愛国学園大学人間文化研究紀要　第22号　37～46頁　2020年３月

〔論文〕

Smart Agriculture and Data Science
－Progress and Innovation in Intelligent Agriculture－

Akira Nissaka, Takemi Machida
Aikoku Gakuen University
Japan

Abstract

ICT-based Smart Agriculture is an innovative method of agriculture in which each phase, from production to marketing, is integrated and improved upon by taking advantage of existing knowledge and techniques in information and communication technology (ICT), as well as by creating and testing new ICT tools and frameworks within which such tools are implemented. The ultimate goal of the approach is to achieve low-input, sustainable agriculture comprised of high productivity, cost reduction and efficiency, and improvements in food and labor safety.

Smart Agriculture is made possible by drawing upon both ICT and human wisdom. Fortunately, we live in an age in which platforms for highly advanced ICT systems are well-established and in which cloud computing environments and multifaceted analyses of Big Data are becoming more readily available. With ease of accessibility to technology and complex data, it is now possible for any of us to understand and implement Smart-Farming techniques. As such, high expectations exist for this powerful approach, and it is apparent that there is a significant need for ICT-based systematization of vast information and the creative development of algorithms for the production of biological systems.

Keywords: Smart Agriculture, Data science, Big Data, Data-driven agriculture

1. Introduction

It has been about ten years since the term “Smart Agriculture” was coined.

Recently, the word often appears in newspapers, and it has been received as a term that represents innovative agriculture. There are other similar terms, such as "IT Agriculture", "Digital Agriculture", and "AgTech", all of which refer to agricultural methods that include IT and efforts to modernize agricultural production sites. However, "Smart Agriculture" is somewhat different from the other terms, and although the use of IT has high importance, the concept includes the farming style and the method to develop farm management. The term generally aims for agricultural and rural innovations, including energy and environmental innovation technologies. The goal of Smart Agriculture is to develop optimal agricultural ecosystems and sustainable agricultural production systems, which differs from other terms that emphasize IT.

2. Smartification of agriculture

The social environment around agriculture has been becoming increasingly complex and diverse, and agriculture in the 21st century is forced to make significant changes to achieve an environmental-friendly and sustainable agriculture. Knowledge-based and record-based methods, the so-called data-centric principles are becoming increasingly crucial in agricultural method and work management, and new concepts and solutions, such as future environmental protection, food safety, farm resource management, and others, are crucial in the field of agricultural management. The implementation of Smart Agriculture is becoming even more critical in situations where the integrated resource management of farming areas and the establishment of agricultural ERP (Enterprise Resource Planning) are needed.

The dissemination of the internet and smartphones have significantly contributed to the development of IoT ecosystems at farming areas, and full-scale use of agricultural big data by using cloud computing is now being implemented. Agricultural cloud services to substitute management and analysis based on the collection of agricultural data have been started, and the use of farming software based on cloud-type SaaS (Software as a Service) has been spreading gradually. Along with the progress of commercialization and large-scale management, there is a growing need to save labor and improve efficiency, which requires efficient data utilization and information systemization that actively utilize the IoT ecosystem.

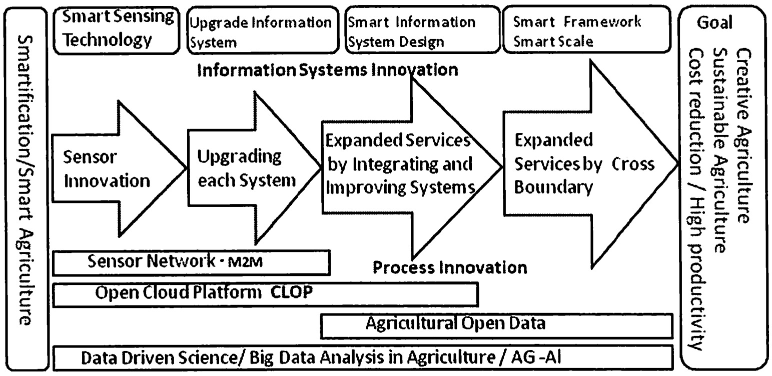

Fig.1 Road map for Smart Agriculture

The underlying technology of smartification is IT, in particular, information systemization. That means to systemize the information on agricultural sites and to fuse and integrate them. As systemization has its own challenges and purposes, the required type and amount of data are different. The methods to obtain and analyze data are also different. The progress in computer technology has enabled the use of machine learning, big data, and others, and the ecosystem of agricultural system construction has evolved to such an extent that is incomparable to that in the past. The spread of mobile IoT ecosystems, such as smartphones and tablet PCs, has dramatically expanded the possibilities and scope of agricultural system construction, where data collection often takes place in fields and outdoors. In particular, low power consumption, low bit rate, and long-distance communication LPWA (Low Power Wide Area), which are now close to their full-scale use, are considered to dramatically increase the efficiency of data collection are useful for agriculture that needs to collect a lot of data outdoor. The information system is in the form of data processing, and it is possible to discover new things by using data that exist in the agricultural field but have not been used due to the difficulties in collecting them. Advances in sensing technology will improve the time density of the measurement systems, and it is expected that new methods of analysis, such as crop and field dynamics, will progress along with the rapid progress of the dynamic analysis methods. The improvements in data collection ecosystem have given rise to data-driven agriculture, which derives the following data from the existing data, allowing the accumulation of many actual and test data and the increasing effectiveness of hypothesis-driven analysis methods. Since machine learning and deep learning require a large amount of high-quality training data, the evaluation of fruit ripening and growth using machine learning can be anticipated. Also, the application range has been expanding, such as for the optimal control of liquid fertilizer and formalization of the technique owned by serious farmers, in addition to image analysis. As shown in Figure 1. the road map for Smart Agriculture and Figure 2. shown in the position of Smart Agriculture.

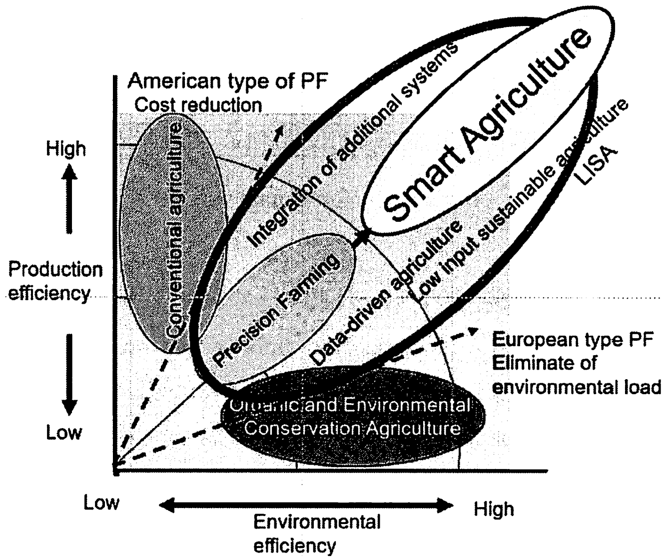

Fig.2 Position of Smart Agriculture

3. Data create agriculture

The 21st century is known as the data century. The IoT connects 50 billion devices around the world, and a vast amount of information is generated from each industrial, social, and personal activity. The total amount of generated data is now at the "Zetta" byte level, and it is predicted to reach 163 zettabytes by 2025. We are in the era of data explosion. These data and information can be considered as new social resources, and they have been attracting much attention as the seeds of new industries through big-data analysis and AI technology. In agriculture, it is vital to use data in a way similar to other industries. The ability to analyze the big data will lead to the knowledge of everything, and it is said that society and new agriculture can be created by controlling the data.

Farming fields and areas are gold mines of data, and a large amount of data are produced daily from the cultivation, production, sales, and business management stages. The data-driven methods of analyzing and accumulating those data are the basic methods of Smart Agriculture and is an agricultural method in which IT and information systems support a series of processes of data collection, visualization, analysis, forecasting, decision making, planning, and others.

Smart Agriculture treats data as essential resources, and all actions are taken based on a data-based approach. Decisions on what kind of data should be collected and analyzed as well as how to design the data and the database should be constructed in consideration of the contents of analysis, and these judgments require the basics of data science. The above are the duties of SEs (system engineers) and data scientists who specialize in agriculture. However, there are very few agricultural SEs and data scientists at present.

Data design and the setting of results and goals for crops and agricultural work require agricultural knowledge and experience, and the education and training of agricultural SEs and scientists are urgently needed.

The agricultural information system needs not only internal data but also external data, such as weather and market data. In particular, since the state and local governments hold various technical agricultural data, it is necessary to open and provide these data to the public. Farm equipment manufacturers that provide cloud services also hold the data on the farmer's farm work and field. An agricultural data collaboration platform was established last year to connect and effectively utilize the data from different organizations, thus improving the usage ecosystem of external data. Agriculture is a complex system and is not possible to be built through a simple information processing system. It is believed that the development of Smart Agriculture will be further promoted through data connectivity, as well as information system connectivity and integration.

4. Expectations toward artificial intelligence

The third explosion of artificial intelligence started with the invention of deep learning in 2006. The breakthrough recognition rate that was achieved in The 2012 ILSVRC Image Recognition Contest was partly due to the adoption of reverse propagation in deep learning (multilayer neural network). Because of the adoption, it became possible to effectively extract high-order feature quantities in image recognition that brought the level of recognition to a human level, causing a breakthrough in AI technology, which was relatively stagnant at that time. The potentials shown by artificial intelligence increased dramatically with the use of deep learning and enforced learning. The AlphaGo Zero, which came out in 2017, was the result of enforced learning and is a surprising result that opened the door to the potentials of artificial intelligence.

The most effective application of deep learning is image recognition, and most commercial AI systems use supervised learning (CNN). Image data are used in many aspects of agriculture, such as crop growth and disease diagnoses, equal-size sorting of fruits, determination of when to harvest, and others. AI-based physical condition management systems, which apply the trajectory of behavioral data of individual animal livestock in a barn and individual image data, have also been developed. The improvement of system accuracy depends on how to obtain and apply a large amount of high-quality training data. AI research in agriculture requires a development environment that is stable for a long term because the measurement factors and time-series elements differ depending on the type of work and division, making intensive studies at the national level and a collaborative study system between the industry, government, and academia necessary. The development of AI chips and chips for inference engines have been progressing rapidly, and the future may see a productive agricultural system developed with a general-purpose AI farming system that can be used for general agriculture and various types of autonomously working robots. Various farming equipment and machines have been further improved to be smart devices that incorporate artificial intelligence with the goal of autonomy of machines as well as autonomous operation of agricultural production facilities and various cultivation systems centering on agricultural machine robots with autonomous functions. For example, there have been progressing developments of a complete robot cultivation system for rice, a complete machine robot system for potato plantation, and the like. We believe that AI farming systems and AI mechanical robot systems are at the center of the solution to securing enough quantity and quality of agricultural products in a time when there is a severe shortage of farmers as the result of population aging.

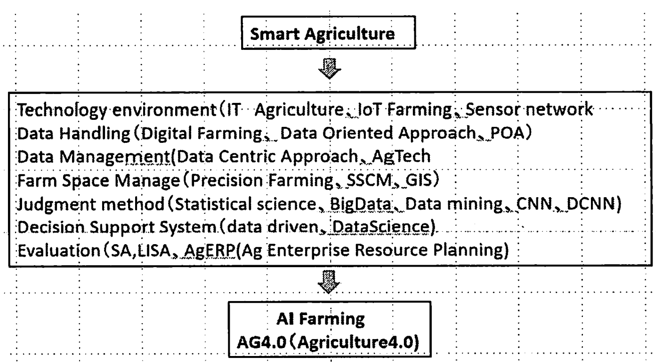

Fig.3　Approach for AG.4.0

5. Agricultural data science

The analysis of a large amount of data enables hyper-observability, which in turn enables accurate prediction and estimation, leading to the said: "data create society". Similarly, it can be said that "data create agriculture", as in Smart Agriculture.

The analysis of big data has emerged as an alternative to statistical methods, and the information processing and computer engineering related to big data are collectively known as "data science", and those who perform it are called "data scientists".

Smart Agriculture can also be called data-centric agriculture, which uses data analysis to make predictions and decisions. Farming methods that depend on such data to optimize the management of crop growth, cultivation environment in real time and the whole farm production, as shown in Figure 3, are considered to be the first technological stage toward Agriculture 4.0.

Data science is the science of data-based decision-making and is one of the tools for problem-solving. It is a science that uses a large amount of data to derive meaningful information, laws, relationships, and others. Through the dissemination and cost reduction of various sensors and improvement of IoT ecosystem that center on the internet, the enormous data generated from every aspect of the industry and society, including agriculture, can be mined, enabling the exploration of new knowledge, which can be considered as the treasure among them.

The knowledge gained through the use of big data covers a large area and is highly explanatory. Moreover, it is highly objective and superior to conventional experience-based methods in the sense that it is independent of subjective interpretation. In the conventional method, experience-based events are modeled by mathematical expressions and others, but in data science, it is possible to create a model that represents the knowledge and events based on a large amount of data. Machine learning and statistics are then used for the classification, clustering, pattern extraction, and prediction processes in model creation.

There are limitations in obtaining useful information among human experiences, and their range of application is incomplete. However, since high-quality data can be collected in significant quantities by using appropriate collection and handling methods, so the coverage does not depend on subjectivity, it is possible to achieve automatic extraction of highly-exhaustive models and obtain unexpected findings.

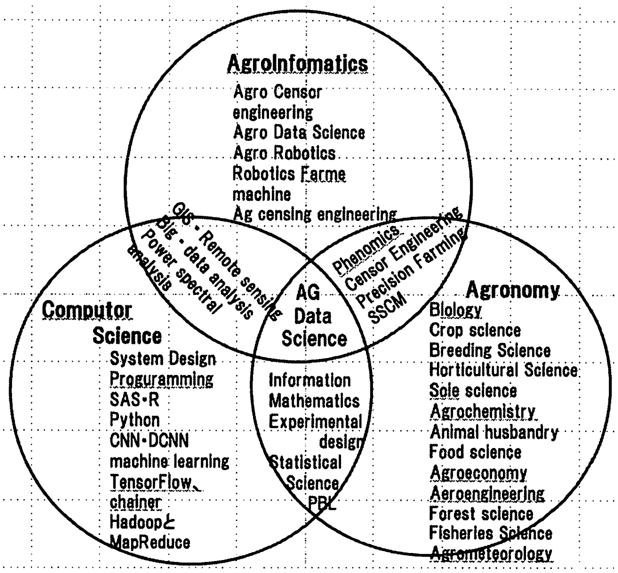

Fig.4　Structure of agricultural data science

Data science requires appropriate technologies for collecting and storing high-quality data in a large amount, statistics for data analysis, computer engineering that can utilize analysis libraries, and adequate business knowledge to evaluate the results.

Since data handling and processing are essential in data-dependent Smart Agriculture, data science can be said to be the primary technology to support farming methods. In addition, the use of data science is crucial because agriculture possesses its own difficulties, which are unique to biological production systems. As shown in Figure 4, requires three fields of knowledge, which are the agricultural informatics for the collection, storage, analysis, and prediction of various data that are specific to the agricultural field, computer engineering that includes big data and AI, and agricultural sciences.

6. Data-driven agriculture

Data-driven methods to analyze and accumulate data are the methods to develop Smart Agriculture. Agricultural systematization has many difficulties compared with the use of ICT in other industries because it deals with living organisms. In that sense, the roles played by the systematization and smartification are even more significant. The sensing and modeling of biological and environmental information, including soil, and the development of algorithms for the prediction of crop growth are the crucial points of systemization, which are yet to be fully systemized, despite years of accumulated efforts.

There are many varieties and types of crops, and each cultivated land has its own regionality. Also, there are many fluctuating factors, such as the differences in the natural environment and soil condition. There is also the difficulty in calculating the variable factors and regionality for systematization. Also, it is necessary to have sufficient knowledge of both biology and informatics to understand the relevant crops and organisms and to systemize the information, respectively, in developing the system.

The cycle of agricultural production is about several times a year, which limit the opportunities to collect big data. In many cases, control relationships and coefficients obtained from statistical processing of mutual relation or correlation developed by the hypothesis-driven methods that use experience and accumulated results from cultivation tests will still be important analytical tools in the future.

7. Smart Agriculture and its challenges

In agriculture, ICT usage and systematization are faced with more difficulties compared with other industries because it deals with living organisms. In that sense, the roles played by the systematization and smartification are even more significant. The sensing and modeling of biological and environmental information, including soil, and the development of algorithms for the prediction of crop growth are the crucial points of systemization, which are yet to be wholly systemized, despite years of accumulated efforts.

The meager number of agricultural system engineers is one of the primary issues in promoting Smart Agriculture. Modeling is the measurement of an event using sensors to obtain a control amount (work amount) with a modeled analysis engine that utilizes numerical data. The use of a variable fertilization system can avoid excess fertilization, which will contribute to cost reduction and environmental preservation. However, the prescription engine that determines the amount of fertilizer for a place in the following year based on the amount of harvest perform the calculations based on models. Thus, the working accuracy of such machines is determined by the degree of model completeness. In addition, the tangerine fruit grading data are analyzed by the prescription model, whose results are calculated by an engine, which is modeled by the prescription model that feeds back the results to the fertilization and grading operations. Until now, the development of prescription and growth models use the control relationships and coefficients obtained from statistical processing of mutual relation or correlation developed by the hypothesis-driven methods that use years of experience and accumulated results from cultivation tests. In this regard, although the data-driven big data analysis is a new method, it is faced with the difficulty to accumulate sufficient amount of agricultural data due to the crop growth cycle that is limited to several times a year. Other than obtaining tens to hundreds of petabytes of data, there are also many issues that need to be overcome, such as the standardization of biological sensing intervals and data acquisition. Nevertheless, the big-data analysis of cultivation data can be expected to become an indispensable technology of Smart Agriculture.

References T.Machida, Smart Agriculture and Data Sciences, Agriculture and Forestry Statistics Publishing Inc.,P11-14, 2019

T.Machida, Agriculture and rural change in smart farming, Agriculture and Horticulture YOKENDO, LTD Vol.93 No.8, P725-728, 2018