한국ITS학회논문지 J. Korean Soc. Intell. Transp. Syst.

교통분야 가명정보의 효율적 처리 및 활용을 위한 통합데이터안심구역 프로토타입

Integrated Data Safe Zone Prototype for Efficient Processing and Utilization of Pseudonymous Information in the Transportation Sector

이 형 근* · 유 기 동**

* 주저자 : 한국도로공사 디지털계획처 AI데이터팀 부장
단국대학교 일반대학원 경영학과 박사과정
** 교신저자 : 단국대학교 경영학부(경영정보) 교수

Hyoungkun Lee* · Keedong Yoo**

* Dept. of Digital Planning, Korea Expressway Corporation
Dept. of Business Administration, Dankook University
** Div. of Business Administration, Dankook University

† Corresponding author : Keedong Yoo, kdyoo@dankook.ac.kr

Vol. 23 No.3(2024)
June, 2024
pp.48~66

pISSN 1738-0774
eISSN 2384-1729
https://doi.org/10.12815/kits.2024.23.3.48

Received 18 April 2024
Revised 3 May 2024
Accepted 8 May 2024

요 약

데이터 3법과 데이터 산업법에 따라 가명정보 결합전문기관 및 데이터안심구역 시스템이 물리적으로 분리되어 운영 중이므로, 가명정보의 처리 및 활용을 원하는 중소기업 또는 스타트업 등의 사용자에게 복잡한 절차와 병목으로 인한 부담으로 작용한다. 또한, 개인정보의 유출 등을 우려한, 지나치게 엄격한 가명 처리 과정은 오히려 데이터의 품질을 훼손하는 역효과가 발생한다. 가명정보의 안전한 처리 및 활용을 위한 일련의 조치는 사용자의 편의와 데이터의 품질을 동시에 보장할 수 있도록 구성되어야 한다. 따라서 본 연구는 기존 가명정보 처리 및 활용의 문제점을 개선한 통합데이터안심구역의 프로토타입 시스템을 제시한다. 이를 위해 기존 BPR 가이드라인을 선택적으로 수정하여 새로운 워크플로우 재설계 가이드라인을 개발 및 적용하며, 핵심성능지표를 도출하여 개발된 프로토타입의 성능을 판단한다. 성능평가 결과 제시된 프로토타입은 기존의 시스템에 비해 시간적 측면에서는 약 6배, 비용적 측면에서는 1.28배, 품질적 측면에서는 1.3배의 향상된 성능을 보임을 확인하였다.

핵심어 : 통합데이터안심구역, 가명정보, 워크플로우 최적화, 재설계 가이드라인, 성능평가

ABSTRACT

According to the three amended Laws of the Data Economy and the Data Industry Act of Korea, systems for pseudonymous data integration and Data Safe Zones have been operated separately by selected agencies, eventually causing a burden of use in SMEs, startups, and general users because of complicated and ineffective procedures. An over-stringent pseudonymization policy to prevent data breaches has also compromised data quality. Such trials should be improved to ensure the convenience of use and data quality. This paper proposes a prototype system of the Integrated Data Safe Zone based on redesigned and optimized pseudonymization workflows. Conventional workflows of pseudonymization were redesigned by applying the amended guidelines and selectively revising existing guidelines for business process redesign. The proposed prototype has been shown quantitatively to outperform the conventional one: 6-fold increase in time efficiency, 1.28-fold in cost reduction, and 1.3-fold improvement in data quality.

Key words : Integrated Data Safety Zone, Pseudonymous information, Workflow optimization, Redesign guidelines, Performance evaluation

Ⅰ. 서 론

우리나라는 2020년 8월, 데이터 산업 육성을 위해 개인정보처리자 등 가명정보의 결합을 신청하는 결합신청자(이하 사용자)가 통계의 작성이나 과학적 연구 등 특수 목적을 위해 개인정보 주체의 동의 없이 개인정보를 가명 처리하여 이용할 수 있도록 관련법(개인정보보호법, 신용정보보호법, 정보통신망법)을 개정하였다. 이에 개인정보의 무분별한 이용을 방지하기 위해 가명정보의 결합, 심사 및 반출을 담당하는 가명정보 결합전문기관(이하 결합전문기관)과 가명처리와 결합을 통해 암호화 키의 생성 및 이를 지원하는 결합키 관리기관을 지정하였다(Personal information Protection commission, 2022). 또한, 2022년 4월, 과학기술정보통신부는 '데이터 산업진흥 및 이용촉진에 관한 기본법'을 제정하여 안전한 분석 공간 및 도구를 지원할 수 있는 데이터안심구역을 별도로 지정하였다(Ministry of Science and ICT, 2022).

그러나, 기존 시스템에서는 개인정보나 민감정보의 가명처리 및 가명정보 결합을 위한 시스템과 데이터안심구역 운영시스템이 개별 기관(개인정보처리자, 결합키관리기관, 결합전문기관 및 데이터안심구역 지정기관)에 의해 물리적으로 구분되어 운영되고 있다. 이에 따라, 가명정보 등 민감한 데이터의 생성, 가공, 분석, 및 활용이 별도로 이루어지고 있다. 이는 데이터 생명주기에 따른 민감정보의 추적과 관리를 어렵게 하며, 오히려 개인정보 유출 가능성이 항시 존재하는 문제점을 가지고 있다(Kim, 2020a). 또한, 사용 절차가 복잡하고, 물리적으로 분리된 기관에서 운영되고 있어, 사용자가 기존 시스템을 사용 시 신청, 검토, 승인, 및 결과반출이 이루어지는 과정에서 처리시간이 과도하게 소요되는 비효율성이 존재한다.

이러한 문제 해결을 위해, 첫째 기존 운영시스템의 논리적 통합이 필요하다. 가명처리 및 가명정보 결합시스템과 데이터안심구역 운영시스템을 물리적으로 분리하는 대신, 이 두 시스템의 통합을 통해 데이터 생명주기 전 과정을 지원하도록 하여 민감정보를 효율적으로 추적 및 관리하고 원활한 활용이 가능케 하여야 한다. 둘째, 사용 절차의 간소화와 쉬운 접근이 필요하다. 기존 시스템들의 통합을 위해 워크플로우를 분석하여 재설계하고, 사용자 친화적인 인터페이스 즉, 사용자 포털을 제공하여 사용자가 여러 시스템을 원스톱으로 접속하여 업무처리의 복잡성을 간소화함으로써 전체 처리시간의 효과적인 단축이 필요하다.

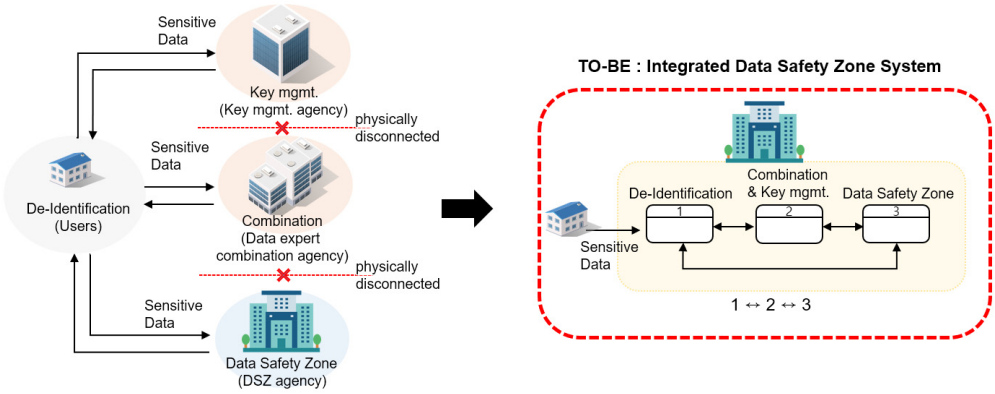
따라서, 본 연구는 Lee and Yoo(2023)가 제안한 통합데이터안심구역 시스템의 분석된 요구사항, 개념 프레임워크 및 아키텍처 등의 개념 설계를 기반으로, <Fig. 1>의 좌측 도식의 기존 운영 시스템(가명처리, 결합키관리시스템, 가명정보 결합시스템 및 데이터안심구역 시스템)을 IT, 제조 등 다양한 산업군의 29개 기업에서 재설계 시 발견한 Best Practice를 반영한 '경험적인 재설계 접근법'인 '비즈니스 프로세스 재설계 가이드라인'(Reijers and Mansar, 2005; Mansar and Reijers, 2007)을 선택적으로 수정하여, '재설계 진단 및 실행 가이드라인'을 개발하여 적용하였다. 이를 통해 <Fig. 1>의 우측 도식과 같이 기존 시스템의 워크플로우를 논리적으로 상호 연계한 차세대형 통합데이터안심구역 시스템으로 재설계하였다. 본 연구에서 재설계한 워크플로우를 기반으로 통합데이터안심구역 프로토타입을 구현하고, 재설계 시 도출된 구현 지표와 성능평가지표를 활용하여 프로토타입의 성능을 평가함으로써, 제시된 아이디어의 유효성을 검증하였다.

Ⅱ. 관련 연구

1. 교통 분야 데이터의 분류 및 특징

4차 산업혁명의 촉발에 따라 AI, 빅데이터와 같은 기술은 데이터 기반 행정과 새로운 비즈니스 창출 등에

대해 공공 및 민간을 가리지 않고, 혁신의 시대적 요구를 더 해 가고 있다. 특히 데이터의 수집에 있어, 모바일, 사물인터넷을 기반으로 교통, 금융, 유통, 의료 등 다양한 분야에서 방대한 개인정보의 생산 및 수집이 가능하게 되었고, 특히 교통 분야에서도 기존의 도로나 시설에 장착하여 데이터를 수집하던 방식에서 벗어나, 차량이나 사람 등의 이동체(Mobility)에 직접 장착하거나 휴대가 가능한 모바일 기기로의 기술 전환이 급속히 진행되고 있다.



<Fig. 1> The concept of redesigned IDSZ (Left : Legacy System, Right : Redesigned IDSZ)

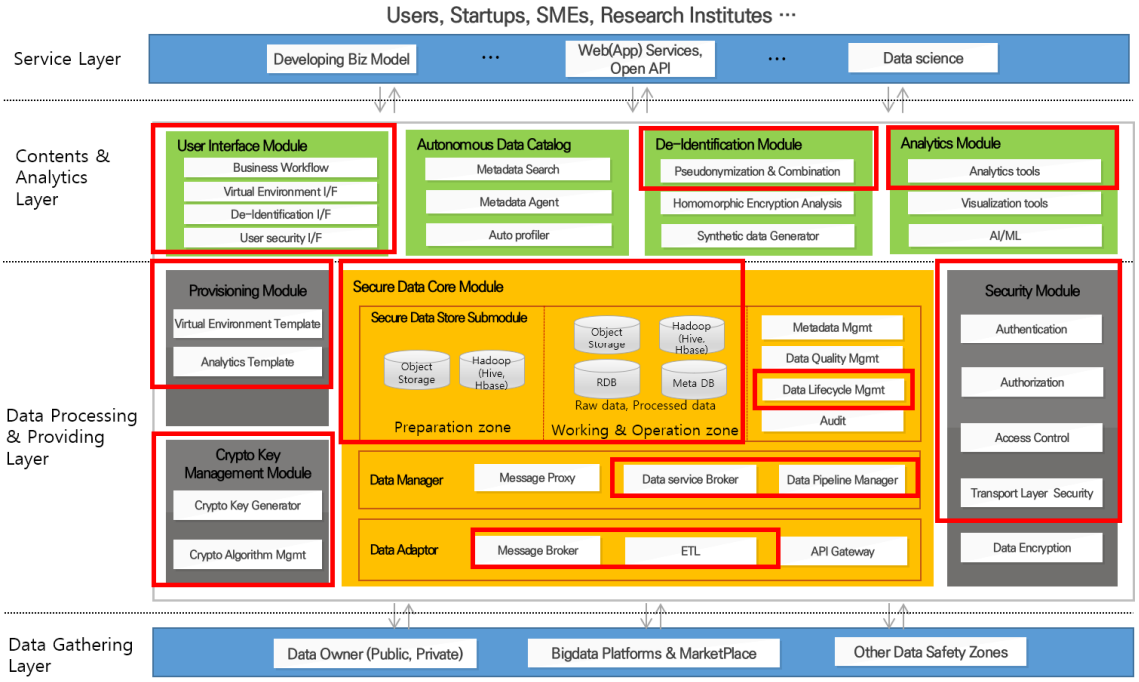
교통 분야의 데이터를 살펴보면, Lee and Jung(2018) 및 The Korea Transportation Institute(2017)는 교통 분야 데이터를 교통 관련 데이터(차량 통행, 주차장, 네비게이션 등)와 비 교통 관련 데이터(이동통신, 신용카드, 정밀도로지도 등)로 구분하여 정의하였고, Korea Data Agency(2024)는 교통 분야 데이터를 교통 및 이동 데이터(이동 및 궤적), 교통 및 이동 데이터(운영 및 시설), 사회경제 데이터(사회경제지표, 지역별 교통량 환산계수 등), 도시 및 공간 정보(GIS 데이터, 건물 위치 데이터 등)로 정의하여 분류하였다.

교통 분야 데이터의 특성은 사람이나 차량 등의 이동궤적 데이터(Movement Trajectory Data)와 교통카드 등 개인의 사회경제적 데이터 등이 포함되어있어, 개인정보보호는 매우 중요한 문제 중 하나이며(Sim, 2023), 다른 데이터와 융합 시 특정 개인을 식별하는데 악용될 수 있다(The Korea Transportation Institute, 2017). 따라서 개인정보를 가명 처리 등의 비식별화 기술을 이용하여 가명 정보를 생성하고, 이를 공인된 기관을 통해 결합, 분석 및 활용할 수 있는 보다 안전한 방안을 모색해야 한다. 이를 통해 사전에 개인정보의 침해요인을 예방하고, 개인정보의 안전한 처리, 사회적 불안 해소 및 보호와 이용 간의 조화를 이룰 수 있도록 하여야 한다(Kim, 2020b).

2. 통합데이터안심구역시스템(Integrated Data Safety Zone System, IDSZ) 아키텍처

Kim(2020b)과 Kim and Kim(2020)은 결합전문기관을 통한 개인정보의 안전한 처리 등을 위해 가명정보의 결합 및 분석, 개방 및 거래를 통합하는 원스톱 활용체계 구축을, Kim and Kwon(2023)은 결합전문기관의 법/제도 등 운영현황 등을 연구해 가명처리와 가명정보 결합을 확장한 안전한 재활용 공동저장소 구축을, Kim(2023)은 가명처리와 가명정보의 결합기능을 데이터안심구역으로 통합할 것을 제안하였다. 이에 Lee and Yoo(2023)는 기존의 가명처리 및 가명정보 결합시스템과 데이터안심구역 운영시스템을 통합한 차세대형 통합데이터안심구역 시스템 프레임워크를 제시하고, 가명정보 등 민감한 데이터의 생명주기 전 과정을 지원하

는 통합데이터안심구역시스템 아키텍처를 <Fig. 2>과 같이 제안하였다. 차세대형 통합데이터안심구역 시스템은 4개의 Layer와 8개의 모듈로 구성되어 있으나, 본 연구에서는 프로토타입 구현을 위한 주요 핵심 기능들을 선별하여 5개의 모듈(User Interface, De-Identification, Combination, Analytics, Data Core)로 구성하여 구현한다.



<Fig. 2> Proposed Architecture of IDSZ (Lee and Yoo, 2023)

3. 비즈니스 프로세스 재설계 가이드라인

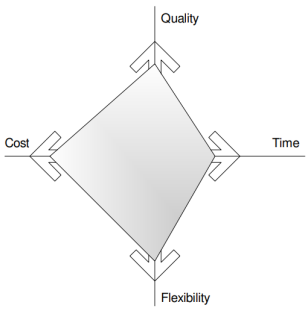
1990년대 마이클 해머가 제안한 비즈니스 프로세스 재설계 원칙(Principles)은 “작업 말고 결과를 조직화하라, 프로세스를 출력하는 사람들이 프로세스를 수행하게 하라, 정보를 생성하는 실제 작업에 정보 처리 작업을 포함하라, 지리적으로 분산된 자원을 중앙 집중화된 것처럼 처리하라, 결과를 통합하는 대신 병렬 작업을 연결하라, 작업 수행의 의사결정 지점을 배치하라(Hammer, 1990)”는 매우 선언적 원칙으로 실무의 적용에 있어 실무자들에게 많은 고려할 사항을 남겼다. 비즈니스 프로세스의 재설계를 연구하는 연구자들이 재설계 원칙이나 가이드라인을 개발하는 것은 재설계 작업의 경험적(Heuristic) 특성상 많은 시간과 노력이 필요하며, 전략적 가치로 인해 대중에게 공표되는 일이 거의 없었다(Yoo et al., 2007). 실무에서 비즈니스 프로세스 재설계는 여러 가지 가이드라인이 동시에 적용되는 특성이 있다고 주장하며, 다양한 산업군에 속하는 기업의 Best Practice 29개의 사례를 바탕으로 기본적인 비즈니스 프로세스 재설계 가이드라인을 이용한 접근법을 제시하였다(Reijers and Mansar, 2005). 이에 본 연구에서는 Reijers and Mansar(2005)의 재설계 가이드라인을 선택적으로 수정하여 본 연구에 맞게 ‘진단 및 실행 가이드라인’을 <Table 1>과 같이 구성하고, 이를 활용하여 기존 시스템을 진단하여 통합데이터안심구역으로 재설계한다.

<Table 1> Redesign Guideline (Reijers and Mansar, 2005; Mansar and Reijers, 2007)

Best Practice	Guideline	BPCode
Control Relocation	· Move controls towards the customer	C1
Contact Reduction	· Reduce the number of contacts with customers and third parties	C2
Integration	· Consider integration with a business process of a customer or a supplier	C2
Task Elimination	· Eliminate unnecessary tasks from a business process	B2
Task Composition	· Combine small tasks into composite tasks and divide large tasks into workable smaller tasks	B5
Resequencing	· Move tasks to more appropriate places	B6
Case assignment	· Let workers perform as many steps as possible for single orders	O1
Numerical Involvement	· Minimize the number of departments, groups and persons involved in a business process	O2
Buffering	· Instead of requesting information from an external source, buffer it by subscribing to updates	I2
Task Automation	· Consider automating tasks	T1
Integral Technology	· Try to elevate physical constraints in a business process by applying new technology	T2
Interfacing	· Consider a standardized interface with customers and partners	E3

4. 비즈니스 프로세스 성능평가 지표 및 모델

성능평가 모델은 조직의 전략적 방향의 개발 및 배포의 통제 수단에 초점이 맞추어 있지만, 비즈니스 개선계획 이전에 요구사항이나 추진 동인에 대한 식별 수단으로 활용된다(Haffey and Duffy, 2001). 성과측정 모델의 접근방법은 BSC(Balanced Scorecard)를 근간으로 하는 조직 차원의 전략, 전술 및 운영 수준의 성과측정 모델과 비즈니스 프로세스 차원의 질적, 양적 성과측정 모델의 접근방법이 있다(Van Den Ingh et al., 2020). Jansen-Vullers et al.(2008)은 비즈니스 프로세스 재설계가 성능에 미치는 영향을 정량화하는 데 중점을 두고, 성능측정모델을 분석하여 비즈니스 프로세스 성능 측정에 적합한 차원을 평가했다. 성능 피라미드, 성능 측정 매트릭스, 결과/결정 요인 매트릭스, Balanced Scorecard, Devil's Quadrangle 및 성능 프리즘과 같은 여러 모델을 검토하여, 이들 모델에서 비즈니스 프로세스와 관련된 주요 차원으로 품질, 비용, 프로세스 시간, 유연성을 식별하였다. 특히 Devil's Quadrangle의 4가지 차원인 시간, 비용, 품질 및 유연성 측면에서 비즈니스 프로세스의 성능을 효과적으로 측정이 가능하다는 결론을 도출했다.



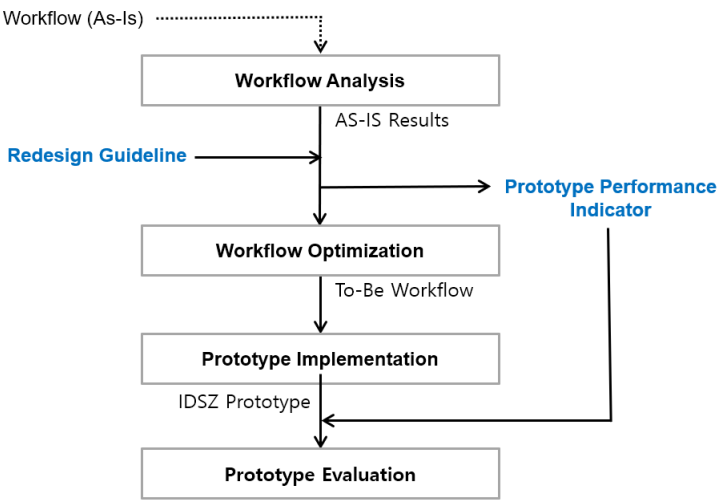
<Fig. 3> Devil' Quadrangle(Mansar and Reijers, 2007)

본 연구에서도 직관적으로 성과의 평가가 가능한 비즈니스 프로세스 성능 측정모델인 Brand & Van der Kolk의 Devil's Quadrangle(Reijers and Mansar, 2005)의 4가지 차원 모델을 제안한다. <Fig. 3>는 Devil's Quadrangle의 예시이다. 예를 들어, 비즈니스 프로세스가 수행되는 시간(Time)을 줄이기 위해서는

프로세스 수행 비용(Cost)의 증가가 수반되며, 프로세스의 품질(Quality)을 개선하기 위해서는 예외 처리 등 프로세스의 유연성(Flexibility)이 감소 될 수 있다. 이는 “비즈니스 프로세스 재설계 가이드라인을 적용하는 과정에서 4가지 차원의 영향 요인 간의 상충관계(Trade-off)가 발생한다”는 것을 의미한다. 이때 4가지 차원은 연구자의 연구 목표 및 방향에 따라 차원을 변경하여 사용이 가능하다(Yoo et al., 2007; Mansar and Reijers, 2007).

Ⅲ. IDSZ 프로토타입

본 연구에서 제시하는 IDSZ 프로토타입의 구현 및 평가 절차는 <Fig. 4>와 같다. 첫째, 가명처리 및 가명정보 결합시스템과 데이터안심구역 운영시스템의 As-Is 워크플로우를 분석한다. 둘째, As-Is의 분석 결과를 바탕으로 비즈니스 프로세스 재설계 진단 및 실행 가이드라인을 이용하여 As-Is 시스템의 핵심 워크플로우를 선정하고, IDSZ의 To-Be 워크플로우를 최적화하여 재설계한다. 재설계 시 진단 및 실행 가이드라인의 Best Practice의 구현 지표를 선별하여 제시하고, 이를 통해 IDSZ 프로토타입의 성능평가지표를 도출한다. 셋째, IDSZ 워크플로우를 기반으로 프로토타입을 구현한다. 마지막 단계로 도출된 성능 평가지표를 이용하여 기존 시스템 대비 IDSZ 프로토타입의 성능을 평가한다.



<Fig. 4> Procedure of redesign and evaluation of IDSZ Prototype

1. Workflow Analysis

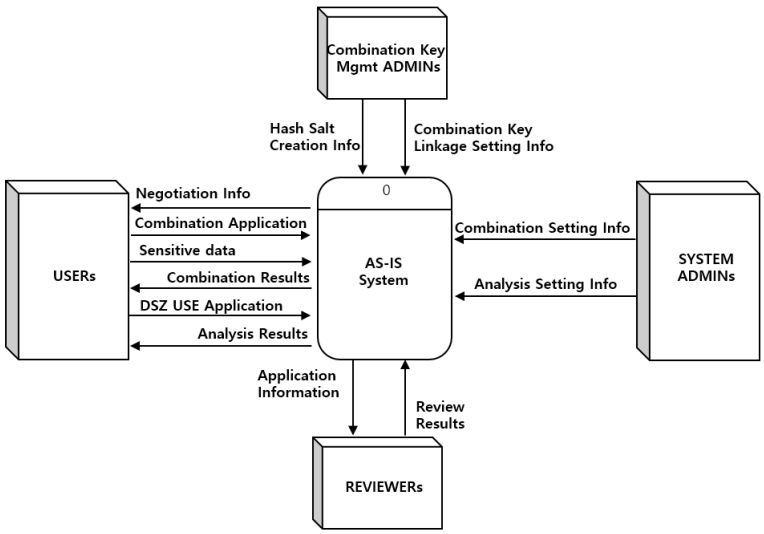
기존 Lagacy 시스템인 가명처리 및 가명정보 결합시스템과 데이터안심구역 분석 시스템이 물리적으로 분리되어 있어 별도의 관리기관에 의해 개별적으로 운영되고 있다. 따라서, 워크플로우 간 상호 단절된 형태로 상호연계하여 운영되지 못하였다. 본 연구에서는 As-Is 시스템의 워크플로우를 분석하여 <Fig. 1>의 To-Be 워크플로우와 같이 논리적으로 통합하였다. 즉 여러 시스템이 하나의 시스템처럼 상호연동이 가능한 통합 워크플로우로 작동되는 IDSZ 프로토타입을 구현한다. IDSZ 프로토타입은 민감한 데이터의 입력과 이를 처리하는 프로세스 및 출력의 연결된 처리 과정인 IPO(Input-Process-Output)의 개념으로 볼 수 있어, 이러한 프로세스를 하향식 접근

방법으로 모델링하는 것에 특화되어있는 구조적 개발방법론의 DFD(Data Flow Diagram)를 이용하여 다음과 같이 분석 및 재설계를 수행하였다(Nam, 2009).

1) As-Is Context Diagram

Context Diagram은 분석하고자 하는 시스템과 상호작용하며 영향을 주는 외부 개체와 이들 개체와 시스템 간에 주고받는 핵심 자료가 무엇인지 직관적으로 분석하기 위해 작성하는 다이어그램이다.

<Fig. 5>는 기존 시스템(가명처리, 결합키관리시스템, 가명정보 결합시스템 및 데이터안심구역 운영시스템)에서 사용자가 가명처리를 위해 '결합키관리기관'으로부터 암호화 키를 할당받고, 가명 처리를 통해 가명 정보, 결합키 및 결합대상정보를 생성하는 과정을 간략하게 보여준다. '가명정보 결합전문기관'은 사용자로부터 결합대상정보를 수령받고, '결합키관리기관'으로부터 결합키연계정보를 전송받아 가명정보를 결합한다. 결합된 가명정보는 가명정보 결합 적정성 평가 심의위원들의 검토를 받은 후 사용자에게 반출된다. Context Diagram을 작성 결과, 기존 시스템에 영향을 미치는 4개의 외부 개체(Users, Key Management Admins, Reviews, Combination System Admins)를 식별하였으며, 외부 개체와 프로세스 간 핵심 자료의 흐름을 파악하였다.



<Fig. 5> The context diagram of Legacy systems

2) As-Is Level 1 Diagram

Level 1 Diagram은 Context Diagram의 As-Is 시스템의 프로세스를 한 단계 세부적으로 분할 하여 도식화한 다이어그램으로 <Fig. 6>의 좌측 도식은 As-Is 프로세스의 Level 1 Diagram이며, D1(De-Identification), C1(Pseudonym Data Combination Processing), S1(Data Safety Zone Analysis) 프로세스는 입력된 데이터를 처리 후, 개별 관리기관이 물리적으로 분리된 각자의 데이터 저장소에 출력하여 활용하는 기존 시스템의 운영 중인 프로세스를 도식화한 다이어그램이다.

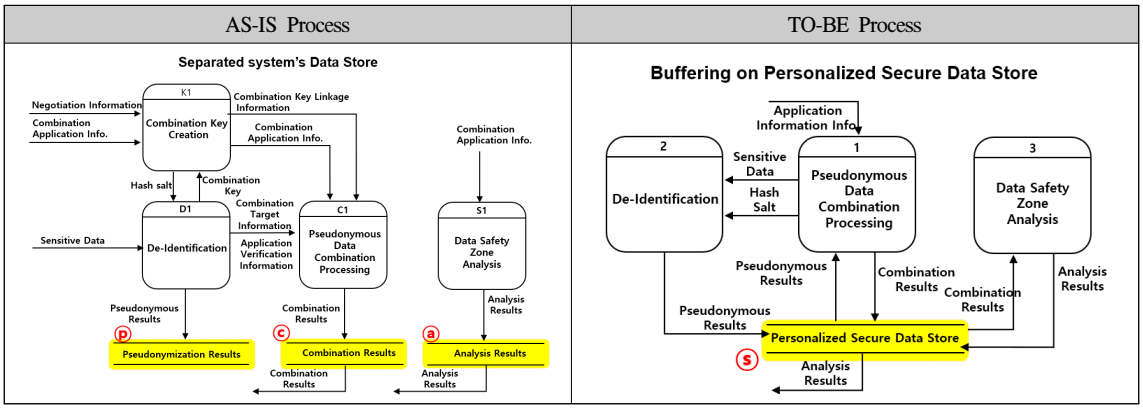
사용자는 가명정보를 생성 및 가공, 결합 및 분석을 위해 D1 프로세스를 이용해 가명정보를 생성하고, 사용자가 관리하는 시스템이나 개인 PC에 저장하여 사용한다. 이후 가명정보를 결합하고자 하는 경우, 사용자는 가명 처리한 가명정보에서 결합키를 제거하고, 일련번호를 생성하여 결합대상정보로 만든 후 결합전문기

관에 인터넷으로 전송하거나 USB 등 저장매체에 담아 직접 방문해 결합대상정보를 제공하게 된다.

결합전문기관의 관리자는 C1 프로세스를 이용하여 결합키관리기관이 전송한 결합키연계정보와 결합대상정보를 연결하여 결합 결과물을 생성한 후 사용자에게 반출한다. 사용자는 반출된 가명정보를 안전하게 분석하기 위해 데이터안심구역 분석 시스템을 이용하는 경우에도 인터넷으로 전송하거나 USB등 저장매체에 담아 직접 방문해 데이터안심구역을 이용하게 된다.

이렇게 D1, C1, S1 프로세스의 입력, 처리 및 출력의 과정이 상호 연결되지 않고, 물리적으로 단절되어 있어 사용자는 활용이 복잡하고, 별도의 시스템에서 가명정보 등 민감한 데이터를 생성 및 가공, 분석 및 활용하게되어, 데이터 생명주기 전 과정의 추적 및 관리가 불가능 할 뿐만 아니라, 데이터의 유출, 훼손 및 오남용 등의 가능성이 상시 존재한다. 이러한 As-Is 프로세스의 복잡한 처리 절차와 데이터의 유출 등의 문제를 해결하기 위해 IDSZ To-Be 프로세스는 D1, C1, S1에서 입력, 처리 및 출력되는 민감한 데이터 저장소를 개인화된 안전한 데이터 저장소 및 데이터베이스로 통합하여 관리(1)하고, 시계열분석 등이 가능하도록 <Fig. 7>의 우측 도식과 같이 재설계하였다.

Guideline BPCode(I2) : (ⓟ, ⓒ, ⓐ) ⊆ ⓢ ····· (1)
➔ 관련 모듈 : Secrue data store sub-module



<Fig. 6> Redesigned Personalized secure data store of IDSZ

3) As-Is Level 2 Diagram

Level 2 Diagram은 Level 1 Diagram의 프로세스(K1, D1, C1, S1)를 하부 프로세스로 한 단계 더 분할하여 기존 시스템을 분석하기 위해 사용하는 다이어그램으로 자세한 설명은 아래와 같다.

■ 결합키관리 프로세스(K1)

결합키관리 프로세스는 암호화키생성(K1.1), 결합신청(K1.2), 결합키연계정보생성(K1.3), 생성정보전송(K1.4) 프로세스로 식별된다. 사용자의 접근 창구를 일원화하여 최적화할 수 있도록 K1.2, K1.3, K1.4 프로세스는 삭제(2)하고, K1.1 프로세스를 Level 1 Diagram의 가명정보 결합 프로세스(C1)의 하나의 기능(3)으로 재설계하였다.

Guideline BPCode(B2) : Eliminate K1.2, K1.3, K1.4 ····· (2)
➔ 관련 모듈 : Secure Data Core Module

Guideline BPCode(C1) : K1.1 ∈ C1 ····· (3)
➔ 관련 모듈 : Pseudonymization and Combination sub-module

Guideline BPCode(C2) : Users directly access to C1 ····· (4)
➔ 관련 모듈 : User Interface Module, Pseudonymization and Combination sub-module

■ 가명처리 프로세스(D1)

가명처리 프로세스는 사용자가 가명처리를 위한 Hash Algorithm 등 가명처리 암호화 알고리듬과 이의 보안 강화를 위해 사용되는 Salt 값을 결합키관리기관으로부터 할당받기 위한 암호화 키 설정(D1.1), 가명처리(D1.2) 및 가명정보 전송(D1.3) 프로세스로 식별된다. 가명처리 프로세스의 재설계 시 D1.1과 D1.2는 하나의 기능으로 통합(5)하여 구성하고, 통합된 D1.1과 D1.2를 Level 1 Diagram의 가명정보 결합프로세스(C1)의 기능으로 흡수(6)하여 사용자가 직접 접근하여 사용토록 제어 권한을 재조정하여 구성하였다. 이에 따라 불필요하게 된 D1.3의 기능은 삭제(7)하였다.

Guideline BPCode(B2) : D1.1 + D1.2 ····· (5)
➔ 관련 모듈 : Pseudonymization sub-module, Crypto Key Generation sub-module

Guideline BPCode(C1) : (D1.1, D1.2) ∈ C1 ····· (6)
➔ 관련 모듈 : Pseudonymization and Combination sub-module

Guideline BPCode(B5) : Eliminate D1.3 ····· (7)
➔ 관련 모듈 : Data Pipeline sub-module

■ 가명정보 결합(C1) 및 데이터안심구역 분석 프로세스(S1)

<Fig. 7>은 가명정보 결합과 데이터안심구역 분석 프로세스에서 유사한 기능은 작업을 재구성(8)하고, 가명정보 결합과 데이터안심구역 프로세스를 통합(9)하였다. 이에 따라 작업 절차를 재순서화(10)하고, 프로세스의 처리 절차를 단순화하였다. 이때 C1.4와 S1.4의 연계를 자동화(11)하고, 인터페이스를 통해 상호 연동(12)하게 하였다. 따라서 C1.1과 S1.1의 재구조화에 따라 사용자의 접속점을 하나(13)로 줄일 수 있도록 재설계하였다.

Guideline BPCode(B5) : (S1.1 ∪ C1.1), (S1.2 ∪ C1.2), (S1.3 ∪ C1.3), (S1.5 ∪ C1.5), (S1.6 ∪ C1.6) ····· (8)
➔ 관련 모듈 : User Interface Module, De-Identification Module, Anaytics Module, Provisioning Module

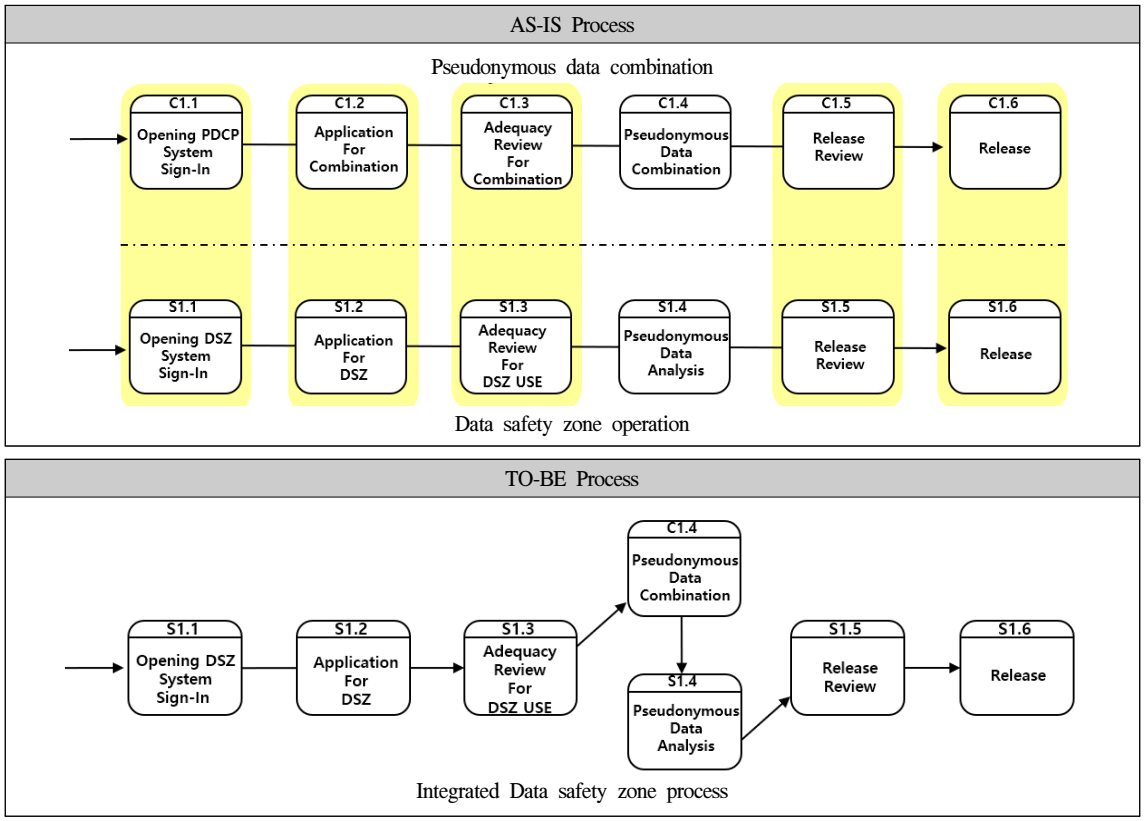
Guideline BPCode(C3) : (DSZ ∪ PDC) ⊆ IDSZ ····· (9)
➔ 관련 모듈 : User Interface Module, De-Identification Module, Anaytics Module, Provisioning Module

Guideline BPCode(B6) : S1.1 → S1.2 → S1.3 → C1.4 → S1.4 → S1.5 → S1.6 ····· (10)
➔ 관련 모듈 : User Interface Module, De-Identification Module, Anaytics Module, Provisioning Module

Guideline BPCode(T1) : Pseudonymous Data automatically moves from C1.4 to S1.4 ····· (11)
➔ 관련 모듈 : Secure Data Core Module, Security Module

Guideline BPCode(E3) : C1.4 interfaces with S1.4 ····· (12)
➔ 관련 모듈 : Secure Data Core Module, Security Module

Guideline BPCode(C2) : Users directly access to S1.1 ····· (13)
➔ 관련 모듈 : User Interface Module

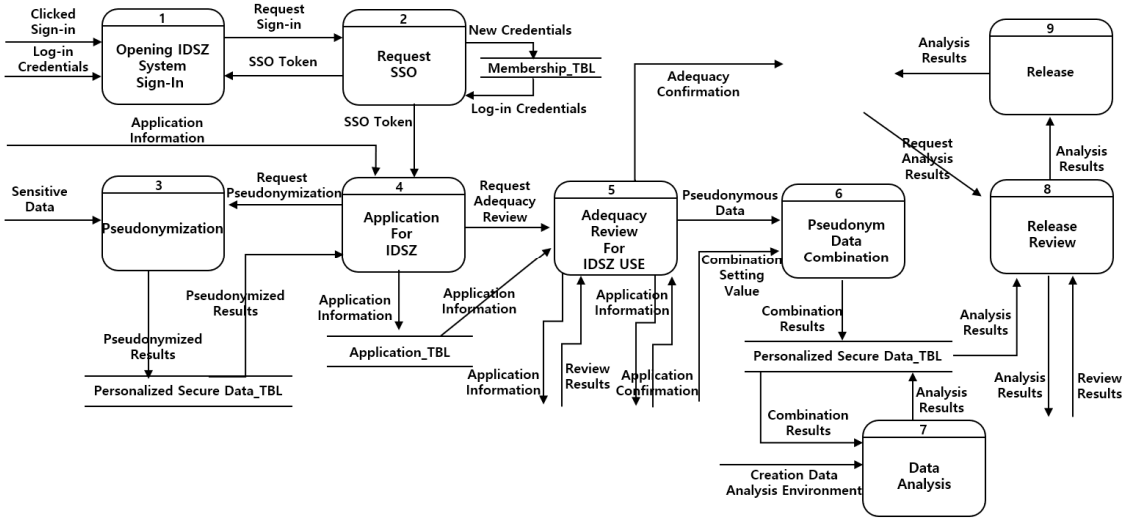


<Fig. 7> Redesigned IDSZ from legacy systems Analysis

2. Workflow Optimization

<Fig. 8>은 재설계 진단 및 실행 가이드라인에 따라 기존 시스템의 워크프로우를 최적화하여 3개의 외부 엔티티(Users, Reviewers, system Admins)와 9개의 워크플로우로 재설계한 IDSZ 프로토타입의 Level 0 Diagram 이다. 재설계된 워크플로우는 다음과 같이 작동한다. 1) 사용자는 논리적으로 통합된 단일 접속점을 통해 IDSZ UI 포탈에 접속하여 회원가입을 한다. 2) 회원가입이 된 사용자는 인증 및 접근권한을 위한 접속 토큰을 받아 IDSZ 시스템에 접속한다. 3) 승인된 사용자는 IDSZ에서 제공하는 가명처리 서브시스템을 이용하여 가명 처리를 진행하고, 이때 생성된 가명정보는 IDSZ의 개인화 안심 저장소 및 데이터베이스로 저장이 된다. 4) 이후 사용자는 IDSZ의 사용 신청을 통해 데이터의 반입, 가명 결합 및 분석환경을 신청하고, 이 정보는 신청정보 저장소 및 데이터베이스로 저장된다. 5) 시스템 관리자는 사용자의 신청정보를 확인하고, 개인화 안심저장소의 가명정보의 적합성 정보와 신청정보를 심사위원에게 제공하여 사용자가 별도의 방문 없이, 가

명정보 결합 및 데이터 안심구역 이용의 적합성을 심사하게 된다. 시스템 관리자는 적합성 심사가 완료되면 필요한 정보를 생성하고, 승인 여부를 사용자에게 알려준다. 6) 승인된 신청 건에 대해 시스템 관리자는 가명정보 결합이 필요한 경우 가명정보 결합 서브시스템을 이용하여 가명정보를 결합하게 된다. 이때 결합된 가명정보는 사용자의 개인화 안심 저장소로 인터페이스를 통해 자동으로 저장된다. 이때, 재설계된 워크플로우에 따라 결합을 위한 결합키 및 결합대상정보, 결합키연계정보의 생성은 불필요하게 된다. 7) 이후 사용자는 요청한 분석 환경, 즉 가상데스크톱을 호출하여 결합 후 사용자의 개인화 안심 저장소에 저장 되어있는 데이터와 가명정보 등을 이용하여 필요한 데이터를 분석하게 된다. 8) 분석이 완료되면 사용자는 분석된 데이터의 반출을 요청하고, 이때 시스템 관리자는 분석 결과를 반출 심사위원을 구성하여 반출 여부를 결정한다. 9) 반출심사가 완료된 데이터는 IDSZ UI 포털에 업로드되고, 사용자는 포털에 접속하여 쉽게 다운로드하여 이용할 수 있게 된다. 반출되는 데이터는 가명처리, 가명결합 및 데이터 분석이 완료된 결과물 또는 학습된 각종 AI모델 소스코드 등으로 개인정보 등 민감정보가 완전 제거 된 형태로 제공된다. 이를 통해 민감한 데이터의 안전한 생성 및 가공, 분석 및 활용의 생명주기 전 과정을 지원할 수 있도록 하였다. 또한, 재설계를 통해 기존 19개 프로세스 86개 활동(Activity)에서 9개 프로세스 30개 활동으로 워크플로우를 최적화하여 사용자의 시스템 활용에 따른 병목을 해소하고, 처리시간을 단축할 수 있게 하였다.



<Fig. 8> Level 0 diagram for redesigned IDSZ

3. IDSZ Prototype Implementation

Lee and Yoo(2023)가 제시한 개념 설계와 재설계된 워크플로우를 기반으로 프로토타입의 주요 핵심 기능들을 선별하여 5개의 모듈, 11개의 기능과 3개의 핵심 데이터베이스로 구성하여 구현하였다. <Table 2>는 각각의 모듈 및 기능과 데이터베이스 간의 내용을 검증한 CRUD Matrix이다.

<Table 2> The main function of IDSZ prototype

Function		Membership	Application	Results
UI portal	Single Sign On	CRUD		
	Application	R	CRUD	
	Adequacy Review	R	RUD	RUD
	Release Review	R	RUD	RUD
	Release	R	RUD	R
De-Identification	Key Creation	R		CRD
	Pseudonymization	R		CRUD
Pseudonymous Data Combination		R		CRUD
Data analysis	Resource Provision	R	CRUD	
	Support Analysis Tools	R	R	CRUD
Data Core	Data Pipeline	R	R	R
	Data Management	R	R	R

각각의 모듈 및 기능에 대한 설명은 다음과 같다.

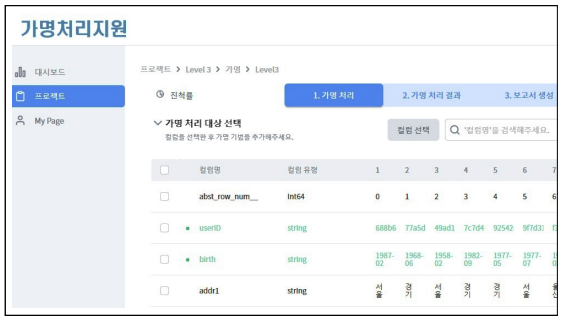
■ IDSZ UI Portal: 정당한 사용자가 접근하여, 허가받은 서비스와 데이터를 조작하여 분석할 수 있는 인증 및 접근권한 관리를 보장하고, 이를 기반으로 가명 처리, 가명 결합과 데이터안심구역 분석 시스템의 사용 신청, 심사, 반출 등의 사용자 인터페이스와 워크플로우를 제공한다. 이때 서브 시스템을 호출하여 상호연동하기 위한 제어 인터페이스를 제공한다. <Fig. 9>는 통합데이터안심구역 프로토타입 로그인의 SSO(Single Sign On)의 Token 발행에 대한 예시이다.

<Fig. 9> Main page and Log-In Procedure

■ De-Identification Module: 개인정보 등 민감정보는 가명 처리를 통해 가명 정보로 변환하여 사용되어야 한다. 데이터의 속성(Attribute) 중 주민번호 등과 같은 직접 식별이 가능한 주식별자(Primary Identifier)와 나이, 성별 등 추론을 통해 식별이 가능한 준식별자(Quasi Identifier) 및 민감 속성 등을 구분하여 비식별화 처리기법이 달리 적용된다. 이때 단방향 Hash 알고리즘 등을 통해 주식별자를 가명처리 하거나 삭제한다. 준식

별자는 총계처리, 데이터 삭제, 데이터 일반화, 데이터 마스킹 처리 등 <Fig. 10>과 같이 비식별화 기능을 포함한다.

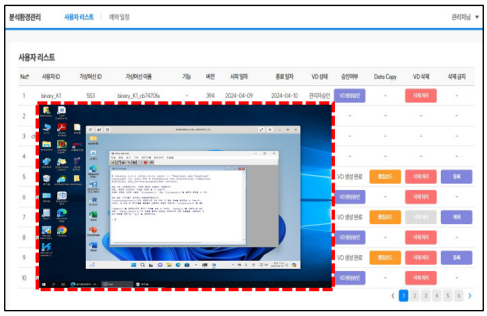
■ Pseudonymous Data Combination Module: 둘 이상의 사용자가 De-Identification Module에서 생성된 가명정보를 융합데이터로 만들기 위해 가명 정보 결합을 신청하는 경우, 결합을 처리하는 시스템 관리자는 가명정보의 결합키를 기반으로 Inner Join, Outer Join 등의 기능을 이용해 가명 정보를 결합하는 모듈이 <Fig. 11>과 같이 제공된다.



<Fig. 10> De-Identification page

<Fig. 11> Combination page

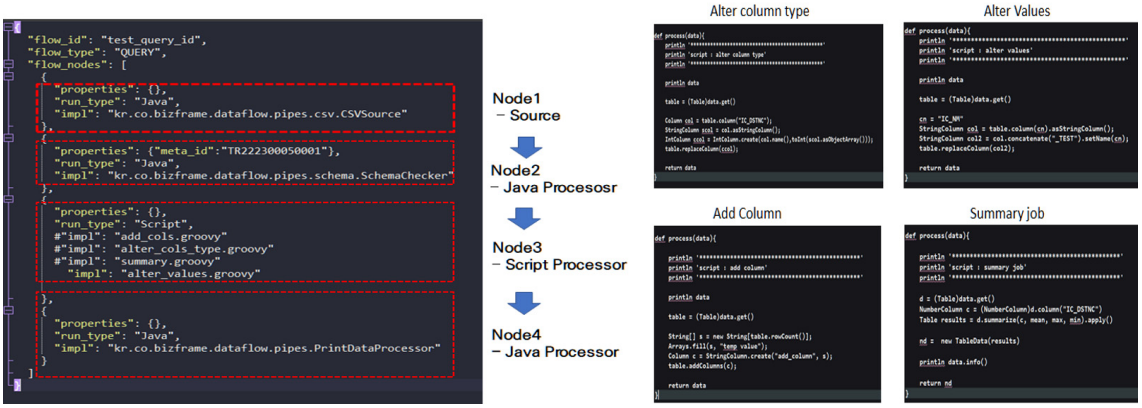
■ Data Analysis Module: UI Portal에서 사용자가 통합데이터안심구역 시스템의 사용을 신청 하는 경우, 사용자가 요청한 데이터안심구역의 분석 인프라(가상데스크톱, 메모리, 개인화 안심 저장소 등)를 <Fig. 12>와 같이 클라우드 기반으로 생성하여 제공하고, Pseudonymous Data Combination 모듈에서 결합 된 가명정보와 사용자가 반입한 미개방데이터 등을 분석하기 위한 맞춤형 분석 도구를 제공하는 프로비저닝 기능을 포함한다.




```
<command index="11" name="Start VM" async="false">
<request>
<element order="0" name="vm_name" type="string" size="32" key="1"/>
<element order="1" name="config_path" type="string" size="128" />
</request>
<response>
<element order="0" name="vm_name" type="string" size="32" />
<element order="1" name="state" type="integer" size="1" />
<element order="2" name="status" type="integer" size="1" />
<element order="3" name="message" type="varchar" />
<element order="4" name="vm_id" type="integer" size="2" />
</response>
</command>
```

<Fig. 12> Virtual Desktop and Analysis Infra (Left: VDI Provisioning page, Right: VDI Start-up XML script)

■ Data Core Module: 각각의 서브 시스템에서 입력, 처리 및 출력되는 데이터가 동기 및 비동기 방식, 실시간 및 배치처리 방식으로 교환되고, 이를 관리하는 데이터 파이프라인과 시계열분석이 가능하도록 데이터를 안전하게 저장 및 추적관리 하는 데이터 관리기능을 포함한다. <Fig. 13>은 메타데이터의 데이터 파이프라인 처리 흐름을 정의한 JSON(JavaScript Object Notation) 파일과 데이터의 처리 기능의 예시 코드이다.



<Fig. 13> Code of Data pipeline (Left: Flow Definition JSON, Right: Data pipeline sample code)

4. Prototype Performance Evaluation Indicator

본 연구에서 제시하는 성능평가지표는 Cho et al.(2017)이 제시한 2단계 성과측정모델을 준용하여 성능 평가지표를 도출하였다. 1단계로 Reijers and Mansar(2005)가 제시한 29개의 Best Practice 재설계 가이드라인을 경험적(Heuristic)인 방법에 따라 기존 시스템의 워크플로우를 분석하는 과정에 대입하여 핵심 성능평가지표 도출을 위한 12개의 구현 지표(BPIs: Best Practice Implementation Indicator)를 식별하였다. 2단계로 식별된 BPIs와 연결되는 성능평가지표(PPIs: Process Performance Indicator)로 구성하였다. 이때 성능평가지표는 Devil's Quadrangle 모델의 Time, Cost, Quality, Flexibility의 4가지 영향 요인에 그룹핑하여 IDSZ 프로토타입의 성능을 평가하였다. 본 연구에서 제시하는 프로토타입의 PPIs와 BPIs의 요약은 <Table 3>과 같으며, 재설계한 IDSZ 프로토타입의 성능평가에 대한 Best Practice와의 비교를 위해 Reijers and Mansar(2005)에 의해 제시된 개괄적인 성능(Positive(+), Negative(-), Neutral(0))도 함께 제공한다.

재설계에 있어 경험적(Heuristic) 접근법은 유용하나, 이러한 접근방법론은 성능에 영향을 미치는 요인을 측정할 양적 데이터를 확보하기가 매우 어렵다. 본 연구에서는 성능 측정에 대한 차원의 특성을 반영할 수 있는 양적인 성능평가지표를 재설계 시 도출하고, 프로토타입에 대한 성능을 평가할 수 있도록 정량화하였다. 또한, Devil's Quadrangle의 4가지 차원(Time, Cost, Quality, Flexibility)을 도입하여 직관적으로 평가를 확인할 수 있도록 하였다. 성능평가지표의 도출에 있어 시간(Time) 및 비용(Cost)의 차원은 측정된 값을 계산하는 것이 매우 직관적이나, 시스템의 프로세스 또는 워크플로우의 품질을 평가하는 것은 어렵고, 심지어는 불가능할 수도 있다(Jansen-Vullers et al., 2008).

본 연구에서는 품질(Quality) 차원에 대해 IDSZ가 민감한 데이터를 입력받고, 이를 비식별화 등을 통해 가명정보 등으로 가공한다는 점과 가공 시 필연적으로 발생하는 비식별화 수준별 가명 정보의 품질이 원본의 속성을 훼손한다는 점에 착안하여, 최종 가공된 가명정보가 데이터의 원본과 얼마나 유사한지를 측정할 수 있는 대리 지표로 데이터 유사성(Data Similarity)을 품질의 평가지표로 사용하였다. 이는 데이터를 지나치게 가명 처리 함으로써 원본과의 유사성을 읽어버리게 하거나 크게 훼손한다면 가명 처리된 데이터는 그저 쓸모없는 바이트로, 사용자는 분석 및 활용에 의미를 잃어버리게 된다. 따라서, 데이터의 효용성을 높이는 것이 IDSZ 프로토타입에 있어 가장 중요한 품질 차원의 성능으로 본다면, 중소기업, 스타트업 등 보안 환경의 구비가 어려운 사용자가 민감한 데이터를 안전하게 생성 및 가공, 분석 및 활용하기 위해 IDSZ를 이용함으

<Table 3> Overview of BPIs and PPIs

Best Practice	BPCode	BPIs	PPIs			
			Time (PPIT)	Cost (PPIC)	Quality (PPIQ)	Flexibility
Control Relocation	C1	· Compare Before/After Process (BPI1)	0	-	+	0
Contact Reduction	C2	· Compare Before/After Process (BPI1)	+	-	+	0
Integration	C2	· Compare Before/After Process (BPI1)	+	+	0	-
Task Elimination	B2	· Compare Before/After Activities (BPI2)	+	+	-	0
Task Composition	B5	· Compare Before/After Activities (BPI2)	+	+	0	-
Resequencing	B6	· Compare Before/After Process (BPI1)	+	+	0	0
Case assignment	O1	· Compare Before/After Resources (BPI3)	0	0	+	-
Numerical Involvement	O2	· Compare Before/After Resources (BPI3)	+	-	0	-
Buffering	I2	· Compare Before/After Process (BPI1)	+	-	0	0
Task Automation	T1	· Compare Before/After Resources (BPI3)	+	-	+	-
Integral Technology	T2	· Compare Before/After Process (BPI1)	+	-	0	0
Interfacing	E3	· Compare Before/After Process (BPI1)	+	0	+	-

로써, 데이터 처리환경의 위험도를 낮추고, 데이터 생명주기 전 과정에서 추적 및 관리를 가능케 함으로써 데이터 구성의 위험도를 낮출 수 있다(Choi, 2020).

또한, 유연성(Flexibility) 지표는 시스템의 변경이나 변화에 얼마나 빠르게 반응 또는 적응하는지 측정할 수 있는 지표로, 가명 정보 등 민감한 데이터의 생명주기 전 과정을 안전하게 지원하는 IDSZ 시스템은 변경의 유연성이 제약되어야 하는 특성을 고려하여, 기존 시스템과 IDSZ의 유연성 지표는 동일한 것으로 고정하여 평가를 진행하였다. <Table 4>는 도출된 최종 성능평가지표이다.

<Table 4> Performance Evaluation Indicators

Dimension	PPI	Definition
Time	• Lead Time	• Compare Before/After Overall Processing Time (BPI1)
Cost	• TCO (Fixed costs) • Execution Cost of Activities(Variable costs)	• Compare Before/After Total Cost Ownership (BPI3) = ∑(Asset + Management + Support) • Compare Before/After # of Activities (BPI2)
Quality	• Data Similarity	• Compare Before/After Data Similarity
Flexibility	Considering the sensitivity of the data handled by the Legacy systems and IDSZ prototype, this indicator has been designated as a fixed metric in the evaluation	

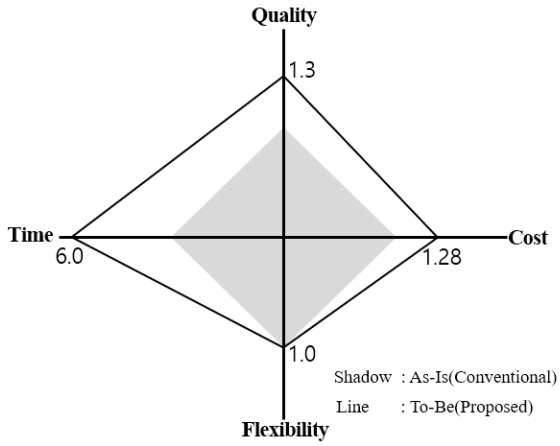
Ⅳ. IDSZ 프로토타입 성능평가

시간 차원의 성능평가를 위해 기존 가명정보 결합전문기관인 A 기관이 2022년부터 2023년까지 진행한 가명정보 결합 사례의 Lead Time(사용자가 사용요청에서 가명처리, 가명정보 결합, 분석 및 최종 결과물 반출까지 걸리는 전체 시간)과 데이터안심구역 운영기관인 B 기관의 2023년 진행한 데이터안심구역 이용 평균 Lead Time을 산출하였다(Jansen-Vullers et al., 2008). 또한, IDSZs 프로토타입의 재설계 시 도출된 워크플로우

상의 활동 감소(86개 → 30개)부분을 산입하여 성능을 비교하였다. 기존 가명처리 및 결합 시스템의 Lead Time은 평균 158일이 소요되었으며, 데이터안심구역 운영시스템의 경우 평균 10일 소요되었다. 이에 반해 IDSZ의 Lead Time은 평균 28일로 기존 시스템 대비 최소 6배의 시간 단축 효과(+)를 확인할 수 있었다.

비용관점에서의 성과측정 모델들로는 TVO(Total Value of Opportunity), TEI(Total Economic Impact), TCO(Total Cost of Ownership) 등이 있다. TVO 모델은 정성, 정량적으로 성과의 측정이 가능하나 과정이 복잡하고, 어렵다는 단점이 있다. TEI 모델은 주관적이며, 비 통계적이라 비용 측면의 지표로 적합하지 않다. TCO는 전략적 목표와의 연계가 다소 어렵다는 단점이 있으나 측정이 쉽고, 통합 전과 후의 차이 분석이 명확하다는 점에서 이점이 있다(Jung and Ra, 2008). TCO 모델은 시스템의 하드웨어, 소프트웨어 구축의 자산비용(Asset), 외부 용역이나 자체 인건비에 따른 관리비용(Management)과 공간임대, 네트워크, 정보보호 등의 지원비용(Assist)으로 나눈다(Jung and Ra, 2010). 본 연구에서는 TCO 모델을 이용하여 비용 차원의 성능을 평가하였으며, 기존 시스템 대비 IDSZ의 효과가 1.28배 비용 절감 효과(+)가 나타났다.

품질 차원의 성능평가는 개인정보의 가명 처리의 엄격한 기준으로 인해, KISA에서 가명 처리 및 결합을 위해 교육용으로 제공하는 가상의 개인정보가 포함되어있는 데이터를 이용하여 Choi(2020)가 제안한 가명처리 위험도 측정 프레임워크의 수준별 가명처리 기법을 적용하였다. 품질 평가 결과, 기존 시스템 환경(Before)에서 가명 처리한 결과와 IDSZ의 안전한 환경(After)에서 가명 처리한 결과의 데이터 유사성(Data similarity)이 1.3배 향상(+)된 것을 확인할 수 있었다.



<Fig. 14> Performance of redesigned IDSZ

<Fig. 14>는 Devil's Quadrangle의 4가지 차원을 이용하여 성능을 평가한 요약이다. 시간 차원의 성능은 워크플로우의 최적화로 기존 대비 6배의 성능향상을 보였으며, 비용 및 품질 차원의 성능은 각각 1.28배, 1.3배의 다소 완만한 성능향상을 보였다. 시간 차원의 극적인 성능향상에도 불구하고, 비용의 효과가 그렇게 크지 않은 것으로 나타나는데, 이는 민감한 정보를 다루는 시스템의 특성상 기본적인 소프트웨어, 하드웨어 외에 운영인력과 보안시스템 등의 적용에 따른 시스템 구축 및 운영의 초기 투자 비용이 높다는 것으로 차원의 영향 요인 간 상쇄 관계(Trade-Off)를 어느 정도 보여준다. 앞장에서 논의한 바와 같이 유연성 지표는 시스템이 보안 요구사항이 엄격하다는 점과 시스템의 잦은 변경에 따른 보안 침해 영향을 최소화하기 위해, 기존 시스템과 IDSZ 프로토타입의 유연성 성능을 같은 수준으로 평가하였다.

Ⅴ. 결 론

최근 우리나라에서는 데이터 산업 활성화를 위한 관련 법제도 정비 등의 노력에 힘입어, 정부는 기업들이 자신이 보유하거나, 제3자로부터 제공받은 개인정보가 포함된 모빌리티 데이터 등 교통 관련 데이터를 가명처리하여 상호결합, 분석 및 활용하도록 하여 기업들이 창의적인 먹거리나 서비스를 창출할 수 있도록 새로운 길을 열어주었다. 하지만, 중소기업이나 스타트업 등의 사용자는 개인정보를 가명 처리해 가명 정보를 생성하고, 결합 및 활용하기에는 이를 처리하는 전문기관들이 물리적으로 분리되어 운영됨에 따라, 매우 복잡한 처리 절차와 많은 시간이 소요된다. 또한, 가명 처리 및 가명정보 결합 과정에서 개인정보의 유출 등을 우려해 과도하게 가명 처리를 요구하는 등 데이터 품질이 지나치게 훼손되어, 가명정보 활용의 실효성이 매우 떨어지는 실정이다.

이에 본 논문은 개인정보 등 민감한 데이터를 가명 처리하여 이용하고자 하는 사용자와 가명정보의 결합을 지원하는 결합전문기관 및 결합키관리기관, 데이터의 안전한 분석환경을 지원하는 데이터안심구역 운영기관이 관리하는 시스템이 물리적으로 분리되어 있어 발생하는 문제점을 해소하고자 하였다. 이를 위해 복잡한 처리 절차를 간소화하여 병목에 따른 처리시간을 단축하고, 데이터 생명주기 전 과정에서 민감한 데이터의 투명한 추적 및 관리로 개인정보가 유출 및 오남용되지 않도록 데이터 생명주기 전 과정을 안전하게 관리하는 통합데이터안심구역 프로토타입을 제안하였다.

이에 Lee and Yoo(2023)가 제시한 통합데이터안심구역 시스템 요구사항, 개념 프레임워크 및 아키텍처를 기반으로 핵심 모듈을 선별하고, 비즈니스 프로세스 재설계 진단 및 실행 가이드라인을 적용하여 워크플로우를 재설계한 후, 프로토타입을 구현하고 성능을 평가하였다. 성능평가 결과, 시간 측면에서는 6배, 비용 측면에서는 1.28배, 품질 측면에서는 1.3배의 성능이 향상되는 것을 확인하였다.

본 연구에서 IDSZ 프로토타입을 구현해 우수한 성능을 검증하였지만, 실무의 적용에 있어서는 한계가 존재한다. 첫째, IDSZ 프로토타입의 검증에 있어 민감한 데이터를 처리하는 시스템의 특성으로 제한된 데이터와 시스템 환경에서 프로토타입을 검증하였다는 한계가 있다. 실제 실무를 위해서는 정부의 규제샌드박스 특례 신청을 통해, 실 환경에서 추가적인 검증이 필요하며, 이를 통해 데이터 3법 및 데이터 산업법 등의 법과 제도적 지원 장치를 보완할 필요가 있다. 둘째, 재설계 진단 및 실행 가이드라인의 경험적(Heuristic) 특성과 본 연구가 프로토타입 구현이라는 한계로 필수 성능평가지표만을 도출하여 검증하였으나, 향후 추가적인 성능지표의 도출과 프로세스 마이닝 기법 등의 적용을 통해 지표를 보다 정량화하고, 지속적인 성능향상 및 개선을 위한 방법론의 연구가 필요하다.

이러한 한계에도 불구하고, 본 연구는 첫째, 기존 워크플로우 재설계 시, 관행적으로 설계하던 방식에서 벗어나 다양한 산업군의 Best Practice에서 발견한 '재설계 진단 및 실행 가이드라인'을 적용하여 개선점을 도출하고, IDSZ 프로토타입의 워크플로우를 재설계하는 과정을 절차별로 제시함으로써 추후 유사한 시스템을 구축하는 실무자에게 유용한 재설계 접근법을 제공하였다. 둘째, 민감한 데이터를 활용하는 시스템에 필요한 성능평가지표를 제시하고, 데이터 생명주기 전 과정을 지원하는 프로토타입을 구현해 성능을 평가함으로써 아이디어의 유효성을 검증하였다는 점에서 가치가 있다.

REFERENCES

Cho, M. S., Song, M. S., Comuzzi M. and Yoo, S. Y.(2017), "Evaluating the effect of best practices for business process redesign: An evidence-based approach based on process mining techniques," *Decision Support Systems*, vol. 104, pp.92-103.

Choi, K. H.(2020), *A Study on policy framework for utilizing risk-based pseudonymized data*, Graduate School Chonnam National University, pp.133-181.

Haffey, M. K. D. and Duffy, A. H. B.(2001), "Process performance measurement support – A critical analysis," *13th International Conference on Engineering Design (ICED)* 01 GLASGOW August, pp.21-23.

Hammer, M.(1990), "Reengineering work: Don't automate, obliterate," *Harvard Business Review*, vol. 68, no. 4, pp.104-112.

Jansen-Vullers, M. H., Kleingeld, P. A. M., Loosschilder, M. W. N. C., Netjes, M. and Reijers, H. A.(2008), "Trade-offs in the performance of workflows – Quantifying the impact of best practices," *Information Systems Management*, vol. 25, no. 4, pp.332-343.

Jung, H. Y. and Ra, J. H.(2008), "A study on development of integration performance measurement model for each stage of information systems integration and measurement indicators of physical integration stage," *Journal of Information Technology Services*, vol. 7 no. 4, pp.247-268.

Jung, H. Y. and Ra, J. H.(2010), "A case study of the economic performance measurement of information system integration in public sector," *Journal of Digital Convergence*, vol. 8, no. 4, pp.185-204.

Kim, J. S.(2020a), "Research on the use of pseudonym data: Focusing on technical processing method and corporate utilization directions," *Journal of The Korea Institute of Information Security& Cryptology*, vol. 30, no. 2, pp.253-261.

Kim, K. B. and Kwon, H. Y.(2023), "Improvement plan to expand the role of expert data combination agency," *Journal of The Korea Institute of Information Security & Cryptology*, vol. 33, no. 1, pp.99-116.

Kim, K. H.(2023), "A study on the improvement of the legal system related the data safe zone for the promotion of small and medium-sized enterprises," *Ajou Law Review*, vol. 16, no. 4, pp.7-30.

Kim, S. G. and Kim, S. K.(2020), "An exploration on personal information regulation factors and data combination factors affecting big data utilization," *Journal of the Korea Institute of Information Security & Cryptology*, vol. 30, no. 2, pp.287-304.

Kim, S. O.(2020b), "A study on the balancing rational use and safe processing of pseudonymous data: In addition to the constitutional evaluation of 3 acts regarding to data," *Public Law*, vol. 49, no. 2, pp.371-407.

Korea Data Agency(2024), *A Report on Sustainable Smart Transportation Systems in the Asia-Pacific Region*, pp.3-5.

Lee, C. K. and Jung, P. W.(2018), "Legal issues arising in developing and operating a transportation platform," *Hongik Law Review*, vol. 19, no. 4, pp.337-370.

Lee, H. K. and Yoo, K. D.(2023), "System architecture of the integrated data safety zone for the secured application of transportation-specific mobility data," *The Journal of The Korea Institute*

of Intelligent Transport Systems, vol. 22, no. 3, pp.88-103.

Mansar, S. L. and Reijers, H. A.(2007), "Best practices in business process redesign: Use and impact," *Business Process Management Journal*, vol. 13, no. 2, pp.193-213.

Ministry of Science and ICT(2022), *Guidelines on the designation & operation of data safety zone*, pp.1-15.

Nam, J. H.(2009), "Process innovation methodology," *Korean Journal of Business Administration*, vol. 22, no. 3, pp.1337-1356.

Personal Information Protection Commission(2022), *Guidelines on pseudonym information processing*, pp.7-64.

Reijers, H. A. and Mansar, S. L.(2005), "Best practices in business process redesign: An overview and qualitative evaluation of successful redesign heuristics," *The International Journal of Management Science*, vol. 33, pp.293-306.

Sim, J. H.(2023), "Federated learning-based route choice modeling for preserving driver's privacy in transportation big data application," *The Journal of The Korea Institute of Intelligent Transport Systems*, vol. 22, no. 6, pp.157-167.

The Korea Transportation Institute(2017), *Studies on developing transport big data platform and its application*, pp.85-86.

Van Den Ingh, L., Eshuis, R. and Gelper, S.(2020), "Assessing performance of mined business process variants," *Enterprise Information Systems*, vol. 15, no. 5, pp.1-18.

Yoo, K. D., Suh, E. H. and Kim, K. Y.(2007), "Knowledge flow-based business process redesign: Applying a knowledge map to redesign a business process," *Journal of Knowledge Management*, vol. 11, no. 3, pp.104-125.