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

Kazuyuki Motohashi

CENTER FOR FRENCH-JAPANESE ADVANCED STUDIES IN PARIS
CEAFJP DISCUSSION PAPER #18-04

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

Kazuyuki Motohashi

University of Tokyo, NISTEP and RIETI, Japan
2017 CEAFJP/Michelin Fellow

November 2018

Center for French-Japanese Advanced Studies in Paris (CEAFJP)
Fondation France-Japon de l'EHESS
54, boulevard Raspail 75006 Paris - ffj@ehess.fr

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

Kazuyuki Motohashi

Abstract

The linked dataset of AI research articles and patents reveals that substantial contribution of public sector is found for AI development. In addition, the role of researchers who are involved both in publication and patent activities, particularly in private sector, increased over time. That is, open science, publicly available by research articles and propriety technology, protected by patents are intertwined in AI development. In addition, the impact of AI, combined with big data and IoT, defined "new" IT on innovation is discussed, by comparing with traditional IT, consisted by technological progress of computers and software developments. Both of new and traditional IT could be understood by using the framework of GPT (general purpose technology), while the organization of new IT innovation can be characterized by emergence of multiple eco-systems, instead of the pattern of platform leadership, found in traditional IT.

Keywords

Patent data, AI, General Purpose Technology

Acknowledgment

This paper was developed while the author visited FFJ-EHESS as a Michelin Fellow from April 2017 to March 2018. The author would like to thank FFJ and Michelin for their financial support. In addition, the author would like to acknowledge the support of OECD for providing an office space and intellectual inputs, particularly Dominique Guellec, head of science and technology policy division. The statistical analysis in section 3 is drawn from the NISTEP Project on innovation process database. The author also thanks the participants of various workshops for their helpful comments, 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 digital innovation era (March, 2018) and the MSI seminar at KUL (May 2018). All views expressed in this article are those of the author and not of his affiliated organizations.

Contents

1. Introduction

There is no doubt that the importance of scientific knowledge for the industrial innovation process is growing. Genome science is significantly changing the pharmaceutical industry's R&D process, while understanding of the physical properties of materials at the nanoscale level has become essential to the miniaturization of the large scale integration circuit (LSI) fabrication process. Information technology has had a huge social and economic impact, and big data analysis has brought deeper insights for business and management activities. In particular, the advance of data science and AI contributes to scientific understanding of business process, which can be applied to whole economy. Therefore, scientific findings used to be applied to only specific sector, such as pharmaceutical industry, while recent trend of economy wide impact of scientification of innovation can be called as the science economy, instead of science based industries (Motohashi, 2014).

As a distance between science and innovation becomes shorter, the relationship between these two activities are shifting from the linear model (scientific knowledge is first gained at research institutions such as universities and then used by companies to develop new products) to co-occurrence model, where scientific and innovation activities occur simultaneously by interacting with each other. With this regards, we have developed a new indicator of science representing the co-occurrence of science and innovation, based on the linked datasets of patents and research articles by the same author/inventor (Ikeuchi et, al, 2016), as a complementary information to the indicator based on research paper citation of patent, a traditional science linkage indicator based on the linear model.

In this paper, the nature of AI (Artificial Intelligence) driven innovation is analyzed, based on the author/inventor linkage data of US patents and research articles. There is a growing attention to AI, not only in scientific sector such as university, but also in industry. AI is perceived as a key technology causing innovation landscape fundamentally by contributing to make what is coming in the future cheaper and more certain (Agrawal et. al, 2018). Therefore, private incentives to capture such potential values are substantial. Another characteristics of AI is that a potential application of AI can be found in wide area of industries. In other words, AI is a killer application of IT as general purpose technology. Furthermore, AI can be called as IMI (the invention as a method of inventing), in a sense that AI can be used in a process of inventing new applications such as autonomous driving, condition based maintenance and new drug discovery (Cockburn et. al, 2018).

This paper is organized as follows. The next section described the concept of AI driven innovation. Here the interrelationship between AI and the complementary elements to innovation, big data and IoT is discussed. Then, an analytical part by using the linked dataset of research articles and patents is provided. This section is followed by the discussion, focusing on the difference of AI driven innovation, or "new" IT innovation. Finally, this paper concludes with some managerial and policy implications.

2. Conceptualization of AI driven innovation

There has been significant progress in the field of AI, such as the construction of new methods of machine learning (e.g., Deep Learning and Generative Adversarial Networks), but the basis of such technologies, neural network, is not new at all. The progress in developing methods has been made possible due to the improved performance and capabilities in computing and the availability of big data. More specifically, the emergence of big data and developments in AI-based science innovation are inseparable. In addition, the concept of "IoT (Internet of Things) is also important for AI technologies that yield results as industrial applications such as AI speakers and automated driving systems. Here, we will discuss the necessary elements for AI-based innovations such as big data, and IoT, and discuss about mutual complementary relationships of these two items and AI.

2.1. Big data

Managerial issues in companies are being solved by data analysis more than ever before. What is new in utilization of big data? Companies have conventionally used data accompanying a business system. While corporate financial data are accumulated in a financial accounting system, data such as that related to personal histories and employee salaries are accumulated in human resource systems. On the other hand, a supply chain management (SCM) system manages stock status regarding materials and products and order records as data. Such data are respectively characterized as being made for a specific reason. Since the 1990s, each business system has been increasingly integrated as an optimum resource management (ERP: Enterprise Resource Planning) system for the whole company. Such systems place an emphasis on improving the operational efficiency of corporate activities.

On the other hand, big data is characterized by not being constructed for a specific use. Amazon is frequently cited as a company that played a pioneering role in use of big data. In this case, one example of big data is users' purchase records. Using this information, Amazon makes book recommendations matching to the interests of individual users. Purchase record data related to merchandise such as books are automatically accumulated by companies that engage in e-commerce. Such data is not specifically collected to make recommendations to customers regarding books they would be highly likely to purchase. In addition, to make highly accurate recommendations, a certain data size is required (including the number of users and the number of purchase histories). This is because the method of making recommendations is calculated based on a stochastic model for estimating what books the customers would be highly likely to purchase based on the characteristics of their purchase histories. As the number of samples increases, the accuracy of recommendation improves. In other words, the value can be said to 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 searching keyword specific advertisement, using their search engine. Data that has been used in IT systems thus far has primarily been numerical data. However, on the internet, there are various forms of data such as audio and visual data, as well as text data. Big data is also characterized by the difference from conventional IT use in the value added by converting data therefrom (Datafication).

Furthermore, these days, in addition to the development and use of the internet, data is also collected using various kinds of sensor information to add economic value. For example, Komatsu has had its construction machinery equipped with GPS systems and fuel gauges to collect location data and operation status data from their construction machinery. Using such data, they can provide services such as anti-theft functionality and make recommendations for fuel cost savings, which will assist in differentiating their products from those of their competitors. The sensors for this type of data collection collect data from various places, such as people's travel data from cell phones and onboard GPS systems, for example, as well as the operational status of all types of industrial equipment. The use of such sensors is being actively developed (Kinukawa et. al, 2014).

As described above, with e-commerce purchase records, online information, and various information obtained from sensors in addition to the data produced from business systems within companies, various types of data are available. Big data are characterized by the "3Vs" (i.e., their large data size (Volume), datafication of various information including text, images, and audio (Variety), and the continuous daily inflow of data from the internet and sensors (Velocity)). In addition, as a method of data utilization, it is possible to see various events related to business management, society, and human behavior in a wider and more comprehensive manner, and analyze the mechanisms in detail at a micro level.

2.2. Artificial Intelligence (AI)

AI is a concept in which technologies are integrated to exploit beneficial information in business management from big data. Applied AI technologies may be categorized into image or text data recognition technologies (conversion into computer readable data), human interfaces (visualization of data and interactive agents), knowledge discovering technologies related to the diagnosis, monitoring, and datamining of various types of equipment devices, and so on. In addition, machine learning, fuzzy control, and mathematical models such as genetic models have been implemented on systems as basic technologies to realize those functions.

In particular, due to the existence of big data, significant technological progress has been made in machine learning methodology. Model estimation methods for machine learning are generally grouped into supervised learning and unsupervised learning. In either method, the 3Vs of big data are important. 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 provides translation services by having their translation system read 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 is based on grammatical structure of sentences and word dictionary. On the other hand, in the models that use machine learning, computers produce rules for translation from a large volume of documents (corresponding documents between English and Japanese for Japanese-English translation, for example), which are provided as inputs. In other words, computers automatically perform language parsing work, which is the basis for translation rules in place of the one developed by linguists. In this case, since human thinking is replaced by a computer, we can consider this to also be one example of AI.

In the field of image recognition, the news that "A computer has recognized a cat." quickly circulated around the world based on a paper presented by Google in 2012. When they performed machine learning by randomly extracting 10 million images from videos uploaded on YouTube, the computer was able to automatically select images containing the features of cats. In this case, rather than a collection of images of cats being given to a computer as training data, this is characterized by the fact that only 10 million images were given as inputs to conduct unsupervised learning. To prepare training data, it is necessary to use human recognition ability. Therefore, in the supervised model of machine learning, the model is developed through cooperation between humans and computers. However, Google's unsupervised image recognition model is revolutionary since the model was developed without human input.

Machine learning models for automatic translation and image recognition are constructed using the deep learning method. Deep learning is a method of machine learning that uses a multi-layered neural network. A neural network is a classical mathematical method with decades of history. There have conventional ideas of deep learning to make a multilayered network layer; however, there was a problem in that it was difficult to estimate parameters, which increase through the creation of a multilayered network. In addition, the abilities and performance of a computer are insufficient. In recent years, deep learning has been re-examined, and AI studies are now hot spots since computer performance has been improved, and a large volume of information has been compiled on the internet, which enables big data to be utilized when estimating models. In recent years, estimation methods have been developed for respective types of data (e.g., image data or text data) and characters, and 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 household appliances such as cleaning robots and AI speakers

2.3. Internet of Things (IoT)

The sensor information for Komatsu's construction machinery described in the big data example is data originating from an object rather than people. Information on physical things has been exchanged over the internet, which is called the IoT (Internet of Things). Those things used to exchange information include electronic devices, automobiles, production equipment, household appliances, and industrial machinery. The point of IoT is that, once various equipment and devices are connected on the internet, opportunities for new business innovations spread.

The realization of IoT requires elements such as the identification of each item with an ID like an IP address (Identification), measurement and datafication of the item (Sensing), data communication (Communication), data analysis related to the item (Computation), and implementation as a specific service (e.g., maintenance and operation of industrial machinery and energy management systems for buildings) (Al-Faqaha et al., 2015). Big data is a concept focusing on the data of things exchanged over the internet, and AI can be considered to be a key component of technologies used for the aforementioned data analysis. Accordingly, the concept IoT can be considered to be a more comprehensive concept including those technical elements and, furthermore, including implementation in society as services.

It has been demonstrated that IoT can provide various solutions with information via the internet by connecting various things respectively through a sensor network. It is believed that one trillion things (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. It is not realistic to exchange all information through the internet, so edge computing, which forms local networks and performs distributed processing, has also attracted attention. Here, through the aggregation of a certain level of information from those local networks and the connection of “things” in wider areas through the internet, even more expanded applications can be enabled. As a result, the world will become a place where information of all kinds will be connected via the internet.

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

While the technical standardization of communications infrastructures has been advanced by standardization organizations such as ESTI and IEC, the service layer has primarily been developed through the activities of a consortium of private companies (collective entities formed through collaboration among companies). For example, in the field of household appliances, Apple Computers has created the HomeKit standard and engages in collaborative activities with manufacturers related to home automation and network devices. Such activities are based on the idea of using their product, the company’s iPhone, as a control hub for illumination at home, for management of home appliances, and as a home security system.

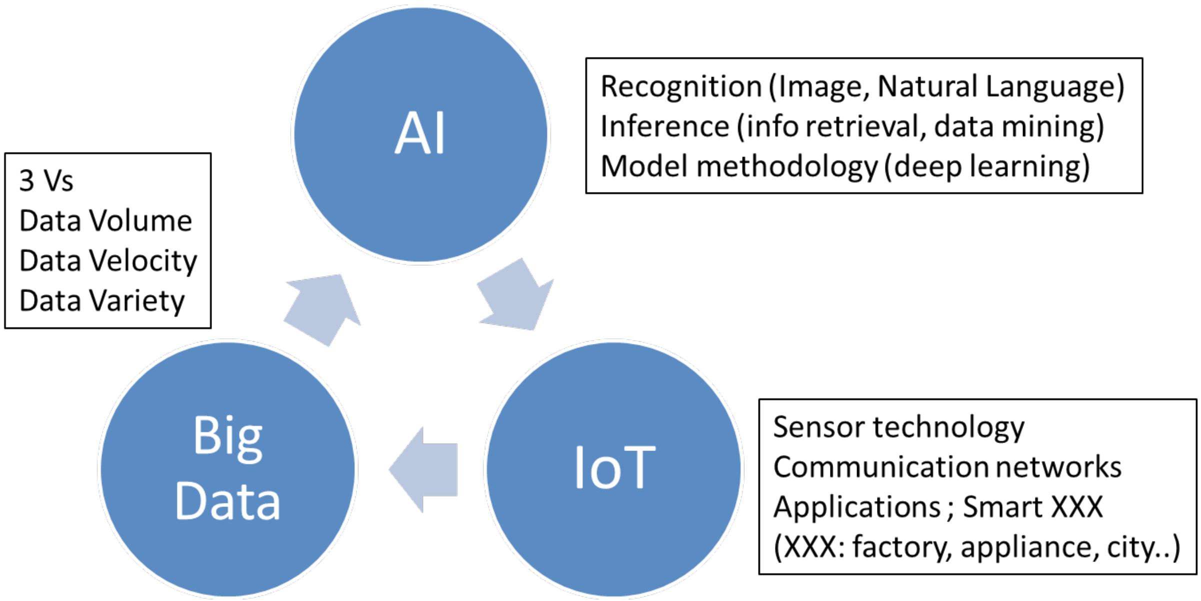
For industrial machinery such as aircraft jet engines and power-generating gas turbines, manufacturers are the ones that are actively engage in such activities. Among them, the General Electric Company (GE) launched GE-Data as a cross-sectional organization for their business departments such as jet engines, energy (wind power generator), healthcare, and railroad systems, and develops and provides a platform for an IoT system, which is called “PREDIX”. The company also started up the IIC (Industrial Internet Consortium) with information communication companies such as IBM, Cisco, Intel, and ATT, and has developed activities to expand their IoT platform to fields other than their business fields.

2.4. Relations of these new IT

AI and IoT, as well as big data as discussed above, are in a mutually complimentary relationship. In addition to the huge volume of information (image data, text data, etc.) that has originated from humans and been compiled on the internet, data originating from things due to the advancement of IoT sensor network technology has also been compiled. This has markedly enhanced the 3Vs of big data (i.e., Volume, Velocity, and Variety). For example, the data volume per day generated during the production process at a factory can be calculated as (number of products in a production line) * (number of production steps) * (data granularity). Here, data granularity refers to the data acquisition frequency per unit of time (e.g., per minute). Accordingly, the IoT service produces data with Volume and Velocity, exceeding the physical capabilities of humans. In addition, from surveillance cameras for security services and self-driving cars, various data are generated such as audio and image data.

As previously mentioned, the expansion of potential uses for such big data has contributed significantly to the development of different varieties of AI technology. By gathering various types of data that are not suitable for information processing as is, such as audio, image, and text data, perceptual and recognition technologies, including audio/image recognition and natural language processing, have been improved. In addition, by using a large volume of data, there have been remarkable advancements in various AI-related basic technologies for knowledge/ discovery techniques and deep learning, such as information searching and data mining.

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



Lastly, these AI-related technologies are key components for realizing every sort of IoT service such as smart factories, smart appliances, and smart cities. A large volume of internet information and sensor information is available, but such big data is not intentionally collected for specific purposes. More specifically, IoT sensors naturally collect such data, which are perceived, interpreted, and realized as economically valuable systems. This is the reason for naming them smart-XXX. In other words, since this is different from the human-led conventional form of processing data provided to computers, computers would be able to provide services with much less human involvement. Innovations enabled by big data AI and IoT are characterized as new services, which would not require service recipients, people, or companies to work and which would bring even more beneficial effects to their respective entities.

More specifically, it 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, for IT applications in corporate finance/accounting and human resources, the primary tasks of their IT systems typically involve financial/accounting computations and human resource databases. However, in the future, such finance/accounting computations may be able to be performed according to revised accounting standards and a recommendation system for personnel allocation according to characteristics of positions in the companies. Moreover, in sales departments, IT has been increasingly used in customer information management. In this case, with the development of new customers and determination of customer segments, IT would support more strategic marketing activities. Furthermore, efforts made for IoT services such as smart household appliances and smart cities can be considered to be the sprouts of service innovations beyond existing business models. Upon improving productivity (output per unit input), conventional IT is often considered to be a tool for reducing input (making the existing work more efficient). However, with technological advancements in big data AI and IoT, IT should be recognized as an enabling technology for expanding output.

Next, it will 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 it is still in the empirical study stage, smart cities represent a typical example, toward which automobile, home appliance manufacturers, and power companies, etc., are jointly working. Examples of practical application include the Nest thermostat, which learns the living patterns of residents and automatically controls the ON/OFF functionality of household appliances/equipment, 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 together with IT companies, and venture companies that specialize in specific technical fields such as artificial intelligence technology. As a result, it will bring new innovations such as those described above, and may also destroy the conventional manufactures' business models, in which profits are made by producing and selling things. Through household appliances, such as the Nest thermostat, being connected to a controller, the manufactures of those household appliances will lose contact points with their end customers. In other words, those manufacturers' statuses will fall to that of suppliers of smart household appliances services (i.e., some parts to a whole system). The status of automobile manufacturers will be the same in regard to becoming a transportation service due to self-driving cars. It is highly likely that the IoT wave will have a significant impact on the business structures of manufacturers that handle "things".

3. Measuring science and innovation of AI

In this section, I would provide some statistical evidences of interactions between science and innovation, focusing on AI. It is clear that new way of using big data, such as deep learning, changes the way that business innovation is organized. In addition, deep learning techniques is used for a tool to invent new products or innovation, as is seen in the case of IBM Watson's application to new drug development (Nayak et. al, 2016). In this sense, AI can serve not only for new business applications but as an invention for method of invention (Cockburn et. al, 2018).

Another special feature of AI is found in the style of its development, where public science sectors are involved in such development. For example, the idea of deep learning (deep neural network) is not new, but actual implementation with new methodology and powerful computer powers was made by academia. Subsequently, a series of development of deep learning algorithms, suited for various kinds of datasets (such as CNN for image data and RNN for text data) are developed by computer science scholars. In addition, a substantial contribution to AI development by private sector is also found. A typical case example is Google Brain's publication of "alpha go" (Silver et. al, 2017). A team at Deep Mind, currently under the Google AI department (Google Brain), not only developed the software to beat the world go (Chinese chess) champion, but also made it in a public as a research paper. At the same time, innovations, i.e., economic valuation of such public technology, are flourishing as is seen in the previous chapter for IoT applications (smart XXX), as well as mushrooming of startup firms providing a specific AI technology with fee. Due to a tremendous speed of technological progress and business development, the public sector and the private sector are co-mingled in this process. Therefore, AI is a typical example of science based economy, in a sense that scientific findings (published for free) and innovation (commercial activities) are co-occurring, by the cross over style of public sector such as universities and public research institutions and private sector formed by firms.

In this paper, I take the methodology of linking research article and patent, more specifically, finding identical author/inventor in the database of research articles and patents (Ikeuchi et. al, 2016), close to the idea of finding the paper patent pair with the similar contents (Lissoni et. al, 2013). Traditionally, the degree of scientific basis, or ‘science intensity’ of industry has been measured using non-patent literature (research article) citations made by patents (Narin and Noma 1985, Schmoch 1997). Non-patent literature citations show the degree of disembodied scientific knowledge that flows into patents, while the patent-publication pair can capture the state of co-occurrence of scientific and invention activities within the same researchers, i.e., interplay of science and technology embodied in human capital.

I use the SCOPUS data from Elsevier for research articles and US patent data from USPTO for patents. In both datasets, I select researchers working for the organizations located in the United States. There are around 8 million papers from SCOPUS and 3 million patents from USPTO data. These two data are linked by author/inventor names as well as his/her affiliates, and about 5% of all authors from SCOPUS and about 13.3% of inventors from USPTO data can be linked (Motohashi, 2018).

From this linked dataset of research articles and patents, AI papers and patents are identified. For AI papers, ASJC (All Science Journal Classification), provided by Elsevier, is used. Concretely, ASJC 1702 is labeled as “Artificial Intelligence” under a broad category of “computer science”, which is used in this paper as AI paper identification. In terms of patents, IPC code, G06N is used, along the line of JPO publication of technology trend survey (JPO 2014). G06N stands for the technology of “computer systems based on specific computational models, such as neural network, inference machine and fuzzy logic”. This definition narrowly defines the concept of AI, in a sense that only basic methodologies for analytics and models are included.[1]

Figure 2 shows the trend of the shares of AI papers and patents to the all. Both of them have a upward trend until 2010, but stable afterwards. It should be noted USPTO discloses the information of only granted patents. Therefore, only data applied until 2011 can be used even the datasets are obtained from USPTO data download site, called patentview.org in 2016. In Figure 1, the trend of AI patent applications, whose truncation bias is relatively small, are also presented. The surge of AI patent applications after 2010 is found. It should be also noted that the shares of AI papers and patents are very small, like less than 1% of total papers and patents, since only core technology of AI is included in both definitions.

1. Another approach is taking a broad category such as G07F, including general purpose computer software and applications (OECD, 2013)

Figure 2: Share of AI papers and patents

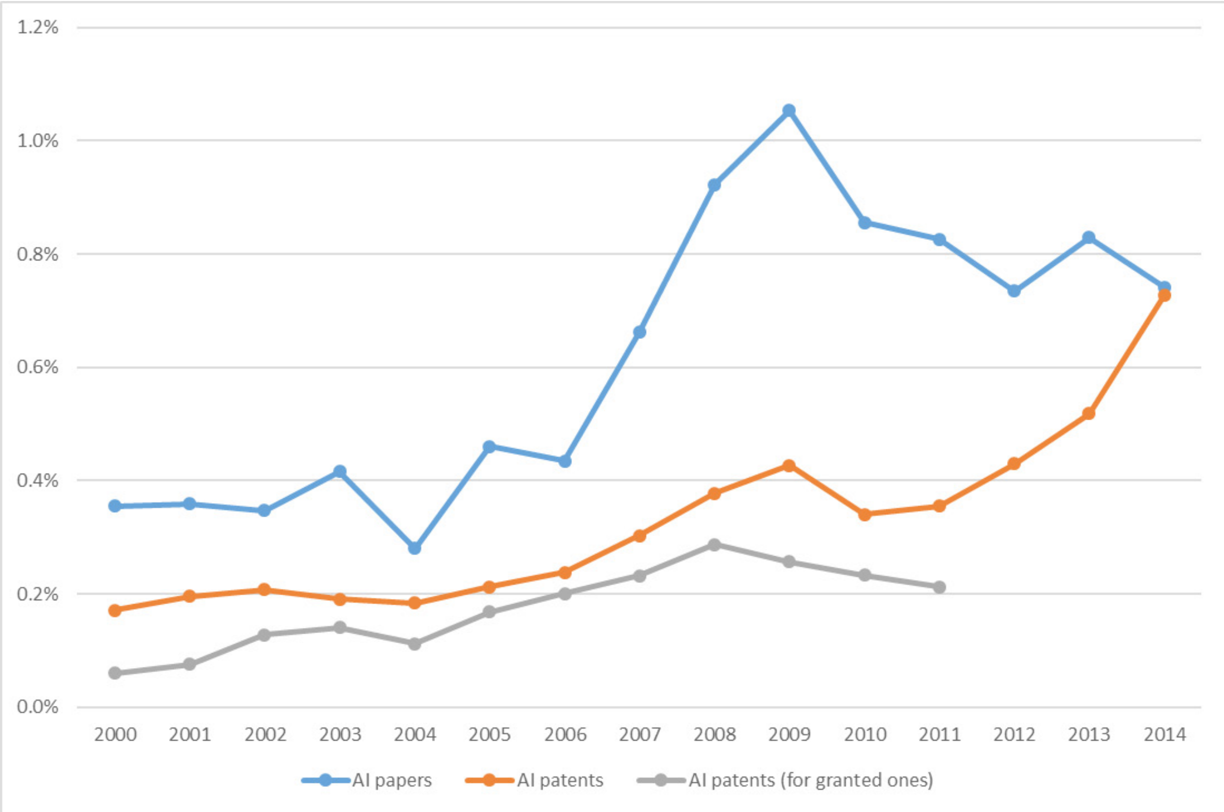
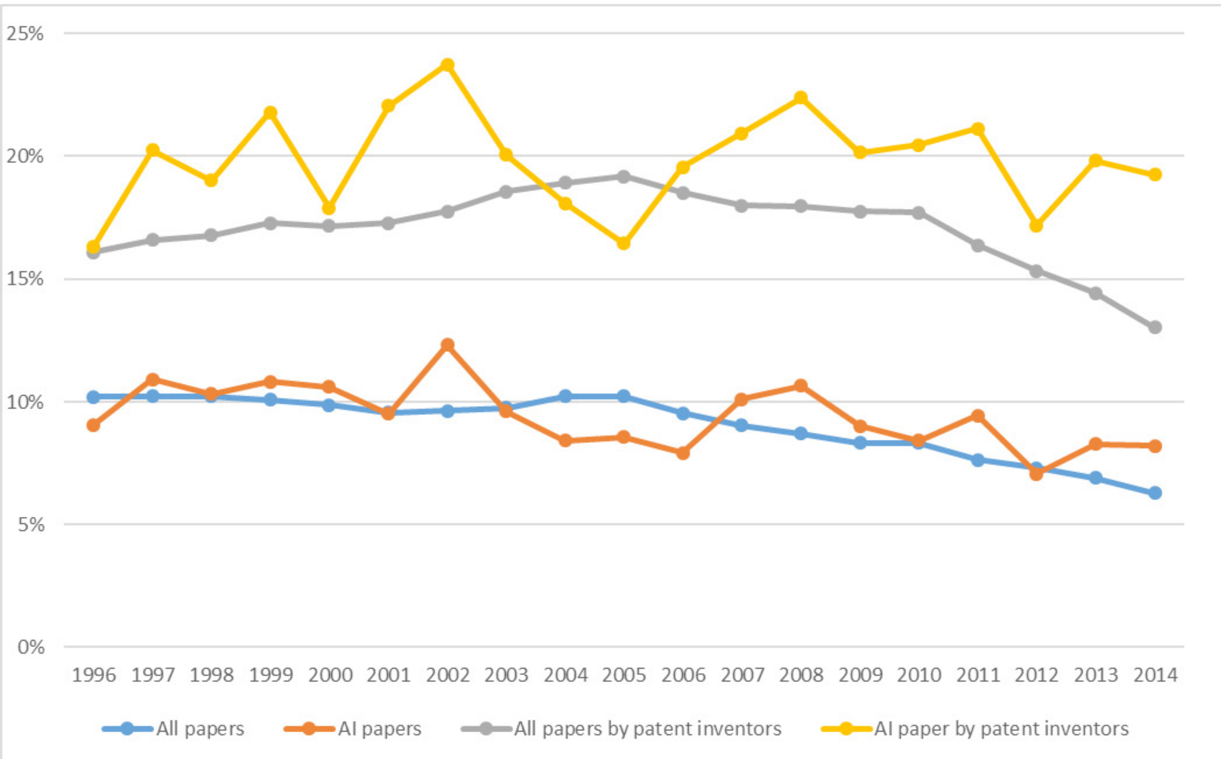


Figure 3: Share of private sector authors



Next, I break down author/inventor by her affiliation, private sector (for profit firm) or public one (non profit research organization such as universities and government laboratories). Figure 2 shows the share of private paper authors. The majority of paper authors are affiliated to public sector, so that the private shares for all papers are less than 10%. In addition, the pattern of private sector's retreating from basic science for private sector is found by its decreasing trend (Arora et. al, 2015). The share of private publications of AI papers is almost same as that of all papers, suggesting that the public sector is a major contributor to basic science of AI. When we look at the papers written by the author who also has some patents as an inventor (an author/inventor researcher), the share of private sector becomes higher (around 15% for all discipline), since the patenting activities are driven by private sector. It should be noted that the share of private sector for AI papers by an author/inventor researcher is not declined over time, while the private share of AI papers decreases. Therefore, more and more firm researchers are involved with both publication and patenting activities.

Figure 4 shows the similar trend from the patent data side, but here the share of public patenting, instead of the private share, is displayed. Since patenting activities are driven by private sector, the share of public sector patenting for all technology fields is around 1.5%. The public share of AI patents is higher, such as around 3.0%, suggesting more academic patenting is found in this technology field. This looks natural, since we take the definition of AI as a basic methodology for models and analytics. It is found that the share of public patenting by those who is an author of research articles is decreasing for both all fields and AI. It should be also noted that the public share of author/inventor patents for AI is lower than that of that of all AI patents, even though the public sector is supposed to be a major contributor to publication. These findings is consistent with Figure 3, suggesting that private researchers become to be involved with both publication and patenting activities.

Figure 4: Share of public sector inventors

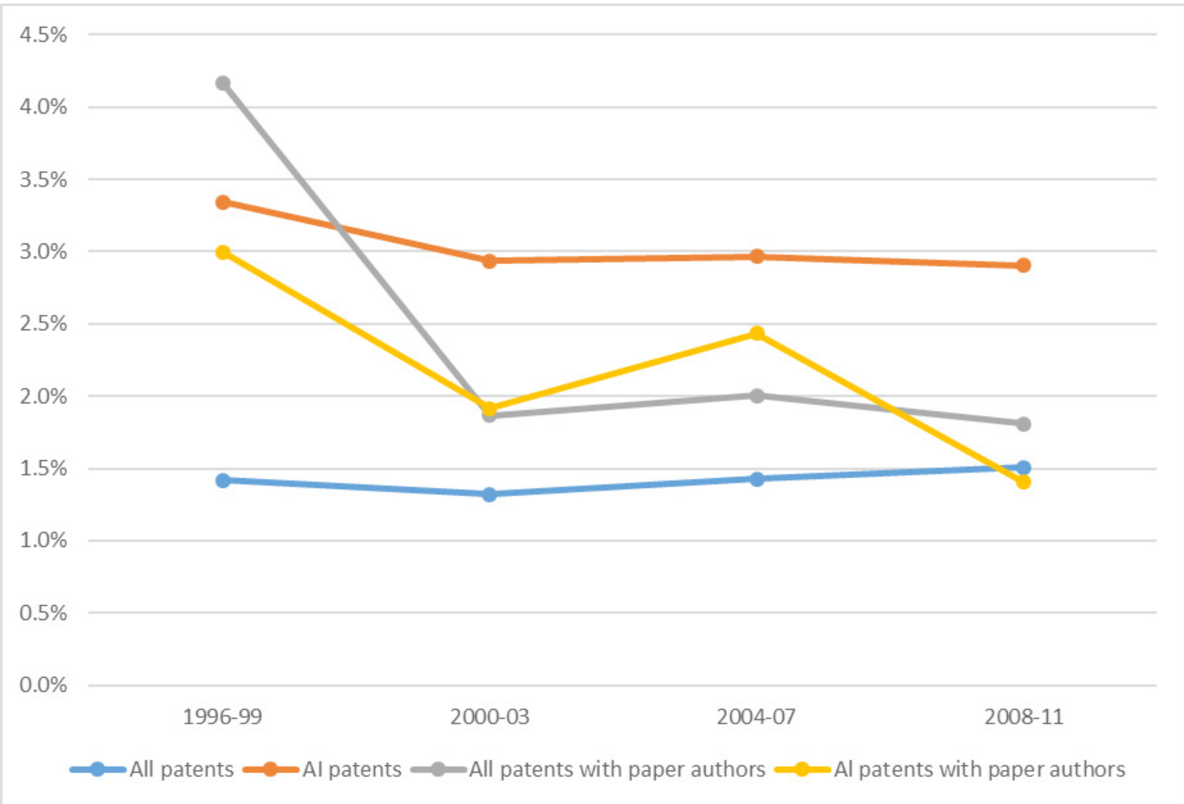


Table 1 provides the share of AI patents/papers by type of researcher affiliation and by author/inventor or not (only author or only inventor). This table reflects the relative importance of researchers by types above for AI patents and papers, normalized by general trends of patents and papers over our observation period. It should be noted that the researchers can be categorized into three types, (1) inventor only, (2) both inventor and author and (3) author only. In terms of private sector, the share of AI patents are higher for the category 1 (0.15%) , as compared to the category 2 (0.11%). When we look at the AI papers share, the larger number is found in category 2 (0.91%), as compared to the category 3 (0.55%). Therefore, inventor only researchers (category 3) is the most significant contributor to AI related outputs for private sector. The same exercise for the public sector reveals that the category 2 researchers (author/ inventor) is the highest, as compared to the category 1 (inventor only) and the category 3 (author only).

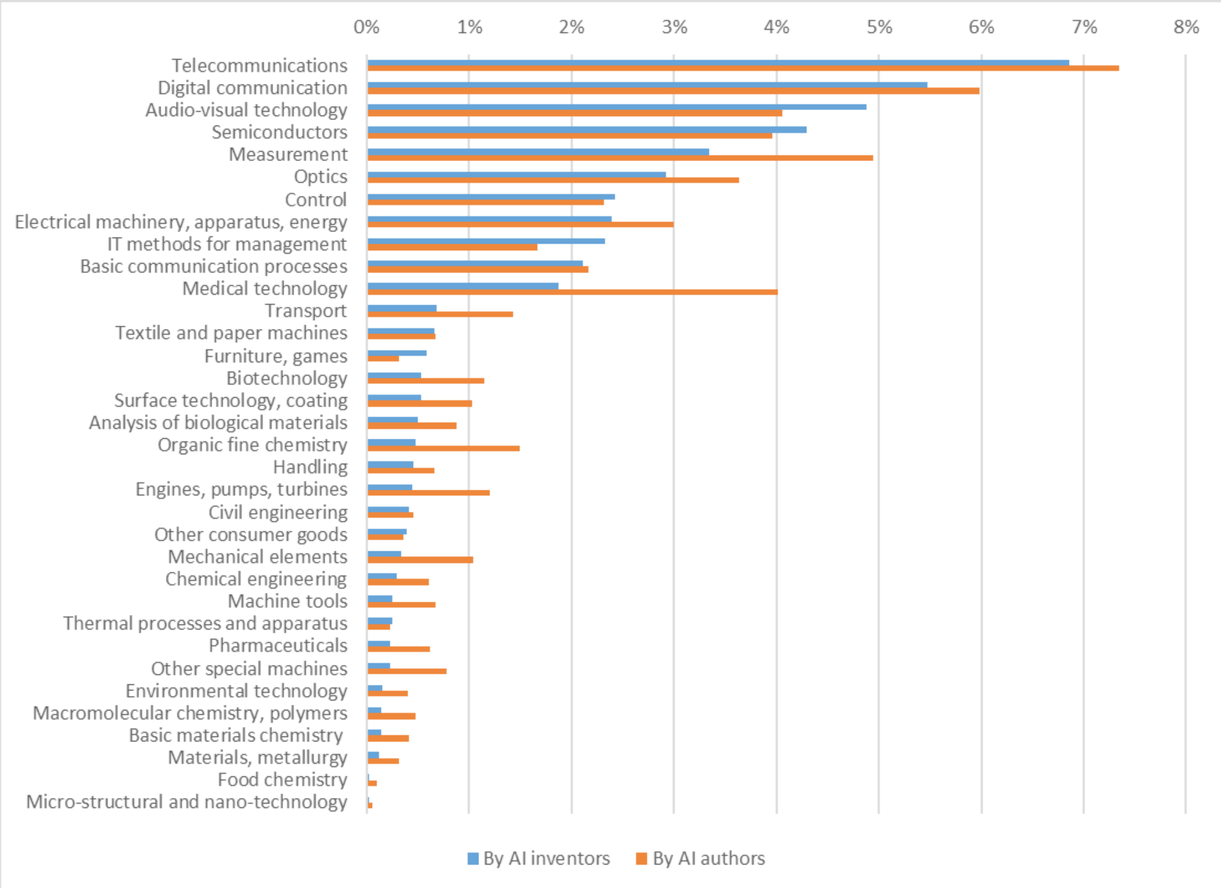
Table 1: Share of AI patents and papers by type

	Share of AI patents		Share of AI papers	
	W_paper	WO_paper	W_patent	WO_patent
Private	0.11%	0.15%	0.91%	0.55%
Public	0.18%	0.08%	0.76%	0.60%

Finally, the impact of AI publication and patenting to other fields is investigated. Figure 5 shows the share of patents by technology field by AI inventors and AI authors. The category of “computer technology”, including G06N, is excluded from this chart, to see how AI technology is used widely across other technologies.[2] It should be noted that the patents invented by AI authors have wider applications across technology field, as compared to AI inventors. The difference between two figures are found particularly in “measurement”, “medical technology”, “transport” and “organic fine chemistry”. These findings reflect that the publication of AI plays an important role in its wide applications by various technology fields.

2. The share of computer technology is 53.2% for AI inventors and 41.6% for AI authors.

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



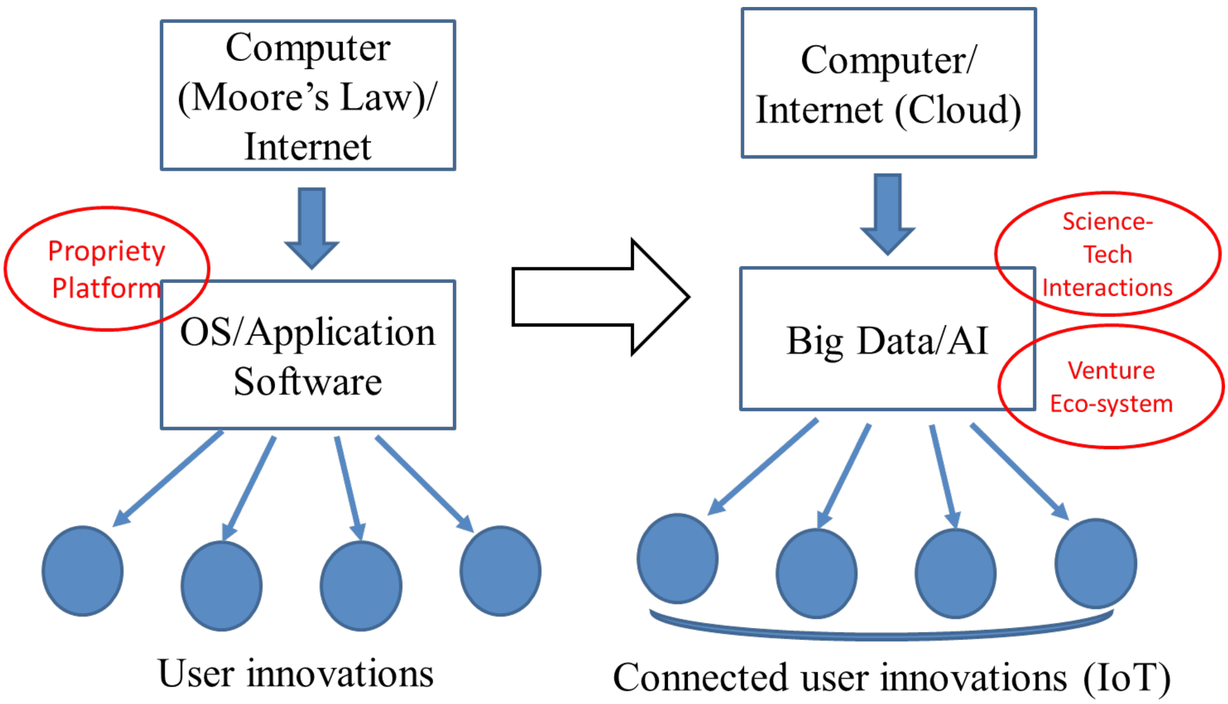
4. Discussion

The innovation that information technology brings is seen across business sectors and in various business areas. The application also extends from a system in supporting functions at firm such as personnel and financial accounting to a wide range of operations including production management, customer management and supply chain management. Artificial intelligence is one of the core technologies that supports applications of these various information technologies. Characteristics as general purpose technology (GPT: General Purpose Technology) of such information technology has been analyzed by economists over 20 years (Bresnahan and Greenstein, 1996; Helpmann, 1998). For example, backbone systems related to factory operations in the manufacturing industry and banking financial transactions are totally different tasks performed by the information system. However, since improvement of computer performance can realize more advanced tasks in both, progress of information technology that is geometrically progressed by Moore's Law will bring great benefit to the entire economy. The new methodology in computer modeling and analytics, called AI, obviously brings new possibilities for business applications, but what is new as compared to information technology in general?

First, let us review the literature discussing about the GPT feature of information technology. The General Purpose Technology is characterized as its "speed", "pervasiveness" and "downside applications" (Helpmann, 1998). A driver of speed of IT is Moore's law in semiconductor, and development of operation and application software makes pervasive use of computer in various industries possible. However, IT investment does not leads to improvement of business performance and productivity automatically. It requires organizational innovations, such as business process changes (Bresnahan et. al, 2002). Therefore, IT is just an "enabling" technology to fruit as innovation that is applied to business and bring about management effect. For example, innovations using general-purpose technology is achieved by a manufacturing user's application to highly controlled production process in its factory and by a bank to use biometric authentication system to improve in security in ATM transactions. Basically, the "new IT", such as AI, Big data and IoT, discussed in the section 2, is on the extension of information technology which is a general purpose technology. In other words, these are the enablers for realizing innovation, not innovation itself. In order to realize innovation, business innovation is needed to assemble it into activities with business value on the user side.

However, there are some differences in its innovation organizations. In traditional IT (left panel of Figure 6), the key technology to link computer technology to user innovation is the bundle of software, and the operation software plays a particularly important role to make various application software developments possible. The propriety operations software constitutes a platform, and the upstream platform leaders, for example, Intel and Microsoft in the case of PC, drive a whole system of innovation (Gawer and Cusumano, 2013). As a results, the division of innovative labor is found between platform leaders and downstream user innovation players (Gambardella and McGahan, 2010).

Figure 6 : Traditonal IT vs AI/BD/IoT



In contrast, such platform leader cannot be found in the new IT (right panel of Figure 6), at least at this moment. A key difference is that open, not propriety, science contributions are found in the development of AI technology. At the same time, the role of private sector increases over time, both in open scientific articles and in propriety patents. These two components are intermingled at researcher level, as is seen by the growing presence of author/inventor in the previous section. At the end, we can say that the interactions between open science and propriety technology are intensified, and there is no clear sign of emergence of a platform leader in the new IT.

Off course, there are some firms acting like a platformer, such as Google and Amazon. However, while Google dominates public internet information, an expansion of IoT makes it difficult for them to play as a platform leader. In IoT business innovation, the critical component is domain specific propriety data, such as driving record for autonomous driving, and the domain expands to variety of field, like consumer electronics, medical treatment, smart factory, financial services etc. In this sense, the value of user innovation comes from inter-connections of open and propriety technology, data and business know-hows. As a whole structure of innovation becomes complex, various eco-systems, instead of platform business, are created as ventures to fill the gaps emerged among incumbent firms' capabilities.

5. Conclusion and implications

In this paper, AI driven innovation is analyzed. The linked dataset of AI research articles and patents reveals that substantial contribution of public sector is found for AI development. In addition, the role of researchers who are involved both in publication and patent activities, particularly in private sector, increased over time. This is, open science, publicly available by research articles and propriety technology, protected by patents are intertwined in AI development. In addition, the impact of AI, combined with big data and IoT, defined "new" IT on innovation is discussed, by comparing with traditional IT, consisted by technological progress of computers and software developments. Both of new and traditional IT could be understood by using the framework of GPT (general purpose technology), while the organization of new IT innovation can be characterized by emergence of multiple eco-systems, instead of the pattern of platform leadership, found in traditional IT.

I would conclude this paper by providing some implications. As for management implications, the ecosystem strategy should be emphasized. The business ecosystem consists of "keystone" which plays a central role in the ecosystem (business inter-company relation network) and other niche players (Iansiti and Levien, 2004). The role of "keystone" is to attract many niche players and to spread the whole ecosystem. On the other hand, "niche players" are with each original technology, contributing to the diversity of the entire ecosystem. With this mutually complementary relationship, the ecosystem holds. For example, the Apple computer has an up store on the iphone and is developing a service that meets the diverse needs of consumer needs. Combined with the common platform (platform) of various services provided by individual application providers (niche players) and Apple computers (Keystone), they will have overall value. Keystone's role is to improve the business value of the whole ecosystem, and it is important to build a win-win relationship with niche players. Strengthening control over niche players and continuing exploitation of value will ultimately destroy the ecosystem. In order to increase the value of the entire ecosystem, in order to attract diverse niche players, it is necessary to make attractive management resources for niche players.

Along this line, the most important strategic question is which way, key stone or niche, to go in new IT ecosystem. A New IT user tends to be a niche player, based on its own business knowledge and propriety data. However, a user firm can be a key stone player, by providing common resources to other firms within a certain business domain. For example, GE chooses the key stone strategy for IoT business in manufacturing industry, by offering PREDIX platform which serves as a basis of IoT industrial applications such as condition based maintenance of large industrial equipment. An IT technology firm typically seeks for the key stone strategy, instead of platformers, by building up win-win relationships with niche players and nurturing an ecosystem as a whole.

A policy maker should be aware of the interplay between open science and propriety technology in AI field. Therefore, a policy support to AI innovation is not just providing a fund to basic science, but a cross over model of public and private sector should be investigated. For example, promoting university startup is more suitable, as compared to the policy instruments based on traditional linear model, such as licensing of university patents. In addition, a non-profit public research institution can be a place where new eco-system development, since such body cannot be a direct commercial competitor to private firms seeking for mutual collaborations. In this regards, the role of university should be expanded from the place for research and education, to the place for entrepreneurial activities and innovation experimentations.

6. 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

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

Previous CEAFJP Discussion Papers

DP 18-03 (November 2018)

“The Yen Exchange Rate and the Hollowing-out of the Japanese Industry”, Ansgar Belke (University of Duisburg-Essen) & Ulrich Volz (SOAS University of London, 2017 CEAFJP/ Banque de France Fellow)

DP 18-02 (October 2018)

"Cross-cultural (France and Japan) and Multidisciplinary Discussion on Artificial Intelligence and Robotics: Tendencies and Research Prospects", Naoko Abe (CEAFJP Research Fellow)

DP 18-01 (July 2018)

"Impact of Shareholder-Value Pursuit on Labor Policies at Japanese Joint-Stock Companies: Case of Nikkei Index 400", Kostiantyn Ovsiannikov (University of Tsukuba, Prizewinner of the 2018 FFJ/SASE Best Paper Award)

DP 17-05 (November 2017)

"Female Board of Directors and Organisational Diversity in Japan", Yukie Saito (CEAFJP Associate Researcher, University of Geneva, Institut de Recherches Sociologiques)

DP 17-04 (August 2017)

"*Keiretsu* Divergence in the Japanese Automotive Industry: Why Have Some, but Not All, Gone?", Akira Takeishi (Graduate School of Economics, Kyoto University; CEAFJP Visiting Researcher) et Yoshihisa Noro (Mitsubishi Research Institute, Inc.)

DP 17-03 (June 2017)

"Globalization and Support for Unemployment Spending in Asia: Do Asian Citizens Want to Embed Liberalism?", Sijeong Lim (University of Amsterdam) et Brian Burgoon (University of Amsterdam) ; Prizewinners of the SASE/FFJ Best Paper Award.

DP 17-02 (April 2017)

"Does ‘Driving Range’ Really Matter? The Hidden Cost of Internal Combustion Engine Vehicles and the Potentially New Value Proposition of Electric Vehicles: Two Cases from Countryside and Urban Side of Japan", Hidetada Higashi (2016 CEAFJP/Valeo Fellow)

DP 17-01 (March 2017)

"How Can We Understand the Differences between France and Japan in the Growth of Shared Mobility Services? The Paradox of Trust and its Social Construction", Naoko Abe (2016 CEAFJP/Renault Fellow)

DP 16-03 (September 2016)

"Parameter Bias in an Estimated DSGE Model: Does Nonlinearity Matter?", Yasuo Hirose (Faculty of Economics, Keio University) and Takeki Sunakawa (Graduate School of Public Policy, University of Tokyo)

DP 16-02 (April 2016)

"Financialization and Industrial Policies in Japan and Korea: Evolving Complementarities and Loss of Institutional Capabilities", Sébastien Lechevalier (EHESS), Pauline Debanes (EHESS), and Wonkyu Shin (Kyung Hee University)

DP 16-01 (April 2016)

"How Do Credit Hours Assure the Quality of Higher Education? Time-Based vs. Competency-Based Debate", Ayaka Noda (National Institution for Academic Degrees and Quality Enhancement of Higher Education (NIAD-QE)

DP 15-04 (December 2015)

"Government Policy and the Evolution of Japan's Photovoltaic Industry, 1961-2014", Maki Umemura (Cardiff University, 2015 CEAFJP/Michelin Fellow)

DP 15-03 (December 2015)

"Japan's Financial Crisis and Lost Decades", Naohisa Hirakata (Bank of Japan), Nao Sudo (Bank of Japan), Ikuo Takei (Bank of Japan), Kozo Ueda (Waseda University, 2015 CEAFJP/Banque de France Fellow)

DP 15-02 (May 2015)

"Can Increased Public Expenditure Efficiency Contribute to the Consolidation of Public Finances in Japan?", Brieuc Monfort (CEAFJP Associate Researcher)

DP 15-01 (May 2015)

"Policy Regime Change Against Chronic Deflation? Policy Option under a Long-Term Liquidity Trap", Ippei Fujiwara (RIETI, Keio University, Australian National University), Yoshiyuki Nakazono (Yokohama City University), Kozo Ueda (Waseda University, 2014 CEAFJP/Banque de France Fellow)