JAST (Journal of Animal Science and Technology) TITLE PAGE

Upload this completed form to website with submission

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Review Article
Article Title (within 20 words without abbreviations)	The vision of Big Data Recirculation for Smart Livestock Farming in South Korea
Running Title (within 10 words)	Data Recirculation for Smart Livestock Farming in South Korea
Author	Seung-Hoon Lee[1, 13], Kyu-Sang Lim[2, 13], Hakkyo Lee[3, 13], Jaeyoung Heo[3, 13], Jaemin Kim[4, 13], Seon-Ho Kim[5, 13], Sung-Hak Kim[6, 13], Jong-Eun Park[7, 13], Dajeong Lim[8, 13], Jae-Don Oh[9, 13], Bu-Min Kim[10,13], Song-Won Yoo[11, 13], Donghyun Shin[12, 13] and Jun-Mo Kim[1, 13]
Affiliation	[1] Department of Animal Science and Technology, Chung-Ang University, Anseong, Gyeonggi-do 17546, Republic of Korea [2] Department of Animal Resources Science, Kongju National University, Yesan, Chungcheongnam-do 32439, Republic of Korea [3] Department of Animal Biotechnology, Jeonbuk National University, Jeonju, Jeonbuk-do 54896, Republic of Korea [4] Division of Applied Life Science, Gyeongsang National University, Jinju, Gyeongsangnam-do 52828, Republic of Korea [5] Department of Animal Science and Technology, Sunchon National University, Suncheon, Jeollanam-do 57922, Republic of Korea [6] Department of Animal Science, Chonnam National University, Gwangju 61186, Republic of Korea [7] Department of Animal Biotechnology, College of Applied Life Science, Jeju National University, Jeju, Jeju-do 63243, Republic of Korea [8] Department of Animal Resources Science, Chungnam National University, Daejeon 34134, Republic of Korea [9] Department of Biotechnology, Hankyong National University, Anseong, Gyeonggi-do 17579, Republic of Korea [10] National Institute of Animal Science, Rural Development Administration, Wanju, Jeonbuk-do 55315, Republic of Korea [11] Korea Institute for Animal Products Quality Evaluation, Sejong 30100, Republic of Korea [12] Department of Agricultural Convergence Technology, Jeonbuk National University, Jeonju, Jeonbuk-do 54896, Republic of Korea [13] The Korean Society of Animal Big Data Research, Korean society of animal science and technology, Seoul 06367, Republic of Korea
ORCID (for more information, please visit https://orcid.org)	Seung-Hoon Lee: 0000-0001-6703-7914 Kyu-Sang Lim: 0000-0001-5406-266X Hakkyo Lee: 0000-0001-5387-4885 Jaeyoung Heo: 0000-0002-9721-8043 Jaemin Kim: 0000-0003-1746-2546 Seon-Ho Kim: 0009-0006-5947-4157 Sung-Hak Kim: 0000-0003-4882-8600 Jong-Eun Park: 0000-0003-0718-3463 Dajeong Lim: 0000-0003-3966-9150 Jae-Don Oh: 0000-0001-7756-1330 Bu-Min Kim: 0000-0001-7836-3360 Song-Won Yoo: 0000-0002-7650-0779 Donghyun Shin: 0000-0002-0819-0553 Jun-Mo Kim: 0000-0002-6934-398X
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	NRF (No. NRF-2022R1A2C4002510)

Acknowledgements	This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2022R1A2C4002510).
Availability of data and material	Upon reasonable request, the datasets of this study can be made available from the corresponding author.
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Lee SH, Lim KS, Lee H, Heo J, Kim J, Kim SH, Kim SH, Park JE, Lim D, Oh JD, Kim BM, Yoo SW, Shin D, Kim JM Writing - original draft: Lee SH, Lim KS, Shin D, Kim JM Writing - review & editing: Lee SH, Lim KS, Lee H, Heo J, Kim J, Kim SH, Kim SH, Park JE, Lim D, Oh JD, Kim BM, Yoo SW, Shin D, Kim JM
Ethics approval and consent to participate	This article does not require IRB/IACUC approval because there are no human and animal participants.

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	**Fill in information in each box below**
First name, middle initial, last name	Jun-Mo Kim
Email address – this is where your proofs will be sent	junmokim@cau.ac.kr
Secondary Email address	
Address	Department of Animal Science and Technology, Chung-Ang University, Anseong, Gyeonggi-do, 17546, Republic of Korea
Cell phone number	+82-10-4026-5644
Office phone number	+82-31-670-3263
Fax number	+82-31-675-3108
First name, middle initial, last name	Donghyun Shin
Email address – this is where your proofs will be sent	sdh1214@gmail.com
Secondary Email address	
Address	Department of Agricultural Convergence Technology, Jeonbuk National University, Jeonju, Jeonbuk-do, 54896, Republic of Korea
Cell phone number	+82-10-9228-3959
Office phone number	+82-63-219-5614
Fax number	+82-504-467-0947

Abstract

A smart livestock farm is a livestock farm where information and communication technology systems are used. Based on the measured data, these systems can make decisions regarding all processes, including stocking, breeding, shipping, and evaluation. The data generated from smart livestock farms have increased the complexity and diversity of phenotypes. Fused data that integrate environmental and phenotypic information from smart livestock farms with genetic data are valuable for detailed applications in breeding and specifications, as they help understand complex and organic phenotypes and environments. However, their effectiveness is limited by restrictions on data sharing and non-standardized formats. This limitation leads to other restrictions against researchers, such as restrictions on the range of projects, the supply of new technologies or farm species, and policy development or application restrictions. Therefore, promoting a recirculating environment to increase productivity, developing climate-adapted livestock, and implementing policies are necessary. We discuss the smart livestock farm from the perspective of '$Phenotype = Genetic\ value + Environment\ value$'. The dissemination of smart livestock big data and essential components, such as data warehouses, is outlined.

Keywords: Smart farm, Livestock, Data recirculation, Big data, Data warehouse, Digital twin

Introduction

The global human population is estimated to reach 9 billion by 2030 [1]. In South Korea, livestock face several issues such as disease, animal welfare, increasing population age, stinking, and excrement treatment. The Paris Agreement, which encourages countries to submit long-term low greenhouse gas emission development strategies [2], is another issue that must be solved to achieve goals in the livestock sector [3]. These problems are not limited to South Korea but are becoming a reality in livestock farming worldwide.

A smart livestock farm (SLF, also known as Precision livestock farm [PLF]) is defined as a livestock farm where information and communication technology (ICT) systems are used [4]. The SLF produces more data than previously measured information at livestock sites, and SLFs are constructed using these data. Based on the measured data, decisions could be made for all processes, such as stocking, breeding, shipping, and evaluation. Before the smart livestock technology is introduced into the livestock farm, workers manually recorded data in an analog form and subsequently digitized them into a simple datasheet. Moreover, the post-slaughter grading data are less compatible because they existed between the farm and the Institute for Animal Products Quality Evaluation, respectively, which limits their availability. An SLF system is an extended concept in SLF and creates new value by utilizing new data produced in SLF and data from the entire existing livestock system. When the SLF system was constructed, the data types that could be collected were increased. Previously, the data could only be recorded when an event occurred or at a specified time; however, as real-time observable data become available as phenotypic and environmental data, they are being built in the form of comprehensive big data to improve the competitiveness of the livestock industry at the individual farm and national level.

Generally, in livestock, phenotype (P) (i.e., appearance, productivity, meat quality, and disease resistance) is obtained by the summation of genetic (G) and environmental (E) values ($P = G + E$) [5, 6]. The development of next-generation sequencing (NGS) has led to the creation of various platforms for estimating genetic effects on phenotypes. Consequently, the data generated from SLFs have increased the complexity and diversity of phenotypes. Additionally, the expanding range of measurable environmental data enables more detailed and accurate effect estimation compared to traditional technologies. Fused data that integrate environmental and phenotypic information from smart livestock farms with genetic data are valuable for detailed applications in breeding and specifications, as they help understand complex and organic phenotypes and environments. However, its effectiveness is limited by restrictions on data sharing and non-standardized formats. This limitation leads to other restrictions against researchers, such as restrictions on the range of projects, the supply of new technologies or farm species, and policy development or application restrictions. Therefore, promoting a recirculating environment to increase productivity, developing climate-adapted livestock, and implementing policies are necessary. In this review, we introduce a dataset from the SLF and some projects for the recirculation of SLF big data. Moreover, a data recirculation system could provide a blueprint for livestock and related industries.

Phenotypic data in smart livestock

In livestock, phenotypic data represent several phenomena such as growth rate, disease information (infection information or viral load), feed intake, average daily gain, milk production, carcass information, and carbon emissions [5]. In the pre-SLF era, phenotypic data were written manually or simply digitalized as a data sheet. With the development of SLF instruments, the types of enhanced phenotypic data, such as carbon data, have increased.

To facilitate data collection using the SLF, standardization of ICT devices and data is being carried out at the national level. Starting with the establishment of the national standard for 'Sensor Interface for Smart Livestock' in 2020, 'Livestock Specification Management Device Data Collection Standards' laid the foundation for revitalizing data utilization services by improving the compatibility and quality of livestock ICT devices until the establishment of national standards (2022-23) and the establishment of group standards for the 'Smart Livestock Data Model' (2023). In addition, we compiled a list of the step-by-step smart livestock technologies, devices, programs, and systems used in Korea to implement smart livestock farming (as of 2023) (Table 1).

What kind of phenotypic data are in smart livestock farm systems?

At the beginning of SLF in South Korea, a robotic milking system (RMS, also known as an automatic milking system) was applied to dairy farms. An RMS without human involvement was first introduced in 1986 in Europe [7]. The adoption and spread of RMS in dairy farms has increased milk yield [8]. However, early models of the RMS system had some restrictions on installation and operation, such as cost, population size, and cow-teat arrangement. The RMS was developed in the form of a combination of artificial intelligence (AI) and internet of things (IoT) technologies. The developed RMS serves as a comprehensive SLF device capable of recording various phenotypes (e.g., milk quality, ruminating data, and somatic cell counts) in dairy farms in real time. Moreover, a combined AI-based image sensing system can be used to measure body shape, such as the height, angle, and width of the hips, by comparing real data [9].

An automated measurement system (AMS) was developed to measure carcass phenotypes and evaluate them to grade abattoirs [10]. In the past, workers were evaluated using a ruler and scale to measure the carcass directly [10]. However, as the number of slaughtered animals increases, this method becomes less efficient because of the time required for measurements [11]. There are two types of AMS, the Fat-O-Meat'er (FOM) and VCS2000 systems, which are distributed in Europe [11]. The FOM measurement system measures the lean meat percentage and fat thickness of the carcass using an ultrasonic instrument [12, 13]. The VCS2000 system, another AMS instrument, uses a video-based image analysis system [14, 15]. The types of phenotypes that were scanned and measured in these AMS were lean percent, back fat thickness, carcass weight, meat cut percentage, and weight [12]. Based on measured phenotype data, the Korea Institute for Animal Products Quality Evaluation (KAPE) evaluates animal products in abattoirs. The measured data were created as a database and uploaded to the server for the SLF system.

AI-based monitoring systems are used to measure the behavior of livestock as a phenotype. The monitoring system installed in the shed tracks the behavioral patterns of livestock 24 h a day and records their physical characteristics, such as body temperature [16]. In the swine industry, there are challenges to improving productivity and disease issues; thus, monitoring body conditions is important for the swine industry. Ear tag sensors are useful devices for tracking swine behavioral patterns. The sensor measures the body temperature in direct contact with the tissue [17]. The tag also includes radio frequency identification (RFID), which can recognize individuals and forward the measured information to an AI-based monitoring system.

Genetic data in smart livestock farms

Genetic data are used to estimate phenotypes based on the genotypes of livestock. Previously, animal breeding used phenotypes and statistical data; however, recent animal breeding studies have used genetic data [6]. The Human Genome Project has led to the development of technologies to detect genomic sequences and identify their functions. Currently, sequencing costs have decreased to less than $ 100 per individual, which has enabled researchers to generate a wide range of genetic data [18, 19]. Biological research on livestock and agriculture has been conducted to improve genetics and reduce environmental stress [20]. Therefore, it is necessary to comprehensively understand their complex interactions because various phenotypes are organically connected [20]. In the field of livestock, genetic data have been collected by many research groups worldwide, including South Korea, and shared at the national or consortium level to secure publicity. In addition, Korea is making great efforts to build a platform for producing large-scale SNP data for genetic and genome selection for paternity testing, preprocessing analysis, and breeding value estimation, and is considered a smart livestock technology (Table 1).

How is the genetic data from smart livestock farms used?

In South Korea, two main public bioinformatics centers operate to save bioinformatics data, including genetic, transcriptomic, and metabolic data. The Rural Development Administration of Korea operates the National Agricultural Biotechnology Information Center (NABIC) to preserve agricultural bioinformatic data, such as biosequence, transcriptome, proteome, variation, and metabolome data, to create a database. The Korea Bioinformation Center (KOBIC) is a national interministerial center that comprehensively manages domestic bio-research data and provides an advanced data research utilization environment. KOBIC has prepared a standardized registration form for collecting various bioresearch data and has efficiently collected data scattered across ministries, businesses, and researchers. In addition, it plans to provide data storage space for each researcher so that they can conduct data-based research and build a virtual research environment that allows data sharing and collaboration among researchers. In addition, the National Bioresearch Resources Information Center is conducting various activities such as information linkage between ministries, signing MOUs for research cooperation between institutions, and bioinformatics education and research support.

Milk containing A2 protein is more digestible than milk containing A1 protein [21]. A2 milk was produced on farms where cows with A2 genotypes were identified through the analysis of cow genetic data, and only these were closed and raised. Kim et al. (2020) reported the large-scale production of African cattle genomes, shared data with research teams from Ethiopia, Sudan, Kenya, Sweden, and the United Kingdom, and discovered the process of African cattle adaptation to the environment [22].

Several consortia have been established globally to share genetic data within the livestock sector. Notable among these are the Functional Annotation of Animal Genomes (FAANG) and the Agricultural Genome-to-Phenome Initiative (AG2PI) [20]. These initiatives aim to enhance connectivity with the crop research community, promote the sharing of genetic data, conduct diverse research, and concurrently operate specialized educational programs (Figure 1) [20]. Other consortia of the genetic field in livestock are presented in Table 2.

Environmental data in the smart livestock industry

The establishment of an SLF has enabled the production of comprehensive environmental data, including real-time measurements of the farm environment. These data, which encompass climate data, smart farm data, and more, can be linked to smart livestock farms from the farm to the regional scale. Furthermore, SLFs integrate disease data affecting the livestock industry to establish appropriate quarantine measures or strategies for feed supply, breeding support, and animal welfare.

As previously explained for genetic data, AG2PI cooperates in livestock and crop research fields for common data feedback because there are methodological similarities in research. From the perspective of livestock farming, crops are applied to livestock feed and affect the phenotype of livestock; therefore, crop research fields interact with each other within the ecosystem [20]. The research methods and goals are consistent, such as increasing productivity and developing new varieties to respond to climate change by revealing the relationships between phenotypic, genetic, and environmental data. In addition, because it utilizes a public resource database rather than forming individual communities, it aims to create an integrated community and educate researchers and stakeholders in the agricultural and livestock fields about the flow of information, thereby increasing the utility value of the data. With the list of smart livestock technologies currently in use, we expect large-scale data production and utilization to be possible (Table 1).

An air-recirculated ventilation system based on ICT technology is another environmental system in the SLF. Existing livestock facilities have problems such as a high energy load, poor breeding environment owing to dense breeding, inflow and outflow of diseases into the air, and odors. These problems increase the likelihood of diseases on farms, leading to lower feed efficiency and daily increases in heat and wheat stress, as well as an increase in civil complaints and legal disputes. An ICT-based air circulation ventilation system minimizes the outflow or inflow of pathogens in the air, reduces odors, analyzes the complex environment inside the shed, and optimizes air circulation. The air circulation ventilation system measures air quality, temperature, and humidity inside the shed, stores the data in a farm information system, and operates a control system using ICT devices [23].

Expanding environmental data: Utilizing public data

Public data, such as climate, disease, market, and economic data, can be used as environmental data for the SLF system. In the past, only the individuals and surrounding environment within the farm were calculated as factors affecting the phenotype; however, with the development of science and technology, the scope of the environment that can be measured is gradually expanding owing to the invention of various means. Because public data include various elements that affect these environmental factors, it can be seen as a factor that can affect the phenotype of the SLF.

The Animal and Plant Quarantine Agency (QIA) provides information on domestic and international animal disease outbreaks. Information conjugated with SLF big data can help farms take proactive measures using the data to prevent or minimize damage from diseases. Weather or feed market information can help make economic forecasts for farm operations. Advanced SLF systems using big data, AI, and machine learning can utilize market information to predict the optimal dispatching time and scale, making farm production plans and profitability predictions more accurate. Thus, if public data can be applied as an environmental factor in the livestock industry, it is expected that the existing labor-intensive livestock industry will develop into a more advanced technology-intensive industry. However, there are still many hurdles to overcome before the commercialization of systems and technologies that can automatically link public data to smart livestock farming. Because no data center can automatically collect and link public data and no software can analyze them for SLF, these two are essential prerequisites for applying environmental factors in the SLF.

A vision of the SLF system via data recirculation

For recirculation of SLF big data

In an SLF system, it is essential to record the data monitored on farms and animals. The recorded real-time field data in the animal uses sensors (i.e., location, accelerometer data, and temperature data). These real-time data are shaped into big data formats for each SLF. In Europe, SLF systems were installed in five broiler houses and 10 pig houses as part of the EU-PLF project [24]. For 3 years, 90 fattening periods for pigs were monitored, which resulted in a total of 5,475 measuring days. This extensive monitoring generated over 120 terabytes of image data and 4,906,000 sound files, each lasting 5 min [24].

Phenotypic, genetic, and environmental data from SLF have been accumulated as independent big data. However, a comprehensive utilization system that integrates them to produce information with new value is still lacking. The potential of SLF big data is immense and can be doubled by combining and utilizing data from livestock farms and the government, public institutions, researchers, companies, and consumers. This realization should motivate stakeholders to actively participate in the integration and utilization of big data by the SLF. Therefore, recirculating big data in SLF means that farms, livestock product quality evaluation institutes, genetic information data centers, and government public data (e.g., weather and disease information) can be gathered for reprocessing and utilization according to the user's purpose (Figure 2).

The concept of recirculation of SLF big data has been suggested in the ‘Future Livestock Forum’ in Korea in 2021. The forum also suggests a data warehouse for recirculating SLF big data. A data warehouse is a more advanced concept than a traditional database and can be seen as more structured and systematic than a user-based, free-form data lake. This type of data warehouse differs from the databases listed in Table 2. Moreover, because SLF big data cannot exclude information that falls under the category of private property, the need for an integrated operating entity has been raised from the perspective that it is necessary to establish a governance system that can be operated by a public institution or consultative body for big data collection and management. Owing to the governance needs, the work related to the SLF big data integrated operation, previously managed by the ‘Korea Agency of Education, Promotion & Information Service in Food, Agriculture, Forestry & Fisheries (EPIS)’, has been transferred to the KAPE. This ensures that information on smart livestock from farms to animal product quality is gathered in one place. Farm environmental information and information on the objects collected from each farm for smart agriculture, including SLF data, are converted into a database and built into the ‘Smart Farm Data Mart’ DB of Smart Farm Korea. However, although individual information, farm information, etc., are collected in one place, the genetic information to be used for genetic improvement is collected separately. Therefore, it is necessary to realize complete SLF big data via a data center, forming a data warehouse. Looking at the list of smart livestock technologies (Table 1), we can see that much effort is being devoted to collecting and utilizing the collected data. Therefore, as smart livestock technologies are developed and distributed in the field, the effects of data feedback are expected to increase.

Digital twin: Future SLF systems via big data recirculation

A digital twin is a representative field application of data recirculation, a process in which data are continuously collected, analyzed, and fed back into a system [4, 25]. It is a digital replica of a real object kept up-to-date with continuous data inflow. Applying digital twins in the livestock industry is expected to be very useful for understanding and improving current complex systems and building new ones. However, digital twinning occurs at the toddler stage because of some challenges such as inadequate communication, data recirculation, and conflicts of interest among stakeholders (Figure 3) [4, 25, 26]. In Europe, some countries have attempted to construct digital twins based on SLFs. Wageningen University in the Netherlands is at the forefront of research on digital twins in the agricultural sector, particularly in livestock farming [16, 25]. The university is defining digital twins as applicable to livestock farming and continuously enhancing the connection between physical and virtual environments, keeping the industry informed [25]. In England, the University of Leeds constructed the National Pig Centre for digital twinning [4]. The center was developed to cooperate with engineering and computer sciences. The center set a goal for achieving net zero production by 2030 via multi-platform SLF technologies, including digital twinning [4].

Several conditions must be met before South Korea adopts digital twinning:

1. SLF, in its total sense, must be implemented for each livestock species. A policy encouraging extensive equipment investment is required to achieve a complete understanding of SLF.
2. It is necessary to recognize the publicity of the data produced by the SLF and cooperate to improve mutual understanding when sharing data. The data produced by each farm and institution often contain sensitive information. However, efforts are required to reassure people about the security of this information through robust legal mechanisms.
3. An SLF big data recirculation center is required. The SLF data produced by many farms are big data, and the information produced by each farm is often too large for the farm to process on its own. Therefore, it is necessary to build and operate a data center that can recirculate and store data for public purposes.

In addition, by combining public data and farm information operated by the state with AI and machine learning technologies and integrating SLF big data, the digital twin is expected to further increase the economic added value of the livestock industry by operating farms or implementing policies through simulations in advance.

Acknowledgments

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2022R1A2C4002510).

References

1. FAO. World livestock 2011-livestock in food security. Rome: FAO (Food and Agriculture Organization of the United Nations); 2011.

2. Agreement P, editor Paris agreement. report of the conference of the parties to the United Nations framework convention on climate change (21st session, 2015: Paris) Retrived December; 2015: HeinOnline.

3. Godfray HCJ, Aveyard P, Garnett T, Hall JW, Key TJ, Lorimer J, et al. Meat consumption, health, and the environment. Science. 2018;361(6399):eaam5324.https://doi.org/10.1126/science.aam532.

4. Collins L, Smith L. Smart agri-systems for the pig industry. Animal. 2022;16:100518.https://doi.org/10.1016/j.animal.2022.100518.

5. Falconer D. Introduction to quantitative genetics: Longman; 1996.

6. Hong K-C, Kim J-M, Kim K-S, Kim J, Do K-T, Do C-H, et al. Animal Molecular Breeding: Korea University Press; 2018.

7. John AJ, Clark CEF, Freeman MJ, Kerrisk KL, Garcia SC, Halachmi I. Review: Milking robot utilization, a successful precision livestock farming evolution. animal. 2016;10(9):1484-92.https://doi.org/10.1017/S1751731116000495.

8. Aerts J, Kolenda M, Piwczyński D, Sitkowska B, Önder H. Forecasting Milking Efficiency of Dairy Cows Milked in an Automatic Milking System Using the Decision Tree Technique. Animals. 2022;12(8):1040.https://doi.org/10.3390/ani12081040.

9. Lee JG, Lee SS, Alam M, Lee SM, Seong HS, Park MN, et al. Utilizing 3D Point Cloud Technology with Deep Learning for Automated Measurement and Analysis of Dairy Cows. Sensors (Basel). 2024;24(3).https://doi.org/10.3390/s24030987.

10. Lee SH, Kim JM. Breeding potential for pork belly to the novel economic trait. J Anim Sci Technol. 2023;65(1):1-15.https://doi.org/10.5187/jast.2022.e118.

11. Kim GT, Kang SJ, Yoon YG, Kim HS, Lee WY, Yoon SH. Introduction of automatic grading and classification machine and operation status in Korea. Food Science of Animal Resource and Industry. 2017;6(1):34-45

12. Font IFM, Gispert M. Comparison of different devices for predicting the lean meat percentage of pig carcasses. Meat Sci. 2009;83(3):443-6.https://doi.org/10.1016/j.meatsci.2009.06.018.

13. Lohumi S, Wakholi C, Baek JH, Do Kim B, Kang SJ, Kim HS, et al. Nondestructive estimation of lean meat yield of South Korean pig carcasses using machine vision technique. Korean journal for food science of animal resources. 2018;38(5):1109

14. Choi JS, Kwon KM, Lee YK, Joeng JU, Lee KO, Jin SK, et al. Application of AutoFom III equipment for prediction of primal and commercial cut weight of Korean pig carcasses. Anglais. 2018;31(10):1670-6.https://doi.org/10.5713/ajas.18.0240.

15. Daumas G, Causeur D. Tests d'homologation des appareils automatiques de classement des carcasses de porc. Journées Rech Porcine. 2008;40:91-2

16. Neethirajan S, Kemp B. Digital Phenotyping in Livestock Farming. Animals (Basel). 2021;11(7).https://doi.org/10.3390/ani11072009.

17. Sellier N, Guettier E, Staub C. A review of methods to measure animal body temperature in precision farming. American Journal of Agricultural Science and Technology. 2014;2(2):74-99.https://doi.org/10.7726/ajast.2014.1008.

18. Collins FS, Green ED, Guttmacher AE, Guyer MS, Institute USNHGR. A vision for the future of genomics research. Nature. 2003;422(6934):835-47.https://doi.org/10.1038/nature01626.

19. Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) 2023 [cited 2024 2024/10/13]. Available from: www.genome.gov/sequencingcostsdata.

20. Tuggle CK, Clarke J, Dekkers JCM, Ertl D, Lawrence-Dill CJ, Lyons E, et al. The Agricultural Genome to Phenome Initiative (AG2PI): creating a shared vision across crop and livestock research communities. Genome Biol. 2022;23(1):3.https://doi.org/10.1186/s13059-021-02570-1.

21. Ramakrishnan M, Eaton TK, Sermet OM, Savaiano DA. Milk containing A2 β-casein only, as a single meal, causes fewer symptoms of lactose intolerance than milk containing A1 and A2 β-caseins in subjects with lactose maldigestion and intolerance: A randomized, double-blind, crossover trial. Nutrients. 2020;12(12):3855.https://doi.org/10.3390/nu12123855.

22. Kim K, Kwon T, Dessie T, Yoo D, Mwai OA, Jang J, et al. The mosaic genome of indigenous African cattle as a unique genetic resource for African pastoralism. Nature Genetics. 2020;52(10):1099-110.https://doi.org/10.1038/s41588-020-0694-2.

23. Kim J-g, Lee I-b, Jeong D-y, Park S-j, Cho J-h, Kim R-w. Development and validation of an air recirculated ventilation system, Part 2: Evaluation of pig productivity in spring and summer seasons including examination of cooling methods. Biosystems Engineering. 2023;230:83-105.https://doi.org/10.1016/j.biosystemseng.2023.04.007.

24. Berckmans D. General introduction to precision livestock farming. Animal Frontiers. 2017;7(1):6-11.https://doi.org/10.2527/af.2017.0102.

25. Neethirajan S, Kemp B. Digital Twins in Livestock Farming. Animals (Basel). 2021;11(4).https://doi.org/10.3390/ani11041008.

26. Pylianidis C, Osinga S, Athanasiadis IN. Introducing digital twins to agriculture. Computers and Electronics in Agriculture. 2021;184:105942.https://doi.org/10.1016/j.compag.2020.105942.

Tables and Figures

Table 1. Step-by-step technologies, devices, programs, and systems for implementing smart livestock farming

Step-by-step		**Data collection and management**		**Optimal breeding management model**	
		Data type	**Data collection device(IoT)**	**Model development (program)**	**Model application (system)**
Genetics		① Breeding information (bloodline) ② Genomic information (SNP) ③ Gene information (STR) ④ Artificial insemination information	① Genomic analysis equipment ② Gene analysis equipment	① Genomic analysis program ② Gene analysis program	① Genomic data utilization platform ② Gene data utilization platform
Breeding Management	**Environment**	⑤ (Inside) Temperature and humidity ⑥ (Gas) CO2, light intensity, ammonia ⑦ (Outside) Temperature and humidity, wind direction, wind speed, rain, solar radiation	③ (Internal) Temperature and humidity sensor ④ (External) Wind speed, wind direction, outdoor weather station ⑤ (Safety) Lightning protection device, fire detector, power outage detector ⑥ (Welfare) Heat stress measurement system	③ Complex environmental information management program (facilities, sensors, weather, satellites) ④ Information and communication and facility control management program	③ Composite environment automatic control system (fan, curtain, sprayer) ④ LED lighting automatic control system ⑤ Air conditioning and heating control system ⑥ Uninterruptible power supply management system

	Feeding (Fattenuation)	⑧ Water supply information (water supply amount, water temperature, water quality) ⑨ Feed information (type, feeding amount, pattern) ⑩ Biological information (history, body temperature, weight, heart rate) ⑪ Eco-friendly livestock product certification information	⑦ Individual identification device ⑧ Feeding amount, drinking amount, weight measurement device ⑨ Feed bin management device ⑩ Video and audio recording device (CCTV, 3D camera) ⑪ Carbon emission measurement device	⑤ Individual information-based feed supply program ⑥ Data-based quality grade/carbon emissions management program ⑦ Production management program	⑦ Feed automation system (mixing amount, feeding amount, drinking amount, additives, disinfection water washing) ⑧ Robot for labor-intensive process (cleaning, milking, lactation) ⑨ Integrated livestock breeding management platform
	Manure (Odor)	⑫ Livestock manure emissions ⑬ Odorous gases (ammonia, hydrogen sulfide, CO_2)	⑫ Livestock manure measurement device ⑬ Odor measurement device	⑧ Livestock manure and odor confirmation management program ⑨ Weather and weather information linkage odor management simulation	⑩ Odor reduction system (biofiltering, fog spraying) ⑪ Livestock manure treatment system (solid-liquid separator, aeration device, livestock manure fermentation drying, agitator, automatic cleaning, purification discharge, biogas)
	Quarantine (Disease)	⑭ Vehicle entry/route information	⑭ GPS device for livestock-related vehicles	⑩ Unmanned livestock quarantine management program	⑫ Automatic preemptive blocking quarantine system

		⑮ Quarantine/disease information	⑮ Blocking quarantine entry and exit devices (barriers, vehicle number recognition devices) ⑯ Vehicle and visitor disinfection facilities	⑪ Individual disease management program (early detection/immediate action) ⑫ Livestock disease management program linked to history system	⑬ Smart disease diagnosis and management system
	Energy Management	⑯ Energy consumption (electricity, gas) ⑰ Energy production (electricity, gas)	⑰ Energy consumption measuring device ⑱ Energy production/storage measuring device	⑬ Energy consumption optimization and self-sufficiency program	⑭ Electricity generation and control (solar, geothermal, wind power) ⑮ Livestock manure energy conversion system
	Reproduction	⑲ Real-time location information ⑳ Breeding management information (estrus/parturition)	⑲ Biological information collection device (history, time, location, body temperature, weight, voice, image, infrared) ⑳ Estrus detection, pregnancy diagnosis device (ultrasound) ㉑ Smart incubator	⑭ Individual reproductive ability management program ⑮ Livestock estrus management program (reduction of empty days) ⑯ Livestock parturition management program (parturition detection and environmental management)	⑯ Livestock estrus artificial intelligence platform ⑰ Parturition management artificial intelligence platform
Quality (Slaughter)		㉑ Carcass information ㉒ Grade information (meat	㉒ History information management device		

	quantity grade, meat quality grade)			
Distribution	㉓ Production price ㉔ Distribution price	㉓ Electronic scales and packaging devices		
Consumption	㉕Consumption price			

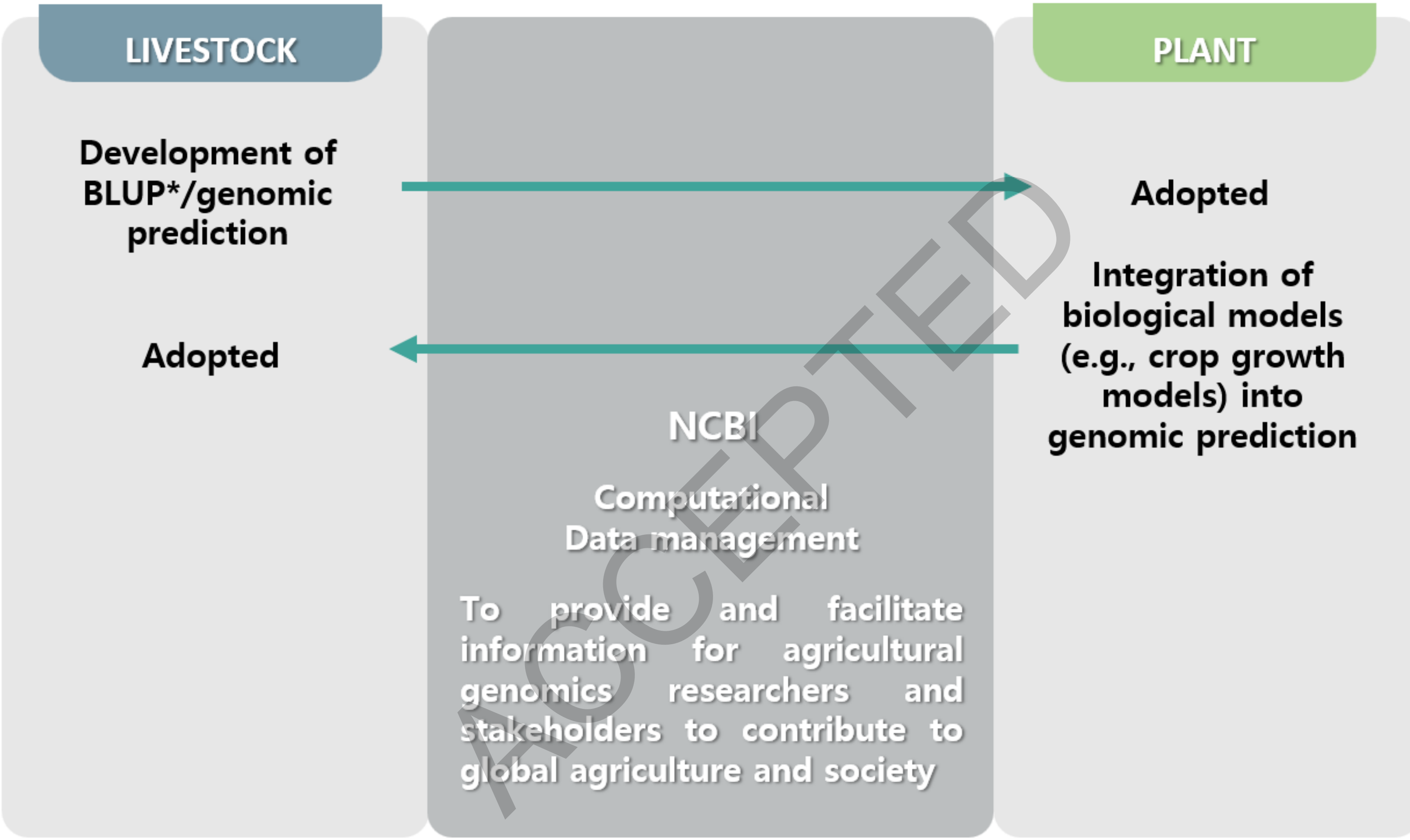
Table 2. Consortia using genetic data with livestock

Consortium	Participations	Contents
FAANG	Various nations	Construction of reference genomes of livestock animals
1000 Plant & animal project	Beijing Genomics Institute (China)	Genomic analysis of 1,000 species of animals and plants
G10K project	Over 50 institutes	Build reference genomes of 10,000 vertebrates
SGSC	USA (USDA), Netherlands, France, UK, Korea Rep., China, Japan etc.	Pig genome sequencing
ABC	Korea Rep., Japan, China	Sharing bioinformatic data

Table 3. Characteristics between data warehouse and database

Characteristics	Data Warehouse	Database
Suitable workload	Analysis, reporting, big data	Transaction Processing
Raw data	Normalized data from various sources	Data captured as is in a single source, such as a transaction system
Data Capture	Bulk writes operations, usually based on a pre-determined bulk batch schedule	Optimized for continuous write operations with new data available to maximize transaction throughput
Data Normalization	Denormalized schemas, such as star schemas or snowflake schemas.	Highly normalized static schema
Data storage	Optimized for simple access and fast query performance using columnar storage	Optimized for high throughput write operations to single-row oriented physical blocks.
Data access	Optimized to minimize I/O and maximize data throughput	Large number of small read operations

Figure legends



*BLUP: Best Linear Unbiased Prediction

Figure 1. Identifying similarities between livestock and plant genomic communities by the agricultural genome to phenome initiative (AG2PI) consortium.

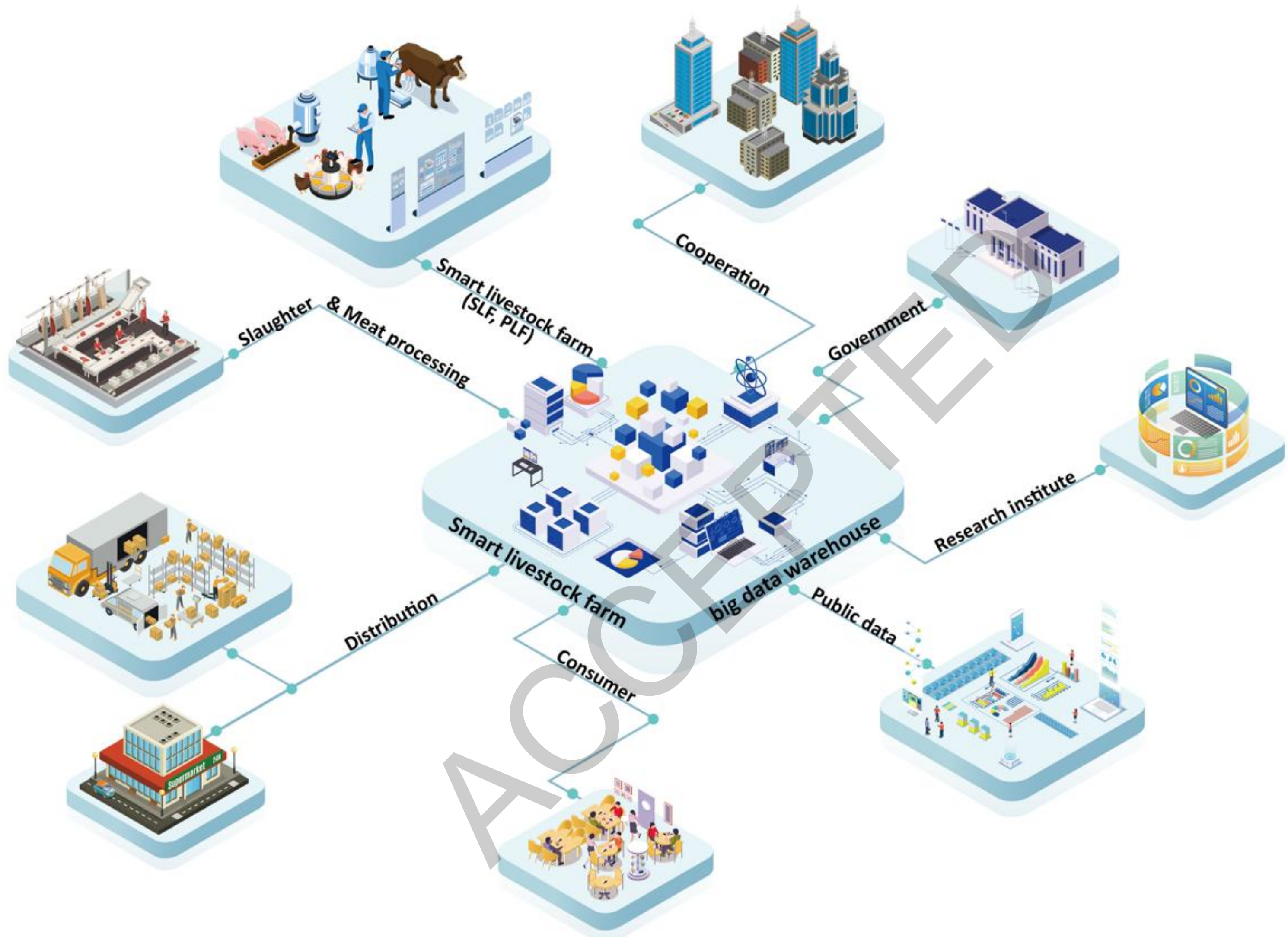


Figure 2. Concept of data recirculation from Smart livestock farm (SLF) big data.

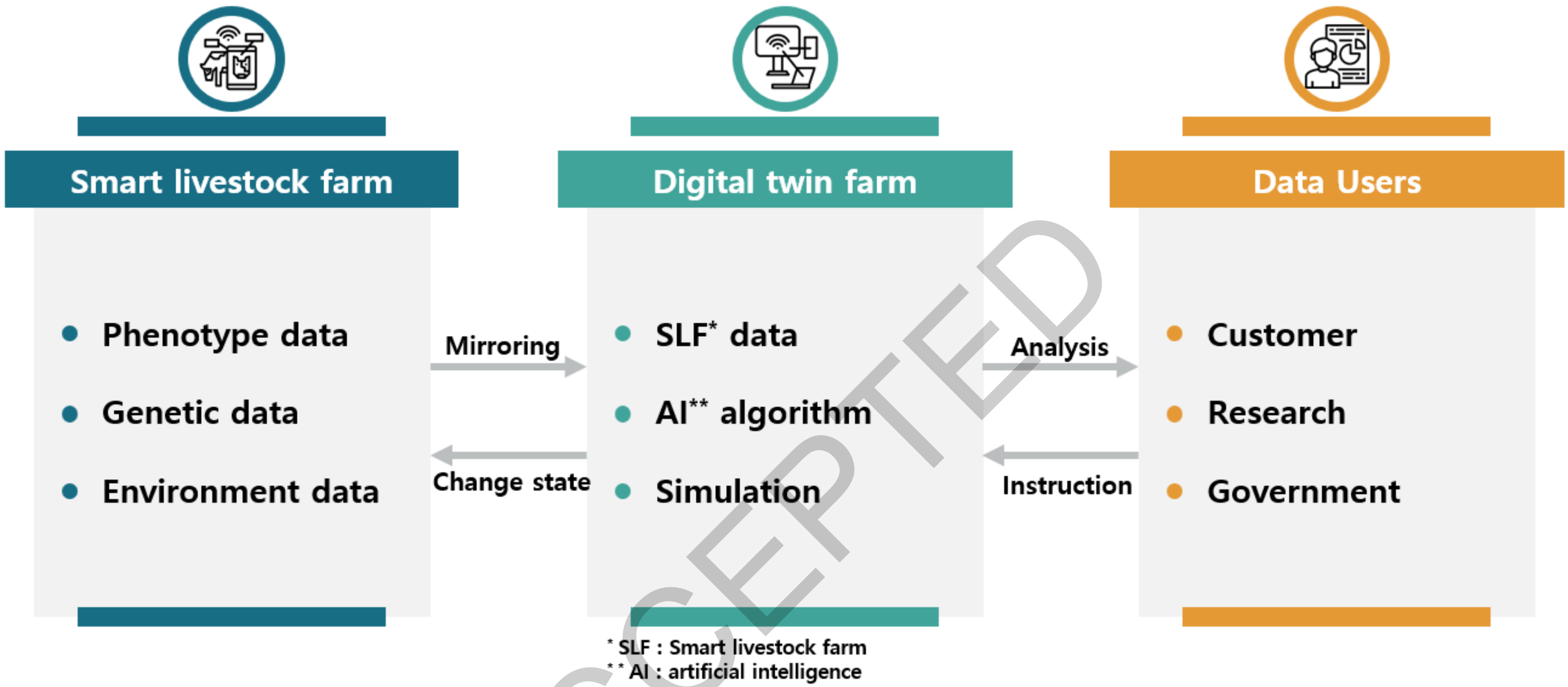


Figure 3. Definition of digital twins in smart livestock farm (SLF) using data recirculation.