MyData Personal Data Store Model(PDS) to Enhance Information Security for Guarantee the Self-determination rights

Seong-hyun Min[1], and Kyung-ho Son[2*]
[1] Department of Computer Science, Kangwon National University
Republic of Korea
[e-mail: tjdgus1277@gmail.com]
[2] Division of Liberal Studies, Kangwon National University
Republic of Korea
[e-mail: khson@kangwon.ac.kr]
[*]Corresponding author: Kyung-ho Son

Received December 7, 2021; accepted January 12, 2022;
published February 28, 2022

Abstract

The European Union recently established the General Data Protection Regulation (GDPR) for secure data use and personal information protection. Inspired by this, South Korea revised their Personal Information Protection Act, the Act on Promotion of Information and Communications Network Utilization and Information Protection, and the Credit Information Use and Protection Act, collectively known as the "Three Data Bills," which prescribe safe personal information use based on pseudonymous data processing. Based on these bills, the personal data store (PDS) has received attention because it utilizes the MyData service, which actively manages and controls personal information based on the approval of individuals, and it practically ensures their rights to informational self-determination. Various types of PDS models have been developed by several countries (e.g., the US, Europe, and Japan) and global platform firms. The South Korean government has now initiated MyData service projects for personal information use in the financial field, focusing on personal credit information management. There is also a need to verify the efficacy of this service in diverse fields (e.g., medical). However, despite the increased attention, existing MyData models and frameworks do not satisfy security requirements of ensured traceability, transparency, and distributed authentication for personal information use. This study analyzes primary PDS models and compares them to an internationally standardized framework for personal information security with guidelines on MyData so that a proper PDS model can be proposed for South Korea.

Keywords: DataSecOps, MyData, Personal Data Self-determination, Personal Data Store.

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1G1A1007540). This study was supported by a 2019 Research Grant from Kangwon National University.

1. Introduction

With the advent of the Fourth Industrial Revolution, the world has entered a new era based on artificial intelligence (AI) where the internet of things (IoT) connects household and business devices, 5G and 6G technology provides high-speed communications, and cloud computing facilitates big data processing. Data, the driving force of the Revolution, are at the center of these changes.

There are various types and forms of data, such as weather data, which require little human involvement to gather and interpret, including various measurement and sensor data, which are collected and transmitted by different means. Numerous types of personal data are also generated by the social and economic activities of human beings. Among these, so-called high-quality personal information include medical, location, credit, and purchase data. Personal information refers to both personally identifiable information, such as names and addresses, and information obtained through pseudonymous processing. When a specific individual can be inferred based on pseudonymous personal information combined with other types of data, such information can be regarded as personally identifiable information (PII). As such data are actively applied in research and industrial fields, the “Three Data Bills” (i.e., the Personal Information Protection Act, the Act on Promotion of Information and Communications Network Utilization and Information Protection, Etc., and the Credit Information Use and Protection Act) were revised and implemented on August 5th, 2020, to improve data use regulations in South Korea.

Multiple methods and technologies for pseudonymous processing and the use of personal information have also been developed to reduce the risk of re-identification. However, these solutions cannot effectively perform pseudonymous processing and the protection of personal information, owing to several problems, such as the difficulty in determining the level of processing to be applied for de-identification, and reputation risks that can be affected by data use, ownership of personal information, data monopolies, and distributions of profits created by data use.

MyData is a novel approach to personal information management where data subjects directly download and use their information or are provided with rights of data portability, including controlling the supply of their data to third parties. This approach converts existing controller-based data application systems to those based on data subjects. See **Fig. 1**.

This approach was designed to establish a platform that provides tools to enable users to select certain data to be closed or open, allowing users to maintain digital sovereignty and to ultimately retain the public profits and social improvements based on the data use.

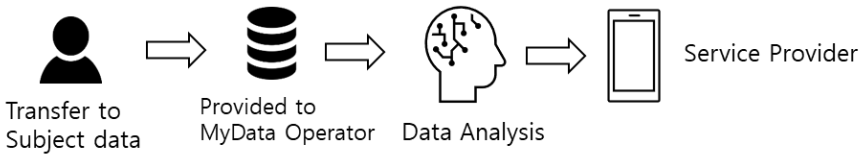

Fig. 1. MyData Overview

This study compares and analyzes various PDS structures applied to MyData models to propose an enhanced PDS model proposal that may ensure the rights of individuals to informational self-determination based on their approval and their rights to check, remove, and control their personal information. This study also examines personal data frameworks established in South Korea and other countries, authentication regulations of information banks in Japan, and MyData guidelines in South Korea to present authentication and security requirements to further ensure rights to informational self-determination.

2. Definition and Models of MyData (PDS)

MyData Global directs that MyData models utilize personal information on data subjects and that the following principles in **Table 1** should be satisfied to facilitate the participation of data subjects in the collection, use, and application of their personal information [3]:

Table 1. MyData Principles

Transparency	Individuals should be notified of which types of their personal information are collected, which methods are used to collect them, what personal information is used, and who accesses the information.
Reliability	A technical and systematic architecture should be established to guarantee that relevant services and systems are designed to protect personal information and that service providers do not utilize the information for malicious purposes.
Right to control	Measures should be prepared to enable individuals to effectively manage the range of methods to share their personal information (data) and the rights afforded to others to access and use such information.
Value	All interested parties should be able to share the value created by personal information use. Firms and individuals should be provided with clear and visible incentives thereof.

PDSs have received attention as models that utilize MyData to fulfill the aforementioned principles. PDS models allow individuals to directly manage, share, and utilize their personal information; furthermore, the models return control of this information to the owners, as distributed by various firms [1]. **Fig. 2** presents the types of PDSs according to their operations [4]. Within the centralized PDS environment, a service provider stores and manages data on a server. Within a decentralized PDS environment, users store and manage data on their own devices. Both enable users to mark data with access rights.

Information trust banks, introduced in Japan, are examples of data transaction models that apply PDS concepts. **Fig. 3** shows that Information banks make contracts with users regarding the use and management of their data within PDS systems and provide data to third parties based on valid user authorization [4].

It is essential for information banks operated as PDS systems to control users' personal data to ensure transparency, reliability, and rights of control. To control the personal data of users, information banks should meet the following requirements [2]. First, Information banks should be able to manage, renew, and adjust individuals' personal data securely and protect their data based on different user settings. Second, Information banks should be able to select

various types of public control, such as the selection of reliable users. Third, Information banks should allow users to check the details of their data use agreements. Fourth, Information banks should allow users to transfer their personal data to other PDS systems and information banks as personally designated.

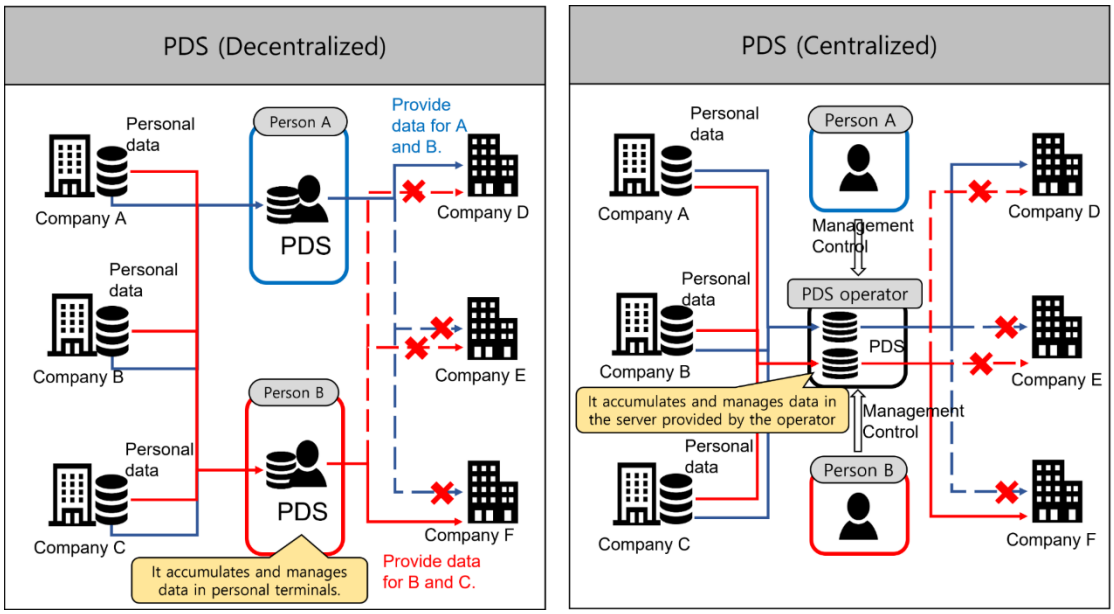

Fig. 2. Types of PDSs [4]

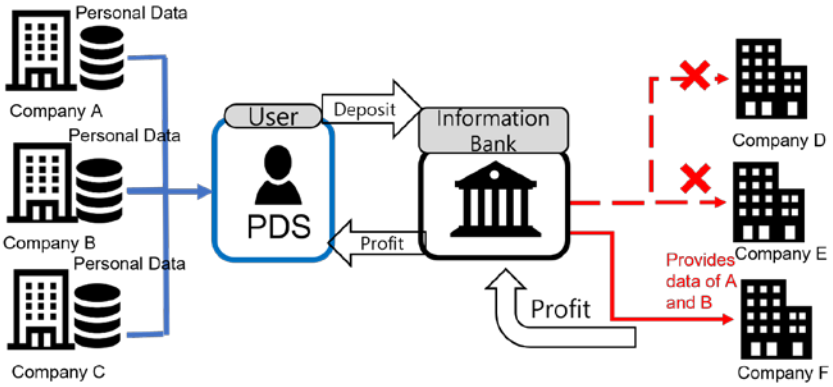

Fig. 3. Information Bank Overview [4]

Information banks should provide services based on user consent and allow third parties to utilize personal data based on user decisions regarding the states of collection, control, and use, while providing traceability, confirmation, and approval. Details of these service services are as follows [5].

3. Analysis of MyData/PDS Models

MyData Global has awarded 16 business operators from the MyData2020 eight of which are currently operating PDS models based on OSS(Open Source Software). Of those, this chapter analyzes Personium, MyDataShare, and OwnYourData so that an appropriate PDS architecture can be proposed for the circumstances in South Korea.

3.1 Personium

Personium is a Japanese open-source PDS server that manages data based on individual preferences. This decentralized service supports data distribution and applications to manage data value [6]. It can read and write data from various applications and devices and provides a backend-as-a-service (BaaS), which can host application data ranging from electronic medical records to games. Personium can establish a PDS server for individuals, firms, and governments. The RESTful application programming interface (API) provides flexible user definitions and compatibility. Its platforms are easy to use by PDS service operators and end users. Furthermore, it leverages uniform resource locator authority to apply access regulations and user authority for storage and control.

Fig. 4 shows that Personium defines basic three-layer objects, within which a unit is a server that hosts multiple cells, a cell is a data store for a data subject, and a box is an application data store.

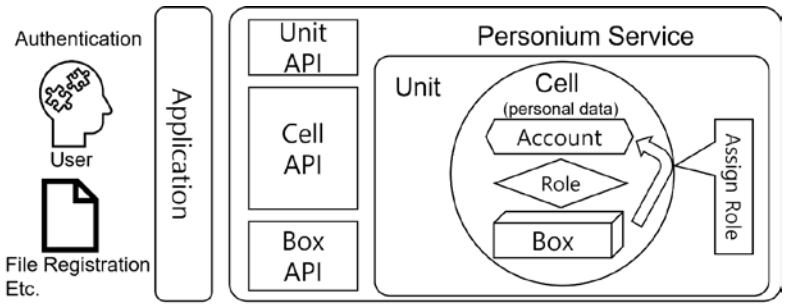

Fig. 4. Personium Service [6]

Table 2 shows that Personium service administrators manage cells based on units that provide a system infrastructure with a unique fully qualified domain name and provide services by connecting different units. Administrators also generate relationships between units according to unique dispersion structures, applying these relationships to provide authority for users and to generate several cells per unit. PDS users can utilize individual cells for authentication and approval, access control, management of boxes serving as application data stores, and event processing. Each cell is independently operated like tenants in a multitenancy model. A directory, a file object, and an OData service can be stored in a box. Users can manage personal data by granting authority for access to cells with open personal data, including writing and reading access according to regulations and accounts.

Table 2. Detailed Description of Functions: Personal Data Management [6]

Name and containment relationship	Outline	Number that can be created	Operating privileges		
Unit	The unit for managing cells	1 unit per service environment	Unit user token	-	-
Cell	The unit for managing users and user data	1 Cell per user		Cell-level permission note: only a unit user token can create a cell	-
Account	The account of the user belonging to the cell	Multiple			-
Role	Defines the user's role and privileges	Multiple			-
Box	The User's data storage area	Multiple			Box-level permission-note: must have cell-level permission or higher to create a box

Users can obtain the rights for public and private control of personal data by defining their relationships with specific access authority as parameters. Hence, they can select personal data to be revealed to the public and to maintain secrets between users.

The three practical uses of Personium are as follows. First, Personium's calendar service enables individuals to efficiently combine and manage several schedule management resources. This service leverages an API to search data, store them in a PDS, and give access authority to other users. Second, Timefiller can be combined with the calendar service introduced in the first case to provide access to personal data and to provide individuals with tasks. Third, Personium Trails allows public medical institutes to access location records of individuals, such as COVID-19 patients.

3.2 MyDataShare

MyDataShare is a software-as-a-service platform developed by Vastuu in Finland to facilitate digital identification (ID) management between individuals and digital services, personal data sharing, and relevant authority processing. Core functions of MyDataShare include identification, which combines MyDataShare services with those of external ID providers, consent managers, and connection services [7]. MyDataShare Wallet can be connected to existing services that provide MyDataShare functions in terms of user interface, context, and branding based on user consent.

The ID function manages MyDataShare identifiers and authenticators, consent-management activating functions, logs, active ID connections, and e-profiles. In the ID system, a user account serves as a master identifier for connected services and maps hosted by service users. The consent management function provides control and inspection services on data systems for lifecycle management, relevant problem solving, and data persistence. Data transfer and API authentication processes depend on the verification of individual consent conditions to

respond to data requests. To fulfill valid consent regulations, users are required to provide all necessary information, including identity and processing purposes, for explicit decision making. The connectivity functions on MyDataShare manage ID providers, various data connection services required for applying data access authority, and a consent management interface. This function manages registration and user IDs by combining ID providers' OpenID services. It also offers user authentication and single-sign-on services summoned for personal information management.

Fig. 6 shows that the platform architecture of MyDataShare consists of a three-layer back end. Each layer contains ID, consent, and connectivity. The ID layer combines services of ID providers into a single function and generates key chains according to users, peers, and pseudonyms. The consent layer manages requests, conditions, storage, and logs related to consent, and the entire consent layer connects user IDs. The connectivity layer provides an access gateway function that converts existing data service operations using an API for connecting the back end via MyData principles. Data consumers can access user data during transactions via consent tokens, which hide individual ID. Users can gain consent with a personal wallet function, which can be applied using applications on the MyDataShare ID layer. The connectivity layer is designed to provide data portability. A function for data export based on personal wallet service is currently being tested.

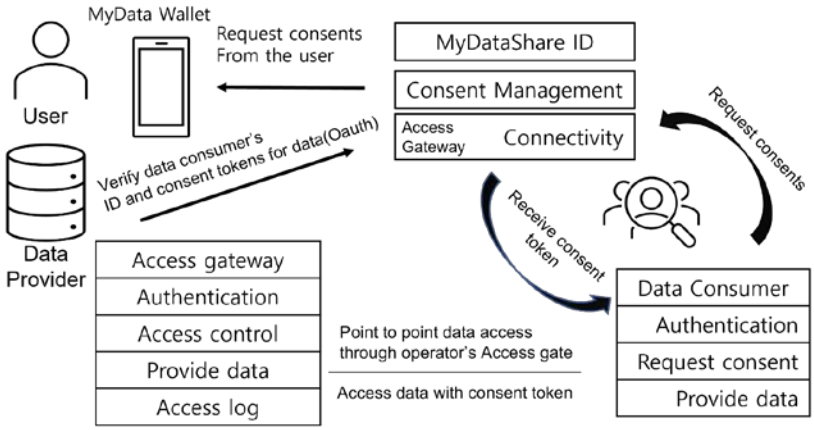

Fig. 6. MyDataShare Architecture [7]

3.3 OwnYourData

OwnYourData is a non-profit organization that grants full open-source licensure in accordance with MyData principles and provides several services and products for free online. They offer insight reports based on stored personal data via their data vault service [8]. The data vault is encrypted end-to-end, and keys are separately stored. Data access is allowed via authentication through a sandbox to prevent leakage to external paths. The MyData Weekly Digest service provides the latest information on MyData services, and the Online Notary service offers safe storage for personal information and documents. The Semantic Container service enables multiple users to exchange data in a safe and traceable environment. This service is provided as a light open-source infrastructure. When users use this service, they can

efficiently access data based on their right-to-control profit model according to the intended use. Three types of input and output levels are available. The Metadata Level provides explanations of data uses allowed for container users and information charges. The Syntax Level ensures data-type compatibility and performs validity tests. The Semantic Level includes semantic explanations of data properties, metadata, and syntax information to ensure data quality.

Fig. 7 presents the semantic container lifecycle. First, the base container is established based on the data definition and policy. Second, a data validity test is conducted to create an audit trail. Data from the test results are transferred to the container based on the API endpoint. Data processing is carried out to convert the base container to one having a specific function, and details of authorized data use and updated audit trails are confirmed. Finally, data sharing is performed, and full access is documented.

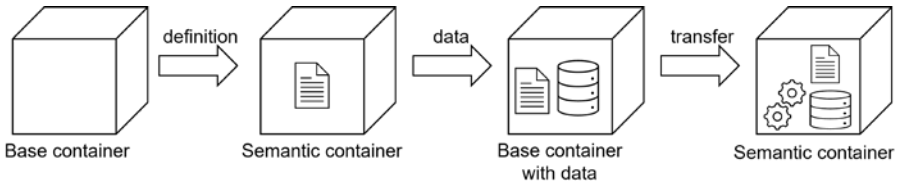

Fig. 7. Semantic Container Lifecycle [9]

Researchers have developed a diabetes processing database using the semantic containers of the OwnYourData model to ensure data portability as supported by the Horizon 2020 Research and Innovation Program, initiated by the EU [9]. **Fig. 8** shows the flow of personal data processed by this service. The service supports data uploading, data encryption, data authentication, data visualization, and insightful analysis, and it supports various standard tools [10].

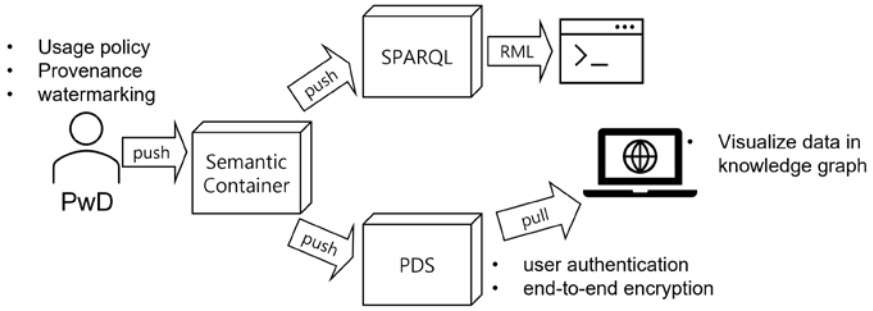

Fig. 8. Example of Semantic Container Use [10]

4. Analysis of Requirements for MyData Authentication and Security

This chapter analyzes security requirements based on international standards, the Information Bank Establishment Committee in Japan, the Financial Services Commission (FSC), and the Financial Security Institute (FSI) of South Korea for safe MyData application services.

4.1 ISO/IEC 29100 – Privacy Framework

ISO/IEC 29100 (Information technology – Security techniques – Privacy framework) protects personally identifiable information (PII) in an information communication technology system [11]. PII can be used to identify an individual and can be derived directly or indirectly from data pertaining to the individual. Service operators are required to consider all methods necessary to identify a person. The ISO/IEC 29100 framework covers 11 personal data protection principles as follows: consent and selection; legality and characteristics of purpose; limitations in collection; data minimization; limitations in use; possession and release; accuracy and quality; openness, transparency, and announcement; personal participation and access; responsibility, data security; and compliance with personal information protection.

4.2 Guidelines on Authentication at Information Banks in Japan

Japan has encouraged the use of private data in cooperation with various entities since it announced its framework act on the promotion of private data use in 2016. Accordingly, the Data Distribution Environment Maintenance Review Committee and the Working Group on Data Utilization in the era of AI and IoT, established in February 2017, conduct ongoing research on revitalizing data transaction markets of PDS-based information banks for smooth distribution of personal data. In 2017, the Ministry of Internal Affairs and Communications and the Ministry of Economy, Trade, and Industry in Japan developed information bank models and prepared standards on authentication and distribution and use of personal data via authentication systems. Subsequently, these ministries announced guidelines on certification of data trust functions in June 2018 [12]. These guidelines describe a range of information banks that are subject to screening for certification, a system of governance for operation, processes of applying for establishment and certification, and standards and terms and conditions on certification.

Chapter 5 in these guidelines states that cyber security and personal data protection standards and solutions should be implemented in information banks for personal data protection under the assumption that personal data providers have established a sufficient system for data security risk management and have prepared data protection measures in reference to international standards, such as the ISO/IEC 29100 frameworks, regardless of the increase in the number of providers.

Standards on cyber security to be implemented in information banks are as follows:

- Establishment of data security management
- Operation, observation, and review of data security management
- Maintenance and improvement of data security management
- Formulation of data security policies
- Data security organizations
- Human resources' data security
- Resources management
- Data security related to technology, physical conditions, environments, operations, and communications

- System acquisition, development, and maintenance
- Supply network relationships
- Data security accident management
- Business continuity management in terms of data security

Standards on personal data to be implemented in information banks are as follows:

- Establishment of basic policies
- Implementation of safety management measures in terms of organizations, human resources, physical conditions, and technology
- Consent and selection systems
- Legitimacy and clarity of purposes of use
- Limitations in and minimization of personal data collection
- Limitations in use, conservation, and release
- Accuracy and quality
- Release, transparency, and announcement
- Access to their own data of individuals

Among the standards of screening for certification of information banks, Check Sheet (v2.0) describes standards of certification related to cyber security and personal data protection, documents to be submitted for certification, requirements indicated in policies, and guidelines on certification in detail[19].

4.3 Guidelines on MyData in the Financial Field in South Korea

The FSC proposed guidelines on MyData services as established by the Korea Credit Information Services and on MyData technology and security as established by the FSI in February 2021 to allow MyData-related firms to implement services[13]. Guidelines on MyData services provide operational measures, responsibilities of service providers, data providers, and intermediary service providers, authentication, and procedures for requesting transfer of personal credit data in accordance with the Credit Information Use and Protection Act. **Table 3** summarizes the FSC's My Data guidelines.

These guidelines also describe processes for self-authentication, authentication provided by data providers, and integrated authentication provided by institutions in the case of personal credit data transfer. Self-authentication processes require the satisfaction of fundamental principles related to purposes, tools, methods, and processes. Regarding individual authentication processes, customers use authentication protocols provided by data providers. Regarding integrated authentication processes, customers use authentication tools provided by authentication institutions to access multiple data providers using one-time authentication. The former processes differ from the latter.

Furthermore, the FSC presented requirements on MyData security for data providers and recipients who possess and collect personal credit data. The interested parties must comply with the Credit Information Use and Protection Act and clauses related to data protection in the Regulation on Supervision of Credit Information Business. In terms of network

segmentation standards and the use of cloud computing services, they should follow the Regulation on Supervision of Electronic Financial Activities. The Credit Information Use and Protection Act is prioritized for personal credit information, whereas the Personal Information Protection Act is prioritized for personal data. To fulfill the requirements of managerial security, personal data providers should designate people responsible for credit information management and protection for inspection and education. They should also describe personal credit information management and methods of collecting, storing, conserving, and removing such information. Furthermore, internal and external access authority and records should be managed and controlled. As MyData services are implemented based on data flow between organizations, verification of certification is required via access tokens, APIs, and cloud services. Data providers should also prepare regulations on personal credit information leakage and countermeasures against disasters to protect MyData service users. To satisfy the requirements of physical security for MyData services, an access control system for facilities and devices should be developed. Furthermore, a management system for passwords, systems, and development should be established to meet the requirements of technical security.

Table 3. The FSC's Guideline

Management Security	Credit Data Management	Design the credit data manager
		Main work of credit data manager
		Check management and protection of the personal credit data
	Educate the Personal Credit data	Educate the personal credit data
	Manage Personal Credit data	Disclosure of credit information utilization system
		Collect personal credit data
		Retention of personal credit data processing records
		Store personal credit data
		Delete personal credit data
	Access control of personal credit data processing system.	Internal access right management
		External access right management
		external access security
		Access record management
		Confidentiality Pledge
	Job Separation	Establishment of criteria for job separation.
		Preparation of supplementary measures for job separation.
	API system Management	Access Token Management
		API system protection
		Anormal API detection
		Cloud Access
	User Protection	Personal Credit data Leakage
	Disaster Preparedness	Backup and recovery system operation
Physical Security	Access Control	Separation of computer equipment
		Access control of auxiliary storage media
		outsider access control
	Physical Security	Establishment of physical security facilities
		document archiving

Technical Security	Password Management	Password Management
		Password encryption
	Password Control	Personal Credit data encryption
		Network encryption
		encryption key management
	System Security	Network Separation
		IDS/IPS Management
		Vaccine Management
	Development Security	Design the Security
		Vulnerability check
	Print and Copy Management	Establish the internal protection system
		Minimize print
		Print/copy management
		External transfer pre-approval

5. Preparation of Measures for Ensuring the Right of Data Subjects to Informational Self-Determination

EU adopted the GDPR in the Personal Information Protection Act in 2016 to provide data subjects with the right to control their personal information. This regulation has been applied since 2018[14]. This chapter analyzes issues of consent for MyData and rights of access, transfer, and erasure related to data ownership mentioned in the GDPR and established by EU.

5.1 Regulations on Consent to Personal Information in the GDPR Established by the EU

Regulations on consent for ensuring the right to informational self-determination in the GDPR established by the EU specify details of the legality of personal data processing, consent conditions, and personal data collected from data subjects.

According to the GDPR, the legality of personal data processing can be classified into six types as follows:

- Data subjects give consent to the processing of their personal data for one or more specific purposes.
- Processing is necessary for the performance of a contract to which the data subject is party or to take steps at the request of the data subject prior to entering a contract.
- Processing is necessary for compliance with a legal obligation to which the controller is subject.
- Processing is necessary to protect the vital interests of the data subject.
- Processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.
- Processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the

interests or fundamental rights and freedoms of the data subject, which require protection of personal data, in particular where the data subject is a child.

According to Item 11 of Article 4 in the GDPR, the consent of the data subject includes any freely given, specific, informed, and unambiguous indication of the data subject's wishes by which a statement or clear affirmative action signifies agreement to the processing of personal data. Item 32 of Article 4 in the GDPR states that consent should be given by a clear affirmative act establishing a freely given, specific, informed, and unambiguous indication of the data subject's agreement to the processing of personal data relating to themselves. It is regarded that data subjects do not consent to the processing of their personal data when they do not clearly express consent or when consent is automatically checked in advance. Moreover, when consent is gained for multiple purposes, each purpose should be satisfied in the process of obtaining consent.

Because MyData services are conducted based on the consent of data subjects, Item 42 in the GDPR should be followed. According to this item, the controller should demonstrate that the data subject has given consent to the processing operation. A declaration of consent pre-formulated by the controller should be provided in an intelligible and easily accessible form using clear and plain language, and it should not contain unfair terms.

Article 13 of Section 2 of the GDPR defines the information that should be provided to a data subject when personal data relating to the data subject are collected. Such information includes the identity and the contact details of the controller and, where applicable, of the controller's representative; the contact details of the data protection officer, where applicable; the purposes of the processing for which the personal data are intended as well as the legal basis for the processing; the legitimate interests pursued by the controller or by a third party; and a range of personal information to be obtained. This article also defines additional information that should be provided to the data subject to guarantee reliable data processing in the process of collecting personal data. Such information includes a period for personal data storage and relevant standards, the right to lodge a complaint, and the right to withdraw consent without affecting the legality of data processing.

Thus, MyData/PDS models should adopt consent methods that enable individuals to consent to provide their personal data by following policies of data provision presented in advance or by analyzing information previously provided by a third-party service provider that is willing to utilize personal data for purposes of data use, data to be used, and conditions for data use.

5.2 Rights to Access, Transfer, and Erase Personal Data Defined in the GDPR by the EU

The GDPR, as established by the EU, provides data subjects with rights of information, access to, rectification, erasure, and portability of personal data to guarantee a consent system and the right to informational self-determination. Item 60 states that the data subject should be informed of the existence of the processing operation and its purposes. The controller should provide the data subject with any further information necessary to ensure fair and transparent processing, accounting for the specific circumstances and context in which the personal data are processed. Furthermore, the data subject should be informed about personal data

processing when data are collected. Item 63 defines the right to access personal data, indicating that the data provider should have access to the results, including the purpose of personal data processing, the period for such processing, and the recipient.

The GDPR provides the data subject with the right to erasure, the right to be forgotten, and the right to rectification. Item 65 states that the data subject can erase or rectify personal data by withdrawing consent when personal data are not processed in relation to the purposes for which they are collected. Item 66 specifies that the right to erasure should also be extended in such a way that a controller who has made the personal data public should be obliged to inform the controllers, which are processing such personal data, to erase any links to, copies of, or replications of personal data to strengthen the right to be forgotten in the online environment. Article 16 of Section 3 states that the data subject shall have the right to obtain from the controller without undue delay the rectification of inaccurate personal data. Article 17 defines the right of the data subject to obtain from the controller the erasure of personal data without undue delay. The controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies: the personal data are no longer necessary in relation to the purposes for which they were collected, the data subject withdraws consent on which the processing is based, and the personal data have been unlawfully collected processed.

Item 68 states that the data subject should be allowed to receive personal data that have been provided to a controller in a standardized format to guarantee the right of data portability, which is regarded as an element for ensuring the right to informational self-determination, the foundation for MyData/PDS services. Article 20 of Section 3 describes details of the implementation of the right to data portability. Article 21 of Section 4 indicates that data subjects shall have the right to object on grounds relating to their situation at any time to the processing of their personal data. When the subjects exercise their rights to object, the data subjects should be informed about relevant details.

Based on the analytic results of the GDPR, it can be concluded that MyData/PDS models should possess a function for identifying which data are shared with whom and a function for implementing decentralized autonomous authentication, approval, and access control to allow the aforementioned functions.

6. Proposal of the MyData Platform for Ensuring the Right to Informational Self-determination

Section 3 of this article defined the PDS and analyzed its models, architectures, and open-source models developed in various countries to establish a safe personal data distribution platform. Sections 4 and 5 examined requirements of authentication and security for MyData and regulations regarding ensuring the right to informational self-determination as stated in the GDPR established by the EU.

This section proposes a PDS model for establishing a safe personal data distribution platform based on the analytical results indicated in previous sections. It also presents technical requirements for data processing and security, which are appropriate for MyData PDS models,

based on ensured data integrity; they can satisfy dynamic consent and the rights to access, portability, and erasure to ensure the right to informational self-determination.

6.1 Proposal of the MyData Platform and the PDS Model

This section proposes a MyData PDS model for establishing a safe personal data distribution platform. To develop this model, an architecture was introduced in the IHAN project [15], which was conducted to reflect architecture requirements for MyData establishment as part of the Fair Data Economy Project initiated in Finland. The proposed model in the **Fig. 9** includes a dynamic consent model based on a smart contract to ensure the right to informational self-determination in a PDS environment.

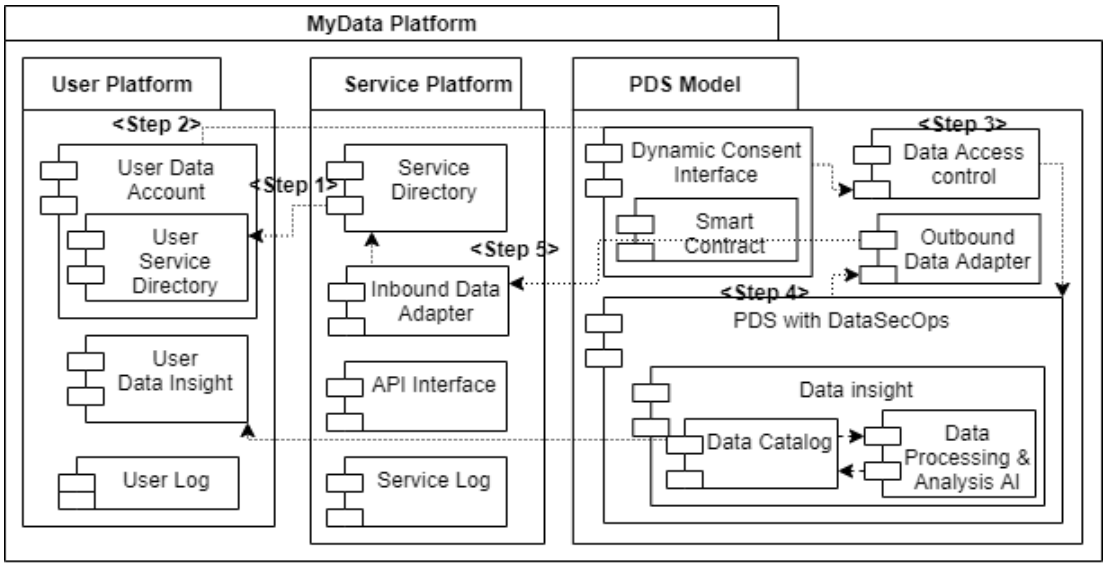

Fig. 9. Presenting the MyData Platform

Regarding the proposed model, the user manages ID and data access for the user data account. The user service directory in the user data account stores the list of services used by the user. The service directory stores the list of services created by the service provider and detailed descriptions of each service. This directory exists in both platforms of the user and the service provider. The inbound/outbound data adapter is the sending/receiving point from which the data provider transmits data to the service provider. Data access control uses the dynamic consent of the PDS model to verify the qualification of consent based on dynamic consent when the service provider approaches data.

The dynamic consent interface of the PDS model allows data subjects to utilize data of the PDS model and to control approvals, rejections, and withdrawals by themselves [18]. Diverse types of consent can be provided. When dynamic consent is used, data access and management are controlled in a semantically defined hierarchical structure. Accordingly, a data management priority is applied to dynamically consent to the information when and with whom data are shared, stored, and used. Data subjects should be allowed to perform

peer-to-peer transactions based on smart contracts, dynamically revise and store consent, define authority for access, and record the history of access, removal, change, and storage.

Furthermore, the proposed PDS structure should comply with DataSecOps: a methodology for improving data quality and analysis within the lifecycle of data use. This methodology is oriented and integrated for data collection and analysis [16][17]. DataSecOps includes security internalization in the DataOps data lifecycle and applies security policies and governance to the stages of data collection, access, and analysis. High-quality user data can be swiftly transferred to the service platform, which accumulates such data based on dynamic consent without post-processing only when a great amount of data are securely stored in the PDS. Data insight accesses user data based on the data catalog. The user's catalog based on AI is used to process and analyze personal data stored in the PDS in real-time. The analytic results obtained are provided in the form of a dashboard that displays the lifecycle of data use according to user requirements.

As consent is frequently gained, a web service consent description is used to match consent forms through actions, such as negotiation and consultation. It is also required to apply automation based on AI technology. Regarding dynamic consent components, the data subject should be clearly informed about methods for processing personal data, managing consent endlessly, and canceling details of previous consents. The data subject should also be allowed to control access to information at the appropriate personal data processing level and have access to methods for managing consent to personal data processing at any time. Thus, a dynamic consent model is needed that can satisfy these requirements. The functional flow of elements of the proposed model is as follows. The model generates a unique user data account that the user registers for MyData. To use a MyData service, the user must search the desired service and add it to their user service directory. To this end, it is necessary to implement technology for ensuring traceability of data by applying technologies such as OpenID, Connect, OAuth, or UMA for decentralized autonomous authentication, approval, and access control.

Table 4 displays the processes of service providers accessing data in the PDS of the user in **Fig. 9**.

Table 4. Process of the MyData Platform

Step 1	The service provider is authorized to access data in the user's account based on the user service directory of the user data account.
Step 2	The user sends his or her request to the dynamic consent interface, which determines the existence of consent.
Step 3	When consent is obtained, the service provider accesses data in the PDS based on data access control.
Step 4	Data requested by the service provider are transferred to the service platform through the outbound data adapter.
Step 5	The service provider can check data transferred by the inbound data adapter in the service platform.

Table 5 compares the core functions of three types of MyData open-source models examined in Section 3 regarding the MyData PDS model that ensures the right to informational self-determination.

The API contains a function for ensuring data portability between. Because the Declaration of the MyData Global Organization aims to guarantee the right to data portability, three models exist for this function. The OwnYourData model provides data insight based on the Weekly Digest service, but it does not include infrastructure for processing and analyzing data in real-time as does the proposed model. The user data account stores information on data access, profiles, logs, and consent in the data store of the data provider. Because this function is essential for the PDS environment, three models support this function.

Dynamic consent, Personium, and MyDataShare manage consent based on the consent management function. However, this process cannot be regarded as dynamic because decisions are made regardless of certain situations. Smart contracts are applied to facilitate traceability of the history of consent change caused by dynamic consent. Hence, three models that include MyDataShare do not correspond to dynamic consent. DataSecOps is utilized to transfer data safely and to increase data quality. This function is not included in the three models analyzed.

Table 5. Comparative Analysis of Important Functions in the MyData Model

Proposed model	Personium	MyDataShare	OwnYourData
API(Data Portability)	Cell	Connectivity	Semantic Container
Data insight	N/A	N/A	Weekly Digest
Dynamic Consent	Not Dynamic Consent	Not Dynamic Consent	N/A
DataSecOps	N/A	N/A	N/A
Smart Contract	N/A	N/A	N/A
User Data Account	Cell	MyDataShare ID	Data Vault

6.2 Proposal of Security Requirements for Establishment of the Safe MyData Platform and PDS Model

This section compares and analyzes the MyData authentication and security requirements analyzed in Section 4 to suggest security requirements for a secure PDS model. **Table 6** shows that security requirements for a safe PDS model are classified as cyber security, personal data protection, and measures for protecting personal data. These categories include 29 sub-categories. Regarding cyber security, this study compares requirements established by ISO/IEC 27001, ISO/IEC 29100, guidelines on certification of trust banks in Japan, and guidelines on MyData in the financial field in South Korea. Hence, this study proposes requirements for enhancing the right to control and implementing MyData/PDS models.

This study presents security requirements to ensure the right to informational self-determination as guaranteed by the GDPR, as analyzed in Section 5 for the protection of personal data, including a range of consent to and selection of personal data stored in the PDS, data collection, and rights of access, erasure, and portability under the PDS structure.

The first requirement for fundamental policies on security is to develop policies on data (cyber) security and personal data (privacy). The Japanese guidelines on security requirements comply with A.5.1 of ISO 27001. However, the South Korean guidelines based on the Credit Information Use and Protection Act do not include policies of personal data protection. ISO 29100 is a framework for privacy, but it does not include information on establishing policies

on personal data protection. The second requirement is related to the preparation of safety management measures in organizations of MyData service providers. ISO 27001, a data protection management specification, covers this requirement from various aspects, and the Japanese guidelines also comply with this system. However, the South Korean guidelines provide only measures on personal data (credit). The third requirement is based on data security for human resources. Particularly, employees who are responsible for use of personal data should participate in data protection training. Only ISO 27001 and the South Korean guidelines need this requirement.

The fourth and fifth requirements are related to details of physical and technical security. Physical security refers to access control to physical facilities, documents, and devices. ISO 27001 and the Japanese and South Korean guidelines reflect these requirements. Technical security refers to the control of passwords or systems. Access control and encryption are required by ISO 27001 and the Japanese and South Korean guidelines. On the other hand, the security of the API system for MyData is required only in the South Korean guidelines.

The sixth requirement is related to the security of personal data. This requirement is applied to manage PII information and shows details of consent to personal data, selection, collection, and limitation. ISO 29100 and the Japanese guidelines reflect details of consent and selection. However, the South Korean guidelines reflect only requirements on storage of personal data and data minimization, management, and elimination without details of consent and selection. A requirement for data storage and minimization is reflected in ISO 29100 and the Japanese and South Korean guidelines. However, a requirement on management measures is reflected in only ISO 29100 and the South Korean guidelines. A requirement on the elimination of personal data is needed only in the South Korean guidelines.

There is also a requirement on quality management for data transferred to MyData. ISO 29100 and the Japanese guidelines define the accuracy of data to satisfy security requirements related to this condition. Additionally, ISO 29100 and the Japanese guidelines determine policies on external personal data security required for releasing personal data based on the consent of data subjects. Furthermore, requirements on the application of measures in the case of obtaining personal data without the consent of data subjects are also needed, although these requirements are not reflected in the four frameworks and guidelines analyzed in this study.

Table 6. Proposed Security Requirements for a Secure MyData Platform and PDS model

			ISO 27001	ISO 29100	Japan Guideline	FSC Guideline
1	Formulation of a basic policy	Establishment of information security policy	A.5.1, 4.3, 5.2, 6.1	N/A	5.2.2.1	5.2.1
		Establishment of personal information security policy	A.5.1	N/A	5.3.2.1	N/A
2	Organizational safety management measures	Operation, monitoring, review of Information security management	5.3, 6.1.2, 9.2, 10.2	N/A	5.2.2.2	5.2.1
		Resources, roles, responsibilities, authority	A.6.1.1	N/A	5.2.2.5	N/A

		Incident management	A.16.1	N/A	5.2.2.14	5.2.7, 5.2.8
		Segregation of duty	A.6.1.2	N/A	N/A	5.2.5
3	Information security of human resources	Personal information security education	A.7.2.2	N/A	N/A	5.2.2
4	Physical security management measures	Physical and environmental information security	A.11.1	N/A	5.2.2.9 5.3.2.4	5.3.1, 5.3.2
5	Technical security management measures	Access control	A.9.1, A.9.2, A.9.4.3	N/A	5.2.2.8 5.3.2.4	5.2.4 5.4.1
		Cryptography	A.10	N/A	5.2.2.8	5.4.2
		API system management	N/A	N/A	N/A	5.2.6
6	Personal identifiable information management	Consent and choice	N/A	5.2	5.3.2.6	N/A
		Disclosure of personal information utilization system	N/A	5.2, 5.3	5.3.2.7	5.2.3
7	Personal information collection	Collection and restriction	N/A	5.4	5.3.2.8	5.2.3
8	Restrictions on use, retention, and disclosure	Storage of personal information	N/A	5.6	5.3.2.10	5.2.3
		Data minimization	N/A	5.5	5.3.2.9	5.2.3
		Security management measure	N/A	5.10	N/A	5.2.3
		Delete personal information	N/A	N/A	N/A	5.2.3
9	Accuracy and quality	Ensuring accuracy	N/A	5.7	5.3.2.11	N/A
		Right of the person regarding personal information	N/A	5.9	N/A	N/A
		Operation confirmation	A.6.1.6	N/A	N/A	N/A
10	Openness, transparency and notification	External personal information security policy	N/A	5.8	5.3.2.12	N/A
		Actions when acquiring personal information	N/A	N/A	N/A	N/A
11	Individual participation and access	Individual participation and access	N/A	5.9	5.3.2.13	N/A
12	Operations information security		A.12.1, A.12.2, A.12.3.1, A.12.4, A.12.6.1	5.11	5.2.2.10	5.4.3
13	Communication information security		A.13.1.1, A.13.1.2, A.13.1.3, A.13.2	N/A	5.2.2.11	5.4.2

7. Conclusion

South Korean and international firms and governments have recently conducted verification projects for services applying various MyData models to safely manage personal data based on the consent of the data subjects. Accordingly, great efforts have been levied to prepare legal and technical solutions for efficiently obtaining the consent of data subjects and transferring personal data from organizations and firms that currently possess such data to other ownership channels. In South Korea, MyData projects have been actively carried out in the financial field related to personal credit information management. Researchers have developed functions for the establishment of MyData platforms, such as those for self-authentication, consent to the supply of personal data, data transfer from providers, standard API development, data connection and collection, and data analysis. However, few studies have been conducted to ensure the right to informational self-determination for practically achieving personal data protection on MyData platforms and in PDS environments, associated with centralized and decentralized MyData/PDS operations, exchange methods, accountability, traceability, secret sharing, smart contracting (e.g., blockchain), dynamic consent, multi-layer pseudonymous processing, encryption, standardization, ontology, and compatibility.

This study analyzed various MyData/PDS models developed in other countries and presented a new direction for developing a secure model based on dynamic consent, which can practically ensure the right to information self-determination. This study also investigated different regulations and requirements governing personal data as established in South Korea and other counties to identify a contiguous method with security requirements that can be applied to MyData and PDS systems and services that will be appropriate for circumstances in South Korea. However, this study has limitations in that the proposed PDS model, consent method, and security requirements cannot be applied to all industries. This study also presented necessary security requirements for MyData platforms, which were not reflected in existing government and international requirements. Furthermore, this study suggested only requirements applicable to models based on dynamic consent. In this regard, further research on model operation processes should be carried out. Follow-up research is also needed to verify a MyData platform that implements processes for checking personal data stored in PDSs, managing details of records and consent, transferring data between MyData platforms, and enhancing the proposed model and security requirements. Furthermore, additional research is planned to analyze decentralized PDS models that integrate centralized PDS models using blockchain, which has been examined by international open-source firms and organizations.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1G1A1007540). This study was supported by a 2019 Research Grant from Kangwon National University.

References

[1] Woongryul Jeon, "A Study on PDS Architecture and Security Functions," *The Journal of Information Technology and Architecture*, vol. 15, no. 3, pp. 345-356, Sep. 2018. Article (CrossRef Link)

[2] Kazushi Ishigaki, and Akio Shimono, "A Concept of Community based PDS and Information Bank as a platform for realizing Engaged Society," Fujitsu Ltd. Japan. Nov. 25, 2017. [Online] Available: http://id.nii.ac.jp/1001/00184563/

[3] Joss Langford, Antti 'Jogi' Poikola, Wil Janssen, Viivi Lähteenoja and Marlies Rikken, "Understanding MyData Operators," MyData Global Apr. 29, 2020. [Online] Available: https://mydata.org/wp-content/uploads/sites/5/2020/04/Understanding-Mydata-Operators-pages.pdf

[4] "Interim report of data utilization working group in the age of AI and IoT," Data distribution environment development study group, Japan. Mar. 15, 2017 [Online] Available: https://www.kantei.go.jp/jp/singi/it2/senmon_bunka/data_ryutsuseibi/dai2/gijisidai.html

[5] "MyData Service Guideline," Korea Data Agency, Korea. Dec. 30, 2019. [Online] Available: https://kdata.or.kr/kr/board/info_01/boardView.do?pageIndex=6&bbsIdx=482&searchCondition=all&searchKeyword= .

[6] Personium Project, Fujitsu. Japan. [Online] Available: https://personium.io/docs/en/README/

[7] "MyDataShare White Paper," Vastuu group, Finland. May 2020. [Online] Available: https://www.mydatashare.com/whitepaper

[8] OwnYourData, Austria. 2015 [Online] Available: https://www.ownyourdata.eu/en/

[9] "White Paper Semantic Container," OwnYourData, Jan. 2019. [Online] Available: https://www.ownyourdata.eu/en/semcon/

[10] "Semantic Containers Diabetes" OwnYourData, EU. 2020. [Online] Available: https://www.ownyourdata.eu/en/diabetes-data-processing/

[11] International Standard Information Technology Security Techniques Privacy framework, ISO/IEC 29100, Dec. 2011.

[12] "Information Bank certification application guidebook," Information Technology Federation of Japan Information Bank Promotion Committee, Jul. 2020 [Online] Available: https://www.tpdms.jp/file/Guidebook_ver2.01.pdf

[13] "Finance MyData Technical Guidelines," Financial Services Commission, Jul. 29, 2020. [Online] Available: https://www.mydatacenter.or.kr:3441/myd/bbsctt/normal1/normal/66c3dd13-6526-4ca1-b424-cebfd069d2e0/22/bbsctt.do

[14] "General Data Protection Regulations," EU, Apr. 2016. [Online] Available: https://gdpr.eu/tag/gdpr

[15] Juhani Luoma-Kyyny, Jyrki Suokas, "IHAN BluePrint 2.5," Sitra, Finland. Jan. 22, 2020. [Online] Available: https://media.sitra.fi/2018/12/22091907/ihan-blueprint-2-5.pdf

[16] Ereth, Julian, "DataOps-Towards a Definition," LWDA 2191, pp.104-112, 2018. [Online] Available: http://ceur-ws.org/Vol-2191/paper13.pdf

[17] Munappy, Aiswarya Raj, et al, "From ad-hoc data analytics to DataOps," in *Proc. of the International Conference on Software and System Processes*, pp. 165-174 Jun. 2020. Article (CrossRef Link)

[18] Mont, Marco Casassa, Vaibhav Sharma, and Siani Pearson, "EnCoRe: dynamic consent, policy enforcement and accountable information sharing within and across organizations," HP Laboratories., USA, Technical Report HPL-2012-36, Feb. 21, 2012. [Online] Available: https://www.hpl.hp.com/techreports/2012/HPL-2012-36.pdf

[19] "Information Bank Certification Examination Check Sheet," Information Technology Federation of Japan Information Bank Promotion Committee, Jul. 2020. [Online] Available: https://www.tpdms.jp/file/CheckSheet_ver2.01.xls

Seong-hyun Min is a postgraduate student in master's course at Dept. of Computer Engineering in Kangwon National University, Republic of Korea. He received the B.E. degree from Kangwon National University, Republic of Korea. His research interests include Information Assurance, ICT Supply Chain Security, Zero-Trust Network, Software Bill of Materials.

Kyung-ho Son received his B.S. degree in received his B.E., M.S., and Ph.D. degree from Sungkyunkwan University in 2001, 2013, and 2015, respectively. He worked at Korea Internet Security Agency from 2001 to 2018 and he has been worked in Kangwon National University since 2018. His research area Information Assurance, Privacy by Design, Design of Security system, IoT·CPS Security