform businesses as ventures to fill the gaps that have emerged in incumbent firms’ capabilities.

The advancement of machine-learning techniques in AI as well as Big Data availability has enabled substantial downstream innovation. In this sense, Figure 5 illustrates such versatility, which can support the perspective of AI as a new GPT. However, it is unclear whether AI could replace

Moore’s Law as computer technology continues to progress. Deep learning could be a one-time event, and new innovation opportunities may cease to exist when all potential applications are exploited. Alternatively, we could observe continuous growth in new methodologies to improve predictions’ rates of precision, which could lead to its increased application. Currently, it is difficult to predict which scenario will occur, but we can at least posit that AI’s momentum in the innovation landscape will not cease in the near future.

5. Conclusion and implications

This paper primarily analyzes AI-driven innovation, and a linked dataset of AI research articles and patents reveals that the public sector substantially contributes in AI development. Moreover, the role of researchers involved in both publication and patent activities, and particularly in the private sector, has increased over time. Namely, open science—which is publicly available through research articles and propriety technology, and is protected by patents—has become intertwined in AI development. Additionally, this paper discussed the impact of AI, combined with Big Data and the IoT, which has defined “new” IT innovation. This contrasts traditional IT, which has consisted of the technological progress of computer hardware and software developments. Both new and traditional IT can be understood by using a general purpose technology (GPT) framework, while new IT innovation can be organized and characterized by the emergence of multiple ecosystems, instead of a pattern of platform leadership as found in traditional IT.

This paper concludes by providing some implications for future research. First, regarding management implications, the ecosystem strategy should be emphasized. The business ecosystem consists of a “keystone,” which plays a central role in the ecosystem, or businesses’ inter-company relationship network, and other niche players (Iansiti and Levien, 2004). The keystone’s role is to attract many niche players and spread the entire ecosystem. Alternatively, “niche” players exist with each original technology, and contribute to the entire ecosystem’s diversity; the ecosystem is maintained with these mutually complementary relationships. For example, Apple Inc. has an app store on its iPhone, and is developing services to meet its consumers’ diverse needs. Combined with the common various services (platform) provided by individual application providers (niche players) and Apple (the keystone), this will create overall value throughout this ecosystem. The keystone’s role is to improve the entire ecosystem’s business value, and thus, it is important to build mutually beneficial relationships with its niche players. Strengthening control over these niche players and continuing to exploit their value will ultimately destroy the ecosystem. Consequently, attractive management resources must be created to attract diverse niche players and increase the value of the entire ecosystem.

Along this line, the most important strategic question involves which direction to pursue in a new IT ecosystem, whether keystone or niche. New IT users tend to be niche players, and operate based on their own business knowledge and proprietary data. However, a user firm can be a keystone player by providing common resources to other firms within a certain business domain. For example, General Electric has chosen a keystone strategy its IoT business in the manufacturing industry by offering its PREDIX platform to serve as a basis for IoT industrial applications, such as the condition-based maintenance of large industrial equipment. An IT technology firm typically seeks this keystone strategy instead of platformers by building mutually beneficial relationships with niche players and nurturing their entire ecosystem.

Policymakers should be aware of the interplay between open science and propriety technology in the AI field. Therefore, policies supporting AI innovation will not only provide funds for basic science, but will also investigate a crossover model between the public and private sectors. For example, it is more suitable to promote university startups than policy instruments based on traditional linear models, such as the licensing of university patents. Additionally, non-profit public research institutions can develop new ecosystems, as such bodies cannot directly and commercially compete with private firms seeking mutual collaborations. In this regard, the university’s role should be expanded, from a place for research and education to a center for entrepreneurial activities and innovative experimentation.

Figure 1: Inter-relationships of AI, IoT and Bigdata

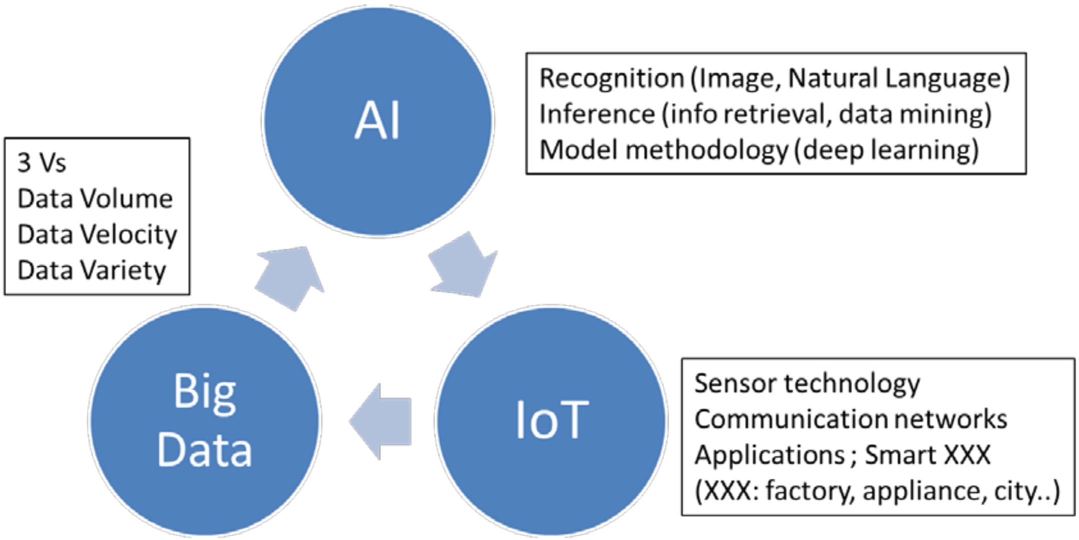


Figure 2: Share of AI papers and patents

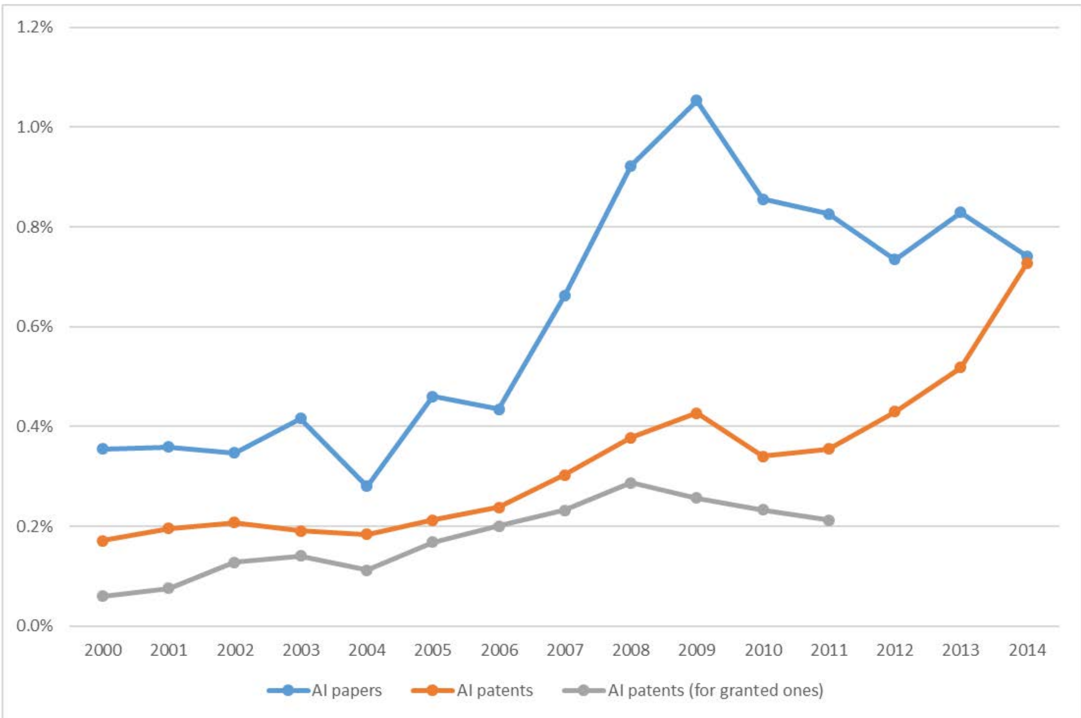


Figure 3: Share of private sector authors

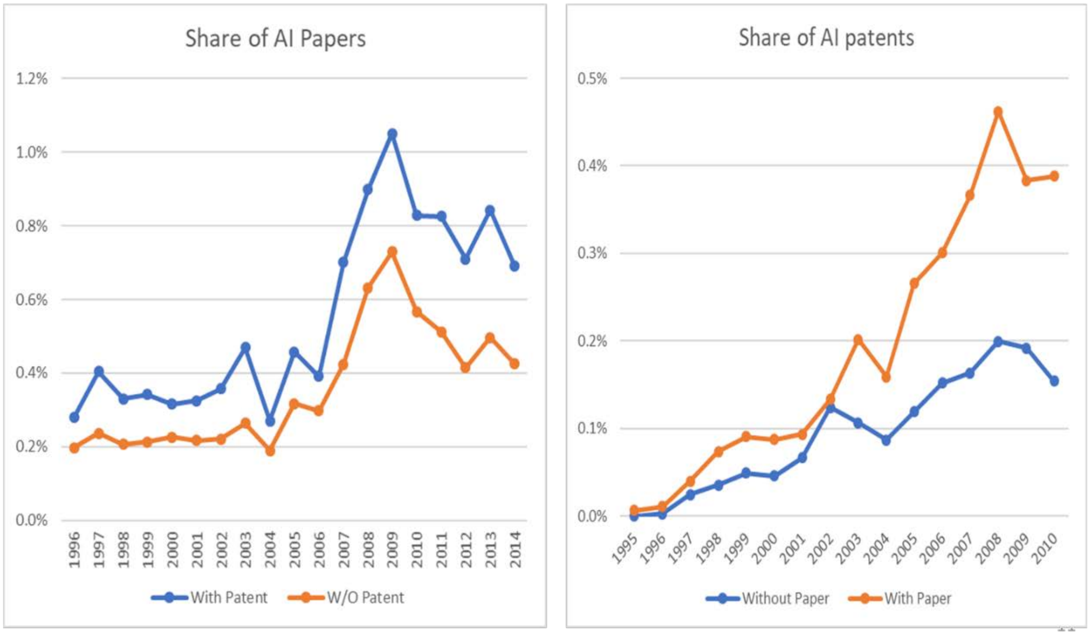


Figure 4: Interaction of author and inventor by type

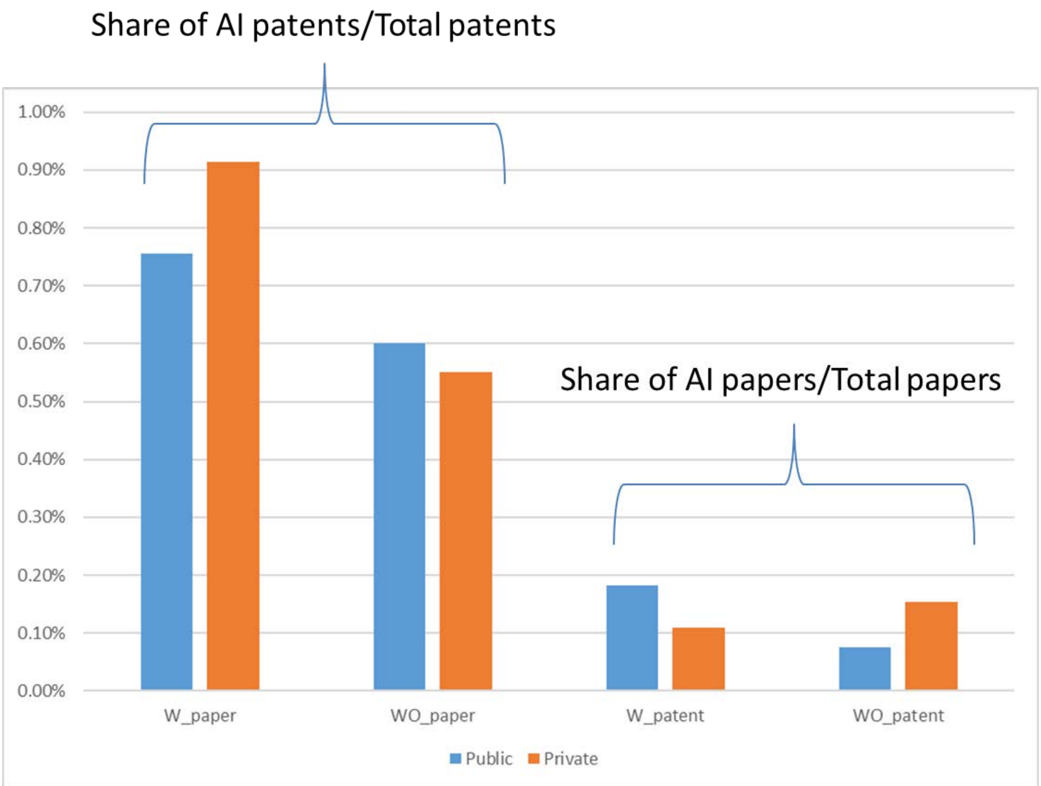


Figure 5: AI author and inventor patents by technology (except for computer technology)

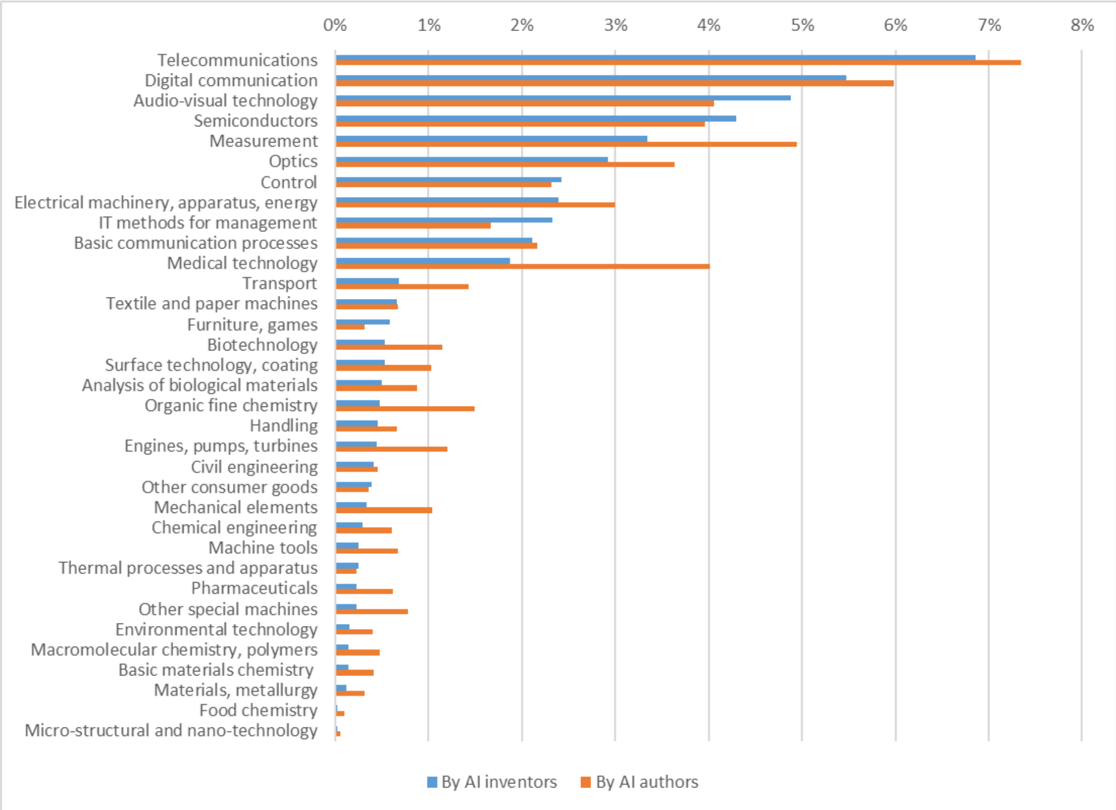
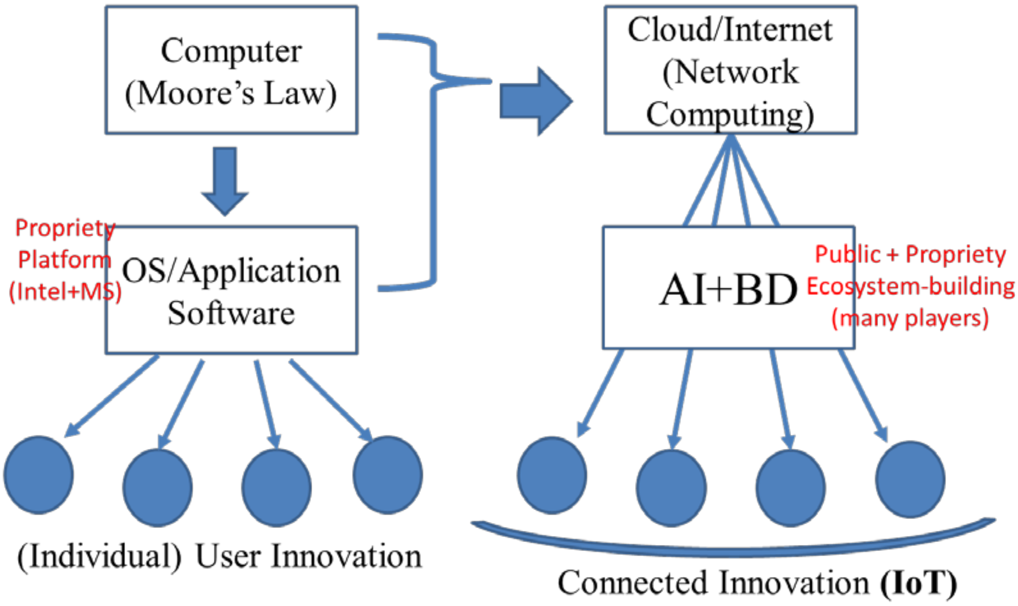


Figure 6 : Traditonal IT vs AI/BD/IoT



References

Al-Fuqaha, A., Guizani, M. Mohammadi, M., Aledhari, M. and M. Ayyash (2015), Internet of Things: A Survey on Enabling Technologies, Protocols and Applications, IEEE Communications Surveys and Tutorials

Agrawal. A., J.. Gans and A. Goldfarb (2018), Prediction Machines: The Simple Economics of Artificial Intelligence, Harvard Business School Press, April 2018

Arora A., S. Belenzon and A. Patacconi (2015), Killing the Golden Goose? The Decline of Science in Corporate R&D, NBER Working Paper #20902

Breshanan, T., Brynjolffson E, 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

Breshanan, T. and S. Greenstein (1996), The Competitive Crash in Large-Scale Commercial Computing. In *The Mosaic of Economic Growth*, edited by Ralph Landau, Timothy Taylor, and Gavin Wright, pp. 357-97. Stanford University Press

Cockburn, I., Henderson, R. and S. Stern (2018), The impact of artificial intelligence on innovation, NBER working paper #24449, March 2018

Cross, T. (2017), Beyond Moore's Law, in Magatech : Technology in 2015, D. Franklin ed. , Economist Book, Profiles Book Ltd

Gambardella, A. and AM McGahan (2010), Business-model innovation: General purpose technologies and their implications for industry structure, *Long range planning*, 43, 262-271

Gawer, A. and M. Cusumano (2013), Industry Platforms and Ecosystem Innovation, *Journal of Product Development and Innovation Management*, 31(3), 417-433

Helpman, E (1998), *General purpose technologies and economic growth*, MIT Press, Cambridge MA

Iansiti,, M. and R. Levien (2004), The Keystone Advantage: What the New Dynamics of Business Ecosystems Mean for Strategy, Innovation, and Sustainability, Harvard Business School Press, Boston, MA

Ikeuchi, K, K Motohashi, R Tamura, and N Tsukada (2017), "Measuring Science Intensity of Industry using Linked Dataset of Science, Technology and Industry", RIETI Discussion Papers Series

17-E-056.

JPO (2015), Artificial intelligence technology: 2014FY Report on Technology Trend, Japan Patent Office, March 2015 (in Japanese)

Kunikawa, S., T. Tanaka, K. Nishio and K. Motohashi (2015), Innovation Trend and Case Studies Using Big Data Analysis, RIETI Policy Discussion Paper 15-P-015, October 2015 (in Japanese)

Lissoni, F, F Montabio, and L Zirulia (2013), "Inventorship and authorship as attribution rights: an enquiry into the economics of scientific credit", Journal of Economic Behavior and Organization, 95, 49-69

Mayer-Schonberger, V. and K. Cukier (2013), *Big Data: A revolution that will transform how we live, work and think*, John Murray Publisher, Great Britain

Motohashi (2018), Co-occurrence of science and innovation in AI : Empirical analysis of paper-patent linked dataset in the United States, NISTEP working paper No. 160 (in Japanese)

Motohashi (2014), The Sun Also Rises High: Regenerating Japan's industrial competitiveness, Nikkei Inc. (in Japanese)

Narin, F, and E Noma (1985), "Is technology becoming science?", Scientometrics, 7, 368-381.

Nayak, V. S., Khan, M. S., Shukla B. K. and P. Chaturvedi (2016), Artificial intelligence in clinical research, International Journal of Clinical Trial, 3(4): 187-193

OECD (2013), Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by Big Data, *OECD Digital Economy Papers*, No. 222

Schmoch, U (1997), "Indicators and relations between science and technology", Scientometrics, 38 (1), 103-116

Silver D., J. Schrittwieser, K. Simonyan, I.Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel and D. Hassabis (2017), Mastering the game of Go without human knowledge, Nature Vol 55, 19 October 2017