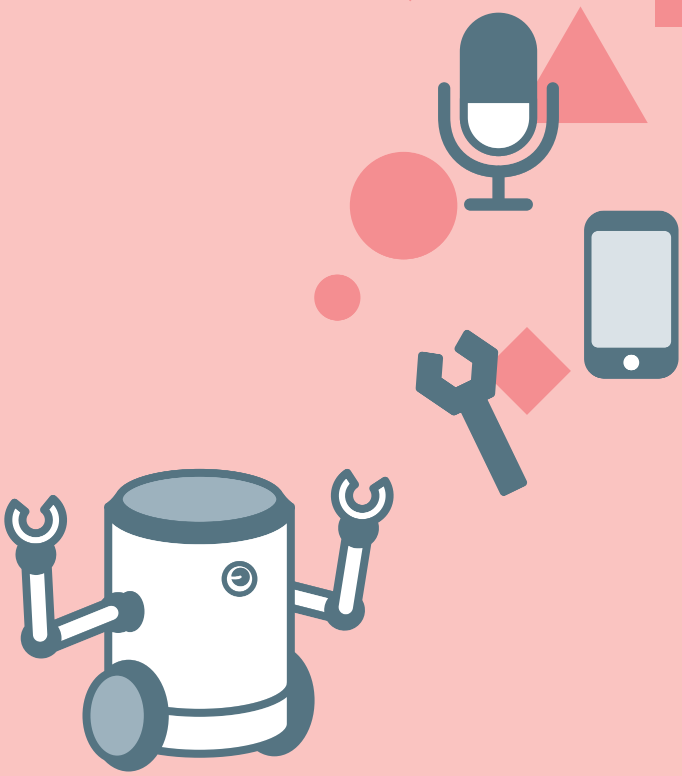
Part 1

Trends in Research and Technology



Part 1: Research and Technology Trends

[Overview]

Through the ongoing artificial intelligence (AI) boom, society's interest in information and communication technology (ICT) is increasing not only in fields driving the boom, such as machine learning, but also in a wide range of related technical fields that have not received keen attention in recent years. AI community has experiences cycles of "AI spring" and "AI winter." Caution is necessary over excessive expectations. It is said of not only ICT including AI but also new technology in general that technology responds to various needs in society while itself changing society, and giving rise to issues in tandem. It is preferable that society respond to these issues. In order that the AI boom not be a transient one, it is important to recognize that AI is developed and used in interaction with society.

Technologies related to AI are fragmented, and it is difficult to address them all. Therefore, Part 1 outlines the following nine research topis in three domains, (1) topics related to knowledge and data processing, (2) topics related to the boundary between human and machine, and (3) topics related to industry and daily life.

① Topics related to knowledge and data processing

In addition to the knowledge processing and machine learning technologies driving the current AI boom, natural language processing and image acquisition / recognition, which have shown remarkable development as a result of this boom, are described.

② Topics related to the human–machine boundary

Through the widespread application of AI and robotics, the importance of the conveyance of information between humans and machines is increasing. The speech interfaces responsible for this, and the human–agent interaction, which handle the interactions between machines and humans are described.

③ Topics related to lifestyle and industry

In addition to robot and IoT (internet of things), which continue to permeate into society, changing industry and daily life, multi-agent systems and crowdsourcing, which are the fundamental technologies for the design and operation of new social structures and systems, are described.

Notable social background, technological trends within Japan and overseas, real-world applications, and promising areas of application for each research topic are described, then the social issues that may arise in the future are presented.

I Knowledge Processing and Machine Learning

1. Historical Background

Looking back at the history[1] of artificial intelligence, several underlying technologies have been studied in each era.[2] In the 1960s, the task of expressing intellectual works by combining various items and how to effectively find a solution from among them (called the "search problem") was addressed as an issue central to AI research.[3] For example, when playing Shogi, from among the various combinations of moves, it is necessary to efficiently search for the move that leads to victory. The method of searching for this solution is called a "search" technique.[4] However, for example, when trying to consider combinations of moves in Shogi, the number of moves is enormous and searching is not straightforward.[5] Moreover, when translating, it is necessary not only to understand simple syntax combination operations but also to have knowledge of the field to be translated. Therefore, a new approach was seemingly required.[6]

Humans possess a variety of knowledge; by applying that knowledge, they can solve accurately and efficiently problems. For example, if you are a professional Shogi player, by applying knowledge of set moves, among others, you can narrow down prospective moves. In AI, the technology to use this kind of knowledge, known as "knowledge processing,"[7] attracted attention in the 1970s and research progressed.[8] This technology is widely used in infectious disease diagnosis systems (MYCIN)[9], and in

---

1 At the Dartmouth conference held in the US in 1956, the academic field of artificial intelligence was established. "History of Artificial Intelligence" The Japan Society for Artificial Intelligence Website <http://www.ai-gakkai.or.jp/whatsai/AIhistory.html> (in Japanese)

* The date of last access for internet information in this paper is February 26th, 2018

2 AI is defined as "cognitive machines, especially the science and technology to create cognitive computer programs." However, this definition varies between researchers, and because "cognition" and "intelligence" are not defined, defining artificial intelligence is still said to be difficult. "Artificial Intelligence FAQ" The Japan Society for Artificial Intelligence website <http://www.ai-gakkai.or.jp/whatsai/AIfaq.html> (in Japanese); Ministry of Internal Affairs and Communications "White Paper on Information Communications Heisei 28 Edition," 2016, pp. 233-234. <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/pdf/n4200000.pdf> (in Japanese)

3 "Assuming the source of intelligent expression to be in searching, and focused on the issue of efficient searching" The Japanese Society for Artificial Intelligence "Encyclopaedia of Artificial Intelligence," Kyoritsu Publishing 2017, pp. 5-6, Ishizuka Mitsuru et al. '"Foundations of Artificial Intelligence," Introduction' (in Japanese)

4 As above, pp. 12-14.（Ishizuka et al. 'Search'）

5 This problem was called "combination explosion," which was a major problem in early AI research. In the UK in 1973, at the symposium of the Science Research Council (SRC), which was discussing government subsidies for scientific research, the lack of understanding of this issue was pointed out, and it was said that the British government had significantly reduced its AI research budget as a result. Stuart Russell and Peter Norvig "*Artificial Intelligence, A Modern Approach*" Second Edition, Prentice Hall, 2003, p. 22.

6 As above, p. 21. In America, machine translation of cutting-edge science and technology papers between English and Russian was vigorously researched; however, the Automatic Language Processing Advisory Committee (ALPAC) of National Academy of Sciences reported in 1966 that "there is no prospect of immediate realisation," and hence, a new direction was being sought.（ALPAC, "Language and Machines: Computers in Translation and Linguistics," 1966, p. 24. <http://www.mt-archive.info/ALPAC-1966.pdf>）

7 "Knowledge processing" is classified into fields of "knowledge expression," which discusses how to express knowledge, and "reasoning," which discusses how to use the expressed knowledge. Refer to "*Artificial Intelligence, A Modern Approach*" Second Edition" part 3, "Knowledge and Inference"

8 The Japanese Society for Artificial Intelligence supra note (3), pp. 6-8, Ishizuka Mitsura et al. "*Foundations of Artificial Intelligence,*" Introduction (in Japanese)

9 Bruce G. Buchanan and Edward H. Shortliffe, *Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Massachusetts: Addison-Wesley, 1984. <http://people.dbmi.columbia.edu/~ehs7001/Buchanan-Shortliffe-1984/MYCIN Book.htm>

systems for estimating the composition of organic compounds (DENDRAL)[10], and was a fundamental technology supporting the AI boom centering on expert systems[11] in the 1980s. In Japan too, expert systems were enthusiastically adopted in the steel industry, etc.[12] However, at the time, for the AI to use knowledge, a huge amount of manual effort was required[13], such as in investigating the knowledge in detail through interviews, and hence, knowledge processing technology was mainly limited to use in large-scale industrial fields.

From these circumstances, the technology required for AI to acquire knowledge by itself, namely "learning" technology, has come to be considered essential for AI.[14] The technology by which a machine automatically learns knowledge, called "machine learning" is a key technology supporting the current AI boom.

As described above, search, knowledge processing, and machine learning are fundamental technologies essential to AI today. Knowledge processing and machine learning, which have been developed remarkably in recent years, are explained below.

2. The State of Knowledge Processing

In the knowledge-processing technology of the 1980s, the methodology of knowledge-building was not yet established, and there were serious issues such as the reusability of knowledge,[15] but through later research, the current situation has significantly changed. Through the application of the concept of organization (ontology)[16] to describe various kinds of knowledge, the theory of knowledge construction was improved and the description and reuse of knowledge has become simple. Furthermore, through the spread of the internet and the popularity of open government and open data[17], even without creating knowledge independently, it has become possible to easily acquire knowledge from the internet.

Currently, a format called "Linked Data"[18] for enabling a computer to understand described knowledge has become popular, and various data are published in this format[19]. Using linked data, such as in association games, one can follow linked items and gather a large amount of related knowledge. For example, NHK trialed providing broadcast information as linked data.[20] In this trial,

---

[10] Robert K. Lindsay et al., "DENDRAL: A case study of the first expert system for scientific hypothesis form ation," *Artificial intelligence*, Vol. 61 No. 2, 1993.6, pp. 209-261. <https://deepblue.lib.umich.edu/bitstream/handle/2027.42/30758/0000409.pdf>

[11] A system making similar judgments as experts by using knowledge

[12] Tsuchiya Shun et al. eds. "*AI Encyclopaedia Second Edition*", Kyoritsu Publishing, 2003, p. 11 (in Japanese)

[13] This problem is called the knowledge acquisition bottleneck. *The Japan Society for Artificial Intelligence* supra note (3) (Hiroshi Motoda et al. "*Machine Learning and Data Mining*" Introduction (in Japanese)

[14] As above

[15] Knowledge used in expert systems should be constructed at each facility of its use, and owing to inefficiency, it was necessary for each facility to share knowledge and reuse it. As above, p. 1256. Riichiro Mizoguchi "*Knowledge Engineering and Expert Systems*" (in Japanese)

[16] As above, pp. 1259-1260. Yoshinobu Kitamura "*Ontology and Schema*" (in Japanese)

[17] Open Government is an effort to open up the government to the public using the internet. In order to increase the transparency of administrative agencies, in Open Government, various data held by administrative agents are made public (Open Data). "What is Open Government?" OpenGovLabWebsite <http://openlabs.go.jp/whatis/>

[18] The Japanese Society for Artificial Intelligence *supra note* (3), pp. 1316-1318. Takeda Hideaki, "*Linked Open Data （LOD） Fundamentals*" (in Japanese)

[19] On the following website, one can check the status of published data. Andrejs Abele and John McCrae, "The Linking Open Data cloud diagram," *Insight Centre for Data Analytics website* <http://lod-cloud.net/>

[20] "NHK Linked Data API Service Information," NHK Broadcasting Technology Laboratory Website

it was possible to connect the region where the user is located to video related to that region, and it was possible to widely use the related information. In addition, Fujitsu Laboratories created a system that collects published link data and facilities, searching the data in aggregate.[21]

By collected linked data, it is possible to create knowledge databases. This kind of knowledge database is called a "knowledge graph."[22] A knowledge graph is a large volume of collected information showing links between, e.g., people, places, and things, and is essential to create question and answer systems for answering various questions and online search systems. For example, in IBM's case, a knowledge graph created from a free online encyclopedia (Wikipedia) is used in the company's AI system, "Watson;" this is a core technology in the system that allows it to answer questions in lieu of people.[23]

IT companies such as Google, Facebook, and Microsoft are creating their own knowledge graphs.[24] To further increase their scale, technology is necessary to automatically create accurate knowledge graphs, and is an important research theme among AI researchers.

3. The State of Machine Learning

Machine learning is one of the fields of research for realizing an AI with learning capabilities.[25] Humans, possessing the ability to learn, can acquire new knowledge. For example, a human child can, upon seeing a cat, being taught by their parents that "there's a cat," and learning that cats have characteristics such as ears, obtain the knowledge required for recognizing a cat upon seeing one the next time. AI is also expected to have the same capability through machine learning. Various machine learning frameworks have been considered; however, among them, "learning from example" is often used.[26] In other words, by providing examples to a machine, common characteristics are extracted from the data as knowledge. In recent years, with the development of the internet and improvement of large-scale databases, in addition to creating an environment in which enormous volumes of data can be easily used, developments in machine learning technology are also drawing attention.

One of these machine learning technologies is deep learning.[27] A learning technology called "Neural Networks," imitating the network of the countless number of neurons comprising the human brain, having been studied for many years, has been developed.[28] In deep learning, data features are learned by expressing the structure of layered hierarchies as a mathematical model, processing that

---

<https://www.nhk.or.jp/strl/ld/> (in Japanese) The service was offered for a limited time for research and investigation purposes

[21] "LOD4ALL." <https://lod4all.net/>

[22] "The following details the Knowledge Graphs announced by Google in 2012." *The Japanese Society for Artificial Intelligence* supra note(3), p. 1314. Takahiro Kawamura "*Metadata Usage*" (in Japanese)

[23] Jose Manuel Gomez-Perez et al., "Enterprise Knowledge Graph: An Introduction," Jeff Z. Pan et al., eds., *Exploiting Linked Data and Knowledge Graphs in Large Organisations*. Cham: Springer, 2017, p. 9.

[24] Heiko Paulheim, "Knowledge Graph Refinement: A Survey of Approaches and Evaluation Methods," *Semantic Web*, Vol.8 No.3, 2017, pp.489-508. <http://www.semantic-web-journal.net/system/files/swj1167.pdf>

[25] The Japanese Society of Artificial Intelligence *supra note* (3), p. 280. Hiroshi Motoda et al., "*Machine Learning and Data Mining*," Introduction (in Japanese)

[26] As above, pp. 281-283. Hiroshi Motoda et al., "*Machine Learning and Data Mining*," Introduction (in Japanese)

[27] As above, pp. 532-534. Hideki Aso, "*Deep Learning*" (in Japanese)

[28] "A method for learning by representing the mechanism by which signals are transmitted between neurones through a combination of simple mathematical expressions, and giving examples to this formula." See also Section 20.5 of Russell and Norvig *supra note (*5)

data as input in a certain layer, and repeatedly processing the output in other layers. In neural networks, up to now, learning could not be performed well when there were many layers (deep)[29]; however, in recent years, technology that can learn even when there are deep hierarchies has been developed[30] and has come to be used as deep learning.

In deep learning, by deepening the hierarchy, it becomes possible to hierarchically extract characteristics present in the data, and consequently, the accuracy of learning is improved.[31] For example, in conventional machine learning, for recognizing an image of a cat, humans must first provide characteristics to the machine—whether it has ears and hair, its color, among others. However, in deep learning, image data is provided to the machine. By extracting characteristics such as ears and hair at lower layers, and synthesizing these at higher layers, the machine can hierarchically extract these characteristics to recognize a cat.

The accuracy of deep learning can be improved if a large amount of data can be made available, and accuracy has improved particularly in the fields of image and voice recognition. Studies on deep learning are being advanced primarily by IT companies holding large amounts of data[32], particularly Google, Facebook, Microsoft etc., and its implementations in Japan are being advanced by companies such as Preferred Networks.

"Reinforcement Learning,"[33] which is another method of machine learning, has also gained traction in recent years. For example, if an animal receives a reward, such as being fed after performing a series of actions, the preference for that behavior is reinforced. Reinforcement learning is a learning method that emulates this mechanism. In recent years, by combining reinforcement learning with deep learning, more advanced learning has become possible. For example, Deep Mind, a subsidiary of Google, has developed an AI that is a stronger Go player than even the top human Go player by combining deep learning, reinforcement learning, and search technologies.[34]

4. Future Issues

At present, advanced AI technology cannot be realized using knowledge processing or machine learning alone, but practical AI technology can be realized by using the two complementarily.[35]

As knowledge processing uses human knowledge, it has the advantage in that humans can easily understand its mechanisms. When using AI technology for decision-making involving responsibility, this kind of technology is necessary because the decision-maker must understand the reason of the

---

29 "Called the vanishing gradient problem." The Japanese Society for Artificial Intelligence, supra note (3), pp. 521-522., Shinichi Asakawa "Recurrent Neural Networks"
30 As above, pp. 532-534. Hideki Aso "Deep Learning" (in Japanese)
31 In 2012, in a general object recognition contest to recognize what is depicted in an image, the significant improvements in accuracy obtained when deep learning was used gathered immense attention, and formed the basis for the deep learning boom today.
32 Companies have been enthusiastic to incorporate outcomes of researches held at universities so that Geoffrey Hinton, emeritus professor at the University of Toronto, who led the development of deep learning, and professor Yann LeCun of New York University have worked at Google and Facebook, respectively.
33 Russell and Norvig *supra note*(5), Chapter 21
34 David Silver et al., "Mastering the game of Go without human knowledge," *Nature*, Vol. 550, 2017.10.19, pp. 354-359.
35 Simple tasks such as recognising an object in an image can be performed by a single AI technology alone; however, when performing complex tasks, such as those performed by IBM's Watson or Deepmind's Go AI, in many cases, multiple AI technologies are combined.

decision. In addition, in knowledge processing, it is possible to explicitly reflect the intent of a human in the AI, and design in advance the handling of special cases not present in the data. For example, in rare events such as disasters, data cannot be obtained in advance; however, by incorporating human knowledge, it is possible to implement responses to such events in an AI. However, as it can be difficult to create and acquire knowledge depending on the field, further technological development is required.[36]

Machine learning has an advantage in that it can automatically construct knowledge where there is a large volume of data. Therefore, where an abundance of data is obtained, it is possible for it to make decisions with a degree of accuracy surpassing human judgment. However, the use of technologies such as deep learning causes difficulties in explaining the reasoning behind the decision-making.[37] Moreover, it is difficult to apply machine learning appropriately to fields where there is scant data or in fields where data is not comprehensively collected, and further technological development is necessary to respond to these issues.[38]

Although some issues remain in current AI technology, its underlying technologies including knowledge processing and machine learning continue to mature, and it is becoming possible to implement advanced AI through the complementary use of these underlying technologies. In the future, it is also necessary to consider architectures to integrate these fundamental technologies.

Ryutaro Ichise, National Institute of Informatics

II Natural Language Processing

Languages used by humans on a daily basis, in contrast to artificial languages such as programming languages, are called “natural languages.” This chapter describes the social background surrounding natural language processing technology as of February 2018, its state both domestically and overseas, promising areas of application, and future issues.

1. Notable Social Background

With the spread of web searches in the latter half of the 1990s, the utility of applications to process and search large-scale data using natural language gained wide recognition. Web searches are systems

---

[36] “Research and development of next-generation artificial intelligence technology capable of mutual understanding with humans,” carried out by the National Institute of Advanced Industrial Science and Technology’s Artificial Intelligence Research Centre (ARIC), wherein research and development of AI that can obtain knowledge understandable to humans from large amounts of inexplicit data is progressing. “NEDO-commissioned project, “Next Generation Artificial Intelligence/ Core Robot Technology Development / Research and Development of Next-Generation Artificial Intelligence Technology Fields / Research and Development of Next-Generation Artificial Intelligence Technology Capable of Mutual Understanding with Humans” National Institute of Advanced Industrial Science and Technology Artificial Intelligence Research Centre website <http://www.airc.aist.go.jp/nedoproject/index.html> (in Japanese)

[37] “There are sceptical views on the application of deep learning in areas requiring explanations based on scientific evidence such as healthcare.” *The Japanese Society for Artificial Intelligence* supra note (3), pp. 1405-1406. Hiroshi Nakajima *“Applications of AI to Health Care Equipment”* (in Japanese)

[38] The RIKEN Innovation Centre’s Advanced Intelligence Project (AIP) is conducting studies on the fundamental research and technology of new algorithms able to learn accurately from small amounts of data. *Artificial Intelligence Technology Strategy Conference* “Artificial Intelligence Technology Strategy,” 2017.3.31, p. 2 *New Energy and Industrial Technology Development Organisation* website <http://www.nedo.go.jp/content/100862413.pdf> (in Japanese)

in which words are input and corresponding webpages are returned. This followed as an extension of full-text searches by searching for matching words (pattern matching), and it was believed that a deep understanding of the input words was unnecessary. However, in 2011, IBM's AI system (Watson) demonstrated performance surpassing that of a human quiz champion in a question answering task,[39] returning knowledge relevant to an input question and becoming a topic of conversation. It came to be recognized that, regardless of the field,[40] it had the ability to understand natural language comparable to humans.

Since 2012, technological innovation centering on deep learning has also been applied to natural language processing, thus changing research in the field dramatically between 2014 and 2015. Subsequently, research and development using deep learning has flourished. Deep learning-based natural language generation outputs fluent text to such an extent that it can scarcely be distinguished from human-written text. Deep learning has brought about a technological breakthrough,[41]achieving drastic improvements in language generation tasks such as machine translation, dialog, and document summarization.

2. Technology Trends

State-of-the-art research and development of natural language processing is being performed in the United States. Companies, universities, and institutions conducting research and development of natural language processing are particularly concentrated in the Bay Area on the West Coast and New York on the East Coast. Characteristic of research and development in America, IT companies typified by Google and Facebook are taking the initiative in the development of natural language processing. Research on information extraction and machine translation began with ACE (Automatic Content Extraction) led by the National Institute of Standards and Technology (NIST) under the Ministry of Commerce, and subsequently, DARPA (Defense Advanced Research Projects Agency) invested a significant amount of national defense expenditure through the GALE (Global Autonomous Language Exploitation), TIDES (Translingual Information Detection, Extraction and Summarization), and BOLT (Broad Operational Language Translation) projects.[42] In universities, military technologies are also being put into use for public benefit. For example, two researchers who founded the company Language Weaver[43] with machine translation technology developed in the above DARPA projects at its core were faculty members at the University of Southern California.

China is carrying out state-led research. In 2016, AI-related technological development was

---

39 "Computer Wins on 'Jeopardy!': Trivial, It's Not" The New York Times < https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html >

40 Supervised by Mamoru Komachi authors Yoh Okuno et al., "*Basics and Techniques of Natural Language Processing,*" Shoeisha, 2016, pp. 14-25 (in Japanese)

41 Yuta Tsuboi et al., "*Natural Language Processing Using Deep Learning*" (Machine Learning Professional Series) Kodansha, 2017, pp. 122-158 (in Japanese)

42 "Collaborations> Past Projects." Linguistic Data Consortium Website <https://www.ldc.upenn.edu/collaborations/past-projects>; "BOLT (Broad Operational Language Translation)" Website <https://www.darpa.mil/program/broad-operational-language-translation>; Japan Science and Technology Agency R&D Strategy Centre, Systems and Information Science and Technology Unit, "Research and development outlook report, Systems and Information Science and Technology Field (2017)" Japan Science and Technology Agency R&D Strategy Centre, 2017, p. 231 <https://www.jst.go.jp/crds/pdf/2016/FR/CRDS-FY2016-FR-04.pdf> (in Japanese)

43 Acquired in 2010 by SDL, which has the top market share for translation globally

headed by Tsinghua University, the Harbin Institute of Technology, and other research institutions such as the Chinese Academy of Sciences, and it is reported that 100 billion Chinese yuan (approximately 1.7 trillion Japanese yen)[44] will be invested over the next three years.[45] Moreover, research and development of natural language processing is also flourishing through IT companies. Both Tencent (腾讯) and Baidu (百度) have research and development sites in the US and have achieved numerous world-leading research results. There is a growing sense of an international community in natural language processing.

Conversely, in Japan, natural language processing has mainly been developed by companies; however, the environment has been changing drastically in the past decade. Until the 1990s, mainstream development originated from manufacturing companies, and research and development of natural language processing came to a brief halt with the collapse of the economic bubble. However, in the early 2000s, with the spread of the internet, the demand for web-related companies increased. Since the advent of deep learning, AI development startups from universities such as Preferred Networks and PKSHA Technology have become active, and these startups have drawn an influx of talent from universities and large corporations. Moreover, the government is carrying out AI research through the National Institute of Advanced Industrial Science and Technology (Artificial Intelligence Research Centre), the National Institute of Information and Communications Technology (Universal Communications Research Centre), and RIKEN (Innovative Intelligence Integration Research Centre), and is reported to be investing approximately 100 billion yen over ten years starting from 2016 (Heisei 28).[46]

3. Real-world Applications

Natural language processing technology has already been used in various situations such as Japanese language input, web searches, and spam filters, and is one of the technologies supporting the information society. The language generation field is predicted to achieve giant leaps in development within the next ten years, specifically in applications that output text, such as machine translation, dialog, and document summarization. In these applications, as fluent language output seems to have been attained through the arrival of deep learning, it appears that some of the roles that have so far been performed by people will be automated.

On the one hand, machine translation using deep learning can output fluently, but on the other hand, it does not always correctly reflect the intent of the original sentence, and close proofing by a human is essential. For example, translations involving deep understanding of the contents, such as literary translation, cannot be substituted by machine translation. However, in industrial translations, quality of that extent is not required for technical literature and medical domain, which constitute a large part of the translation market, and widespread adoption of machine translation is expected. In

---

[44] 1 Chinese yuan was converted as 17 Japanese yen (The official rate reported in December 2017)

[45] "China unveils three-year program for artificial intelligence growth," *China Daily*, 2016.5.24. <http://www.chinadaily.com.cn/business/tech/2016-05/24/content_25442308.htm>

[46] "Strategic Organisation through Artificial Intelligence 3 Agencies' Collaboration, Accelerating Business and Research" Nihon Keizai Shinbun, 2016.2.27, evening edition, p. 1 (in Japanese)

the future, the necessity for humans to take charge of the entire translation will hold for only some fields.[47]

Another application currently gathering significant attention is dialog.[48] In the fields of FinTech (financial services using IT), LegalTech (legal services using IT), e-commerce etc., humans with specialist knowledge have performed customer support tasks so far. In these fields, it is believed that systems that automatically generate responses to customers will become popular as part of support work. Such systems have already become widespread first in the US and Western Europe, but in Japan, an emphasis on the quality of support has meant that adoption has not progressed. In the future, however, even in Japan, the proportion of generations with no resistance to the use of information communications devices will increase, and these generations will become able to not only understand the difference in quality from human responses, but also master the use of these systems, and substitution for these systems is expected to advance.

Moreover, a promising area for the application of natural language processing is the support provided to people with disabilities. So far, voice recognition and speech synthesis technologies have been studied, and have even been loaded on Macs and iPhones to support those with hearing and speech impairments. However, since 2015, the fusion of language and image areas has been studied enthusiastically and methods for attaching captions to images and videos (text description) have been proposed. By combining caption generation for images and video with voice synthesis, visually impaired people can understand image contents. The combination of image, voice, and language processing technologies contributes to the realization of a world in which anyone can access information.

4. Future Issues

As research on language generation becomes more popular, issues of copyright and privacy with regard to generated sentences are likely to manifest, and how to respond to these is a further issue.[49] Ethical issues—for example, the handling of personal information in medical information processing, issues related to being able to delete past information from the internet (the right to be forgotten), and the issue of bias toward minorities on the basis of gender, race, etc.—are being discussed.[50] In Japan, The Japan Society for Artificial Intelligence launched an ethics committee in 2014 (Heisei 26).[51]

From the technology perspective, prior to deep learning, language creation rules and reference dictionaries were used to produce descriptions in a human-readable form, and hence, it was possible, to a certain extent, to prevent inappropriate language generation; however, when deep learning is used, control over the output is technically difficult. Until this is resolved in future research, it should

---

[47] Yuta Tsuboi et al., *supra note* (41) (in Japanese)
[48] Supervised by Mamoru Komachi, authors Yoh Okuno et al., supra note (40), pp. 212-236. (in Japanese)
[49] Koji Okumura "Content Created by Artificial Intelligence and Copyright - Focusing on Copyright Works" *Patents*., Vol. 70, No. 2, 2017, pp. 10-19, <https://system.jpaa.or.jp/patent/viewPdf/2742> (in Japanese)
[50] In machine learning, if there are many ethically problematic expressions concerning minorities included among the data used for learning, there will be problems of the AI using those expressions from which it learned.
[51] Yutaka Matsuo et al., "Efforts of the Artificial Intelligence Society Ethics Committee" "Artificial Intelligence," Vol. 30 No. 3, 2015.5, pp. 358-364 (in Japanese)

be used with a careful understanding of the utility and risks of high-accuracy language generation.

Mamoru Komachi, Tokyo Metropolitan University

Ⅲ Image Acquisition and Recognition

1. Overview of Image Acquisition and Recognition

The progress of AI technology in recent years has brought significant reforms centering on industry, and has also had profound effects on the field of image recognition and processing. As technical background, in addition to high-accuracy cameras and the miniaturization and cost reduction of various cameras, machine learning and recognition technologies such as deep learning have been developed. As social background, through the spread of smartphones, cloud computing,[52] and social media, images can be shared simply, and consequently, the targets and needs of image recognition have expanded. In other words, everyone can easily obtain images, and those images can be shared on social media, and image recognition needs—for example, recognizing people or objects and identification of people—can be said to have increased. This chapter describes the progress of image sensors playing the role of "eyes" in obtaining images, and image recognition and processing playing the role of the "brain" in interpreting these images.

2. Image Sensor Development

An image sensor is an element that converts light in the physical world to digital formats, and Japan has contributed significantly to its development. Image sensors developed by Japanese manufacturers remain competitive even today. Although image sensors are parts that correspond to eyes in people, by capturing information that humans cannot see, they can be applied to various purposes. While they have been used so far in specialist manufacturing and military applications, through cost reductions associated with mass production, high-capability image sensors are also being introduced in devices for general consumers.

(1) Color Image Sensors

Color image sensors capture red, green, and blue light in the same manner as human eyes, and over the past 10 years, their resolution has improved by 5–7×; even the resolution of the cameras used in smartphones can exceed 10 megapixels.[53] Further improvements in resolution are limited by storage capacity issues and lens performance (definition). As an alternative, development of high-dynamic-range compatible elements that can capture lighter and darker regions is advancing.

---

52 Refer to Section 1 (2) Improvement of Infrastructure Environment of Chapter VII, IoT

53 For example, the number of pixels in the rear-mounted sensor on the first-generation iPhone (released June 2007) was 2,000,000 (2 megapixels), but in the latest model (released November 2017), that number is 12,000,000 (12 megapixels). "Guides and Sample Code: iOS Device Compatibility Reference: Cameras." Apple Developer Website <https://developer.apple.com/library/content/documentation/DeviceInformation/Reference/iOSDeviceCompatibility/Cameras/Cameras.html> (in Japanese)

(2) Multispectral Image Sensor (infrared sensor)

If an infrared sensor, which captures infrared light, is used in lieu of a color sensor, the reflection of light is suppressed, and facial recognition performance is improved. It also has the advantage that authentication can be performed using eye luster, and the popularity of this sensor is growing.[54] These sensors have so far been limited to use in night-vision equipment, but are now being incorporated into personal computers (PCs) and smartphones with the spread of facial authentication.

(3) Distance Image Sensors

The range image sensor is a device for measuring depth information around the camera, and it can be used to improve the performance of image recognition and automatic operation. In 2009, Microsoft released a piece of video game equipment equipped with a range image sensor called the "Kinect," which rapidly dropped in price and has been popular.[55] Moreover, in vehicles with driving support functions such as automatic braking, range image sensors are indispensable for determining the distance between the car and objects.

Range image sensors can be divided into several types according to their measurement method. There are those that use two image sensors, and those that use one image sensor combined with infrared radiation. They measure using the infrared light emission and its reflection, and each has its advantages and drawbacks depending on the measurement distance, application environment etc. (light/dark, weather). In the future, methods of combining multiple sensors will also be studied.[56]

(4) Light-field Cameras

Light-field cameras record not only light, dark, and color, but also information on the incident angle of light and include three-dimensional data. A micro lens array is installed on the front face of the image sensor (numerous microscopic lenses arranged together), recording not only the light/darkness of rays of light but also the incident angle information of the light ray.

Unlike conventional cameras, which record only two-dimensional information of brightness and color, the focal position (focus) of the image can be moved after capturing the image, and the depth of field (the in-focus region) can be changed. Thus, these cameras open up new possibilities. While there are drawbacks such as insufficient resolution and requirement of a large amount of light, practical applications such as the American company Lytro's launch to the public are being attempted.[57]

(5) Omnidirectional Cameras

In recent years, omnidirectional cameras are gathering attention owing to the spread of virtual reality (VR) technology; they can capture all-round images across 360° degrees. Products aimed at

---

[54] "Iris / Facial Authentication Optical Sensor Market, 24% CAGR growth 2017-2025" 2017.3.13. LEDs Magazine Japan website <http://ex-press.jp/ledj/ledj-news/ledj-biz-market/17594/> (in Japanese)

[55] "As understood from patent analysis, dominating by volume the strongest company in distance imagery, Microsoft has a decisive lead," Nikkei Electronics No. 1135, 2014.5.26, pp. 49-57. (in Japanese)

[56] "What combination of components is best for automatic operation? DENSO Technical Development Briefing Session," 2016.8.10, Response website <https://response.jp/article/2016/08/10/279956.html> (in Japanese)

[57] Lytro website <https://support.lytro.com/hc >

consumers are being sold both in Japan and overseas.[58] The omnidirectional camera was developed by applying earlier technologies combining multiple components and utilizing fish-eye lenses that can capture wide angles; however, as it is possible to capture images using these cameras in ways different from conventional cameras and as they offer a high degree of affinity to VR content,[59] their demand is increasing worldwide.[60]

(6) Wearable Cameras

As cameras become smaller and less expensive, wearable cameras worn on one's person have been developed with the aim of recording everyday activity. Among wearable cameras, first-person sensors that record the same images as the wearer's gaze are expected to be used in content production in sports, policing, data collection, and identification of criminals.[61] Moreover, there are also cameras on the market that are attached to glasses, allowing the shutter to operate with a wink.[62]

3. Progress of Image Recognition Research

Image recognition research has undergone rapid development in the past 10 years, especially in the past five years with the progress of research on deep learning. Research and development trends in the various fields of application of image recognition are explained as follows.

(1) General Object Recognition

General object recognition is a technique for sensing and recognizing where people and objects are present in a photographed image. Circa 2007, by understanding local features in images, and adopting a technique called "bag of features"[63] to differentiate images based on their composition, its performance improved. Moreover, thanks to the accumulation of sample image data necessary for learning in this method,[64] competitive development has progressed worldwide.

Furthermore, through the adoption of deep learning "convolution neural networks"[65] in image recognition from 2012,[66] the accuracy of image recognition has improved drastically. This method

---

58 RICHO THETA website <https://theta360.com/en/>; "SP360 Action Camera." Kodak Digital Cameras Website < https://kodakpixpro.com/Americas/cameras/actioncam/sp360/ >

59 For example, Toppan Printing produces VR content such as Azuchi Castle and Yomeimon at Nikko Toshogu Shrine. "Main Toppan VR Products" Toppan VR Digital Archive website <http://www.toppan-vr.jp/bunka/action.shtml> (in Japanese); "Next Generation Technical Features, AR developments, Excitement in Five Senses, Relief Printing, Tourist Spot Experiences," Nikkei Business Daily 2017.11.27, p. 10 (in Japanese)

60 "360° Cameras, Hot Competition" Nikkei Business Daily 2016.10.31, p. 7. (in Japanese)

61 "Panasonic develops wearable cameras for police around the world to combat terror" Nihon Keizai Shinbun, 2016.1.5, p. 12. (in Japanese)

62 E.g., Blincam Website <https://www.blincam.co/>

63 Gabriella Csurka et al., "Visual categorization with bags of keypoints," *Proceeding of Workshop on Statistical Learning in Computer Vision*, *8th European Conference on Computer Vision (ECCV)*, 2004. <http://www.europe.naverlabs.com/content/download/23625/171397/version/1/file/2004_010.pdf>

64 The PASCAL Visual Object Classes Website <http://host.robots.ox.ac.uk/pascal/VOC/> Implemented with funds received from the EU. "Pattern Analysis, Statistical Modelling and Computational Learning 2," European Commission CORDIS Website <http://cordis.europa.eu/project/rcn/85729_en.html>

65 A method for processing by integrating (convolution) data obtained by dividing an image so that image recognition can be performed regardless of the position of the identified portion.

66 Alex Krizhevsky et al., "ImageNet Classification with Deep Convolutional Neural Networks," *Proceedings of the 25th International Conference on Neural Information Processing Systems*, Vol. 1, 2012, pp. 1097-1105. <https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>

is also applied to object recognition—for example, the identification of plant and animal species—and the race for development continues.

（2）Facial Recognition

There is immense variety among faces; however, they have partially common characteristics. Focusing on these common characteristics, a method of high-speed facial identification was conceived in 2000,[67] and facial recognition capabilities began to be incorporated into digital cameras.

Deep learning has also applied to this field in recent years, and it has become possible to recognize not only faces, but also sex, age, facial expression, and various other attributes. Furthermore, facial recognition technology, which can identify an individual from their face, has also been developed, and has begun to be applied in immigration control and the identification of suspicious individuals.[68] Technology to identify individuals from their faces is expected to be applied to a variety of purposes; however, this raises privacy concerns described later.

（3）Medical Applications

In the medical field, doctors identify lesions from X-ray, CT, and MRI images, and research on the application of image recognition to this field is advancing.[69] For example, it is supposed that candidate regions with abnormalities can be detected using deep learning technology, and doctors can make a final decision on them.

4. Future Issues

Image recognition has been developed rapidly in recent years; however, it has simultaneously given rise to social and ethical problems, which may affect its widespread application.

（1）Use of Data for Deep learning

Deep learning has led to significant developments in image recognition, but has also brought about technological limitations and social issues. In image recognition using deep learning, performance can be improved by learning from larger image sample sets, but there are limits to the number of image samples that can be used. Some internet companies collect personal image data from the internet and are using them in machine learning cases. However, legal concerns such as copyright and privacy, and ethical concerns regarding the use of data created or published by individuals are being raised for machine learning.[70]

---

[67] The detection speed was increased in evaluating parts of an image that are not a face by using a simple calculation to quickly determine whether or not it was a face. Paul Viola and Michael J. Jones, “Robust Real-Time Face Detection,” *International Journal of Computer Vision*, Vol. 57 No. 2, 2004.5, pp. 137-154.

[68] “High Accuracy Facial Authentication Software by Panasonic using Deep Learning” Nikkei Business Daily 2018.2.22, p. 4 (in Japanese); “Facial Authentication, Evaluation in 10 seconds, Ministry of Justice, and launched at Haneda Airport, Improving Immigration Efficiency for Japanese” Nikkei Business Daily, 2017.10.14, p. 38 (in Japanese); “In anticipation of 2020 the Tokyo Marathon takes measures against the heat” Sankei Shimbun 2018.2.26, p. 30 (in Japanese)

[69] “The current status and outlook for the use of AI in diagnostic imaging” 2017.6.29. MedTec Japan website <http://www.medtecjapan.com/ja/news/2017/06/29/1936> (in Japanese)

[70] Asakawa Naoki, “Can you use data gathered online for AI learning?” 2017.6.22. Nikkei CrossTech website <http://tech.nikkeibp.co.jp/it/atcl/column/17/051800199/062000008/> (in Japanese)

(2) Use of Image Recognition Cameras

As cameras are mounted on various devices and as the accuracy of image recognition is improved, privacy issues will also arise. As images from which individuals can be identified correspond to personal information, considerations such as the observance of the “Law concerning the protection of personal information” (Law number 57 of Heisei 15 (2003)) and giving advance notice of use are required.[71]

Specific examples of privacy and ethical issues are as follows. A small camera is incorporated into the eyewear-type mobile device built by Google, “Google Glass.”[72] From a privacy perspective, Google decided not to use facial recognition technology capable of identifying individuals for the time being;[73] nonetheless, owing to concerns around covert filming etc., sales to individuals were halted.[74] Moreover, issues of racial discrimination from the use of deep learning in image recognition have also arisen.[75]

(3) Issues Surrounding Application to Medical Fields

Image recognition in the medical field, in addition to assisting the identification of lesions as above, is also expected to include the use of wearable cameras to capture and analyze images during communication with patients and in medical treatment applications (for example, the psychiatric field).[76]

However, when applied to the medical field, it is assumed that concerns about errors, missed lesions, misdiagnoses, privacy etc., will arise, and with the widespread application of image recognition technologies in these fields, response to these issues is necessary.

Atsushi Nakazawa, Kyoto University

IV Speech Interfaces

1. Notable Social Background

A speech interface is a system in which a person can interact with a computer using their voice. In general, this includes both when speech is used as computer input (where a human voice is processed by the system) and when it is used as output from the system (where the voice output by the system is heard by a person); however, the focus in this chapter is the former—the handling of

---

71 IoT Development Consortium et al. “Camera Image Usage Guidebook” 2017.1, pp. 16-20. METI website <http://www.meti.go.jp/press/2016/01/20170131002/20170131002-1.pdf> (in Japanese)

72 Thad Starner, “Project Glass: An Extension of the Self,” *IEEE Pervasive Computing*, Vol. 12 No. 2, 2013.4, pp. 14-16.

73 “Google announced that the Google Glass facial recognition application will not be certified for the time being,” 2013.6.1. ITmedia NEWS website <http://www.itmedia.co.jp/news/articles/1306/01/news012.html> (in Japanese)

74 “Privacy halts Google’s Glass Device, ceases sales to individuals” Nikkei Keizai Shimbun 2015.1.17, p. 13 (in Japanese).

75 Jana Kasperkevic, “Google says sorry for racist auto-tag in photo app,” 2015.7.1. Guardian Website <https://www.theguardian.com/technology/2015/jul/01/google-sorry-racist-auto-tag-photo-app>; “Gorillas and faulty image recognition, Google apologises” Asahi Shimbun, 2015.7.3, p. 11 (in Japanese).

76 “Special Feature: Prescriptions for Communication,” Sogo Shinryo, Vol. 27 No. 5, 2017.5, pp. 568-623 (in Japanese).

speech recognition.

(1) Development of Speech Recognition Technology

Speech recognition is a process of considering voice signals in human speech as input, and outputting a corresponding character or word string. Speech recognition research has been undertaken for a long time, and in Japan, research has been ongoing as a field of pattern recognition[77] since the 1960s.[78]

Until the 1990s, this was limited to experiments inside laboratories, but with the widespread usage of computers, around 2000, IBM developed the speech recognition software "ViaVoice" targeting general users, and also sold a Japanese language version.[79] Being able to recognize human speech became a hot topic, and it even appeared in TV commercials, but did not achieve widespread use. The reasons for this are that, prior to its use, it was necessary for each user to record themselves speaking several hundred sentences,[80] called enrolment, in addition to limitations in recognition performance and the size of recognizable vocabulary. In any case, the technology at the time had not been developed sufficiently for practical use. In addition, at that time, desktop PCs, which were difficult to carry around, were popular, and the situations where these PCs could be used to perform speech recognition may also have been a limiting factor.

(2) Arrival of Applications for General Users

From the second half of the 2000s through to the 2010s, through improvements in communication speeds, the development paradigm changed dramatically[81] through the advent of cloud-based speech recognition based on cloud computing.[82] Until then, speech recognition processing was performed on the user's computer; however, in cloud-based speech recognition, speech recognition is performed on a server by sending speech data to that server. The advantages of operating in this way are profound. Users can effortlessly use speech recognition from existing voice devices (e.g., smartphones). Moreover, as the dictionary used for speech recognition is also stored in the server, it is possible for system developers to update it at any time. Furthermore, as speech data is sent to the server, data collection becomes extremely easy, and the collected data can be used for performance improvement through machine learning.

---

77 The recognition of which predefined classes input patterns of characters or sounds belong to (e.g., what alphabet character a set of line drawings represents)

78 Toshiyuki Sakai and Shuji Doshita, "An Automatic Recognition System of Speech Sounds," *Studia phonologica*, Vol. 2, 1962, pp. 83-95. <http://hdl.handle.net/2433/52632>

79 "ViaVoice millennium speech recognition software latest version, improvements to recognition accuracy and ease of use (in Japanese)" 1999.9.28. IBM website <http://www-06.ibm.com/jp/press/1999/09287.html>

80 "IBM Japan announces "ViaVoice 8" with reduced recognition errors (in Japanese)" 2000.9.21. ITmedia website <http://www.itmedia.co.jp/news/0009/21/viavoice.html>

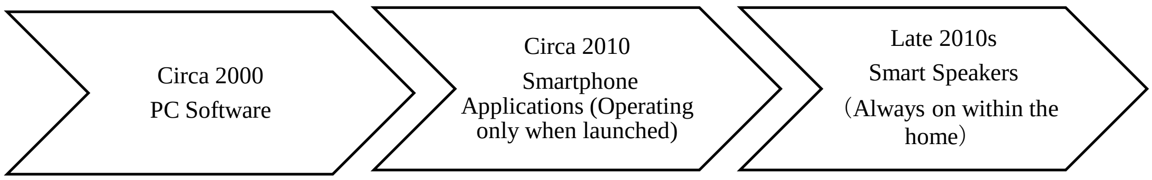
81 "Market Trends Edition: Leaps using the Cloud, Applications Expanding beyond Smartphones (in Japanese)," Nikkei Electronics, No. 1098, 2012.12.24, pp. 26-33.

82 Refer to 1 (2), "Improvement of Infrastructure Environment" of Chapter VII, IoT

From the paradigm shift above, speech recognition applications targeting general users that can be used on their smartphones appeared. Google's "Voice Search"[83] and Apple's "Siri"[84] are typical examples. These applications perform more than just speech recognition; they can perform web searches or respond to questions using the results of speech recognition. In addition, these applications are both free to use, and the aim of the application development is not sales, but an up-front investment based on a long-term vision—display of the companies' technical prowess, collection of customer data including voice, and introduction to the company's other services.[85] Prior to the voice search, Google had been pursuing data collection through the "GOOG-411" telephone directory assistance service.[86]

In recent years (from the latter half of the 2010s), household electronics appliances called "smart speakers" have been introduced, making it possible to listen to music and news and control home appliances through voice operations by connecting to the internet. In Japan, sales of Google's "Google Home"[87] and Line's "Clova WAVE"[88] began in October 2017, and of Amazon's "Amazon Echo"[89] in November 2017.

Figure 1. Changes in the state of application of speech interfaces



(Source) Created by the author(s)

2. Technology Trends

In 1990, the technology used in speech recognition was primarily machine learning based on statistical methods; however, in recent years, the technology has proceeded to that using deep learning, which has dominated the field.[90] In the speech recognition field, there is the clear criterion for the speech recognition accuracy (the ratio of correct words among all recognition results), and shared speech databases are also being constructed, whereby if a superior method is found, it tends to spread at once, both domestically and overseas. In other words, speech recognition performance improves as the amount of training data increases and these data (real data) are obtained from an

---

83 Published in the 2009 Japanese Language Edition. "Japanese-language version of Google Voice-Search launched, accessible from iPhone and Android (in Japanese)" 2009.12.7, InternetWatch website <https://internet.watch.impress.co.jp/docs/news/333996.html>

84 Published in the 2012 Japanese-language edition. "Apple releases the iOS 5.1 update, iPhone 4S supports Siri in Japanese (in Japanese)," 2012.3.8. ITpro website <http://itpro.nikkeibp.co.jp/pc/article/news/20120308/1043362/>

85 Kobayashi Masakazu "From the Cloud to AI, the next battlefield for Apple, Google and Facebook (in Japanese)," Asahi Shimbun, Asahi Shimbun Publishing, 2013, pp. 115-118

86 "Google launches 411-Service," 2007.4.7 TechCrunch Japan website < https://techcrunch.com/2007/04/06/google-launches-free-411-business/> This service was closed in November 2010.

87 "Google Home" Google website <https://store.google.com/product/google_home>

88 "Clova WAVE" (in Japanese) LINE website <https://clova.line.me/>

89 "Amazon launches "Echo" AI speaker at ¥11,980 (in Japanese)" Nikkei Sangyo Shinbun, 2017.11.9, p. 2

90 Tatsuya Kawahara, "Development of Speech Recognition Technology (in Japanese)," IEICE Tech.,PRMU2015-111, 2015.12, pp. 111-116. refer to <http://sap.ist.i.kyoto-u.ac.jp/members/kawahara/paper/KAW-prmu15-12.pdf>

environment close to practical use (namely, actual users). Therefore, performance improvement accelerates where the positive spiral “user usage -> data collection -> performance improvement -> user usage” begins.

In deep learning, a much larger amount of data and computational resources are required than in conventional methods, and hence, speech recognition research and development seems in practice to be limited to companies having large-scale data and computational resources (Google, Apple, Amazon, Microsoft, etc.). In Japan, speech recognition research has been carried out in universities and laboratories of enterprises (NTT, IBM Japan, Hitachi, Toshiba, NEC, etc.); however, research is believed to be shifting more toward areas of application rather than speech recognition accuracy itself. For example, in the speech interface field, which is one of the applications of speech recognition, NTT DoCoMo’s “Shabette Concier”[91] and Yahoo’s “Voice Assist”[92] applications, which allow one to set alarms and make phone calls simply by speaking to one’s smartphone have been released, as well as a multitude of startups providing similar services (Inago, etc.). [93]

In the US, Amazon launched the “Alexa Prize”[94] development contest in 2016, with companies sharing their real data and involving the academic communities such as universities of each country, which is a trend of advancing technological development in this manner. This is not a short-term trend with the aim of collecting on an investment, but rather is viewed as an up-front investment based on a long-term vision. Moreover, research on technically difficult issues makes high demands for real data, and as it is difficult for academia to collect large amount of real data, it can be said that these efforts are an ideal form of industry–academia collaboration. This kind of effort rooted in long-term vision is seldom observed in Japan.

Besides a speech interface, the use of speech recognition technology in business applications is still advancing considerably. In the Japanese House of Representatives, speech recognition technology developed by Kyoto University has already been introduced to create Diet Session Proceedings.[95] Moreover, it is being used to check whether call center recordings contain inappropriate language[96] and to transcribe a doctor’s speech when interpreting X-ray images[97] (obtaining diagnostic findings by examining images). The latter is a situation in which it is difficult to use a PC for input as it is done in a darkroom; furthermore, since speech recognition performance does not degraded because the room is quiet and it does not disturb people working together, it is a

---

91 “Shabette Concier (in Japanese)” NTT DoCoMo website <https://www.nttdocomo.co.jp/service/shabette_concier/>
92 “Voice Assist (in Japanese)” Yahoo! JAPAN website <https://v-assist.yahoo.co.jp/>
93 Inago website <http://www.inago.com/>
94 A contest competing to develop conversation applications that can be used on Amazon’s speech interface “Al exa.” For university, the total prize balance is 2.5 million dollars (approximately 280 million yen) (Converted at 1 dollar equal to 113 yen based on official rate reports for December 2017). “The Alexa Prize.” Amazon. com Website <https://developer.amazon.com/alexaprize>
95 “Introduction of Kyoto University’s speech recognition technology in the new proceedings recording system i n the House of Representatives (in Japanese),” 2011.5.12. Kyoto University website <http://www.kyoto-u.ac.jp/static/ja/news_data/h/h1/news6/2011/110512_2.htm>
96 “Sales begin of the “Contact Centre Business Efficiency Improvement Support Service (in Japanese),” which converts and utilises the contents of calls between customers and operators to text in real-time.” 2017.3.23., Hitachi Ltd., website <http://www.hitachi.co.jp/New/cnews/month/2017/03/0323a.html>
97 “AmiVoice Ex7 Rad and AmiVoice Ex7 RadClient for Radiology Readings Reports (in Japanese),” Advanced Media, website <http://medical.amivoice.com/product/rad.html>

good practice due to an excellent match between needs and application. Companies such as Nuance[98] in America, and Advanced Media[99] and FueTrek[100] in Japan are known for their development and sales of this technology.

3. Real-world Applications

（1）Car Navigation

Voice-operated in-car navigation is a field of application that has been focused upon in Japan. As a forward-gaze is necessary when driving, the need for voice operation is high.[101] Voice navigation began to be equipped in-car navigation systems around 1995,[102] and Japanese automobile manufacturers began research and development; however, the evaluations of the speech recognition capabilities of built-in type car navigation systems were unfavorable.[103] One of the reasons for this, with a (highly demanding)[104] background of an automobile industry that attaches importance to safety and a corporate culture where “we cannot release imperfect products,” is that the systems have been developed as the systems that did not use the Cloud (i.e., a system operating even where communications cannot be ensured), and hence, it has not been possible to update speech recognition dictionaries nor map data, and also to collect data from users. Conversely, car navigation applications that run on smartphones continue to widely spread.[105]

（2）Robots

In Japan, which has a high cultural affinity for humanoid robots,[106] with the release of the voice-interactive robot “Pepper,” there are great expectations for robots that combine speech recognition technology and robot engineering. Moreover, conversation robots are also expected to be applied to nursing services, etc.[107] Owing to the growing issue of declining birthrates, it is a technology Japan requires.

---

98 Nuance Communications website <https://www.nuance.com/>

99 Advanced Media website <https://www.advanced-media.co.jp/english>

100 FueTrek website <http://www.fuetrek.co.jp/en/>

101 However, similar to the robots described later, the degree of difficulty of performing speech recognition in the environment inside a car is greater than with a smartphone. For example, speech is often input to a smartphone very close to the user’s mouth, and hence, the mixing of noise is comparatively small, but inside the vehicle, engine noise, wind noise, music inside the car etc., are mixed.

102 Ichiro Akahori “Development of Speech Recognition for Car Navigation System (in Japanese)” Information Processing Society of Japan Research Report Audio Language Information Processing, 2005-SLP-058, 2005.10, p. 31. <https://ipsj.ixsq.nii.ac.jp/ej/?action=repository_uri&item_id=56972&file_id=1&file_no=1>

103 According to the survey company J.D Power, 19% of voice operation users felt it to be faulty, and the share of users using voice operation fell from 29% (2016) to 26% (2017). Customer satisfaction survey of automobile manufacturer’s own navigation systems, Lexus, Subaru ranked number 1. Operability when connecting a smartphone and improvement of speech recognition capability performance are important issues (in Japanese). 2017.10.6, p. 2. J.D. Power website <http://japan.jdpower.com/sites/default/files/17_jp_navi-oem_pr_j_fnl.pdf>

104 For example, a developer of a vehicle-related manufacturer, even where the average speech recognition rate is 99%, for a specific speaker with a low recognition rate, the product is unsatisfactory, and it is stated that it is difficult to accept it as a product. Akahori, *supra note* (102), pp. 31-32.

105 “How much can you rely on free smartphone navigation applications? Review of two popular applications (in Japanese),” 2015.1.28, Toyo Rubber Industries, <http://www.toyo-rubber.co.jp/ontheroad/drive/978/>; Tsunetake Noma “Navigation by smartphone, traffic information on the latest maps, tunnels a weak point,” Nikkei Style, 2017.6.13. <https://style.nikkei.com/article/DGXMZO17221990S7A600C1000000>

106 “Robots, AI and the Future, Awaji Symposium (in Japanese)” Nihon Keizai Shinbun, 2017.8.9, p. 9.

107 “Wide-scale proof of concept test of the usability of “communication robots” in nursing care, introducing 1,

However, the technical difficulty of speech recognition is dramatically different between speech recognition with a microphone near the speaker’s mouth, as in a smartphone, and speech recognition when speaking to a robot located away (remote speech recognition). Specifically, there are issues of reverberations in the room, reflection of sound, inclusion of background noise, etc. Therefore, beyond laboratory-level conversation robots, further research is necessary to realize conversation robots used by a wide range of real users.

Note that the smart speakers described above may be a possible pivot toward realizing this kind of conversation robot. This is because, through smart speakers, large amounts of speech data in the environment of the usage of general users and those from distant locations can be collected. However, the extent of usage of the smart speakers is as yet unknown, and it is unclear whether it will produce data useful for application to a conversation robot.

(3) Speech Translation

Speech translation combining machine translation and speech recognition is a further promising area of application. Although speech translation of typical sentences has already been implemented,[108] consideration for the contexts included in culture in translation is difficult even for humans, and it is not believed that this can be achieved simply by improvement through deep learning. To realize this kind of translation, steady academic research is necessary, including pragmatics in linguistics and the ways in which humans use language.

4. Privacy Issues

Privacy issues arise from always-on smart speakers.[109] This is because, unlike in smartphone applications where speech recognition (i.e., recording) is only performed when the user launches the application, speech data can be uploaded to servers of service providers at any time. With regard to this kind of problem, in Japan, only the risk aspects tend to be emphasized without careful discussion and enough understanding, which hinders development and adoption. This results in widespread of services originating overseas.

Kazunori Komatani, Osaka University

V Human−Agent Interaction

1. Notable Social Background

(1) Basic Definition

---

000 robotic units, supporting the collection and analysis of various data (in Japanese).” 2016.3.17. Japan Agency for Medical Research and Development website <https://www.amed.go.jp/news/release_20160317.html>

108 “First-aid speech translation application “First Aid Voice-Tra” for fire departments nationwide! (in Japanese)” 2017.4.18., National Institute of Information and Communications Technology website <https://www.nict.go.jp/info/topics/2017/04/170418-1.html>

109 “Concerns for privacy over AI speakers listening to family conversations (in Japanese),” 2017.9.4. Nikkei CrossTech website <http://tech.nikkeibp.co.jp/it/atcl/watcher/14/334361/083100911/>

Human–agent interaction (HAI) is a field of research addressing the interaction between human beings and man-made objects, using information processing, engineering mechanical control technologies, and applied psychology techniques to personify man-made objects; it is a design method developed to realize cordial and smooth communications akin to conversation between people.[110] Specifically, it handles humanoid robots and virtual agents.[111]

For systems to replace humans, there are AIs that substitute for brains, and materials, configuration designs, and machine controls that substitute for bodily function. HAI research aims to realize AI with humanity, including emotions, and systems with a physical body and expressive ability through gestures and expressions accompanying an internal "heart" state. Thus, HAI can be described as the psychological / physical interface, a container for encapsulating an artificial intelligence and producing external interaction, necessary to create the ultimate artificial human being. It is an important field of research for the wide dissemination of AI to general users.[112]

（2） Social Background

Since the 1960s, dialog system technology, as the human interface handling the point of contact between people and computers and in the field of human–computer interaction, has progressed from text conversations[113] to spoken conversations[114] and ultimately face-to-face conversation, imparting physicality to spoken conversation.[115] In particular, between the 1990s and 2000s, influenced by the development of robot control technologies, various humanoid robots appeared. For example, Kismet,[116] developed by the American Massachusetts Institute of Technology, (MIT), is a robot that combines sociability and physicality and can recognize human emotions at the same time as expressing its own.[117]

Through the development of this kind of interactive conversation system combining physicality and sociability, the system is highly anthropomorphic, can convey easy-to-understand information in lieu of people, and appears familiar to humans.[118] Such a system, by providing substitute services, in interpersonal work such as secretarial or call-center work, as conversation partners for the elderly, as counselors, and as everyday partners or pets in the home, can enrich human life. Moreover,

---

110 "Overview." Human-Agent Interaction Website <http://hai-conference.net/>
111 A dynamic, interactive system for the presentation of virtual characters through computer animation
112 The Japanese Society for Artificial Intelligence *supra note* (3), pp. 870-873.（Seiji Yamada, "Human-Agent Interaction"）(in Japanese)
113 Joseph Weizenbaum, "ELIZA: a computer program for the study of natural language communication between man and machine," *Communications of the ACM*, Vol. 9 No. 1, 1966.1, pp. 36-45.
114 Tatsuya Kawahara, Masahiro Araki, The Japanese Society for Artificial Intelligence, "Speech Dialog System (Cognitive Science)," *Ohmsha*, 2006
115 Shigeki Sagayama et al., "Galatea: An Anthropomorphic Spoken Dialogue Agent Toolkit" *Information Processing Society of Japan, Spoken Language Information Processing*, 2002-SLP-045, 2003.2, pp. 57-64. (in Japanese) <https://ipsj.ixsq.nii.ac.jp/ej/?action=repository_uri&item_id=57254&file_id=1&file_no=1>
116 "Through changing its facial appearance, it is possible for the robot to create expressions through movement of the eyebrows, eyes, mouth and ears." Cynthia Breazeal and Brian Scassellati, "A Context-Dependent Attention System for a Social Robot," *Proceedings of the 16th International Joint Conference on Artificial Intelligence*, 1999, pp. 1146-1153.
117 For details about physicality and sociality, see the following literature. eds., Michio Okada et al. "Physicality and Computers" *Kyoritsu Publishing*, 2001; Cynthia Breazeal, "Toward sociable robots," *Robotics and Autonomous Systems*, Vol.42 No.3-4, 2003.3, pp. 167-175. (in Japanese)
118 Seiji Yamada, "Design between Humans and Robots," *Tokyo Denki University Press*, 2007 (in Japanese)

research on such systems, clarifying cognitive frameworks from the perspectives of human cognition and thought, includes the possibility of contribution to the development of computer science.[119]

2. Research Fields

In this section, fields of research of particular importance to sociality and physicality in HAI are described.

(1) Sociality

Research on sociality is to analyze human psychology when aiming for mutual understanding while recognizing oneself and others, and to reflect that knowledge in the system. As a field of research, there is fundamental research related to the mechanisms by which people perceive a system as being capable of conversation, and recognizing the system as another entity distinct from themselves. These are systems that recognize the other party as having a mind, or a mechanism to gauge the intent of the other party (understanding of others),[120] or a mechanism to understand what kind of ways of thinking and beliefs form the basis of the other party (other-party models)[121] etc., and this kind of research is strongly linked to fundamental research related to explaining the "mind." Moreover, based on developmental research of recent years, it is also related to research fields in cognitive scientific analysis[122] and the explanation of mechanisms that make interpersonal communication difficult[123].

Research on sociality is also related to research on participation structures,[124] i.e., the ways in which conversation partners behave during conversation, and research on multi-agent systems[125] clarifying decision-making mechanisms in group discussions.

These studies are important for high-compatibility AI interfaces with people, which alongside basic communications research on reading, understanding, and sympathizing with the other party, are important for incorporating these into all of society's systems.

(2) Physicality

Research related to physicality has the aim of realizing various physically embodied expressions

---

119 Yuichiro Anzai, "Cognitive Science and Artificial Intelligence" Computer Science / Software Technology Course 17, Kyoritsu Publishing, 1987 (in Japanese)

120 E.g., Masuo Koyasu et al., 'Theory of Mind," The process before and after of understanding others' minds - Investigation through case analysis, *Bulletin of the Graduate School of Education, Kyoto University*, No. 46, 2000.3, pp. 1-25. (in Japanese) <http://hdl.handle.net/2433/57383>

121 Kim Bartholomew, "Avoidance of Intimacy: An Attachment Perspective," *Journal of Social and Personal Relationships*, Vol. 7 No. 2, 1990.5, pp. 147-178.

122 Junichi Yamamoto, Chieko Kasumoto "Development and Support for Autism Spectrum Disorders," Cognitive Science, Vol. 14, No. 4, 2007, pp. 621-639. (in Japanese) <https://doi.org/10.11225/jcss.14.621>

123 Watanabe Ayako et al., "2A1-E24 Discovery of Facial Characteristics and Preference Acquisition - Learning model for autistic children through applied behaviour analysis - (Cognitive Robotics), *Robotics & Mechatronics Lecture Presentation Summary*, 2008.6. (in Japanese)

124 Enomoto Mika, Yasaharu Tera, "Analysis of Non-verbal Behaviours Involved in 3-Person Conversation when Substituting Roles," Language and Speech Understanding and Dialogue Processing Research Society, The Japanese Society for Artificial Intelligence Document, No. 38, 2003.7, pp. 25-30. (in Japanese)

125 See Chapter VIII, Multi-Agent Systems. Besides refer to Masahiro Takahashi, Nakamori Akira, "Multi-Agent Consensus Formation under Individual Irrationality," Transactions of the IEICE, D1, Information Systems., I, Information Processing, Vol. 82 No. 8, 1999.8, pp. 1093-1101. (in Japanese)

in communication with people through touch and sight, obtaining information during exchanges with other parties, understanding the situation, changing internal states such as emotions, and performing autonomous actions, through a system connected to a "mind," namely, "interface agents."[126]

Researches on the means by which HAI can express emotions, perform robot control, computer graphics, and analysis, recognize facial expressions etc., are advancing. It is well-known that human beings, when communicating with a particular target, are strongly influenced by the target's outer appearance and behavior when making assumptions about the intelligence and mental capabilities of the target.[127] If the physical properties of the HAI are inadequate, the system may be regarded as a mere AI lacking a human-like mind, thus affecting its acceptance, and hence, research related to physicality is important.

3. Real-world Applications

Medicine, education, nursing, sports, and interpersonal services are highlighted as fields in which applications of HAI are being considered. In these fields, cordial, smooth communications with people may increase the evaluation of these services, and future humanoid robots and virtual agents incorporating HAI may be accepted as a new mode of "human resources." As HAI is expected to have the ability to understand the mindset of the other party, it is expected to have particular application to fields relating to the analysis and treatment of the mind: psychology, cognitive science, psychiatry, and mental health counseling. The use of agents possessing emotional intelligence and emphasizing sociality in light of the other party's emotions, in listening, counseling, or as pets is expected to bring spiritual stability through relations with others, amidst a social environment in which familial nuclearization and societal aging are advancing.

Domestic robots can be cited as an application arising from the early application of HAI. This is because, when compared with applications in medicine etc., less stringent technological development is considered permissible. Japan is strong in content creation technologies for animation and game industries, and is also strong in character design for anthropomorphic content. Moreover, with many Japanese researchers participating in the international HAI development community, Japan is positioned as a world leader.[128] Therefore, it is important to advance research and development of HAI and its application while fully utilizing Japan's fields of specialism.

4. Future Issues

There are technological and ethical issues related to HAI. With regard to technological issues, for society to begin to accept humanoid robots and virtual agents in lieu of people, improvement of

---

126 Kenji Mase et al., "A Basic Study on Interface Agents," *Information Processing Society of Japan Research Reports Human Interface Society Report*, 1996-HI-069, 1996.11, pp. 55-60. (in Japanese) <https://ipsj.ixsq.nii.ac.jp/ej/?action=repository_uri&item_id=37311&file_id=1&file_no=1>

127 Jeremy N. Bailenson et al., "The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments," *Presence: Teleoperators and Virtual Environments*, Vol. 14 No. 4, 2005.8, pp. 379-393.

128 "Steering Committee." Human-Agent Interaction Website <http://hai-conference.net/steering/>

both the physical expressive ability (physicality) and ability to comprehend others (sociality) are cited.

With regard to ethical issues, where it is viewed that humanoid robots and virtual robots live like humans, there is the issue of a need to distinguish between the ethical attitudes required by people and by equipment and systems such as these, and those developed up to now. When devices and systems deemed human are treated improperly or inhumanely, it may form impressions on third parties or influence their psyche, and there is the possibility that this may adversely influence the morals of human society. Furthermore, from the equipment / system usage side, there is the possibility that the ability to understand human psychology could be exploited to influence human behavior through fraud, brainwashing etc. Not only must the personal ethics of equipment and system designers be questioned, but a basic framework must also be constructed for the expression of emotion and appropriate designs related to ethics for such systems and equipment must be developed.[129]

Tomoko Yonezawa, Kansai University

VI Robots

1. Notable Social Background

(1) Definition of “Robot”

A broad definition of the technical concept of “robot” is most appropriate, without being limited to the robots drawn in anime or manufacturing machines used in factories, rather as “an information machine that interacts with physical space.” In the past, robots were, in many cases, defined with a focus on underlying technologies such as sensors, controls, and autonomous movement; however, with the rapid development of information technology, this definition does not capture the full breadth of robots that can be developed. With regard to automobiles, for example, automated driving is an apparent application, but intelligent functions such as preventing accidental pedal application, automatic braking, and automatic wipers, are also applications of robot technology. Moreover, intelligent rooms, i.e., rooms equipped with sensors (sensory devices) and actuators (motion devices) providing various services in response to human actions are also a type of robot.

Furthermore, there may be cases where the functionality of robots is not limited to physical entities. For example, some functions such as speech recognition, as in smartphone speakers[130] using AI, also implement processing using cloud computing[131] over the internet, and the various functions of sensors, cognition, control, and movement are physically dispersed. It is increasingly common to refer to these in general as robots.

---

129 Kukita Minao, “Possibility of Robotic Ethics,” Bulletin of Kyoto University Faculty of Letters Department of Philosophy, No. 11, 2009.3. <https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/71114/1/prospectus11_Kukita.pdf>

130 Speakers that allow operation of household appliances, and listening to news and music through voice operation while connected to the internet. See Chapter IV Speech Interfaces and Chapter VII IoT.

131 See Chapter VII IoT, 1(2) Improvements to Infrastructure Environment

Thus, technology, machine control, information processing, and intelligent processing contribute to a robot.

（2） Recent Trends around Robots

Robotic technology was almost entirely limited to manufacturing equipment applications; however, in recent years, a wide range of applications has been developed,[132] and robots have gradually come to be regarded as familiar things. The technologies making broad contributions to this field of application are primarily computer miniaturization, faster computing, semiconductors, micro-electro-mechanical systems (MEMS),[133] and similar technologies, whose development has led to the miniaturization and greater functionality of devices. The acceleration of computing, in particular, allows the real-time digital processing of complex behavior controls, making robots operable even outside of factory environments constructed specifically for robots. Indeed, with the progress of AI and interfaces, it has become possible to execute complex tasks and functions. Consequently, possible applications requiring interaction with humans, such as household and nursing applications, are expanding.

Furthermore, the application of robotics technology to various specific needs is expected to address issues of aging, population decline, and growing international competition. Nursing robots and autonomous vehicles are expected to be direct solutions to an aging society. The use of robots in the fields of infrastructure, civil engineering, and construction is believed to be essential for compensating shortages in human labor. In addition, work in high-radiation environments, such as the cleaning up after the nuclear accident caused by the Great East Japan Earthquake, cannot be performed by means other than robots, and a wide range of technological developments are desirable.

Even internationally, the development of robotic technology is advancing rapidly, focusing on the fields of manufacturing, distribution, and logistics, such as “Industrie 4.0”[134] promoted by the German federal government and Amazon’s automated warehouses. In China, Korea, and Southeast Asia, robot technology is being treated as a source of international competitive advantage, and in each country, technological development is being strengthened.

Moreover, robot competitions are also drawing attention as catalysts for technological development and education. Various international robot competitions have been held over the past 20 years, starting with the RoboCup (since 1997),[135] various competitions held by the DARPA,[136] the European Robotics League (since 2016)[137] held in Europe, etc. In Japan too, the World Robot

---

132 “Future market forecast of robotics market released, expected to grow to ¥9.7tn by 2035,” 2010.4.23. New Energy and Industrial Technology Development Organisation (NEDO) website: <http://www.nedo.go.jp/news/press/AA5_0095A.html> (in Japanese)

133 A microelectronic device where the electronic circuits, mechanical structure, and sensors are integrated on a single circuit board

134 Industrie 4.0 Website <http://www.plattform-i40.de/I40/Navigation/EN/Home/home.html> See Chapter VII IoT, 3. Practical Social Applications

135 RoboCup Federation Website <http://robocup.org/>

136 The autonomous vehicle races “Grand Challenge” (2004,2005), “Urban Challenge” (2007), and the “Robotics Challenge” (2012) in which humanoid robots compete in tasks, were held.

137 European Robotics League Website <https://www.eu-robotics.net/robotics_league/>

Summit is planned for 2018 and 2020[138]. Amazon has hosted the Amazon Robotics Challenge, encouraging technological development to advance its automated warehouses.[139]

2. Technology Trends

As robotic technology is a combination of various underlying technologies and their integration in addition to design technologies, technological trends for each technology are described.

(1) Sensing Technology

Sensing is a technology used by robots to recognize states in the outside world. Identification technology using machine learning techniques such as deep learning has been a remarkable development in recent years.[140] In particular, image recognition, video recognition, and speech recognition performance corresponding to robotic sight and hearing has advanced rapidly in the past 10 years.[141]

The most important factors in machine learning performance are the volume and quality of data,[142] and data collection and aggregation are of utmost importance to the improvement of sensing technology.

(2) Locomotive and Handling Technology

Flight technology, as represented by rotary-wing type flying robots (so-called "drones") has progressed remarkably in recent years as a technology handling robot movement.[143] The history of flying robot research is long; however, in recent years, the development of flying robots has advanced rapidly and their use is expanding through the miniaturization and acceleration of control computers owing to advances in computing technology. Currently, technological development of more sophisticated control technology and controls for multiple flying robots is being undertaken.

While wheels account for the majority of terrestrial movement methods, research and development of legged robots, including bipedal robots that can operate in the same work spaces as people, is also advancing. In particular, America's Boston Dynamics and Japan's SCHAFT's legged robots are complete, and have attracted much attention.[144]

For actuators, currently, motors are used in nearly all cases; however, research is advancing for

---

[138] World Robot Summit website <http://worldrobotsummit.org/>
[139] Amazon Robotics Website <https://www.amazonrobotics.com/#>
[140] See Chapter 1, Knowledge Processing & Machine Learning
[141] For image recognition and video recognition see Chapter III, Image Recognition & Processing, for Speech Recognition see Chapter IV, Speech Interfaces
[142] Kazunori Ohno et al., "Data Engineering Robotics - Sensor Data Produce New Robot Intelligence," Transactions of the Robotics Society of Japan, Vol.33 No.2, 2015.3, pp.97-99. <https://www.jstage.jst.go.jp/article/jrsj/33/2/33_33_97/_pdf> (in Japanese)
[143] "Report: Frontiers of Robotic Flight Research (1)," Journal of the Robotics Society of Japan, Vol. 34 No. 1, 2016.1, pp. 1-32 (in Japanese); "Report: Frontiers of Robotic Flight Research (2)," Journal of the Robotics Society of Japan, Vol. 34 No. 2, 2016.3, pp. 1-60 (in Japanese); Kenzo Nonami, "State of the art and issue of drone technology and business frontier," Information Processing, Vol. 59, No. 11, 2017.2, pp. 755-763. <https://www.jstage.jst.go.jp/article/johokanri/59/11/59_755/_pdf/-char/ja> (in Japanese)
[144] Both companies were under the umbrella of Alphabet, the holding company of Google; however, it was reported in June 2017 that it was acquired by Japan's Softbank Group. "RoboVB2 acquired by Softbank from Google" Nihon Keizai Shimbun, 2017.6.9, evening edition, p.1. (in Japanese)

next-generation actuators achieving smoother, flexible movements.[145] Particularly in cases where robots operate while in contact with human beings, research on artificial muscles that can achieve movements with a high affinity to human movements is advancing.

(3) Behavior Planning (Action Planning)

Methods for intelligently planning robotic movements are being developed. In this field, to enable complex processing and high-efficiency actions, research and development, in addition to implementation, of design planning for the coordinated movement of multiple robots and multiple actuators are being advanced.

Cooperative action planning is also important for cooperation between robots and people. In recent years, manufacturing robots that operate in cooperation with people have been commercialized,[146] and it is expected that the development of technology for controlling robots in synchronicity with human movements will advance in the future.

Moreover, as autonomous vehicles and flying robots become commonplace, negotiation technology for coordinating movements between robots will also become necessary. The negotiation technology between AIs necessary for coordination between robots has also been undertaken as a development theme by the Council on Competitiveness-Nippon (COCN).[147]

(4) Integration and Design Technology

In robotic technology, balanced integration technology and design technology are also important. “Design” as used herein refers not only to looking good, but also to a comprehensive design including ease of maintenance and operation. Compared with general information devices, in robots, it is important to reconcile the size, weight, and shape of each part. Operational ease, such as a balanced integration combining purpose and application, maintenance of moving parts, adjustment of sensors, etc., are also important technical requirements. Therefore, many companies are primarily engaged in the integration of robotic technology according to its purpose and application.[148]

Moreover, as the internet and cloud computing are becoming commonplace, it is no longer necessary to limit robots to only robotic function. For example, by performing speech recognition, image recognition, action planning etc., on the cloud over the internet, advanced and complex processing and operation can be achieved. The integration of this kind of internet technology is also an important element of integration technology.

---

[145] “Report: Future Images Drawn by Next-Generation Actuators” Journal of the Robotics Society of Japan, Vol. 33 No. 9, 2015.11, pp. 1-34 (in Japanese).

[146] “DENSO WAVE’s first human-cooperative robot, “COBOTTA” begins accepting orders Wednesday 29th No vember” website <http://www.denso-wave.com/ja/info/detail__1000.html> (in Japanese); “Release of human-coop erative robot MOTOMAN-HC10,”2017.6.8. Yaskawa Electric Corporation website <https://www.yaskawa.co.jp/n ewsrelease/product/31940> (in Japanese); “Human-robot collaboration (HRC).” KUKA Website <https://www.ku ka.com/en-us/technologies/human-robot-collaboration> (in Japanese)

[147] Council of Competitiveness - Nippon (COCN) “Interim report for 2017, AI negotiation, coordination and co operation,” 2017.10.13. <http://www.cocn.jp/thema98-M.pdf> (in Japanese) COCN, consisting of volunteers fro m industry, make policy recommendations for science and technology policy and industrial policy.

[148] Yushi Segawa, “An Introduction to Industrial Trends on Robot System Integration Overseas,” Journal of the Robotics Society of Japan, Vol. 33 No. 5, 2015.6, pp. 306-309 (in Japanese).

(5) Software Standardization

Robots have diverse underlying technologies and functions, and software for their flexible combination according to purpose and use; the standards that form the basis for development of such software are gaining importance. The robot operating system (ROS)[149] and robot technology middleware (RTM)[150] are software standards for robotics use. ROS is being developed with a focus on the US, and is currently the most widely used standard. Japan is primarily leading the development of RTM, and has strengths in fine control requiring rapid responses. Moreover, tools to bridge ROS and RTM are also being developed.

Simulation software is also important. "Gazebo," based on ROS,[151] and "Choreonoid,"[152] based on RTM, are softwares used to perform general robot simulation. Both of them can be applied to various robots, and can be used for the development of robot software. In addition, "SIGVerse," which can simulate interactions between robots and people, is also being developed.[153] As the number of robots working together with people is expected to increase in the future, the existence of these kinds of simulation software is important.

3. Real-world Applications

(1) Manufacturing Fields

In the manufacturing field, it is believed that the use of robots involved in the manufacturing process cooperating with people will increase in the future. In particular, to facilitate the transition to high added-value, low-volume multi-product production, it is possible that the processing of products by robots and humans simultaneously will become commonplace.

Moreover, items requiring fine pressure adjustment such as gear meshing, and soft objects such as textiles and food are difficult for robots to handle. Research to overcome this is currently being undertaken.

(2) Application to Harsh Environments

Much of the work performed in harsh environments such as disaster area restoration and nuclear accident clean-up cannot be performed without robots, and hence, their wide use in these applications is expected. Especially in the decommissioning process of the reactor at Fukushima Daiichi Nuclear Power Plant, as the majority of this activity is to be undertaken in environments where humans cannot enter, robot development is an urgent priority.

Moreover, to deal with future population decline and labor shortages, the development of robots to replace labor in harsh environments is further important. Specifically, applications to infrastructure maintenance, public works construction sites, and agriculture are expected.

---

149 ROS.org Website <http://www.ros.org/>
150 OpenRTM-aist Website <http://www.openrtm.org/openrtm>
151 Gazebo Website <http://gazebosim.org/>
152 Choreonoid homepage (From Shinichiro Nakaoka, Senior Researcher at the National Institute of Advanced Industrial Science and Technology) <http://choreonoid.org/ja/> http://choreonoid.org/>
153 SIGVerse website (From Tetsuya Inamura, Associate Professor at the National Institute of Informatics) <http://www.sigverse.org/wiki/en/>)

(3) Traffic and Transportation

Research and development of practical applications of autonomous vehicles is being actively promoted by enterprises and universities in various countries. Technologies such as automatic braking and automatic parking have already been adopted; however, there are expectations that, in the future, autonomous vehicles with a higher degree of autonomy will be put into use, and investigations into social acceptability and methods of usage are becoming a topic of conversation.[154]

The use of robots at transportation sites has already been partially realized in Amazon's automated warehouses, and the realization of fully automated warehouses is expected in the near future.[155] In the future, technological development including automation outside warehouses, cooperating with autonomous vehicles etc., is expected to become a topic of research.

(4) Nursing and Welfare

In the fields of nursing and welfare, robots supporting the movements of the elderly and disabled have reached a practical level and several products have already been developed.[156] In addition, the practical use of robots in particular nursing care scenarios, such as meal support, is progressing.[157] However, such robots support only a single function, and the excessive expectations of robots' ability to reduce the burden of nursing care must be treated with caution.[158]

(5) Homes and Offices

The use of robots in homes and offices is also increasing. With particular regard to cleaning and security, many commercial robots have already been put into practical use. With regard to guidance robots, in particular, owing to the opportunity afforded by events such as the 2020 Tokyo Olympics and Paralympics, there is the possibility to expand the use of robots equipped with multilingual speech interfaces aimed at foreign visitors.

Moreover, many types of robot partner products[159] are being put to practical use in communicating with people and supporting human activities, and examples of practical applications are expected to expand rapidly in the future.

---

154 "Trends and Challenges in Automated Vehicle Technology - Research Projects on Science and Technology 2017 Report, "(in Japanese) Investigation Report April 2018, National Diet Library Research and Legislation Examination Bureau, 2018 addresses technical trends in automated vehicles and addresses various issues.

155 "STARTupX (3) Logitech, Warehouse robots perform the most unmanned sorting, unlimited package shapes using 3D recognition," Nikkei Sangyo Shinbun, 2017.11.18, p. 1; "Robots for extracting and packing products, UK online supermarket," 2017.12.10. Nihon Keizai Shinbun Electronic Edition <https://www.nikkei.com/article/DGXMZO24362230X01C17A2000000/>_(in Japanese)

156 E.g., "The world's first cyborg robot 'HAL,'" CYBERDYNE website <https://www.cyberdyne.jp/products/HAL/>_(in Japanese)

157 E.g., "Meal support robot 'My Spoon,'" Secom website <https://www.secom.co.jp/personal/medical/myspoon.html>_(in Japanese)

158 Shigeru Yamauchi "Nursing Robots, The Current Situation and Challenges," Conference materials for the 2nd Bureau of Social Welfare and Public Health study on forms of regional comprehensive care systems towards implementation in Tokyo, 2015.7.30, pp. 4, 14., Tokyo Bureau of Social Welfare and Public Health website <http://www.fukushihoken.metro.tokyo.jp/kourei/shisaku/chiikihoukatsukaigi/02chiikihoukatsukeakaigi.files/06yamauchisama.pdf>_(in Japanese)

159 "Partner Robot" Toyota Motor Corporation website < http://www.toyota-global.com/innovation/partner_robot/>

4. Concerns

Although not necessarily limited to discussions of robots, improvements to the legal system concerning necessary safety protections and determining responsibility for issues arising from the applications of cutting-edge technologies are slow, and there are concerns that this will hinder their popularization. For example, in the case of traffic accidents caused by autonomous vehicles, no social consensus has been reached as to who bears the burden of responsibilities so far attributed to the driver, or how to proceed in this regard, nor are there legislative provisions. As autonomous driving technology is the core of national next-generation industry, it is necessary to develop legislation as soon as possible to maintain international competitiveness.[160]

Itsuki Noda, National Institute of Advanced Industrial Science and Technology

Ⅶ Internet of Things

1. Notable Social Background

The term “Internet of things” (IoT) appeared around 2000, and originally indicated a device with internet connection capabilities; however, nearly 20 years later, the term has assumed various meanings, as a concept, as a service, and as equipment, and it is becoming difficult to define it in a single word. The air-conditioning control system made by the American company Nest is a famous typical example of IoT. The air-condition control device, connected to the internet, has given rise to new services not only automatically controlling the air-conditioning in each household, but also controlling the power consumption of areas during peak times.[161] Here, devices (IoT devices) with internet connection capabilities are described, in addition to the services using them. Similar terms include cyber physical systems (CPS),[162] machine-to-machine (M2M),[163] and ubiquitous computing.[164] The following four factors are aspects of social background where IoT has gathered attention.

（1）Communications Cost Reduction

To reduce both energy consumption and cost, IoT devices connect to surrounding network connections using Wi-Fi (wireless LAN) or BLE (Bluetooth low energy) rather than by LTE, and these are necessary to send data to the cloud. With the increase in the adoption rate of devices equipped with Wi-Fi and BLE such as smartphones, smart watches (wrist-watch type wearable devices), and activity trackers (wearable devices measuring daily activity), the cost of

---

160 “Responsibility hurdles prevent autonomous driving, disputes over accidents, steep roads,” Nikkei Sangyo Shinbun, 2018.2.22, p. 2 (in Japanese).

161 “Nest’s Smart Home, Advances in the field of Solar Power,” Forbes Japan, 2015.5.8. <https://forbesjapan.com/articles/detail/3942>

162 A broad concept that collects data from sensors installed in the real world (physical space) in the cloud (cyberspace) via the internet, and feeds the results of analysis back into the real world.

163 A term mainly concerned with factories, indicating that machines are interconnected via a network and coordinate movements between one another.

164 Ubiquitous computing refers to installing processors in various things around us, connected to the internet.

communication modules has decreased.

As such, it is becoming easier to equip devices thus far not equipped with communications capabilities with such capabilities, and an environment for the installation of IoT devices is in place.

(2) Improvement of the Infrastructure Environment

One of the tasks for the spread of IoT has been the accumulation of data from IoT devices, and in 2006, Amazon's Amazon Web Service (AWS) was released into the market as a cloud environment in which to accumulate this data, resulting in dramatic changes since then.[165] The pricing of AWS is based on a volume-of-use system, which makes introducing IoT services into the market easier because it is possible to flexibly increase performance when expanding the service, while reducing the cost of cloud computing during the dawn of the service. Subsequently, Google and IBM released similar services, and it became possible to offer IoT services without building a cloud environment oneself. Moreover, the computational resources necessary for deep learning, etc. are also available in the cloud, and barriers to market entry have become extremely low.

(3) Spread of Digital Fabrication and Crowdfunding

When doing business, money and objects are important factors. Devices such as 3D printers[166] have become widespread, and based on design data, objects can be created inexpensively with a high degree of freedom, making ideas easier to implement. This technology is called digital fabrication. Through the widespread usage of crowdfunding,[167] fundraising for the commercialization of prototypes created through digital fabrication has become straightforward. These trends are one of the factors making it easy to commercialize IoT services.

(4) AI Technology Development

Using data gathered from IoT devices, it is believed that it is possible to create new value. For example, with the recent progress of AI technology, there are increasing expectations for new IoT services applying AI that has learned from data gathered from IoT devices to produce various insights and judgments.

2. Technology Trends

(1) Communications Technology

The majority of household IoT devices use Wi-Fi as their communications method, because of their high degree of market penetration and inexpensive communications chips. However, the communications range of Wi-Fi is inadequate for IoT devices performing communications outdoors or across wide areas. Therefore, wireless communications technology covering wide areas using low power consumption, known as a low-power wide area (LPWA) network, is continuing to expand.[168]

---

[165] Ministry of Internal Affairs and Communications, *supra note* (1), p. 91. <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/pdf/n2200000.pdf> (in Japanese)

[166] A machine that produces three-dimensional models by stacking layers one by one, based on 3D design data.

[167] A mechanism through which funding is solicited from many people over the internet.

[168] Ministry of Internal Affairs and Communications, "White Paper on Information Communications, 2017 Editio

LoRaWAN, SIGFOX, and NB-IoT can be cited as representative examples of LPWA. LoRaWAN offers communication speeds of approximately 250 bps and a communications range of approximately 10 km. In Japan, the three major mobile telecommunications companies have launched services used by multiple businesses, and this can be said to be the most widespread LPWA. The biggest feature is that because a radio operator's license is not required (as with Wi-Fi), anyone can set it up, and in such cases, communication fees are unnecessary. SIGFOX, like LoRaWAN, is a wireless communications technology that does not require a radio operator's license. Compared with LoRaWAN, it has slower communication speeds of approximately 100 bps, but a wider communications range of approximately 50 km. It is mainly used for IoT devices using small volumes of data such as water meters, etc. NB-IoT is an extension of the LTE[169] communications method used by mobile phones aimed at IoT, with communication speeds of approximately 100 kbps and communications range of approximately 20 km. As it can use existing LTE base stations, it has the advantage of being able to cover the entire country. However, as it requires a radio operator's license, its adoption has been mostly limited to telecommunications companies.

（2） IoT Devices

In IoT devices, a control device for regulating the device over the internet is required. Circuit boards,[170] which form the basis of the development of plans, and prototypes of these kinds of control devices have already become widespread, and sensors and communications modules that can connect to them are sold inexpensively. Thus, it has become possible for anyone to prototype IoT devices easily.

Instances where IoT devices are being sold as finished goods are also increasing. Locks that can be controlled remotely from smartphones (smart locks), surveillance cameras, and internet-connected, voice-operated (a person's voice, etc.) smart speakers, which enable one to listen to news and music and operate home appliances, are typical examples. Apart from these devices, communications-capable models of conventional household appliances such as air-conditioning units, refrigerators, and TVs are gradually expanding.

In particular, as the accuracy of speech recognition has increased dramatically with advances in AI technology, commercialization of smart speakers has progressed[171] and according to research by the American company eMarketer, the number of smart speaker users in the US in 2017 reached 35.6 million people.[172]

3. Real-world Applications

（1） Factories

---

n," 2017, pp.128-130. <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h29/pdf/n3300000.pdf>(in Japanese)

[169] A radio communications system compliant with the IMT-Advanced standard established by the International Telecommunication Union. It is a higher-speed communications standard as compared with the third-generation mobile communications system (3G). NB., LTE is an abbreviation of long-term evolution.

[170] Control boards such as Raspberry Pi and Arduino are known.

[171] Representative examples include Amazon's "Amazon Echo" series and Google's "Google Home" series. See Chapter IV, Voice Interfaces.

[172] "Alexa, Say What?! Voice-Enabled Speaker Usage to Grow Nearly 130% This Year," 2017.5.8. eMarketer W ebsite <https://www.emarketer.com/Article/Alexa-Say-What-Voice-Enabled-Speaker-Usage-Grow-Nearly-130-This-Y ear/1015812>

In Germany, the “Industrie 4.0”[173] technical policy has been promoted since 2011. This aims to achieve an autonomous, distributed intelligent manufacturing system by networking devices within factories, and combining these with sensors and AI. For example, at Daimler AG, which is adopting Industrie 4.0, network creation is being undertaken in every stage of design and production.[174] The German Research Centre for Artificial Intelligence[175] creates simulated factories within its laboratories, and conducts experiments in collaboration with various companies.[176]

（2）Traffic

Telematic vehicle insurance, linking dash cams, and automobile insurance can be cited as IoT in automobiles. Starting in January 2018 (Heisei 30), Toyota Motor Corporation and Aioi Nissay Dowa Insurance have begun selling insurance with discounts of up to 80% on driving premiums among its insurance fees, based on driving information obtained from vehicles connected to the internet.[177]

Moreover, the American Uber Technologies Inc., which operates a dispatch service, operates a system that changes the route of the ride-sharing service according to past usage and traffic conditions.[178]

（3）Housing

As a representative example of IoT in housing, the American company Nest’s[179] (acquired by Google in 2014) thermostat is famous. Nest’s thermostat, which uses sensors and AI (machine learning) to automatically control the air-conditioning in homes, can reduce electricity bills by 20% while maintaining comfort. Furthermore, using the data accumulated by these thermostats, electricity suppliers can predict the demand for power, and some of the cost reductions achieved through this system are being passed back to consumers.[180]

The smart locks described before can be cited as another example.[181] In addition to private use, they are used as a means of delivering keys in holiday rental services such as AirBnB, and are

173 *supra note* (134)
174 “Production is becoming smart. Industry 4.0 and the networked factory.” Daimler Website <https://www.daimler.com/innovation/case/connectivity/industry-4-0.html>
175 "A central research organisation of German AI research, and takes the form of a non-profit limited public-private partnership." Dengel, Andreas（Translated by Koichi Kose), "20 Years of the German Artificial Intelligence Research Centre (DFKI) - A path to success and those who made it possible," Information Processing, Vol. 49 No. 7, 2008.7, pp. 810-817.
176 “SmartFactory Laboratory.” Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) Website <https://www.dfki.de/web/living-labs-de/living-lab-smartfactory>
177 Toyota Motor Corporation, Aioi Nissay Dowa General Insurance Co., Ltd, “Development of Japan’s first telematic vehicle insurance reflecting drivers’ behaviour for Toyota’s connected car,” 2017.11.8. Toyota Global Newsroom website <https://newsroom.toyota.co.jp/jp/detail/mail/19575777>; “Tough, Connected Car Insurance,” Aioi Nissay Dowa General Insurance website <http://aioinissaydowa-tsunagaru.jp/tough-tsunagaru/> (in Japanese)
178 “ETA Phone Home: How Uber Engineers an Efficient Route,” 2015.11.3. Uber Engineering Blog Website <https://eng.uber.com/engineering-an-efficient-route/>
179 Nest Website <https://nest.com/>
180 “Nest’s Business Model — Flagship IoT startup aims to collect all data in the domestic space,” 2015.8.31. nomad journal website <https://nomad-journal.jp/archives/530>
181 In Japan, “Akerun” and “Qrio” have been commercialised. Akerun website (PhotoSync) <https://akerun.com/>; Qrio website <https://qrio.me/> (in Japanese)

expanding rapidly alongside the expansion of these holiday rental services.[182]

（4）Health and Welfare

IoT devices are being used in elderly care services. In Japan, in serviced accommodation for the elderly, absence detection and fall recognition using AI and sensors are being adopted[183] to prevent the overlooking of falls and simultaneously lower management costs.

Various wearable IoT devices have become popular in the field of personal health information management (personal healthcare). Welfare services supporting walking using activity meters, which measure daily human step-count activity, have already begun[184]; however, in the future, with the diversification and advancement of sensors, it is believed that their use will expand into applications in mental health care and disease prevention.

4. Concerns

In Japan, from past experiences of success focusing on manufacturing, the idea of planning industrial revitalization through "monozukuri" is firm, i.e., there is a tendency to focus on the "T," the things, of IoT. However, services creating new value using data arising from "things" together with AI are necessary.

Moreover, as IoT devices spread to households, privacy issues also frequently arise. From the "Insecam Incident,"[185] where 73,000 units had not had their default password changed, resulting in their cameras being viewable from anywhere in the world, information leaks from teddy bears equipped with IoT functionality,[186] to secret data collection by sex-toy manufacturers,[187] the increasing popularity of IoT devices in households has become a "trojan horse," and cases of private information being distributed are abundant. Therefore, it is important to consider privacy in the usage of IoT devices.

However, overseas, there is a tendency to first attempt to do something, and change the rules when issues arise, whereas in Japan, there is a strong tendency to thoroughly investigate all issues prior to launching services. To not trail behind the West in the development of IoT services for the sake of privacy considerations, a balanced strategy is necessary.

Yutaka Arakawa, Nara Institute of Science and Technology

---

182 "Gemalto, a smartlock solution where your smartphone can be a key," 2017.6.20., MyNavi News <https://news.mynavi.jp/article/20170620-a062/>

183 "Launch of a safety confirmation service for elderly housing using ICT," 2017.9.20. Family Net Japan website <http://www.fnj.co.jp/news/pdf/2017/20170920.pdf> (in Japanese)

184 For example, "Everyone's Health Site 'KENPOS,'" Ewel website <https://www.ewel.co.jp/category/service/kenpos/p796/> (in Japanese)

185 "A Creepy Website Is Streaming From 73,000 Private Security Cameras," 2014.11.6. Gizmodo Website <https://gizmodo.com/a-creepy-website-is-streaming-from-73-000-private-secur-1655653510>

186 "Cyber kid stuns experts showing toys can be 'weapons'," 2017.5.16. Phys.org Website <https://phys.org/news/2017-05-cyber-kid-stuns-experts-toys.html>

187 "This sex toy tells the manufacturer every time you use it," *Splinter*, 2016.8.9. <https://splinternews.com/this-sex-toy-tells-the-manufacturer-every-time-you-use-1793861000>

Ⅷ Multi-agent Systems

1. Notable Social Background

Through the widespread application of various information systems together with the development of the internet, huge amount of diverse social data have been produced in the distributed environment so that information systems based on conventional centralized architecture have had difficulties in their operation. Accordingly, distributed manner of constructing information systems had attracted many researchers. Then, researches on multi-agent systems has been a part of main stream of AI research since 1980s[188].

For example, by simulating human society, groups of robots, etc., as groups of multiple agents (multi-agent systems: MAS), each of which is modeling an human, a robots, etc. with small-scale AI program (agent), it is possible to apply MAS to solve various problems.

In multi-agent systems, each agent, having its own diverse characteristics, abilities, and purposes, makes decisions independently at the same time as influencing other agents, cooperating with others to avoid collisions, etc. This not only makes it possible to analyze the nature and behavior of groups considering both the diversity of actual people and robots and their interactions, but can also be of use to the design of methods to control such groups.

Researches on multi-agent systems have progressed with theoretical research at its core, pursuing computational models to simulate the behavior of groups in society. For example, a broad range of research activities has been developed with various societal actions and mechanisms as the target, such as negotiation and cooperation protocols, modeling and analysis of competitive markets, and matching a wide variety of service providers and users. However, the number of theoretical works linking to practical systems has scarcely increased, and it has not attained widespread use as a technology for developing distributed systems.

In recent years, societal implementation of the IoT has advanced rapidly, and by 2020, approximately 30.4 billion IoT devices are predicted to be connected to networks.[189] Algorithmic trading (automated trading using computers) is now being used in the stock markets, and it is becoming realistic that automated vehicles will intermix with human-controlled vehicles in the near future. It is intractable to expect the result of integration of information technologies into society, nor its effect to the whole of society. Therefore, designing human society is expected to become a very hard task.

Multi-agent systems enable us to individually model a wide variety of entities/elements, such as humans, robots, sensors, various information systems, etc., Multi-agent systems enable us to individually model a wide variety of entities/elements, such as humans, robots, sensors, various information systems, etc., and to calculate or analyze their continuous interactions. Hence, the utility of MAS is once again being recognized for modeling increasingly complex societies.

---

188 Initial research on multi-agent systems, called distributed artificial intelligence (DAI), was held in the first distributed intelligence workshop in the US in 1980. Randall Davis, "Report on the Workshop on distributed AI," *SIGART Newsletter*, No. 73, 1980.10, pp. 42-43.

189 Ministry of Internal Affairs and Communications, "2016 Information Communications White Paper," 2016, p. 80. (Figure 2-1-1-1), <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/pdf/n2100000.pdf>(in Japanese)

2. Technology Trends

(1) Simulation Technology

Modeling complex systems, such as real-life human society, as multi-agent systems by calculating the continuous interactions between agents, the method for performing measurement and analysis of the system's behavior is called "multi-agent social simulation" (MASS). MASS is a powerful technique for reproducing complex social phenomena, and is expected to support the design and validation of new systems, mechanisms, services, etc. in society.

One of the typical application domain for MASS is economics. The "U-Mart Project," [190] begun in 1999 and implemented over more than 10 years, has been actively conducting public experiments both domestically and overseas using virtual artificial markets (U-Mart) constructed based on MASS, as an effort to attempt to design systems for financial markets, and many researchers have participated in the representative artificial market research in Japan. .[191] U-Mart is effective not only for research, but also for obtaining teaching materials for information science and economics, and it is used in instruction at numerous universities and graduate schools.[192]

Rescue ina disaster is another suitable application domain for MASS. In response to the Great Han-Shin Awaji Earthquake, the "RoboCup Rescue"[193] was conceived to foster competition for technologies useful for disaster relief using AI and robotics. Since 2001, an international competition has been held using MASS to simulate rescue activities by firefighters, rescue-team, and rescue robots. Over the course of the competitions, attempts to create disaster prevention plans and to predict disaster damages using multi-agent models have been undertaken, thus improving the technology. This effort has been expanded to competitions using actual robots, and their results are contributing to actual activity support under disasters, such as the development of Quince, a robot sent into the reactor building at Fukushima Daiichi Nuclear Power Plant.[194]

(2) Negotiation Technology

Negotiations are one of the main research areas for multi-agent systems. Research on agent negotiations, defined as "a process of exploratory calculations to identify potential consensuses,"[195] began in the 1990s. At the stage, it has never been used outside the area of theoretical research using multiple agents to solve distributed search problems. Subsequently, with the design of functions representing human's preferences which is to be assigned to agents, and the development of research protocols for multi-issue negotiations, technological development aiming to perform negotiations

---

190 Hiroshi Deguchi, "Laboratory Creation for Economic Advancement - The Significance of Participant Simulation in Virtual Markets," Evolutionary Economics Proceedings, Vol. 3, 1999, pp. 251-252 (in Japanese).

191 Hajime Kita et al. eds., *Realistic Simulation of Financial Markets: Analyzing market behaviors by the third mode of science*, Tokyo: Springer, 2016.

192 Hajime Kita et al., "Educational Activities of Artificial Market Project U-Mart," Systems, Control, Information, Vol. 49, No. 7, 2005.7, pp. 271-276 (in Japanese), <https://doi.org/10.11509/isciesci.49.7_271>

193 RoboCup-Rescue Technical Committee et al. eds., (Satoru Tadokoro, supervised by Hiroaki Kitano), "RoboCup-Rescue: Emergency Large-Scale Disaster Relief Challenge," Kyoritsu Publishing, 2000 (in Japanese).

194 Keiji Nagatani et al., "Emergency response to the nuclear accident at the Fukushima Daiichi Nuclear Power Plants using mobile rescue robots," *Journal of Field Robotics*, Vol. 30 No. 1, 2013.1・2, pp. 44-63.

195 Nick. R. Jennings et. al., "Automated Negotiation: Prospects, Methods, and Challenges," *Group Decision and Negotiation*, Vol. 12, No. 2, 2003, pp. 58-73.

automatically using these agents as human proxies is now being undertaken.

As an attempt to accelerate the development of negotiation technology using agents, the International Automated Negotiating Agents Competition (ANAC)[196] has been held since 2010 and research and development of agents with efficient strategies under varied negotiating conditions has been facilitated. At ANAC, agents conduct negotiations without knowledge of the other side's information. Although closer to negotiations in the real world, as the negotiation partner's information is insufficient, it is not sufficient to simply adopt strategies based on game theory[197] solution concepts and models. In other words, game theory is useful for the analysis of the nature of negotiation problems such as whether a solution (consensus) or equilibrium point exists, but it is necessary to develop new negotiation technologies with consideration for the elements of game theory that are hard to address, such as the complex preferences of negotiators, uncertainty over information about the negotiation partner, and learning and predicting from the negotiation partner's behavior.[198]

Negotiations at ANAC can be viewed as a type of incomplete information game in which all information about the game is not opened to all participants. Conversely, games such as Shogi and Go, where one can obtain all information about the game including information about the other party, are called complete information games. In recent years, AI has shown better performance than humans in Shogi and Go, and interest has now shifted toward incomplete information games. This trend has also occurred in the multi-agent field and technical applications are progressing, with agent players researched and developed from game theory research beating professional human players at poker,[199] an incomplete information game. Moreover, research and development of agents automatically playing a party game "Werewolf," in which players progress through conversation with other players, and wherein players deceive one another and sometimes persuade and cooperate, has also begun in Japan.[200]

3. Real-world Applications

Not limited to the technical fields above, there is a steady accumulation of the number of cases where the results of theoretical research lead to real-world applications. Here, representative examples of applications and promising areas of application are introduced.

---

196 "Automated Negotiating Agents Competition (ANAC)." Negotiation Website (Interactive Intelligence Group, Faculty of Electrical Engineering, Mathematics, and Computer Science (EEMCS), Delft University of Technology (TU-Delft)) <http://ii.tudelft.nl/negotiation/index.php/Automated_Negotiating_Agents_Competition_(ANAC)>

197 "Game theory is a theory studying situations in which multiple decision-making agents (humans and agents) interact with one another in their decision-making. Takahiro Watanabe, "Seminar: Introduction to Game Theory," Nihon Keizai Shinbun Publishing, 2008 (in Japanese).

198 Tatayuki Ito, "Applications of Multi-Agent Automatic Negotiation Model," Information Processing, Vol. 55 No. 6, 2014.6, pp. 563-571 (in Japanese).

199 "Carnegie Mellon artificial intelligence victorious in heads-up poker tournament," 2017.2.01. ZDNet Website < https://www.zdnet.com/article/carnegie-melon-artificial-intelligence-victorious-in-heads-up-poker-tournament/>; "How an AI took down four world-class poker pros," 2017.2.10. Engadget Website <https://www.engadget.com/2017/02/10/libratus-ai-poker-winner/>

200 Daisuke Katagami, Fujio Toriumi, Hirotaka Osawa, Michimasa Inaba, Kosuke Shinoda, Hitoshi Matsubara "Project AI Wolf," Artificial Intelligence, Vol. 30, No. 1, 2015.1, pp.65-73 (in Japanese)

(1) Applications in the Security Field

There are successful examples of improvement of the security of various locations by applying game theory for modeling negotiations and cooperation among people.

When designing security plans, it is necessary to solve the difficult issue of how to position limited numbers of security guards while factoring in the actions of attackers targeting poor security. Professor Milind Tambe of the University of Southern California in the US proposed an algorithm to calculate security behavior that incorporates randomness to prevent security plans being predicted by attackers, while considering the differences in the importance of security targets. The "ARMOR" system for determining the patrol routes of security guards and dogs at Los Angeles International Airport has been developed and adopted based on this research.[201]

This algorithm has been utilized for security at various locations, including for the flight scheduling system "IRIS"[202] at the Federal Aviation Security Bureau and the "Protect"[203] system, which determines the patrol routes of the US Coast Guard.

(2) Applications in the Medical Field

Matching, which appropriately determines the combination of providers and beneficiaries of a given service, is one of the primary fields of multi-agent research. There is a case where the developed theory was applied to the development of an algorithm for organ transplantation.

In the US, over 70,000 kidney disease patients are waiting for a transplant, and only 10,000 patients will undergo transplant surgery each year; 4,000 patients will lose their lives. Professor Tuomas Sandholm of Carnegie Mellon University in the US developed a high-efficiency algorithm matching kidney donors to kidney disease patients based on the proximity of hospitals, and in 2008, it was adopted into the Pan-American United Network for Organ Sharing.[204] Since 2010, this algorithm has been in continuous use, automatically generating plans for kidney transplants. For example, it has been reported that it produced and executed a particularly long organ transplant plan involving 60 donors and recipients.[205,206]

---

201 Marc Ballon, "A Safer World," 2015 Spring. USC Viterbi Website <http://magazine.viterbi.usc.edu/spring-2015-2/features/a-safer-world/>; James Pita et al., "Deployed ARMOR Protection: The Application of a Game Theoretic Model for Security at the Los Angeles International Airport," *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems: industrial track*, 2008, pp. 125-132. <http://teamcore.usc.edu/papers/2008/AAMASind2008Final.pdf>

202 Jason Tsai et al., "IRIS - A Tool for Strategic Security Allocation in Transportation Networks," *Proceedings of 8th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2009)*, 2009.5. <http://teamcore.usc.edu/kiekintveld/papers/2009/trkot-IRIS.pdf>

203 Eric Shieh et al., "PROTECT: An Application of Computational Game Theory for the Security of the Ports of the United States," *Proceedings of 26th AAAI Conference on Artificial Intelligence (AAAI-2012)*, pp. 2173-2179, 2012. <http://teamcore.usc.edu/people/eshieh/shieh_protect_aaai_talk_20120414.pdf>

204 David J. Abraham, et. al., "Clearing Algorithms for Barter Exchange Markets: Enabling Nationwide Kidney Exchanges," *Proceedings of the 8th ACM Conference on Electronic Commerce (EC-07)*, 2007, pp. 295-304. < https://www.cs.cmu.edu/~sandholm/kidneyExchange.EC07.withGrantInfo.pdf>

205 Website (The New York Times) < http://www.nytimes.com/2012/02/19/health/lives-forever-linked-through-kidney-transplant-chain-124.html>, "60 Lives, 30 Kidneys, All Linked", 2012.2.18

206 Website (The New York Times) <http://www.nytimes.com/interactive/2012/02/18/health/record-chain-of-kidney-transplants.html?action=click&contentCollection=Health&module=RelatedCoverage&region=Marginalia&pgtype=article>, "A Record Chain of Kidney Transplants", 2012.2.18

(3) Promising Areas of Application

As already stated, in recent years, implementation of IoT within society has advanced rapidly. Future social systems are becoming more complex, and their design is becoming more difficult. MASS, which predicts and analyzes the behavior of complex systems, is a promising technology for supporting the design of these complex social systems. In Japan, in the 2009 Transdisciplinary Science and Technology Roadmap produced by the Transdisciplinary Federation of Science and Technology,[207] the expectations for the usefulness of MASS in the design of social systems were described.[208] This trend is gaining momentum through the use of Big Data and high-performance computing, and the following budding attempts have begun.

Professor Kiyoshi Izumi of the University of Tokyo is undertaking systems design for economic markets using artificial market simulation with multiple agents. There are ongoing attempts to extract knowledge related to market design through wide-scale analysis of the influence of changes in tick size (minimum units of price fluctuation) on the Tokyo Stock Exchange using MASS, and to establish how to protect the market share (or conversely, acquire) in the face of competition between domestic and foreign stock markets.[209]

In the transportation field, in the design of transport systems using MASS, future ride-sharing companies will combine the advantages of taxis (demand type) and route buses (shared transport), and a system of vehicle operation "smart access vehicle," (SAV),[210] which can operate vehicles along optimal routes in response to the demand communicated between the customer and driver, is being developed. SAV technology is based on the MASS developed by Dr. Itsuki Noda at the National Institute of Advanced Industrial Science and Technology to verify the effects of the introduction of on-demand buses.[211] A good example arising from the actual development and operation of a system is validating the feature requirements of new traffic systems in advance using MASS, and it is believed that similar attempts can be applied to fields outside of traffic.

4. Future Issues

Multi-agent research is not expected to remain theoretical, but to continue to contribute to the construction of new social mechanisms, and cases of development accompanying these practical efforts are expected to increase. In such a case, the issue becomes how to model the actual society targeted for support. In other words, we confront the issues of agent modeling, i.e., with what granularity the detail of each agent should be modeled, and whether the solutions and predictions obtained from the multi-agent system have practical meaning. For this problem, how to secure a

---

207 Union of academic societies specialising in general science and technology applicable in transdisciplinary fields (Designated Non-Profit Organisation)

208 Transdisciplinary Federation of Science and Technology "Transdisciplinary Science and Technology Academic Road Map Report" 2009.3, pp. 94-95. (in Japanese) <http://www.meti.go.jp/policy/economy/gijutsu_kakushin/kenkyu_kaihatu/20fy-pj/oudan2.pdf>

209 Takanobu Mizuta et al., "Relationship analysis of tick size and transaction volume in stock markets using artificial market simulation," JPX Working Paper, Vol. 2, 2013.1.30. (in Japanese) Japan Exchange Group website <http://www.jpx.co.jp/corporate/research-study/working-paper/tvdivq0000008q5y-att/JPXWorkingPaperVol2.pdf>

210 "SAV (Smart Access Vehicle)" (in Japanese) Mirai Share website <http://www.miraishare.co.jp/sav/>

211 Noda Itsuki et al., "Do On-Demand Buses Pay?" IECIE Technology Research Report, Vol. 2003 No. 8, 2003.1.29, pp. 31-36. (in Japanese)

sufficient amount of data related to the human behavior to be modeled (or group thereof), and how to learn about human behavior from the data, are the keys to the development of technology related to data collection and learning.

Hiromitsu Hattori, Ritsumeikan University

IX Crowdsourcing

1. Notable Social Background

Crowdsourcing is a general term for the mechanism of making requests of an unspecified number of people for work, and an online platform to realize it. The presence of compensation for crowdsourcing work varies; however, Amazon's crowdsourcing platform, Mechanical Turk (MTurk), has prompted attention toward crowdsourcing.[212] Using MTurk, clients can easily outsource simple tasks, and workers can receive compensation working entirely online. MTurk has achieved growth as a platform that can capture a large workforce, and many similar platforms have appeared both in Japan and overseas.

To imitate human judgment, AI technology requires correct answer data created by humans. For example, to build an algorithm to determine whether a dog or a cat is depicted in a picture, "pictures of dogs" and "pictures of cats" as classified by humans are necessary. Crowdsourcing has been widely used as a method for expanding the scale of this kind of work by adding meaning to data. Furthermore, to use AI technology in solving social problems, cooperation with human knowledge, perception, judgments, and decision-making is indispensable. The concept of "human computation" has been proposed, incorporating human queries into AI systems, and crowdsourcing is garnering attention as the basis for achieving this purpose.

2. Technology Trends

Here, basic technological trends for improving the usability of crowdsourcing, quality control, workflow control, and mechanism design are discussed.[213]

（1） Quality Control

Unlike conventional forms of employment in which employers and workers can establish ongoing relationships of trust, it can be difficult for employers to gauge the ability, reliability, motivations etc., of workers when crowdsourcing, and hence, quality control of the output of such work is an important issue. As a countermeasure, a method called "redundancy" is often used, i.e., requesting the same work from multiple people to ensure quality. Studies into statistical machine learning methods to estimate the reliability of each worker and the degree of difficulty of each task in such cases are ongoing.

---

212 Amazon Mechanical Turk Website <https://www.mturk.com/>
213 This article refers to Hisashi Kashima et al., "Human computation and Crowd Sourcing" （Machine Learning Professional Series） Kodonsha, 2016, pp.24-68 (in Japanese)

(2) Workflow Control

In crowdsourcing, it is possible to combine multiple tasks, such as handling the results of one task in another. Using this, workflows in which the output by a worker is evaluated and corrected by another worker are being studied.

(3) Mechanism Design

To improve the speed and quality assurance of work in crowdsourcing, the participation of workers having the ability and motivation to carry out the task is indispensable. To encourage the participation of such workers, research is being undertaken into “mechanism design,” i.e., setting the contents of tasks and compensation appropriately. For example, a method of preventing the involvement of dishonest workers by splitting compensation into a participation fee and a performance fee has been proposed.

3. Real-world Applications

(1) Examples of Real-world Applications

(i) Large-scale Character Recognition

Large-scale character recognition is an important application example, having been undertaken since the early stages of crowdsourcing. To get humans to confirm images of characters that AI recognition found difficult, Luis von Ahn, associate professor at Carnegie Mellon University developed the “reCAPTCHA” system, using crowdsourcing to achieve a high degree of accuracy in character recognition when digitizing the contents of books.[214] When determining whether a user accessing a website is human or a computer program attempting unauthorized access, a character recognition test is given to the user. reCAPTCHA can have that user perform character recognition work by including images of characters from books in the test.

Large-scale character recognition through crowdsourcing is also used in industry. For example, Sansan, which provides a service converting business cards to digital forms, uses crowdsourcing for character recognition of those business cards.[215]

(ii) Large-scale Annotation

Image, voice, and video data are created and stored in large quantities on a daily basis. To efficiently classify and organize this enormous amount of data, in addition to making it easy to search, adding annotations to the data is essential.[216] Large-scale annotation through crowdsourcing is also contributing to AI development.

For example, Associate Professor Fei-Fei Li of Stanford University produced annotations for over 10 million images using crowdsourcing, publishing it as the “ImageNet” dataset.[217] At an image

---

[214] Luis von Ahn et al., “reCAPTCHA: Human-Based Character Recognition via Web Security Measures,” *Science*, Vol.321 Issue 5895, 2008.9.12, pp.1465-1468.

[215] “Data Strategy & Operation Center” Sansan Inc. website <https://jp.corp-sansan.com/dsoc/> (in Japanese)

[216] To add annotations (metadata) to data such as images, voice and video etc.

[217] Jia Deng et al., “ImageNet: A Large-scale Hierarchical Image Database,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248-255.

recognition competition held in 2012, an image recognition method applying deep learning to the Image Net image data achieved a high level of accuracy, and catalyzed the subsequent deep learning boom.

(iii) Real-time Applications

Real-time applications (applications processing tasks in real time) incorporating crowdsourcing are being developed. Associate Professor Jeffrey P. Bigham of Carnegie Mellon University developed a smartphone application, "VizWiz," to support the visually impaired.[218] For example, when "I want to select a can from a cupboard filled with cans," the visually impaired person takes a picture of the cupboard with cans and simultaneously records the spoken question in the smartphone; the image and speech data are sent to VizWiz and an answer is returned in tens of seconds. The creation of an answer according to the photo and question is performed by crowdsourcing workers. As a visually impaired person's question is automatically sent to numerous workers, the question and answer process is performed in nearly real time. It is an example of a real-time question and answer task applied to image and speech performed through crowdsourcing, a difficult task for artificial intelligence today.

(2) Promising Areas of Application

(i) Translation, Dialog, and Q&As

It is currently difficult for language processing technology to perform all translation, dialog, and Q&A tasks. Therefore, efficient project advancement through the division of work into projects that can be achieved using language processing technology, and work that cannot be performed except by people, thereafter using division of labor to crowdsource the latter, is being investigated.[219]

Moreover, depending on the contents of the translation, dialog, or Q&A, expertise of the worker may be required. In such cases, using the quality assurance technology described above, it may be possible to predict the expertise and reliability of the worker, and to choose workers appropriate for the project accordingly.

(ii) Gathering Ideas for Problem-solving

There are already practical examples of the use of crowdsourcing to collect ideas for problem-solving, which can be considered gathering the collective wisdom of mankind. For example, the US DARPA has designed military vehicles through crowdsourcing.[220] In addition, the "InnoCentive" platform, which collects ideas to solve issues in research and development, is being used by P&G

---

[218] Jeffrey P. Bigham et al., "VizWiz: Nearly Real-time Answers to Visual Questions," *Proceedings of the 23rd annual ACM Symposium on User Interface Software and Technology*, 2010, pp. 333-342. <http://up.csail.mit.edu/other-pubs/vizwiz.pdf>

[219] Atsushi Nakada, "Using Machine and Human Specialisation in Division of Labour in Translation" 2016.1.5. Nikkei Business Online <http://business.nikkeibp.co.jp/atcl/report/15/061700004/122100063/> (in Japanese)

[220] "After Successful Design Challenge Competition and Testing, DARPA Begins Early Transition of Adaptive Vehicle Make Technologies," 2014.2.5. Defense Advanced Research Projects Agency website <https://www.darpa.mil/news-events/2014-02-05>

America, etc.[221]

In the future, it is expected that problem-solving through crowdsourcing will be applied to a wide range of fields, through the brainstorming of ideas and the development of technologies related to the process of efficiently evaluating and refining ideas.

(iii) Employing the Elderly

Crowdsourcing, which makes it possible to work at home, is suitable for elderly people who are motivated to work. However, it is necessary to assume that there are limitations on the amount of time and work that can be handled owing to physical restrictions among the elderly. Therefore, techniques are being studied to enable multiple elderly people to work as virtual teams through crowdsourcing.

Professor Michitaka Hirose of Tokyo University, collaborating with IBM Japan, is advancing research on the concept of the "Senior Cloud." This method performs matching based on the experience and skill of the elderly to specific projects, and when multiple elderly people work as a virtual team, it realizes efficient matching and schedule organization.[222]

4. Future Issues

As the use of crowdsourcing is advanced, there are concerns as to whether operators may move toward exploitation with unfairly low wages. For example, it can be pointed out that labor exploitation might arise from economic disparities in the crowdsourcing market.[223] This issue can be said to be disadvantageous not only for workers, but also for employers owing to long-term deterioration in the quality of work and decrease in the number of participants. In response to this issue, improvements in the relative standing of workers by introducing a mechanism for the mutual evaluation of employers and workers, and by sharing information about workers on multiple platforms, etc., are being considered as a solution.

Yukino Baba, Kyoto University

221 Jeff Howe "*Crowdsourcing - Why the Power of the Crowd Is Driving the Future of Business,*" Crown Business, 2008.

222 Research and Development of the "Senior Cloud," website <http://sc.cyber.t.u-tokyo.ac.jp>

223 Rowland Manthorpe, "Gig Economy, the actual situation that is causing "a new type of poverty: investigation results," *WIRED*, 2017.4.4 <https://wired.jp/2017/04/04/gig-economy-jobs-benefits-dangers/>