

Journal of STI Policy and Management

Publication details, including instructions for authors and subscription information: <http://www.stipmjournal.org/>

## Evolutionary Path of Development of Artificial Intelligent (AI) and Patterns of Knowledge Convergence over the Second and Third AI Booms

Kumiko Miyazaki<sup>1</sup>, Santiago Ruiz Navas<sup>1</sup>, Ryusuke Sato<sup>2</sup>

<sup>1</sup> Dept. of Technology and Innovation Management, Tokyo Institute of Technology, Tokyo, Japan

<sup>2</sup> Global Business Services Watson Delivery No.2, IBM, Tokyo, Japan

Version of record first published: 15 December 2019

To cite this article: Miyazaki, K., Navas, S.R., and Sato, Ryusuke. (2019). Evolutionary Path of Development of Artificial Intelligent (AI) and Patterns of Knowledge Convergence over the Second and Third AI Booms. *Journal of STI Policy and Management*, 4(2), 125–142

To link to this article: <http://dx.doi.org/10.14203/STIPM.2019.172>

ISSN 2540-9786 (Print); ISSN 2502-5996 (online)




Accreditation Number: 21/E/KPT/2018

Full terms and conditions of use: <https://creativecommons.org/licenses/by-nc-sa/4.0/>

You are free to:

- Share : copy and redistribute the material in any medium or format
- Adapt : remix, transform, and build upon the material
- The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

-  Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
-  NonCommercial — You may not use the material for commercial purposes.
-  ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

- You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.
- No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.
- If you copy the dataset merely to extract the uncopyrightable data elements would not need permission to do so. However, if you republish the full dataset or using the copyrightable data layers require a permission from Research Center for STIPM, Indonesian Institute of Sciences.

**JOURNAL OF SCIENCE, TECHNOLOGY AND INNOVATION  
POLICY AND MANAGEMENT (STIPM JOURNAL),  
Volume 04, Issue 02, December 2019**

**FOREWORD by EDITOR-in-CHIEF**

We are pleased to present the STIPM Journal Vol 4, No. 2, December, 2019. This issue brings together research findings related to science, technology, and innovation policy and management from Japan and Indonesia.

First article was written by **Djisman Simanjuntak *et al.*** entitled *Exploring the Transition to Eudaimonic Tourism: A Case Study of Bali*. This article discusses innovation in tourism focus on the dynamics of tourism grows. As tourism grows, carrying capacity is stretched or even overstretched in some places and industries. A shift toward more eudaimonic tourism is needed, and the innovative elements of eudaimonia include geographical treasure, biodiversity, and local deep culture.

**Taeko Suehiro and Kumiko Miyazaki** present an article entitled *Accumulation of Knowledge by Strategic Public Procurement through Public-Private-Partnership for Service Innovation in Japan*. This study focuses on how governments strategically procure public service through Public-Private Partnership (PPP)—or more specifically, Private Finance Initiative (PFI) arrangements.

**Erman Aminullah** presents *E-Cigarette as Disruptive Innovation: Forecasting of Conventional Cigarette Substitution in Indonesia*. This article intends to forecast conventional cigarette substitution by e-cigarette in the context of disruptive innovation. E-cigarette as disruptive innovation has been driven by technology innovation to create e-cigarette products for global market. The advancement of e-cigarette technology innovation would continue to create smart and less harmful e-cigarette as alternative tobacco products in future.

**Kumiko Miyazaki, Santiago Ruiz Navas, and Ryusuke Sato** present the fourth article entitled *Evolutionary Path of Development of AI and Patterns of Knowledge Convergence over the Second and Third AI Boom*. AI has been through several booms and we have currently reached the 3rd AI boom which followed the 2nd AI boom centering mainly on expert systems. The current AI boom started around 2013 and AI is beginning to affect corporate management and operations. AI has been evolving over six decades but it seems that the current boom is different from the previous booms.

The fifth article entitled *Predicting Potential Co-Authorship using Random Forest: Case of Scientific Publication in Indonesian Institute of Sciences* by **Rizka Rahmida, Asep Saefudin, and Bagus Sartono**. Co-authorship network is one of the proxies to evaluate the emerging research collaborations. Co-authorship that happens for the first time among a pair of author plays an important role as the key of success for their co-authorship in the future.

Finally, **Hiroki Idota *et al.***, present an article entitled *Conducting Product Innovation by Using Social Media among Japanese Firms*. This article based on a study that attempts to conduct an empirical

analysis of how social media use promotes product innovation in Japanese firms by collaboration with consumers based on survey data from Japanese firms using probit analysis. This study finds that collaboration with consumers by using social media is important for innovation, particularly in developing concepts and devising methods of use.

The STIPM Journal is indexed by Google Scholar, ISJD, IPI, DOAJ, BASE, and OCLC World Cat. This make the journal dissemination wider. We would like to thank all the reviewers for their excellent work and the authors who kindly contributed their papers for this issue. We are also indebted to the *STIPM Journal* editorial office at P2KMI-LIPI and the publishing and production teams at LIPI Press for their assistance in preparation and publication of this issue.

We are expecting that STIPM will always provide a higher scientific platform for the authors and the readers, with a comprehensive overview of the most recent STI Policy and Management research and development at the national, regional dan international level.

Happy New Year 2020 to all of you...

Jakarta, December 2019

Editor-In-Chief



# JOURNAL OF STI POLICY AND MANAGEMENT

Volume 4, Number 2, December 2019

## LIST OF CONTENTS

Exploring the Transition to Eudaimonic Tourism: A Case Study of Bali <b>Djisman Simanjuntak, Alvin Desfiandi, Erica Lukas, Isti Setiawati, Nakita Sabrina, and Stanley Makalew</b> .....	77–99
Accumulation of Knowledge by Strategic Public Procurement through Public-Private Partnerships for Service Innovation in Japan <b>Taeko Suehiro and Kumiko Miyazaki</b> .....	101–112
E-Cigarette as Disruptive Innovation: Forecasting of Conventional Cigarette Substitution in Indonesia <b>Erman Aminullah</b> .....	113–124
Evolutionary Path of Development of Artificial Intelligent (AI) and Patterns of Knowledge Convergence over the Second and Third AI Booms <b>Kumiko Miyazaki, Santiago Ruiz Navas, and Ryusuke Sato</b> .....	125–142
Predicting Potential Co-Authorship using Random Forest: Case of Scientific Publications in Indonesian Institute of Sciences <b>Rizka Rahmaida, Asep Saefuddin, and Bagus Sartono</b> .....	143–152
Conducting Product Innovation by Using Social Media among Japanese Firms <b>Hiroki Idota, Sheikh Abu Taher, Teruyuki Bunno, and Masatsugu Tsuji</b> .....	153–166



## Evolutionary Path of Development of Artificial Intelligent (AI) and Patterns of Knowledge Convergence over the Second and Third AI Booms

Kumiko Miyazaki<sup>1\*</sup>, Santiago Ruiz Navas<sup>1</sup>, Ryusuke Sato<sup>2</sup>

<sup>1</sup> Dept. of Technology and Innovation Management, Tokyo Institute of Technology, Tokyo, Japan

<sup>2</sup> Global Business Services Watson Delivery No.2, IBM, Tokyo, Japan

### ARTICLE INFO

#### Article History:

Received : 01 August 2019

Revised : 08 December 2019

Accepted : 08 December 2019

Available online : 15 December 2019

#### Keywords:

Co-words,

Co-authorships,

Knowledge convergence,

Bibliometrics

### ABSTRACT

Although artificial intelligence (AI) was coined by John McCarthy 60 years ago, AI has been confined to the academic and scientific research domain. AI has been through several booms and we have currently reached the 3rd AI boom which followed the 2nd AI boom centering mainly on expert systems. The current AI boom started around 2013 and AI is beginning to affect corporate management and operations. AI has been evolving over six decades, but it seems that the current boom is different from the previous booms. In this paper, we attempt to elucidate the evolutionary path of development of AI and the structural patterns of knowledge convergence in the current and previous booms. For this purpose, we have set two main objectives: 1) to characterize the first (1B), second (2B) and the current, third (3B) AI boom; and 2) to analyze the structure of knowledge convergence around AI. The research question is how have the key technologies and the applications of AI changed over time, in the 2B and the 3B? An innovative method has been used to identify the characteristics of AI and the evolutionary path of knowledge convergence over these booms.

©2019 PAPPITEK-LIPI All rights reserved

## I. INTRODUCTION

John McCarthy from Stanford University coined the term Artificial Intelligence in 1956 at a conference in Dartmouth. According to him, AI is the science and engineering of making intelligent machines (McCarthy, 2007).

AI is said to have reached its 3rd boom (Matsuo, 2016). The first boom (1B) which

continued up until 1969 was based on search and deductive reasoning. AI programs were able to solve only simple problems. In the 2nd boom (2B) which took place in the 1980s, AI was mainly related to expert systems. The knowledge of the experts were used to build the knowledge bases which enabled expert systems to be constructed. For example, one of the authors worked at Schlumberger in the 1980s and did research on reservoir modeling systems used in oil exploration. The knowledge of the geologists was used

\* Corresponding Author.

E-mail: [miyazaki@mot.titech.ac.jp](mailto:miyazaki@mot.titech.ac.jp)

to build the expert system to estimate the exact location and the amount of oil in a reservoir. AI in 2B was highly specialized in specific domains leading to limited generalization of applications.

Since 2013, AI has entered a new phase, the 3rd boom (3B), characterized by machine learning and deep learning. Although AI has been evolving over six decades, it seems that 3B is fundamentally different from the previous booms in certain ways, with definite signs of AI adoption in a broad range of applications, affecting society (Matsuo, 2016).

AI is beginning to affect corporate management and operations. IBM carried out a survey on AI adoption where 65% of 600 companies responded that AI will play an important role in corporate management: 58% responded that cognitive computing is essential to digital transformation; and 58% responded that having AI capability would be absolutely necessary to remain competitive in the future (Kumar, 2017).

AI is based on a range of scientific disciplines and technologies which are interrelated. In this paper, we attempt to elucidate the underlying structure of the evolutionary pattern of development of AI over the previous and current booms, especially focusing on knowledge convergence. The various technologies and scientific disciplines have been converging in AI, at the same time AI has been converging with other scientific disciplines and technologies to open up new opportunities. Furthermore, in the second methodological approach, we seek to contribute to the literature about the detection of early stages of industrial convergence by providing empirical evidence of the usefulness of using keywords as a proxy of knowledge convergence by testing it on the field of AI.

To achieve this goal, we have set two objectives:

- 1) To characterize the first, second and the current, third AI boom
- 2) To analyze the structure of knowledge convergence around AI

Research Question:

How have the key technologies and the applications of AI changed over time, in the 2B and the 3B?

The method proposed to solve these research questions consists of two independent parts. Part 1, uses scientific papers' author keywords and other metadata validated with interviews to descriptively analyze the applications existing in the 1B and 2B and those unique to the 3B. Part 2, uses a combination of quantitative analysis—for more information see Ruiz-Navas and Miyazaki (2018); to identify the keywords proxy of knowledge converging around AI for the 1B, 2B and 3B. These two independent techniques characterize the three booms. Finally, the descriptive and convergence information obtained for each boom is qualitatively analyzed and the research question is answered.

## II. TECHNOLOGICAL CONVERGENCE

The concept of technological convergence has been helpful to study the dynamics of technological change, e.g., how technological change affects the organization of industries and how new entrants and incumbents can benefit from the process of technological change. The concept of technological convergence was reported by Rosenberg (1963), who pointed out how industries which seemed different were implementing the same technologies to solve common problems in their production lines. Further research had developed the concept explaining it as a progressive accumulation of knowledge which lead to the creation of new technologies (Gaines, 1998; Hacklin, Raurich, & Marxt, 2005).

Furthermore, the concept of technological convergence has been used to analyze strategies and the influence on industry competitiveness. In Lee, Yun, and Jeong (2015) the concept was implemented to explain the change of industrial competitiveness on the Korean textile industry, showing the positive and negative effects of adapting to technological convergence. Also, Miyazaki and Giraldo (2015) analyzed the technological competence building strategies implemented by Nippon Telegraph and Telephone (NTT) while adapting to the convergence of internet and telecommunications technologies.

In the context of the creation of new industries, Hacklin (2008) proposed the evolutionary



process of industrial convergence. The framework intended to study the process of industry emergence, such as functional foods, packaging solutions, camera phones, and intelligent buildings; in the first decade of 2000. The evolutionary process of industrial convergence explains that industries are created in four stages, knowledge, technology, application, and finally, industrial convergence.

The definition of the four stages of convergence is presented based on Hacklin, Marxt and Fahrni (2009):

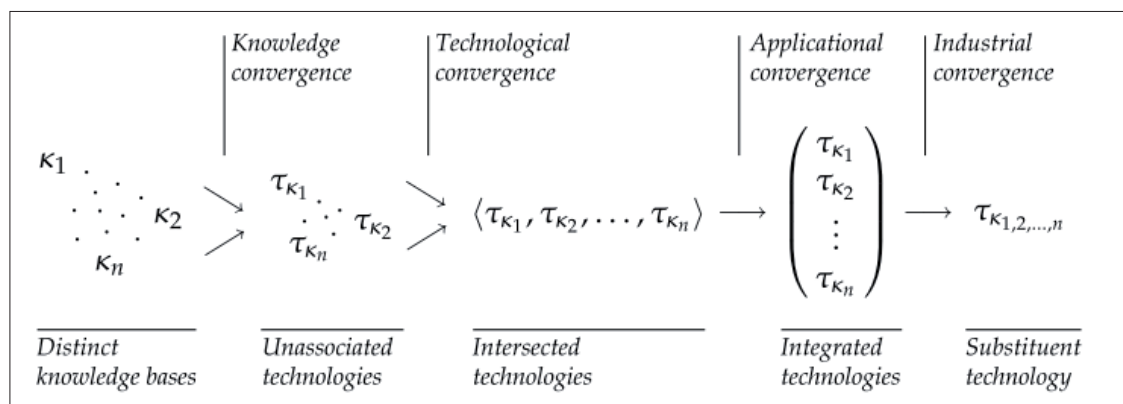
- 1) Knowledge convergence: “*Knowledge convergence denotes the emergence of serendipitous coevolutionary spill-over between previously unassociated and distinct knowledge bases, giving rise to the erosion of established boundaries that isolate industry-specific knowledge.*”
- 2) Technology convergence: “*Technological convergence denotes the transition of knowledge convergence into a potential for technological innovation, allowing inter-industry knowledge spill-overs to facilitate new technological combinations.*”
- 3) Application convergence: “*Applicational convergence denotes the transition of technological convergence into opportunities for new value creation in such a way, that it with respect to the majority of metrics outperforms the sum of the original parts.*”

- 4) Industrial convergence: “*Industrial convergence denotes the transition of applicational convergence into the shift of industry boundaries in such a way, that firms from previously distinct industries through the emergence of common applications suddenly become competitors.*”

The four stages are interlinked and present a simplification of the complex process of industrial convergence over time, starting in knowledge and ending in the creation of new industries. An overview of the process is presented in Figure 1.

From Figure 1 and the definition of the stages of convergence, it is possible to understand that by identifying a phase of convergence the subsequent stages of convergence can be forecasted. Therefore, if knowledge convergence is identified in a given industry, it will be possible to grasp an idea of which technologies will be developed or emerge from that convergence and become part of the stage of technological convergence.

The four stages framework has been used to study technological change in the process of industrial convergence and technology forecasting. On the one hand, In Lalitnorasate and Miyazaki (2016), the framework was used to analyze the capability learning strategies implemented by companies from the food and pharmaceutical sector that were adapting to the emergence of the functional food industry. On the other hand, in Hacklin, Marxt and Fahmi (2009), the four stages framework was used to provide a set of



Source: Hacklin (2008)

**Figure 1.** Relationship over time of the four stages of convergence.



forecasted strategical scenarios for the companies in the nano-and bio-technologies (NBT) industries. Furthermore, the four stages framework was used to predict technological convergence in all fields of technologies reported by Scopus using patent data, network link prediction, and association rules (Lee, Han, & Sohn, 2015). Also, the four stages framework has been applied to detect technological convergence around ICT standards (Han & Sohn, 2016) and the convergence of phytosterols into cosmeceutical, nutraceuticals, and functional foods (Curran, Broring, & Leker, 2010; Curran & Leker, 2011).

The evolution of the methods and techniques to forecast technology convergence, based on the four stages framework, have evolved from descriptive and qualitative approaches to quantitative and predictive. The techniques to detect early stages of convergence requires the extensive use of knowledge and technology convergence proxies, e.g., International Patent Codes (IPC) co-occurrence, keyword co-occurrence, co-authorships and scientific paper's subject categories co-occurrence (Miyazaki & Giraldo, 2015; Lee, Yun, & Jeong, 2015). Among these proxies, keywords are proposed as an ideal indicator to detect knowledge convergence (Hacklin, 2008; Curran, Broring & Leker, 2010). However, keywords had not been widely used because the difficulties provided by semantic characteristics such as homonymy, synonymy, hyponymy, and hypernymy. These translate into false positive relationships when using word-based quantitative approaches (Curran, Broring & Leker, 2010).

Recent research had explored the benefit of keywords as proxies of knowledge convergence. In one case, words were used to identify potential areas of application convergence for Malaysian small and medium enterprises generated from the convergence of big-data and cloud computing (Chan & Miyazaki, 2015). A framework to detect knowledge convergence using keywords was proposed and tested for big-data (Ruiz-Navas & Miyazaki, 2018). These two publications showed that the use of keywords as a proxy of knowledge are useful as they provide detail of the units of knowledge that are converging and also allows to identify papers which contain an extensive description of these units of knowledge.

However, to validate the importance of keywords as a proxy for knowledge convergence, more empirical evidence is required.

### III. METHODOLOGIES AND DATA SOURCES

The methodology consists of two independent parts. In Part 1, analyses of the characteristics of the 1B & 2B and the 3B is carried out. In Part 2, analyses of the dynamic structure and the patterns of knowledge convergence in AI has been carried out.

#### Part 1

The methodology is based on a 3 step approach. In the first step, a bibliometric analyses of AI was carried out to analyse the technological accumulation over the 2B and 3B. In the first phase, having done the desk work and interviewed several experts on AI, we divided the two AI booms into 2 periods, 1990–2012 and 2013–2016. In the second stage, we did a bibliometric analyses of countries and author keyword co-occurrence<sup>1</sup> over the 2 periods. In the third stage, interviews were carried out with corporate members to do a qualitative analysis on the possible application areas of AI and identify the issues to be solved for AI adoption. However, in this paper, we will focus mainly on the first two steps.

One of the main challenges in doing an empirical analysis of AI is that the term AI is quite broad and general so that one has to first identify the key technologies underlying AI in order to apply techniques such as bibliometric analysis. We took into account a set of application keywords related to AI made by Jiqiang Niu and others in Niu, Tang, Xu, Zhou and Song (2016).

We revised the list of keywords based on a range of information sources searched including a report published by the Japanese Patent Office (2016) and by studying the practical applications of AI in IBM where one of the authors is employed (Kumar, 2017). In addition, we interviewed some experts on AI such as Professor Katsuo Nitta and Professor Takao Terano at Tokyo Institute of

<sup>1</sup> See Annex Section 2 for a detailed explanation of why we use author keywords.

Technology. In addition, the knowledge of one of the authors was used as she worked as a computer scientist in the field of knowledge based systems in industrial applications.

As a result, we revised the list by adding some keywords, such as Web science, deep learning, machine learning, neural networks, linear classifiers, logical regression, perception, support vector machine, quadratic classifiers, K-Means clustering, boosting, decision tree, random forest, Bayesian networks, and VR (virtual reality).

Bibliometric analysis was carried out using the search terms,

(“machine learning” or “neural network” or “Linear Classifiers” or “Logical Regression” or “Naive Bayes Classifier” or “Perceptron” or “Support Vector Machine” or “Quadratic Classifiers” or “K-Means Clustering” or “Boosting” or “Decision Tree” or “Random Forest” or “Bayesian Networks” or “Deep Learning” or “reasoning system” or “knowledge base” or “knowledge representation” or “fuzzy reasoning” or “fuzzy control” or “genetic algorithm” or “chaos theory”) AND (“speech recognition” or “image recognition” or “facial recognition” or “character recognition” or “natural language processing” or “data pattern recognition” or “visualization” or “auralization” or “Dialogue” or “agent” or “knowledge discovery” or “information retrieval” or “data mining” or “information recommend” or “Monitoring” or “diagnosis” or “control” or “optimization” or “design” or “prediction” or “web science” or “VR”)

Using Web of Science Core Collection, between the period 1990 and 2016, 210,576 papers were retrieved in January 2017. In the next stage, co-occurrence analysis of the author keywords was carried out over the 2 periods and we were able to extract the application keywords which only existed prior to the 3B and also the new application keywords which appeared in the 3B.

## Part 2

The second objective is achieved by implementing a simplified version of the knowledge convergence framework proposed by Ruiz-Navas and Miyazaki (2018) (Annex 1).

The simplified version of the method consists of ten steps, the first nine steps are applied for the 1B and 2B as well as for the 3B. Step 9 of analysis is a comparison of the results obtained in the 1 to 8 steps (Figure 2).

- 1) Defining the knowledge base.
- 2) Creating the author keyword co-occurrence network.
- 3) Identification of Keywords Proxy of Knowledge Convergence (KPKC).
- 4) Simplifying the KPKC and obtaining the simplified KPKC (S-KPKC)..
- 5) Detecting the S-KPKC related to knowledge to obtain the S-KPKC related to knowledge (S-KPKC-K).
- 6) Obtaining the S-KPKC-K’s Wikipedia categories.
- 7) Creating the network of S-KPKC-K Wikipedia categories co-occurrence.
- 8) Detecting the knowledge converging into AI. and
- 9) Detecting the knowledge converging into AI further convergence potential.

The last step is the analysis which consists of a qualitative comparison of the knowledge topics converging into AI detected for the 1B, 2B, and 3B. The comparison analyses included a descriptive analysis of each network (1B, 2B, and 3B) and qualitative analysis of the 3B network.

- 1) The descriptive analysis consisted of identifying the number of clusters detected in each network and compare them. The identification of a cluster consisted of determining the number of papers linked to the keywords in the cluster. Furthermore, the estimation of the eigenvector centralities of the keywords in the cluster. The cluster comparison was made to detect the unique knowledge converging keywords present in each network (1B, 2B, and 3B).
- 2) The qualitative analysis consisted of analyzing the papers linked to the most central keyword of the most average eigenvector central cluster of the unique keywords of the 3B.

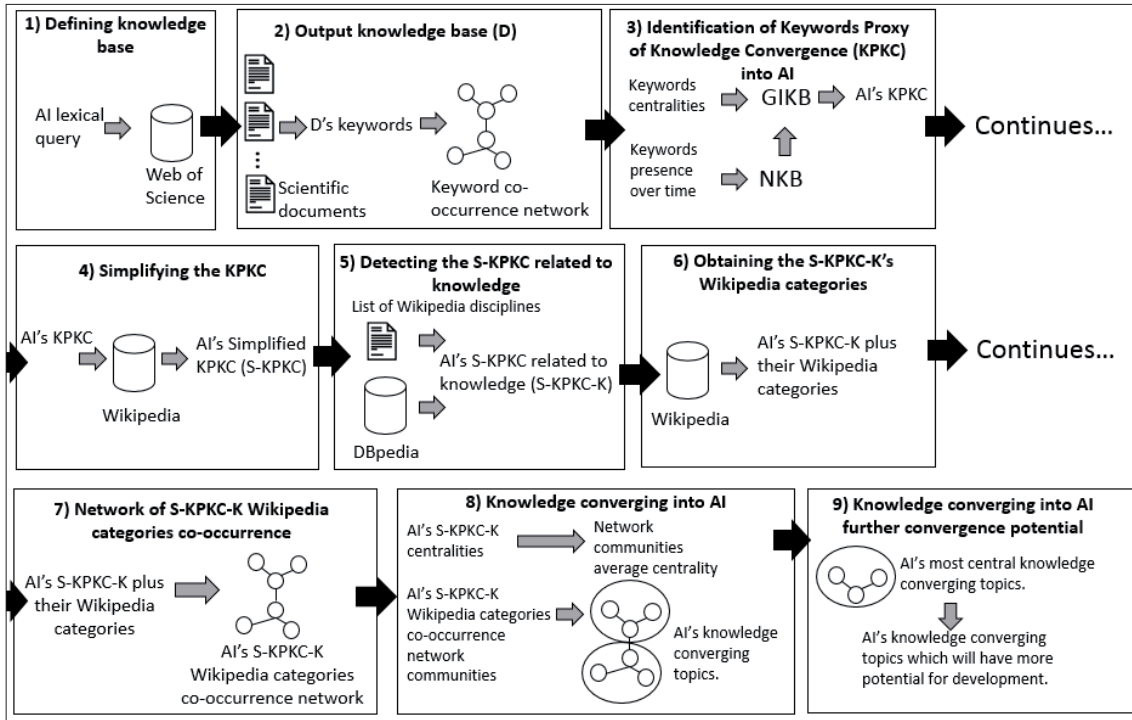


Figure 2. The overview of the first nine steps of the method.

The same lexical query shown above was used to obtain the documents set to represent the AI knowledge base. This lexical query was executed in the Web of Science from 1990 to 2017. Only conferences papers and journal articles were selected. Under these conditions, the data set is comprised of 249,300 documents. The dataset was divided for the periods of 1990–2013 (1B & 2B) and 1990–2017 (3B), to represent the knowledge convergence given in the 1B & 2B and the 3B of AI, respectively.

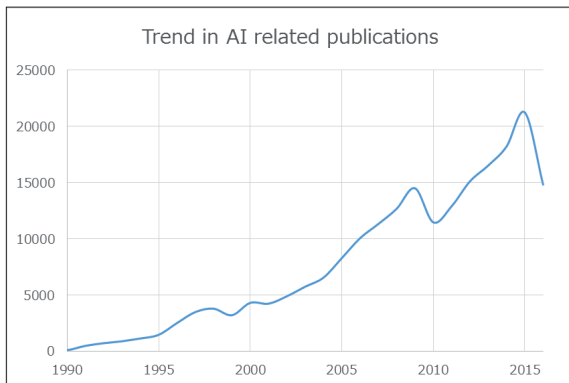


Figure 3. Trend in publications using the search term defined by the authors

## IV. FINDINGS

### Part 1

#### Trend in AI publications

The trend of AI related publications using the keyword ‘artificial intelligence’ led to 25,841 publications (Figure 3).

The trend in publications using the specific search terms defined by the authors resulted in 210,576 publications (Figure 3). The shape of

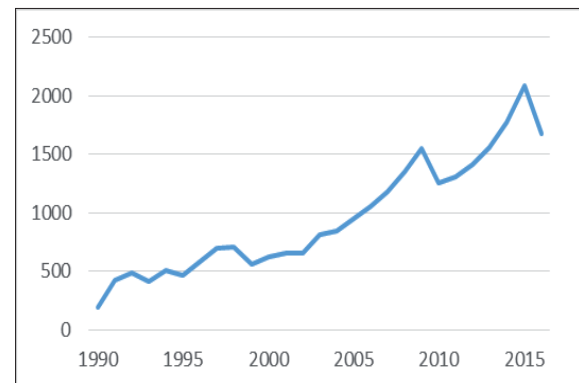


Figure 4. Trend in publications using the term including ‘artificial intelligence’

the trends in publications in Figure 3 and 4 are similar, confirming that the search terms defined above represent the field of AI. From Figure 3 the papers published prior to the 3B were 139,795, compared to the papers published during the 3B (2013–2016) which were 70,781, indicating that 34% of the publications have been published during the last 4 years, confirming the presence of the 3B of AI.

**Analyses of key technologies and applications in the 2 periods**

The occurrences of the author keywords before the 3B and during the 3B were analyzed. Table 2 shows the keywords which existed prior to the 3B but were not retrieved in the 3B. Keywords such as Fuzzy systems, Fuzzy controller, Fuzzy reasoning, Chaos, Knowledge discovery, Nonlinear systems have disappeared in the 3B. Table 1 contains a list of author keywords which were only retrieved in the 3B such as Deep learning, deep neural network. Also, keywords such as smart grid, big data, cloud computing, energy efficiency have emerged, indicating that in the 3B, AI is beginning to converge with other fields which are related to applications of AI.

Co-occurrences of author keywords were analyzed and mapped using R’s free software called KH Coder. The lines linking the keywords indicate the co-occurrences of the author

keywords in a paper. The red nodes indicate a higher degree of co-occurrences with the various keywords than the blue nodes. For example, as shown in Figure 5, Genetic algorithm, Neural networks and Optimization are the 3 main keywords which represent the components of AI prior to the 3B. On the other hand, Classification has co-occurrences with many different keywords such as Machine learning, Support Vector machine, Decision tree, Feature selection, and Data mining, which represent algorithms underlying AI.

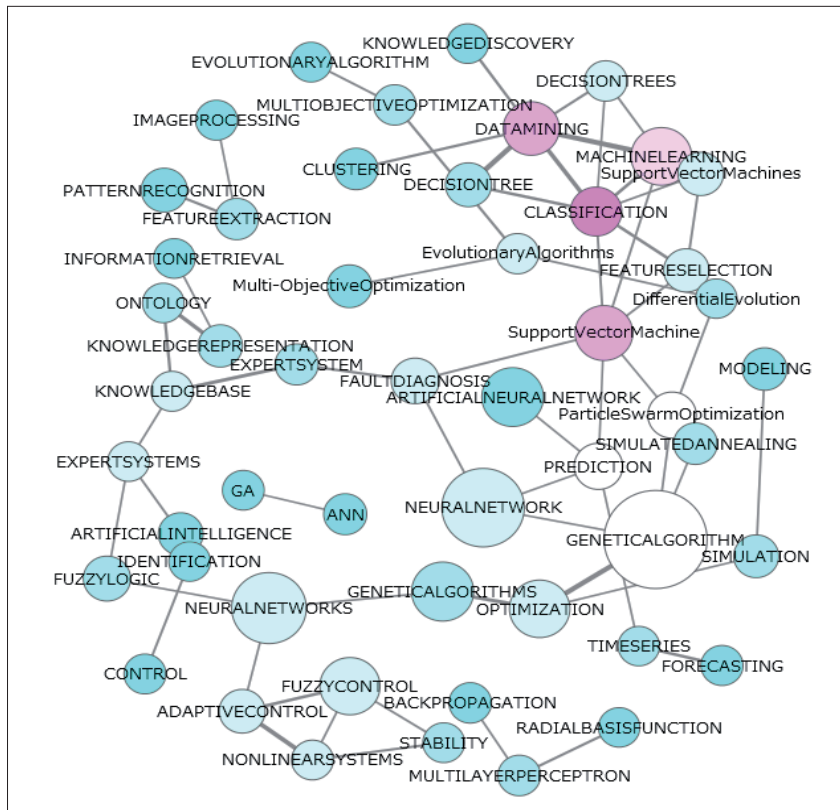
Figures 5 and 6 show the co-occurrence network based on author keywords during the 3B. The size of the circle indicates the frequency of co-occurrence. Machine learning has become as prominent as Genetic algorithm indicating that the 3B is strongly linked to Machine learning. We also note that Machine learning is linked with Data mining, suggesting that the two are closely related. The map also highlights the strong connection between Deep learning and Convolutional Neural Network, as well as ANFIS and ANN. Similar to the previous analysis based on the keywords which existed prior to the boom and during the boom, this analysis also revealed the emergence of new applications related keywords such as Energy efficiency, Big data and Cloud computing, and Scheduling, confirming that AI is beginning to be adopted by different applications.

**Table 1.**  
Keywords which existed only in the 3B

NSGA-II	Deep neural network
Deep learning	Local search
Support vector regression	Multi-objective
ANFIS	Swarm intelligence
Random forest	Parameter estimation
Response surface methodology	Regression
Big data	Taguchi method
Cloud computing	Smart grid
Reliability	Wavelet neural network
Extreme learning machine	Energy efficiency
Metaheuristics	Remote sensing
	Surface roughness
	PID controller

**Table 2.**  
Keywords which were retrieved PRIOR to the 3B

Intelligent control	Recurrent neural networks
Nonlinear systems	Fuzzy systems
Control	Backpropagation
Modelling	Nonlinear control
Chaos	Induction motor
Learning	Robustness
Identification	Fuzzy controller
Knowledge discovery	Bioinformatics
Process control	Predictive control
Tabu search	Fuzzy logic control
Robust control	Computer-aided diagnosis
	Heuristics
	Fuzzy Reasoning
	Nonlinear system



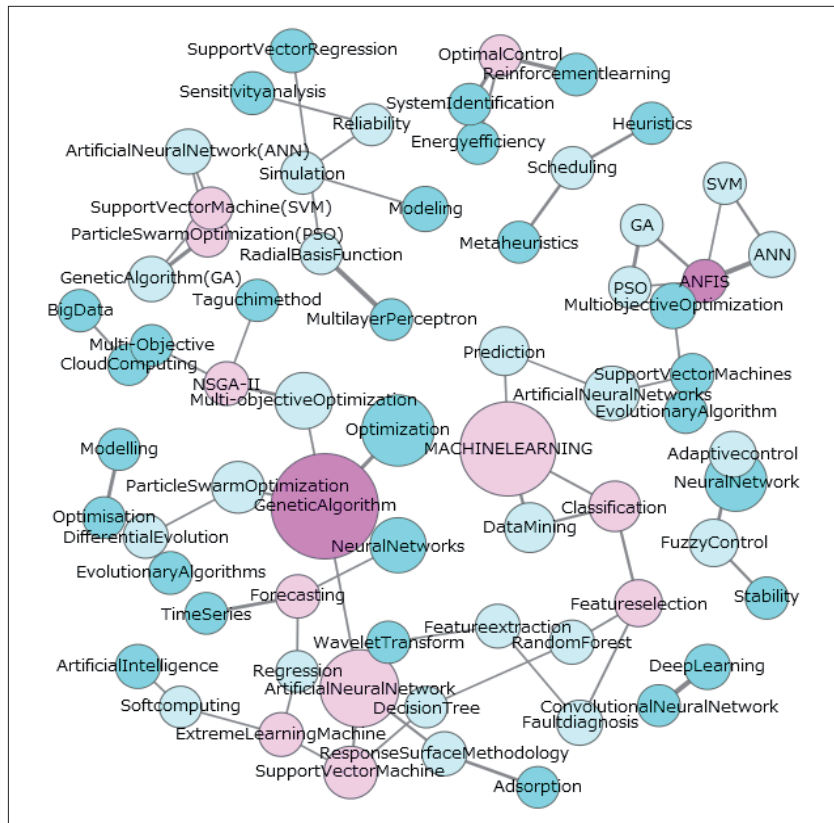
**Figure 5.** Co-occurrence Network based on author keywords prior to the 1B&2B AI boom (1990–2012)

**Table 3.**

The rate of occurrence of application keywords prior to the 3B and during the 3B

Rank	Application Keywords	Stage	The rate of occurrence of keywords prior to the 3rd (occurrences/period)	The rate of occurrence of the keywords during the 3rd
1	Expert system	0	27.95	34.25
2	Optimization	0	166.68	528.50
3	Prediction, Forecast	1	56.91	174.75
4	Machine learning, Classification, Data Mining, Recognition, Pattern Recognition, Deep Learning	1	191.32	647.75
5	Design	0	15.23	34.25
6	Simulation	0	43.09	84
7	Diagnosis, Fault Diagnosis	1	68.09	150.25
8	Decision support system (DSS)	0	-	-
9	Knowledge-based system(KBS)	0	-	-
10	Case-based reasoning(CBR)	0	-	-
11	Knowledge representation(KR)	0	39.68	60.25
12	Management	1	-	-
13	Identification	1	14.45	^
14	Decision making	1	-	^
15	Navigation, Self-driving	1+2	^	^
16	Game	2	-	-
17	Matching, Scheduling	2+3	-	59.75
18	Robot, Automation, Remote Sensing	2+3	-	29.50
19	Smart Grid, Energy Efficiency	3	-	62.00





**Figure 6.** Co-occurrence network based on author keywords during the 3B (2013–2016)

In the next stage, the list of applications keywords was revised by two steps. First, taking into account the suggestions by the experts interviewed, and second, IBM's cases of cognitive adoption (Kumar, 2017), and also using the results of the analyses in the previous section where keywords related to applications such as Smart Grid, Energy efficiency, Remote sensing, Scheduling were identified. The final list is shown in Table 3. The keywords in bold letters are the newly added keywords. Stage 0, means that the keyword existed in the original applications keyword list put forward by Niu, et al. (2016). Stage 1 refers to the keywords which were suggested by the experts interviewed. Stage 2 refers to the keywords found in IBM's cases of cognitive adoption or in the news; Stage 3 refers to the newly identified keywords in the previous section. The figures in Table 3 in the right hand column indicate the rate of occurrence of keywords prior to (or during) the 3rd boom, which is the total number of occurrences divided by the number of years.

The analyses of keywords has confirmed that during the 3B, applications such as Optimization, Prediction, Forecasting, Machine learning, Classification, Data mining, Pattern recognition, Deep learning have become high priority applications. Less important applications include Fault diagnosis, Knowledge representation, Expert systems.

## RESULTS from PART 2

### *Descriptive communities of cluster analysis*

In Figure 7, we present an overview of the principal keywords of the clusters detected for the first and second AI boom knowledge converging keywords network.

In Figure 7, the radius of the circles represents their Eigenvector Centrality, e.g., the node "Natural Language Understanding (NLU)" is the biggest node, and this means that this node is closely connected with all the other nodes of the network. The smaller nodes represent those which are highly connected to the other nodes of the network and other highly connected nodes such





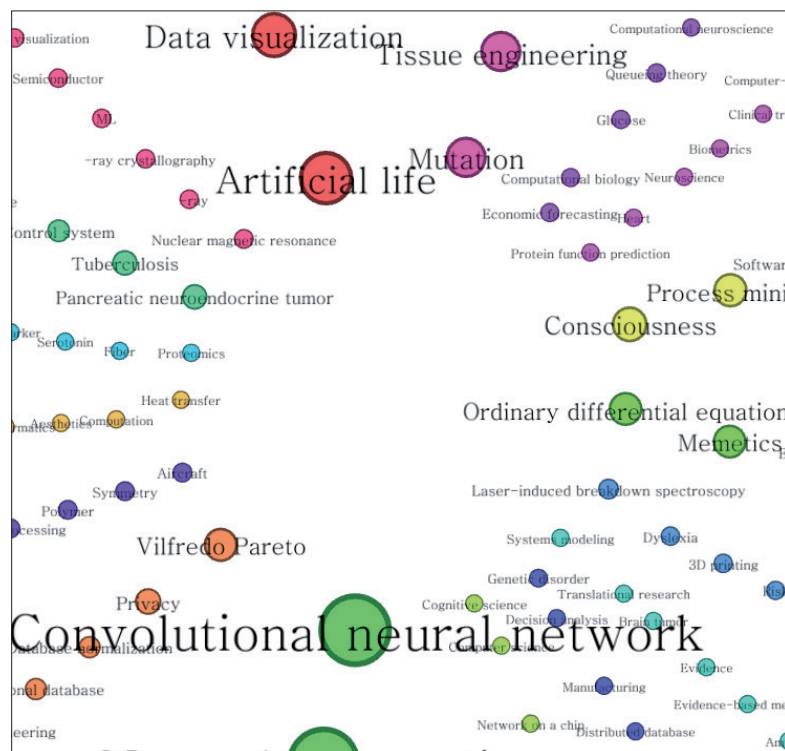
**Table 4.**

Characterization of the clusters of knowledge convergence detected for the 1B and 2B AI booms.

CN	NW	L1	L2	NP	AVGEGC
1	23	Pancreatic neuroendocrine tumor   Tuberculosis	Control system   Network theory   Prostate cancer	3204	0.03148
2	38	Model of computation	Computer network   Image analysis   Soft error	3178	0.02926
3	13	Consciousness	Visual system   Chronic obstructive pulmonary disease   Systems modeling	374	0.02730
4	14	Evolution	Object-oriented design   Biometrics   Clinical trial	753	0.00004
5**	44	Dyslexia	Retina   Digital signal processing   Combinatorial optimization	18354	0.00239
6*	35	Natural language understanding (NLU)   Recycling	Privacy   Vilfredo Pareto   Web Ontology Language	1634	0.42593
7	34	Forging	Aircraft   Atrial fibrillation   Metabolism	350	0.00715
8**	13	Database normalization   Relational database	Statistical model   Drug discovery   South Africa	29900	0.02162
9	12	Steel	Morphology   Health   Color	312	0.00124
10	26	Statistics	Information technology   Evolutionary computation   Automation	6616	0.19856
11	22	Nursing	Electric vehicle   Phenotype   Sound localization	1286	0.00096
12	25	Artificial life	Vaccine   Design   Tourism	2598	0.35439

\*most prominent clusters regarding average eigenvector centrality.

\*\*most prominent clusters regarding number of papers.



**Figure 8.** A manually edited version of a radial layout of the 3B knowledge converging keywords network.

**Table 5.**

Characterization of the clusters of knowledge convergence detected for the 3B.

CN	NW	L1	L2	NP	AVGEGC
1	25	Artificial life	Supramolecular chemistry   Cervical cancer   Exploration	2380	0.31231
2	5	Genetic disorder	Decision analysis	1020	0.00218
3	6	Systems modeling	Translational research	162	0.00002
4**	43	Laser-induced breakdown spectroscopy	Design of experiments   Vibration   Risk	28224	0.00615
5	44	Ordinary differential equation	Energy conservation   Chemometrics   Geographic information system	8771	0.02143
6	56	Consciousness	Model of computation   Computer network   Computer programming	5037	0.02454
7	14	Protein function prediction	Clinical trial   Neuroscience   Melanoma	1219	0.00016
8	21	Economic forecasting	Assistive technology   Object-oriented design   CT scan	1734	0.00408
9**	28	Mutation   Titanium dioxide	Evolutionary robotics   Automation   Conceptual design	11995	0.15833
10	18	Nuclear magnetic resonance	Genetics   Semiconductor   X-ray	4660	0.01272
11**	33	Pancreatic neuroendocrine tumor   Tuberculosis	Game   Statistical learning theory   Statistical mechanics	12285	0.04304
12	15	Proteomics	Liver disease   Climate change   Fiber	1311	0.00019
13	24	Heat transfer	Espionage   Thin film   Cirrhosis	1147	0.00014
14	58	Aircraft	Electric vehicle   Inductive logic programming   Forging	3797	0.00496
15	58	Vilfredo Pareto	Blog   Database normalization   Relational database	3814	0.01738
16***	42	Convolutional neural network   Neurodegeneration   Optimizing compiler	Methylphenidate   Neural network   Artemisinin	36337	0.68031
17	14	Cognitive science	Telecommunication   Virtual reality   Network on a chip	495	0.00006

\*most prominent clusters regarding average eigenvector centrality.

\*\*most prominent clusters regarding number of papers.

\*\*\*most prominent clusters regarding number of papers and average eigenvector centrality.

In Table 5, the columns have the same meaning as in Table 4. What we can see from Table 5 is the significance of the clusters:

- **4** (Laser-induced breakdown spectroscopy);
- **11** (Pancreatic neuroendocrine tumor | Tuberculosis); and
- **9** (Mutation | Titanium dioxide).

In 3B, the most prominent cluster in terms of both number of papers and average eigenvector centrality was the number 16 (Convolutional neural network | Neurodegeneration | Optimizing compiler).

### ***Comparison between unique keywords of the 1B and 2B and 3B***

We calculated the unique keywords for each network, identified their clusters and their centrality values. In Tables 6 and 7, we present the percentage of unique keywords classified by their cluster for the 1B and 2B and 3B, respectively. In Tables 6 and 7, the column CN explain the cluster number. #UKkeywords stands for the number of unique knowledge convergence keywords identified for a given cluster. % column describes the relation of the number of unique knowledge converging

**Table 6.**

Top ten, clusters organized by their percentage of unique knowledge convergence keywords for the 1B and 2B.

CN	#UKeywords	%
3▲	4	23.53%
7	3	17.65%
2	2	11.76%
10	2	11.76%
11	2	11.76%
4	1	5.88%
5**	1	5.88%
6*	1	5.88%
12	1	5.88%
1	0	0.00%

▲ Cluster with more unique keywords

\*most prominent clusters regarding average eigenvector centrality.

\*\*most prominent clusters regarding number of papers.

\*\*\*most prominent clusters regarding number of papers and average eigenvector centrality.

keywords present in each cluster concerning the total of unique knowledge converging keywords.

From Table 6, we can see that the cluster that possessed the higher values of unique knowledge convergent keywords for the 1B and 2B AI booms are number 3 (Consciousness) and 7 (Forging).

From Table 7, we can see that the cluster that possessed the higher values of unique knowledge convergence keywords for the 3B are:

- 14(Aircraft),
- 6 (Consciousness),
- 5 (Ordinary differential equation),
- 4 (Laser-induced breakdown spectroscopy), and
- 15 (Vilfredo Pareto).

In Table 8 and 9, we present the unique knowledge converging keywords (KCKW) with the highest eigenvector centrality for the cluster that were found to be unique to each AI boom. In Tables 8 and 9, the column KCKW describes the highest eigenvector centrality, knowledge convergence keywords for a given cluster.

**Table 7.**

Top ten, clusters organized by their percentage of unique knowledge convergence keywords for the 3B.

CN	#UKeywords	%
14▲	25	11.26%
6	24	10.81%
5	23	10.36%
4**	22	9.91%
15	21	9.46%
16***	16	7.21%
17	14	6.31%
13	12	5.41%
1	11	4.95%
11**	10	4.50%

▲ Cluster with more unique keywords

\*most prominent clusters regarding average eigenvector centrality.

\*\*most prominent clusters regarding number of papers.

\*\*\*most prominent clusters regarding number of papers and average eigenvector centrality.

**Table 8.**

Unique to the 1B and 2B AI boom, Knowledge converging keywords with higher values of eigenvector centrality for each cluster.

CN	L1	KCKW
2	Model of computation	Multi-agent system
3▲	Consciousness	Alzheimer’s disease
4	Evolution	Evolution
5**	Dyslexia	Solar cell
6*	Natural language understanding   Recycling	Reliability engineering
7	Forging	Carbon nanotube
10	Statistics	Wireless sensor network
11	Nursing	Nonlinear control
12	Artificial life	Forecasting

▲ Cluster with more unique keywords

\*most prominent clusters regarding average eigenvector centrality.

\*\*most prominent clusters regarding number of papers.

\*\*\*most prominent clusters regarding number of papers and average eigenvector centrality.

TABLE 9. UNIQUE TO THE 3B, KNOWLEDGE CONVERGING KEYWORDS WITH HIGHER VALUES OF EIGENVECTOR CENTRALITY FOR EACH CLUSTER.

CN	L1	KCKW
1	Artificial life	Neuroanatomy
2	Genetic disorder	Genetic disorder
3	Systems modeling	Translational research
4**	Laser-induced breakdown spectroscopy	Laser-induced breakdown spectroscopy
5	Ordinary differential equation	Memetics
6	Consciousness	Process mining
7	Protein function prediction	Protein function prediction
8	Economic forecasting	Economic forecasting
9	Mutation   Titanium dioxide	Tissue engineering
10	Nuclear magnetic resonance	X-ray crystallography
11**	Pancreatic neuroendocrine tumor   Tuberculosis	Game
12	Proteomics	Fiber
13	Heat transfer	Aesthetics
14▲	Aircraft	Metabolic engineering
15	Vilfredo Pareto	Probability
16***	Convolutional neural network   Neurodegeneration   Optimizing compiler	Artemisinin
17	Cognitive science	Cognitive science

▲ Cluster with more unique keywords

\*most prominent clusters regarding average eigenvector centrality.

\*\*most prominent clusters regarding number of papers.

\*\*\*most prominent clusters regarding number of papers and average eigenvector centrality.

### Conclusions from the analysis

The 3B has more clusters, and this reflects the growing knowledge base of the field and the continuous development of the third boom. The diversity of fields of study of the words found in Table 5, (computer science, omics, genetics, neuroscience, engineering among others) reinforces the results obtained in the first part of this paper.

Concerning the process of knowledge convergence, it was expected that the most central cluster from the 1B and 2B, cluster 6 (Natural language understanding | Recycling), was going to benefit from knowledge convergence into AI, and be further developed. In the 3B, we could confirm this expectation. The keywords “Natural language understanding | Recycling” representative of the cluster 6 of the 1B and 2B, was found to have been encompassed by the cluster 16 (Convolutional neural network | Neurodegeneration | Optimizing compiler) of the 3B.

Cluster 5 (Dyslexia) from the 1B and 2B boom, in the 3B, it was encompassed by cluster 4 (Laser-induced breakdown spectroscopy), which is found to be a topic of interest in the 3B as well.

Cluster 8, (Database normalization), from the first and second boom, became encompassed by cluster 15 (Vilfredo Pareto) in the 3B, which is not central and also covers fewer papers than its previous version, indicating that this concept lost attraction in the 3B.

### Comparison of unique keywords

In the 3B, there is a significant number of unique keywords, again reflecting the findings from the first part, regarding an increased application of AI in different fields. As an additional analysis, we revised the papers belonging to the top unique cluster from the third boom.

The most central keyword of the unique cluster 14 (Aircraft) of the 3B was “metabolic engineering.” We looked into the papers linked

to this keyword and found applications such as optimization of bacteria's metabolic pathways to improve a defined function. However, "metabolic engineering" was also linked to documents not related to AI, because, of the use in the lexical query of the word "boosting" and "knowledge base." Taking into account this issue we then analysed the second most central keyword, which was "sound." All the papers related to sound were relevant. Among the applications related to sound, we report:

- 1) Optimizing the design of structures to reduce sound by vibration,
- 2) Detection of gas leaks using sound under different sound environments,
- 3) Monitoring tool wear using different variables among sound from the tool,
- 4) Computer-aided orchestration sound samples,
- 5) Generating sound summaries of environments to be used in architecture modeling (simulating sound on a theatre or stadium),
- 6) Gear system fault detection, using the sound of the system. and
- 7) Modeling gas emissions using diverse sources of data, among them traffic sounds.

Cluster 14 (Aircraft) has 25 unique keywords, which link with other applications, indicating many other 3B only applications of AI. The objective of the previous example showed a sample of the broader type of problems that are being solved implementing AI-related techniques.

## V. DISCUSSION

The goal of this paper was to elucidate the evolutionary path of development of AI and the structural patterns of knowledge convergence in the current and previous booms.

The bibliometric analysis in the 3B of Part one identified Machine learning and deep learning as crucial technologies of AI. The main focus of the research has been shifting towards AI applications in the 3B, such as Energy efficiency, and Scheduling. AI can be used in many applications depending on the purpose. AI has been used in optimizing, in the 1st boom, expert systems in the

2nd boom and in the current 3rd boom; AI has evolved to the stage of using machine learning and deep learning creating capabilities in sensing, learning, identifying, tracking, forecasting by combining with other technologies such as big data and cloud. It will continue to evolve, having a significant impact on our society.

The analysis has shown some keywords which were only retrieved in the 3B. Among them were big data and cloud computing. Compared to the previous AI booms, the current AI boom is beginning to have a real impact on companies and society. AI is a technology (or a group of technologies) which could be adopted in a variety of sectors and applications.

Concerning the method used in the second part, we discuss the number of papers linked to each cluster and the implications of the change in their eigenvector centrality over time (1B, 2B, and 3B).

We observed from the information provided from the number of papers linked to a cluster in both periods of times that:

- 1) Cluster 8 (Database normalization | Relational database) from 1B and 2B in the 3B was encompassed by Cluster 15 (Vilfredo Pareto) which was not central, nor linked to a considerable amount of papers; and
- 2) Cluster 5 (Dyslexia) from the 1B and 2B in the 3B was covered by Cluster 4 (Laser-induced breakdown spectroscopy) which was linked to a considerable amount of documents.

From this observation, we argue that a high number of documents in one period will not guarantee a high number of documents in the following period, i.e., the number of papers is a useful indicator to see a picture (static analysis) but not to predict what will happen.

The measure of average eigenvector centrality was used to represent the cluster of knowledge convergence which would have more potential to continue further development. This hypothesis was consistent in the evolution of Cluster 6 (Natural language understanding | Recycling) of the 1B & 2B to cluster 16 of the 3B. The keywords in cluster 6 (Natural language understand-



ing | Recycling) of the 1B & 2B were present in Cluster 16 (Convolutional neural network | Neurodegeneration | Optimizing compiler) of the 3B. Therefore, indicating the development of the knowledge represented by the keywords over the 1B & 2B to the 3B.

Also, following the relationships between knowledge convergence and technological convergence, we argue that the topics covered in Cluster 16 (Convolutional neural network | Neurodegeneration | Optimizing compiler) of the 3B, might be a signal of the existence of technological and application convergence around them. Therefore, the keywords in cluster 16 represent AI applications with possible innovation potential.

The method proposed in the second part showed us a structural view of the knowledge convergence around AI, and to explore it in detail by using keywords as indexes to the documents related to each keyword. In other words, for further works, it would be possible to zoom into the detected covering topics and expand the description of such issues and applications related to them.

The second part confirmed the results obtained in the first part, showing how, in the third boom, a relative high number of unique knowledge converging keywords were present. Further research may explore how AI can drive innovation in different sectors or to evaluate how can current policies or national systems of innovation get value from such transversal potential that AI will have in our society.

We highlighted three limitations of this study.

- 1) It used one search strategy (lexical query), other search strategies such as citation, lexical query expansion and specialized journals could have been useful to extend the coverage of the AI data set studied.
- 2) The implementation of author keywords provided advantages for a quick analysis. However, the analysis can be enriched by integrating to the analysis words of other sections of the paper, such as title, abstracts, full content and subject categories.
- 3) The techniques used explored in a general sense the scientific papers related to AI, thus, studies on the specific applications

of AI such as military, fighting cancer and geriatrics; could provide detailed insights on how AI is being applied.

## REFERENCES

- Ávila-Robinson, A. & Miyazaki, K. (2014). Assessing nanotechnology potentials: Interplay between the paths of knowledge evolution and the patterns of competence building. *Int. J. Technol. Intell. Plan*, 10(1), 1–28.
- Ávila-Robinson, A. & Miyazaki, K. (2011). Conceptualization and Operationalization of Emerging Technologies : A Complementing Approach. in *Technology Management in the Energy Smart World (PICMET), 2011 Proceedings of PICMET '11*, 2011, pp. 1681–1692.
- Ávila-Robinson, A. & Miyazaki, K. (2013). Dynamics of scientific knowledge bases as proxies for discerning technological emergence - The case of MEMS/NEMS technologies. *Technol. Forecast. Soc. Change*, 80(6), 1071–1084, July.
- Callon, M. (1990). Techno-economic Networks and Irreversibility. *The Sociological Review May*. 38 (Suppl-1): 132–161. <https://doi.org/10.1111/j.1467-954X.1990.tb03351.x>
- Callon, M., Laredo, P., Rabeharisoa, V., Gonard, T. & Leray, T. (1992). The management and evaluation of technological programs and the dynamics of techno-economic networks: The case of the AFME. *Res. Policy*, 21(3), 215–236.
- Chan, S. K. & Miyazaki, K. (2015). Knowledge convergence between cloud computing and big data and analysis of emerging technological opportunities in Malaysia. *Portl. Int. Conf. Manag. Eng. Technol.*, 2015-Septe, 1501–1512.
- Curran, C.S., Bröring, S. & Leker, J. (2010). Anticipating converging industries using publicly available data. *Technol. Forecast. Soc. Change*, 77(3), 385–395, March.
- Curran, C.S. & Leker, J. (2011). Patent indicators for monitoring convergence-Examples from NFF and ICT. *Technol. Forecast. Soc. Change*, 78(2), 256–273.
- Cucerzan, S. (2007). Large-scale named entity disambiguation based on Wikipedia data. *EMNLP-CoNLL 2007, June*, 708–716.
- Fujimoto, M., Miyazaki, K. & von Tunzelmann, N. (2000). Technological fusion and telemedicine in Japanese companies. *Technovation*, 20(4), 169–187, Apr.
- Gaines, B. R. (1998). The learning curves underlying convergence. *Technol. Forecast. Soc. Change*, 57(1–2): 7–34, January.

- Hacklin, F. (2008). *Management of convergence in innovation: Strategies and capabilities for value creation beyond blurring industry boundaries*, 1st ed. Heidelberg: Physica-Verlag.
- Hacklin, F., Raurich, V. & Marxt, C. (2005). Implications of technological convergence on innovation trajectories: The case of ICT industry. *Int. J. Innov. Technol. Manag.*, 2(3), 313–330, September.
- Hacklin, F., Marxt, C. & Fahrni, F. (2009). Coevolutionary cycles of convergence: An extrapolation from the ICT industry. *Technol. Forecast. Soc. Change*, 76(6), 723–736, July.
- Han, E. J. & Sohn, S. Y. (2016). Technological convergence in standards for information and communication technologies. *Technol. Forecast. Soc. Change*, 106, 1–10.
- Hou, H., Kretschmer, H. & Liu, Z. (2008). The structure of scientific collaboration networks in Scientometrics. *Scientometrics*, 75(2), 89–202.
- Islam, N. & Miyazaki, K. (2009). Nanotechnology innovation system: Understanding hidden dynamics of nanoscience fusion trajectories. *Technol. Forecast. Soc. Change*, 76(1), 128–140, Jan.
- Japanese Patent Office. (2016). 特許出願技術動向調査報告書（概要） [Online]. Available: [https://www.jpo.go.jp/shiryoku/pdf/gidouhoukoku/26\\_21.pdf](https://www.jpo.go.jp/shiryoku/pdf/gidouhoukoku/26_21.pdf). [Accessed: 11-Nov-2017].
- Kumar, K. (2017). How Honda is using cognitive search to drive real changes in quality assurance. *IBM ECM Blog*, 2017. [Online]. Available: <https://www.ibm.com/blogs/ecm/2017/08/30/honda-using-cognitive-search-drive-real-changes-quality-assurance/>. [Accessed: 11-Nov-2018]
- Lalitnorasate, P. & Miyazaki, K. (2016). Convergence in functional food: Technological diversification and path-dependent learning. *Int. J. Technol. Intell. Plan.*, 11(2), 140.
- Lee, W.S., Han, E. J., & Sohn, S.J. (2015). Predicting the pattern of technology convergence using big-data technology on large-scale triadic patents. *Technol. Forecast. Soc. Change*, 100, 317–329.
- Lee, K., Yun, J.J., & Jeong, E-S. (2015). Convergence innovation of the textile machinery industry in Korea. *Asian J. Technol. Innov.*, 23(sup1), 58–73.
- Leydesdroff, L. (1989). Words and co-words as indicators of intellectual organization. *Res. Policy*, 18(4), 209–223.
- Matsuo, Y. (2016). Future of AI what is beyond Deep Learning. *Technol. Econ.* 9(595): 10–25
- McCarthy, J. (2007). Branches of AI. [Online]. Available: <http://www-formal.stanford.edu/jmc/whatisai/node2.html>. [Accessed: 11-Nov-2017]
- Milne, D. & Witten, I. H. (2008). Learning to link with Wikipedia. *Proceeding 17th ACM Conf. Inf. Knowl. Manag. (CIKM '08)*, pp. 509–518.
- Miyazaki, K & Giraldo E. (2015). Innovation strategy and technological competence building to provide next generation network and services through convergence – the case of NTT in Japan. *Asian J. Technol. Innov.*, 23(sup.1), 74–92.
- Niu, J., Tang, W., Xu, F., Zhou, X. & Song, Y. (2016). Global research on artificial intelligence from 1990–2014: Spatially-explicit bibliometric analysis. *ISPRS Int. J. Geo-Information*, 5(5), 66.
- Ruiz-Navas, S., & Miyazaki, K. (2018). Developing a framework to track knowledge convergence in big data. *Int. J. Technol. Intell. Plan*, 12(2), 121–151. DOI: 10.1504/IJTIP.2018.096101
- Rosenberg, N. (1963). Technological change in the machine tool industry, 1840–1910. *J. Econ. Hist.* 23(4), 414–443.
- Velden, T., Boyack, K. W., Gläser, J., Koopman, R., Scharnhorst, A. & Wang, S. (2017). Comparison of topic extraction approaches and their results. *Scientometrics*, 111(2), 1169–1221.
- Yoon, B. & Park, Y. (2004). A text-mining-based patent network: Analytical tool for high-technology trend. *J. High Technol. Manag. Res.*, 15(1), 37–50, Feb.
- Wikipedia, “Wikipedia,” *wikipedia about*, 2017. [Online]. Available: <https://en.wikipedia.org/wiki/Wikipedia:About>. [Accessed: 07-Apr-2017].



## Annex 1, Part 2 Methodological Details

### 1. Network as a representation of the knowledge convergence and eigenvector centrality

We proposed a framework to identify knowledge convergence in Ruiz-Navas and Miyazaki (2018); the framework is based on the following hypotheses:

- Keywords are a proxy of knowledge (Callon, 1990; Callon, et al., 1992; Fujimoto, Miyazaki & Tunzelmann, 2000; Yoon & Park, 2004; Islam & Miyazaki, 2009; Avila-Robinson & Miyazaki, 2014).
- Network of keyword co-occurrence can provide a structural representation of knowledge (Hou, Kretschmer, & Liu, 2008; Avila-Robinson & Miyazaki, 2011).
- The eigenvector centrality of a keyword in a keyword co-occurrence network is a proxy of the importance of that keyword in the network (Avila-Robinson & Miyazaki, 2013).

### 2. Keywords as proxy of knowledge

Author keywords were selected over index keywords and automatically generated keywords from title, abstract or full body of the text. On the one hand, automatically generated keywords from title, abstract or full body text will depend on the efficiency of the algorithms used to obtain them and not necessarily reflect the main ideas of the documents. On the other hand, index keywords can be an object of indexing and are provided by professional indexers who have a general view of the content of the paper rather than an expert's view (Leydesdroff, 1989). Therefore, author keywords were selected because the original authors provide them and deliver full n-grams that describe the content of the article.

### 3. Wikipedia as a tool to disambiguate word meaning

Wikipedia was selected as an established knowledge source to help in the word disambiguation process (Velden, et al., 2017; Cucerzan, 2007; Milne & Witten, 2008; Wikipedia, 2017).

### 4. The knowledge convergence networks

In this section, we will briefly introduce the concepts of keywords proxy of knowledge convergence and characteristics of the networks presented in Figures 7 and 8 of this paper.

#### a. Characteristics of the knowledge convergence networks

The network presented in Figures 7 and 8 are networks of Wikipedia categories co-occurrence. The nodes of the network are Wikipedia categories of keywords proxy of knowledge convergence. The edges are co-occurrence of the Wikipedia categories in each keyword proxy of knowledge convergence. Furthermore, the nodes are aligned by cluster and their size is proportional to their normalized eigenvector centrality.

#### b. Keywords proxy of knowledge convergence

We proposed two characteristics to identify keywords proxy of knowledge convergence in a keyword data set over time. The two characteristics are:

- Newness to the knowledge base (NKB): the converging knowledge has to be previously absent in the output knowledge base.
- Growing importance in the output knowledge base (GIKB): the converging knowledge has to erode established boundaries that isolated the original knowledge bases; in other words, become one with the output knowledge base.

Considering the previous descriptions, the networks in Figs 6 and 7 allow us to see a graphical representation of the knowledge converging around AI at the keyword level of detail. Furthermore, a few of the most eigenvector central keywords of the clusters of knowledge converging around AI.