

빅데이터 플랫폼 및 센터

2021

데이터 품질관리 가이드

제 1권 데이터 관리체계 >>



I

총론 >>

1장	추진 배경 및 목적	
	1. 추진배경	07
	2. 목적	08
2장	품질관리 가이드 구성 및 활용	
	1. 품질관리 가이드 구성	09
	2. 데이터 품질관리 가이드 활용	10

II

데이터 관리체계 >>

1장	개요	
	1. 데이터 관리 정의	15
	2. 데이터 관리 목적	17
	3. 데이터 관리 대상	18
	4. 데이터 관리 영역	20
	5. 데이터 관리체계 구성	21
2장	데이터 품질관리체계	
	1. 개요	41
	2. 모델 수립	44
	3. 체계 구성	47
3장	데이터 연계관리체계	
	1. 개요	67
	2. 모델 수립	71
	3. 체계 구성	78
4장	데이터 정보보호 관리체계	
	1. 개요	102
	2. 모델 수립	105
	3. 체계 구성	125

I

총론 >>

1장 추진 배경 및 목적

2장 품질관리 가이드 구성 및 활용

01

추진 배경



디지털 기술의 발달은 기존에 볼 수 없었던 새로운 비즈니스 모델이 등장하는 것을 가능하게 했다. 디지털 기술 시대의 새로운 비즈니스 기반인 데이터는 데이터 경제를 가능하게 하고 있어 양질의 데이터 확보가 기업뿐 아니라 국가 경쟁력을 좌우하게 되었다. 애플, 구글, 아마존 등 데이터 보유 기업들이 글로벌 시장을 선도하는 등 이미 주요 선진국은 데이터로 경제적 효과를 창출하는 데이터 경제로 산업 구조를 바꾸고 있다.

데이터 경제에서 자본은 곧 데이터이며, 기존 경제에서 금융 자본의 역할은 이제 데이터 자본이 대신해 가고 있다. 데이터 경제의 효과는 데이터 선순환으로 경쟁력이 확대·재생산되는 데이터 네트워크 효과 발생 및 수확 체증의 법칙 작용과 파괴적 혁신의 일상화로 성장하고 있다.

빅데이터의 가치는 데이터가 자산으로서 갖는 특성에 따른 가치와 기술 발전으로 과거에는 처리하지 못하던 것을 처리 및 분석함으로써 경제적, 사회적 가치 창출에 활용되고 있다.

빅데이터는 데이터를 기반으로 사회현상 및 비즈니스의 불확실성 및 리스크의 감소, 또 생성되는 데이터를 활용하여 새로운 비즈니스의 창출 등의 스마트한 경쟁력을 갖추는데 가치를 가지고 있다.

우리나라는 폐쇄적 데이터 축적과 접근성 제약으로 인해 데이터에 대한 인식과 데이터 품질관리 등 기반체제 마련이 미흡한 실정이다. 정부는 데이터 산업 활성화 전략 수립(2018.06), 기획재정부의 혁신성장 전략 투자 분야 선정(2018.08) 등 데이터의 중요성을 강조하고 있으며, 데이터 및 AI 경제 활성화 계획을 발표(2019. 01)하여 글로벌 데이터 경쟁에 본격적으로 뛰어 들었다.

또한 과학기술정보통신부에서는 2025년까지 15개 분야별 대표적 빅데이터 플랫폼 추가 구축, '통합 데이터지도'로 다양한 데이터 플랫폼 연계, 데이터 플랫폼을 분석·서비스 중심으로 혁신하는 민·관 협력 기반 데이터 플랫폼 발전전략(2021.06)과 비식별처리, 동형암호, 차등 정보보호 등 데이터 보호 핵심기술 중점 투자 등이 포함된 데이터보호 핵심기술 개발 전략(2021.11)을 발표하여 데이터 기반 분석 및 서비스체계를 마련하고 있다.

이와 같이 정부에서는 빅데이터 플랫폼 및 센터 구축을 통해 데이터 축적 및 유통을 활성화하고 데이터 수집, 분석, 유통 체계를 마련하여 데이터 생태계의 경쟁력을 한 단계 높이려 하고 있다.

02

목적



최근 정부는 데이터 축적 및 유통 활성화를 위해 빅데이터 플랫폼 및 센터 구축 사업을 추진하여 데이터 수집, 분석, 유통 체계를 마련하여 데이터 생태계의 경쟁력을 한 단계 높여 하고 있으나, 아직도 많은 기관 및 기업에서 데이터베이스의 품질 향상을 위해 무엇이 필요하고, 어떻게 해야 하는지 명확한 이해가 부족하다.

특히, 정보화 사업을 계획하거나 추진하는 데 있어 데이터 품질 관점에서 무엇을 고려하고 어떻게 해야 하는지에 대해 전반적인 인식이 부족한 상황이다.

이러한 상황에서 빅데이터 플랫폼 및 센터에서 구체적인 가이드 없이 데이터 품질관리를 수행할 경우, 데이터 품질관리에 대한 이해 부족으로 인한 시행착오 및 양질의 데이터 미확보, 신규 사업 추진 부담 등 여러 가지 부작용이 발생할 수 있어 각 기관에서 편리하게 활용할 수 있는 구체적인 데이터 품질관리 가이드가 필요하다.

그러므로 품질관리 가이드는 데이터 품질관리에 대한 이해를 증진시키고 구체적인 품질관리 기법을 제공하여 빅데이터 플랫폼 및 센터 관련 기관 및 기업들이 효율적으로 데이터 품질관리 활동을 수행할 수 있도록 하는 목적으로 제공된다.

본 품질관리 가이드는 빅데이터 플랫폼 및 센터의 데이터 품질관리 담당자가 데이터 품질관리 활동을 수행하는데 활용할 수 있도록 하였으며, 사업에 참여하는 구축 사업자, 운영 사업자 및 감리업자도 품질관리 측면에서 고려해야 할 사항 등을 참조할 수 있다.

01

품질관리 가이드 구성



품질관리 가이드는 크게 5개 영역으로 구성되어 있다.

- I 총론**에서는 품질관리 가이드의 개발 배경과 목적을 제시하고, 품질관리 가이드에 대한 전반적인 구성과 기관의 담당자가 본 품질관리 가이드를 어떻게 활용할 수 있는지에 대해 설명한다.
특히, 데이터 품질관리에 대해 이해도가 높지 않은 점을 고려하여, 데이터 품질관리가 무엇인지에 대한 개념을 설명한다.
- II 데이터 품질관리체계**에서는 빅데이터 플랫폼 및 센터가 수행하는 데이터 품질관리의 모델과 체계를 제시하고, 각 요소별 의미와 주요 내용을 설명한다.
데이터 품질관리체계는 빅데이터 생애주기를 고려하여 빅데이터 품질관리를 위한 수집, 가공·분석, 활용 단계별 품질관리 기능 및 프로세스를 도출하고, 데이터 품질관리 수행사항에 대하여 설명한다.
- III 데이터 품질관리 기준**에서는 데이터 품질관리의 관점이 되는 데이터 품질관리 지표를 제시하고, 각 지표의 의미 주요 내용을 설명한다.
데이터 품질관리 기준은 빅데이터 환경을 고려하여 도출된 데이터 품질관리 지표와 각 지표별 세부지표 및 진단기준, 진단기법을 정량적, 정성적으로 설명한다.
- IV 데이터 품질진단 방법**에서는 빅데이터 특성을 반영하여 플랫폼 및 센터에서 가시적인 데이터 품질을 향상시키기 위한 품질 진단 및 개선에 대한 단계별 절차를 설명하고, 품질관리 담당자가 각 단계별 활동에서 고려해야 할 사항과 주의할 점 등을 설명한다.
- V 부록**은 데이터 품질진단을 수행할 때, 작성해야 하는 데이터 품질진단 산출물과 데이터 품질진단 규칙을 도출할 때 참고할 수 있는 다양한 예시를 제시하고 설명한다. 또한 데이터 연계기, 연계표준사전, 가명정보처리 가이드, 소유권 및 이용약관 개선 정 의서 등을 제시한다.

본 품질관리 가이드는 각 장별로 품질관리 담당자의 이해를 돕기 위해 안내 팁(Tip), 고려사항, 사례, 예시 등을 포함하였으며, 필요 서식 및 품질관리 활동에 도움이 되는 자료들을 붙임으로 첨부하여 필요할 때 참조할 수 있도록 하였다.

02

데이터 품질관리 가이드 활용



데이터 품질관리 가이드는 빅데이터 플랫폼 및 센터의 품질관리 담당자를 주요 대상으로 하며, 업무 담당자, 정보화 담당자, 빅데이터 플랫폼 구축 및 운영 사업자, 감리사업자 등 품질관리 관련 업무를 수행하는 실무자들이 데이터 품질관리를 효율적으로 수행할 수 있도록 실무 내용 중심으로 구성하였다.

- I. 총론은 데이터 품질관리에 대한 기본적 내용으로 데이터 품질관리 수행에 대한 개념을 파악하기 적합하다.
- II. 데이터 관리체계는 빅데이터 환경을 고려한 주요 관리 대상 요소인 품질, 연계, 정보보호와 관련하여, 데이터 관리를 위해 기본적으로 구성하고 수행해야 하는 요소를 파악할 수 있다.
- III. 데이터 품질관리 기준은 데이터 품질관리를 위한 데이터 진단 관점의 기준을 제시하고 관점인 지표별로 데이터 품질관리 및 진단을 어떻게 수행하는지 체계적으로 파악할 수 있다.
- IV. 데이터 품질진단 방법은 실제 데이터 품질진단을 수행하는 방법에 대한 설명으로, 데이터 품질 대상 선정에서 진단 수행 및 오류에 대한 분석, 개선 수행 등 데이터 품질진단에 대한 실무적인 내용을 자세하게 확인할 수 있다.
- V. 부록은 데이터 품질진단을 수행할 때, 필요한 산출물의 종류와 내용을 파악하고 작성할 때 참고하며, 진단규칙을 도출하는 예시를 통해 품질진단 규칙을 도출에 필요한 주요 내용을 자세하게 확인할 수 있다. 또한 데이터 연계키, 연계표준사전, 가명정보처리, 소유권 및 이용약관 개선 등에 활용할 수 있다.

데이터 품질관리 가이드는 다양한 사례를 포함하여 실무자들이 이해하기 쉽도록 구성하였으며, 핵심적인 내용이나, 고려사항, 추가 설명이 필요한 사항은 별도로 구성하여 데이터 품질관리 업무에 쉽게 적용할 수 있도록 구성하였다.

다양한 지원 표현

- 활용 예시 : 내용, 용어의 설명이나 이해를 돕기 위한 부가 설명
- 체크포인트 : 주의사항, 고려사항, 전제조건, 선결사항 등 안내
- 사례 : 해당 본문과 관련되어 설명이 필요한 관련 사례, 자료 등 제공
- 관련근거 : 관련 법령, 지침, 규정, 등 관련 근거 제시

II

데이터 관리체계 >>

1장 개요

2장 데이터 품질관리체계

3장 데이터 연계관리체계

4장 데이터 정보보호 관리체계

01

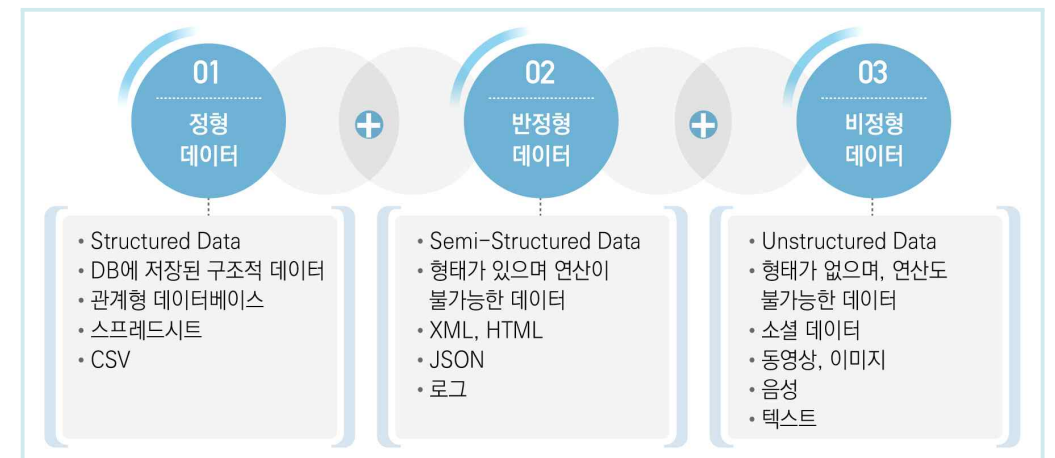
데이터 관리 정의



1.1 데이터

데이터(Data)는 “이론을 세우는 데 기초가 되는 사실, 자료”라고 정의되어 있으며, IT 분야에서는 “프로그램을 운용할 수 있는 형태로 기호화, 숫자화한 자료”라고 정의하고 있다. 대상 데이터는 프로그램을 운영할 수 있는 형태, 즉 디지털화 한 기호 또는 숫자로 구성되어 있는 자료로, 빅데이터 환경에서는 정형데이터와 비정형 데이터를 모두 포함한다.

그림 1 데이터의 유형



데이터의 관리는 “조직의 목표 달성에 필요한 데이터를 확보하고 데이터 사용자가 원하는 시간에 원하는 형태로 안정적으로 서비스를 제공하도록 데이터를 효과적으로 운영하기 위한 활동”으로 정의할 수 있으며, 데이터 관리 원칙, 데이터 관리 조직, 데이터 관리 프로세스를 통해 데이터 관리 활동을 수행한다.

1.2 데이터 품질관리

데이터 관리의 구체적인 형태로서 데이터 품질(Data Quality)은 “데이터를 활용하는 사용자의 다양한 활동 목적이나 만족도를 지속적으로 충족시킬 수 있는 수준”(Larry P. English)으로 정의할 수 있다.

데이터 품질은 데이터의 전반적인 관리 수준을 높임으로써 달성되는 것으로, 데이터 관리로 얻을 수 있는 구체적인 결과이다. 이러한 데이터 품질 수준을 확보하고 지속적으로 유지하기 위해서는 체계적인 관리 활동이 필요하다.

데이터 품질관리(Data Quality Management)는 “데이터의 품질을 확보하기 위한 품질 목표 설정, 품질 진단 및 개선 등 일련의 활동”으로 정의할 수 있으며, 이를 지원하기 위한 원칙, 조직, 절차와 관련 도구가 포함된다.

그림 2 데이터 품질관리 정의



데이터 품질관리는 조직의 목표와 비전을 달성하기 위한 조직의 정보화 전략의 중요한 부분으로 데이터 품질관리를 어떻게 하는지 정의하는 원칙, 이를 수행하는 조직체계, 원칙을 기반으로 한 품질관리 기능과 이에 따른 절차 등과 같은 요소로 구성된다.

02

데이터 관리 목적

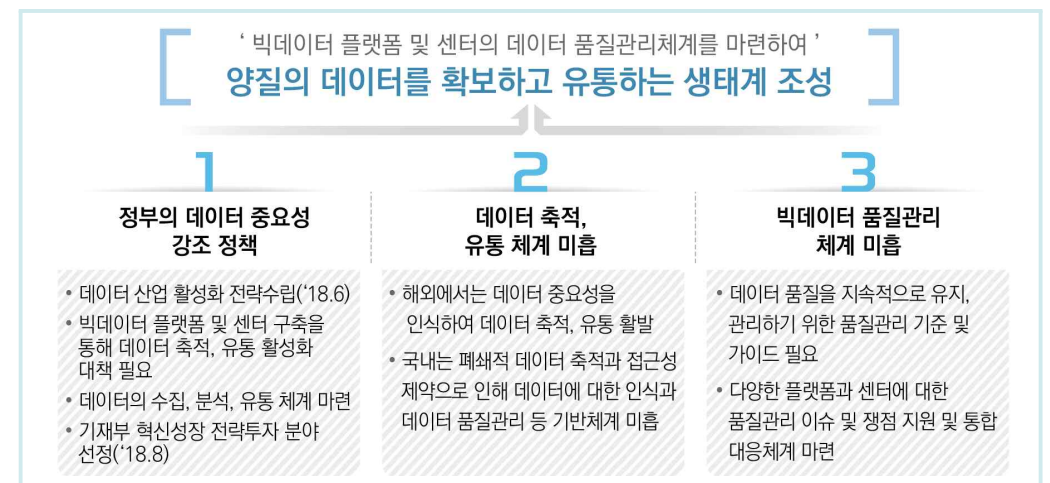


일반적인 데이터 관리의 목적은 정의에 나와 있는 바와 같이, “안정적인 서비스 제공 및 데이터의 효과적인 운영”이라고 볼 수 있다. 하지만, 데이터의 중요성이 강조되고 빅데이터 시대로 진입하고 있는 최근에는 변화된 의미로 데이터 관리 목적을 정의할 필요가 있다.

데이터를 활용하기 위한 요구가 높아지는 상황이나 폐쇄적인 데이터 축적 및 접근 제약 등 데이터를 축적하고 유통하는 환경의 미숙, 기존 정형데이터 중심의 관리 방법을 적용하는 등 빅데이터 트렌드를 반영하지 못한 데이터 관리 문제가 존재한다. 이에 정부는 데이터의 중요성을 강조하는 정책을 추진하고, 민간 분야의 빅데이터 환경을 반영한 데이터 관리에 관심을 높이고 있다.

이러한 최근 데이터 환경을 반영하여 데이터 관리의 목적은 “양질의 데이터를 확보하고 유통하는 생태계 조성”으로 정리할 수 있다. 양질의 데이터 확보를 위해서 빅데이터 환경을 고려한 데이터 관리체계, 데이터 관리기준 수립이 필요하고, 빅데이터 플랫폼 및 센터에 수립된 관리체계 및 기준을 제공하는 표준 가이드를 배포하여 활용될 수 있도록 해야 한다.

그림 3 데이터 품질관리 배경 및 목적



03

데이터 관리 대상



데이터 관리 대상은 데이터 관리 활동의 목적으로 무엇을 관리할 것인가의 관점에서 정의할 수 있다. 본 데이터 품질관리 가이드에서는 데이터 관리 대상을 값, 구조, 표준, 관리체계로 정의한다.

데이터를 정보로 처리하는 과정을 살펴보면, 데이터는 DBMS(Database Management System) 내에서 관리되며, 실세계 업무를 정확하게 반영하기 위해 개념 데이터 모델링, 논리 데이터 모델링, 물리 데이터 모델링으로 구조화하고 구조화된 데이터는 DB에서 테이블 객체(Object)로 정의된다.

모델링 단계의 엔티티나 DBMS 내의 테이블 이름은 데이터 일관성 유지를 위해 표준 데이터(표준 단어, 표준 용어, 표준 도메인)를 사용하고 데이터값도 표준데이터(표준 코드, 표준 도메인)의 영향을 받는다.

품질진단을 위한 핵심품질항목(CTQ: Critical to Quality)은 데이터 구조인 데이터 속성 또는 컬럼을 대상으로 도출하고, 핵심품질항목의 값은 업무규칙(BR : Business Rule)으로 검증한다.

이러한 데이터 처리는 지속적으로 수행할 수 있도록 규정, 역할, 절차를 포함하는 관리체계를 수립하여 적용하고 점검하며, 그 수준을 평가한다.

비정형 데이터는 표준 데이터(표준 용어, 표준 도메인)로 구성된 메타데이터의 정보 값을 검증하며 특히, 현실세계의 지리적 상태를 데이터화하여 표현하는 공간정보는 특정 도메인에 특화되어 있다.

이와 같이, 데이터를 정보로 처리하는 과정에서 데이터 구조, 표준, 값, 관리체계를 데이터 관리 대상의 주요 요소로 정의한다.

표 1 데이터 관리 대상

항 목	내 용
값	<ul style="list-style-type: none"> 업무 및 비즈니스를 수행하면서 생성되거나 필요한 제반 데이터로 데이터베이스에 실제로 저장되어 있는 정보항목의 값 수집된 값을 기반으로 하는 데이터를 가공하여 사용자에게 정확하고 적시에 제공하는 것이 데이터 품질관리의 핵심 수집된 데이터는 전자화되어 DB에 저장되며 전자화된 데이터는 정형데이터, 비정형 데이터(이미지, 동영상, 음성, GIS 등)로 구분할 수 있음
구조	<ul style="list-style-type: none"> 데이터가 담겨있는 틀로서 데이터를 취급하는 관점에 따라 구조가 변경됨 일반적으로 개념 데이터 모델, 논리 데이터 모델, 물리 데이터 모델 및 데이터베이스 객체로 구분하며 이러한 모델들의 계층적 구조를 데이터 아키텍처(Data Architecture)라고 함 데이터 아키텍처의 각 구성요소는 계층적 관계를 가지며, 각 계층간의 관계정보를 얼라인먼트(Alignment)라 함 계층 구조는 주제영역 > 개념모델 > 논리모델 > 물리모델 > DB 객체 구성요소로 계층화됨
표준	<ul style="list-style-type: none"> 정보시스템에서 사용하는 용어, 도메인, 코드 및 기타 데이터 관련 요소에 대해 공통된 형식과 내용으로 정의·사용하는 표준 관련 데이터를 의미 데이터 표준화 : 시스템 별로 산재해 있는 데이터 정보요소에 대한 명칭, 정의, 형식, 규칙에 대한 원칙을 수립 데이터 표준화는 데이터 모델 및 DB에서 정의할 수 있는 모든 객체를 대상으로 주로 관리할 필요가 있는 객체를 대상으로 데이터 표준화를 하는 것이 효율적이며, 일반적으로 데이터 표준으로 관리하는 대상에는 단어, 도메인, 용어, 코드 등이 있음
관리체계	<ul style="list-style-type: none"> 데이터 관리 활동을 효과적으로 수행하기 위해 품질관리 활동에서 준수해야 할 내용을 정의한 지침, 품질관리 활동을 수행할 조직, 품질관리 활동을 구체화한 절차 품질관리 활동을 지원하는 도구 및 관련 시스템 등 품질관리 인프라 활용 방법 이러한 데이터 관리 활동을 지원하는 요소들이 잘 갖추어져 있는지 점검

04

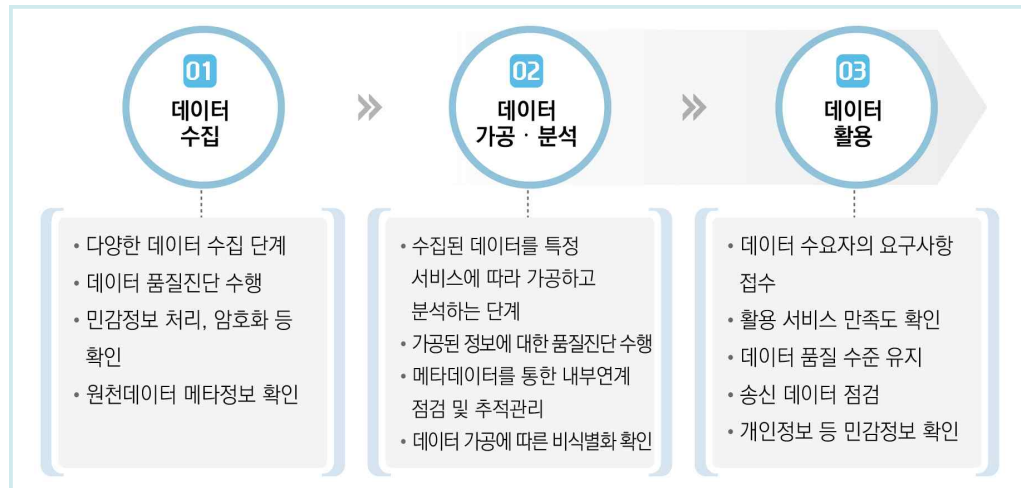
데이터 관리 영역



빅데이터 플랫폼 및 센터의 데이터 관리체계의 관리 영역은 데이터 수집, 데이터 가공·분석, 데이터 활용 단계 등 빅데이터 수 생애주기를 대상으로 한다.

- ▶ 다양한 데이터를 수집하고, 데이터 품질 진단을 수행하며 민감 정보 처리, 원천데이터 메타정보를 확인하는 수집단계
- ▶ 수집된 데이터를 특정 서비스에 따라 가공하고 가공된 데이터의 품질 진단을 수행하여 메타데이터를 통해 추적관리를 하는 가공·분석단계
- ▶ 데이터 수요자의 요구사항을 접수하여 데이터 활용 만족도를 확인하고 데이터 품질 수준을 유지하며 송신 데이터를 점검하는 활용 단계

그림 4 빅데이터 수 생애주기 단계 구분



05

데이터 관리체계 구성

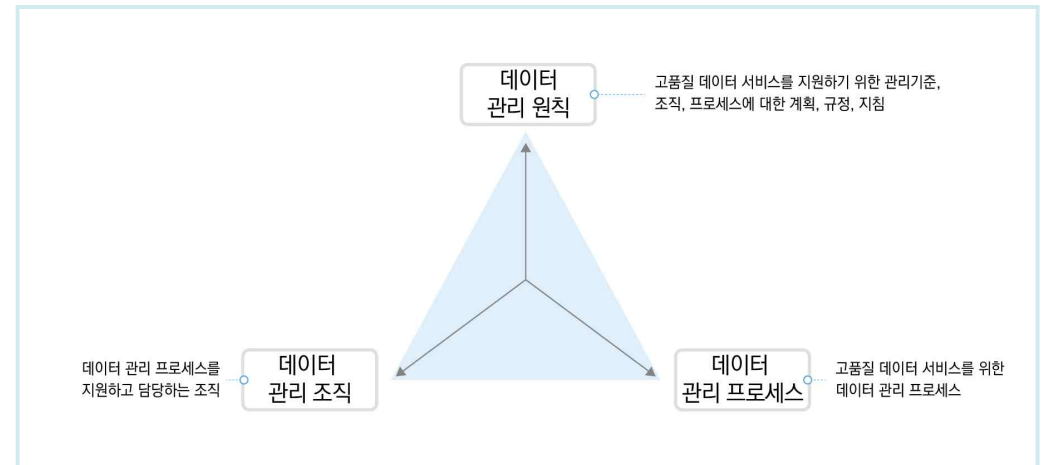


5.1 개요

데이터 관리를 위해서는 데이터 관리 활동을 정의하는 원칙을 기본으로 관리 활동을 정의한 관리 기능이 정의되어야 하며, 데이터 관리 기능을 수행하는 조직과 조직에 역할, 책임이 명확하게 정의되어야 한다. 또한, 어떠한 절차로 데이터 관리 기능을 수행할지 데이터 관리 기능에 따른 프로세스가 정의되어야 한다.

데이터 관리를 효율적이고 원활하게 수행하기 위해서는 데이터 관리 원칙, 데이터 관리 조직, 데이터 관리 프로세스 등 3가지 구성 요소를 기반으로 데이터 관리체계를 구성한다.

그림 5 데이터 관리체계 3가지 구성 요소



데이터 관리 요소는 데이터를 체계적으로 수행하기 위해 데이터 관리 영역을 구분한 것으로 일반적으로 데이터 표준관리, 데이터 구조관리, 데이터 품질관리를 대상으로 하며 빅데이터 특성을 반영하여 데이터 연계관리, 데이터 정보보호 관리를 강화하였다.

데이터 연계관리와 데이터 정보보호 관리는 빅데이터 환경을 고려한 데이터 관리 영역으로 대량 데이터를 다양한 원천에서 수집하는 빅데이터 특성을 고려하여 연계관리를 주요 데이터 관리 요소로 정의하였다.

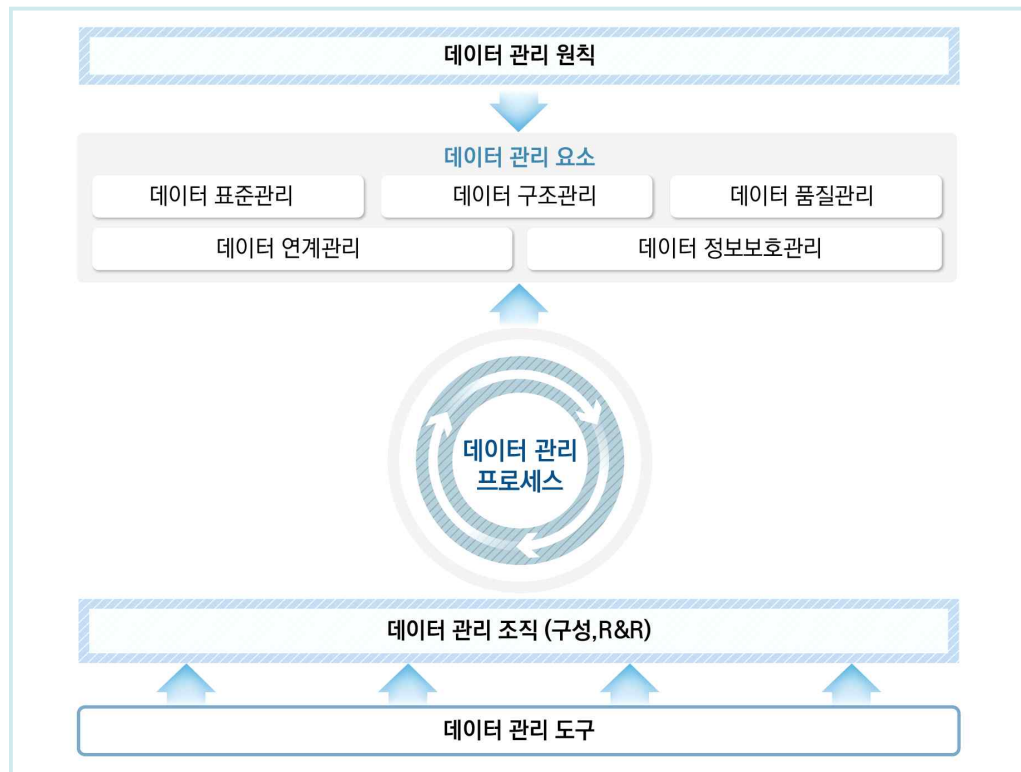
또한 최근 빅데이터 분야에서 이슈가 되는 개인정보보호를 보안관리 기능과 통합한 정보보호 관리를 주요 데이터 관리 요소로 정의하였다.

따라서 빅데이터 플랫폼 및 센터의 데이터 관리요소는 데이터 표준관리, 데이터 구조관리, 데이터 품질관리, 데이터 연계관리, 데이터 정보보호 관리 등 5개 관리요소로 정의한다.

본 가이드에서는 5개 관리요소 중 데이터 품질관리, 데이터 연계관리, 데이터 정보보호 관리 등 3가지 관리요소를 대상으로 작성되었다.

데이터 관리체계 3가지 구성 요소와 빅데이터 플랫폼 및 센터의 데이터 관리요소를 포함하는 빅데이터 플랫폼 및 센터의 데이터 관리체계는 다음과 같이 구성된다.

그림 6 데이터 관리체계 구성



▶ 데이터 관리 원칙 : 데이터의 효과적인 확보, 유지, 관리를 위해 수립된 규정이나 계획, 지침 등에 포함된 데이터 관리 내용

▶ 데이터 관리 프로세스 : 데이터 관리 기능별 데이터 관리를 위해 수행되고 있는 절차

▶ 데이터 관리 조직 : 데이터 관리를 담당하여 수행하는 이해관계자의 구성 및 역할

5.2 데이터 관리체계 원칙

5.2.1 원칙의 정의 및 목표

데이터 관리체계 원칙은 데이터 관리를 위한 기준 및 규정, 그 규정을 준수하기 위한 조직의 역할, 수행 절차의 기본이 되는 선언, 법률 등을 의미하며, 원칙이 정책과 프로세스 및 조직에 잘 반영되도록 하는 것을 목표로 한다.

그림 7 원칙 정의 및 목표



▶ 원칙(Principle)

- 데이터 관리체계 수행을 위한 기본적이고 종합적인 법, 선언, 정의를 의미하며 데이터 관리체계 활동 및 정책의 지침으로 활용

▶ 정책(Policy)

- 데이터 관리체계 활동 및 의사결정을 위한 기준 및 규정을 의미

▶ 프로세스 및 조직(Process/Organization)

- 원칙에 기반을 둔 정책을 수행하는 조직과 역할, 수행 절차를 의미

▶ 개발(Development)

- 데이터 관리 활동이 모두 인정하는 공통된 방향으로 개발되도록 지침을 제공해야 함

▶ 준수(Compliance)

- 데이터 관리 활동이 사전 정의된 방향으로 수행되는지 검증하는 기준을 제공하여야 함

5.2.2 데이터 관리 원칙 수립

데이터 관리 원칙은 빅데이터 플랫폼 및 센터의 데이터 관리를 위한 기본 규정으로, 데이터 관리와 관련된 법률, 정의, 선언 등과 같은 요소와 데이터 관리를 수행하는 중요 내용을 담고 있다.

따라서 데이터 관리 원칙을 수립하기 위해서 빅데이터 플랫폼 및 센터와 관련된 정책적 환경, 기술적 환경, 내부 이슈 등을 고려해야 한다. 초기 데이터 관리 원칙을 수립한 이후 정기적으로 내·외부 이슈를 검토하여 데이터 관리 원칙으로 반영해야 한다.

빅데이터 플랫폼 및 센터 구축 시점의 데이터 품질관리와 관련된 동향 및 정책을 검토하여 빅데이터 환경 특성 및 내부 이슈, 요구사항을 고려하여 데이터 관리 원칙을 수립하였다.

본 가이드의 데이터 관리 원칙은 데이터 품질관리, 데이터 연계관리, 데이터 정보보호 관리 등 3가지 데이터 관리요소 기준으로 도출하였으며, 기본 내용으로 정리하였다.

표 2 데이터 관리 원칙

구 분	원칙	세부 원칙
품질 관리체계	데이터를 핵심 자산으로 한다.	<ul style="list-style-type: none"> 데이터 품질관리 담당자가 데이터 전반적인 관리를 수행한다. 데이터 가치를 높이고 효율적 공유를 위해 데이터 품질을 높이는 정책 및 절차를 수립한다.
	빅데이터 플랫폼 및 센터를 통합적으로 관리해야 한다.	<ul style="list-style-type: none"> 통합 관리를 통해 데이터 중복을 최소화하고 표준화된 기준에 의해 데이터를 관리한다. 통합 관리를 위한 통제 역할의 의사결정기구를 마련하며, 단위 기능별로 관리와 책임 소재를 명확히 한다.
	지속적으로 품질관리를 수행해야 한다.	<ul style="list-style-type: none"> 데이터 변경은 관련 조직간 협조로 수행하며 데이터 품질관리 담당자가 통제한다. 측정 가능한 지표를 설정하며 데이터 품질관리 담당자를 통해 지속적으로 관리를 수행한다. 데이터 품질관리 담당자의 전문성 확보를 위해 지속적인 교육을 실시한다.
	품질관리 활동은 PDCA 사이클을 준수하고 오너십을 부여해야 한다.	<ul style="list-style-type: none"> 데이터 생애주기 전 단계의 데이터 품질을 관리해야 한다. 품질관리 활동은 목적달성과 지속적인 개선을 위한 접근체계인 PDCA 사이클 기반으로 수행한다. 품질주체들 간의 상호협력을 강화하기 위해 역할 정의와 오너십을 명확히 해야 한다.

구 분	원칙	세부 원칙
연계 관리체계	데이터 생애주기 전 단계에 걸쳐 송수신 단절이 없어야 한다.	<ul style="list-style-type: none"> 데이터 생애주기 중단계에서 데이터 송수신 단절이 없도록 관리해야 한다. 장애 또는 오류 발생 여부를 모니터링하여 조치 및 개선이 가능해야 한다.
	데이터 활용처 정보를 제공해야 한다.	<ul style="list-style-type: none"> 데이터 이용 실적 및 집계 등을 위해 흐름정보를 제공할 수 있어야 한다. 데이터 생애주기 중단계에서 활용처 정보를 관리해야 한다.
	데이터 출처 정보를 제공해야 한다.	<ul style="list-style-type: none"> 품질 이상 발생 시 원인 파악과 영향도 분석 등을 위해 추적 정보를 제공할 수 있어야 한다. 데이터의 생애주기 중단계에서 출처 정보를 관리해야 한다.
	연계 메타데이터를 관리해야 한다.	<ul style="list-style-type: none"> 데이터 생애주기 중단계에서 연계 메타데이터(연계목록, 연계관리 항목, 출처정보)를 관리해야 한다. 연계 메타데이터와 메타데이터는 정합성을 유지해야 한다.
	연계 활동은 PDCA 사이클을 준수하고 오너십을 부여해야 한다.	<ul style="list-style-type: none"> 목적달성과 지속적인 개선을 위한 접근체계인 PDCA 사이클 기반 연계활동을 수행한다. 연계주체들 간의 상호협력을 강화하기 위해 역할 정의와 오너십을 명확히 해야 한다.
	표준화 WG의 표준을 준수하여 데이터 이용 활성화에 기여한다.	<ul style="list-style-type: none"> 표준 WG의 메타데이터 항목 및 데이터 연계규격을 준수하여 연계 및 데이터 이용활성화를 제고한다. 오픈 포맷의 데이터 제공을 지향한다.
정보보호 관리체계	기밀성, 무결성, 가용성의 3요소가 보장되어야 한다.	<ul style="list-style-type: none"> 오직 인가된 자만이 접근 할 수 있다. 정보는 고의, 비인가된 우연한 변경으로부터 보호해야한다 정보는 사용자가 필요로 하는 시점에 접근 가능해야 한다.
	플랫폼 정보보호 표준을 준수하여 관리되어야 한다.	<ul style="list-style-type: none"> 이해관계자들의 참여, 의사결정 또는 업무 실행자, 감사 및 감독 등의 활동과 구조 모두 포함해야 한다. 개인정보보호, 정보보호 관련 법률 및 규제, 가이드라인을 우선적용 으로 관리해야 한다. 정보보호기준은 기술적, 관리적 조치 시 적용해야 한다.
	CSUD* 사이클을 준수한 보호 관리가 실행되어야 한다.	<ul style="list-style-type: none"> 목표달성과 지속적인 보호 관리를 위하여 CSUD 사이클을 기반으로 관리 수행해야 한다. 정보 수집, 수집 목적 내 이용, 필요 시 제3자 제공, 보유 기간 경과 후 파기하는 절차가 부합해야 한다.

* CSUD(Collection, Storage, Utilize, Deletion: 수집, 저장·활용, 파기)

5.2.3 데이터 관리 기능 도출

데이터 관리 기능은 데이터 관리 원칙을 기반으로 양질의 데이터를 확보하는 목표를 달성하기 위해, 목표 달성 중심의 접근체계인 PDCA 사이클을 기반으로 데이터 관리 기능을 도출한다.

PDCA (Plan-Do-Check-Action) 사이클은 1950년에 W.Shewhart의 아이디어로 Edward Deming에 의해 개발된 품질 개선을 위한 방법으로, 서비스 및 제품을 지속적으로 개선하기 위한 반복적인 4단계 접근 방식이다. 실적에 근거하여 계획(Plan)을 세우고, 그 계획에 따라 실행(Do)을 한 후 실행 상황이 계획대로 잘 진행되는지 평가(Check)를 하고 계획에 벗어났다면 개선을 하는 절차를 반복한다. 제조업 분야에서 생산관리와 품질관리 등에 사용되었으나 현재 일반적인 경영 및 업무의 기본 관리체계로 많이 활용되고 있다.

그림 8 PDCA 사이클



- ▶ 사용자 중심의 데이터 품질관리
- ▶ 프로세스 중심의 데이터 품질관리 접근체계
- ▶ 데이터 품질계획~개선 활동이 PDCA Cycle 기반으로 전개
- ▶ 지속적인 데이터 품질관리로 데이터 품질향상의 목표 달성

빅데이터 플랫폼 및 센터의 데이터 관리 요소별로, 데이터 관리 기능을 PDCA 사이클에 기반을 두어 도출한다.

그림 9 PDCA 사이클 적용 예시

관리기능	관리대상	세부 관리기능			
		P Plan(계획)	D Do(실행)	C Check(모니터링 및 통제)	A Act(개선)
연계 메타데이터 관리	연계 메타데이터 (표준화 WG '메타데이터 항목 및 데이터 연계규격')	연계메타데이터 정의 • 연계 메타데이터 표준정의 • 메타데이터와 정합성 유지항목 정의 • 표준화 WG의 표준 참조	연계메타데이터 변경 관리 • 사유별 연계 메타데이터 변경 관리 • 변경 영향도 분석	연계메타데이터 정합성 진단 • 연계 메타데이터 및 메타데이터와의 정합성 진단 • 표준화 WG의 표준 준수여부 진단 • 오류 원인 분석	연계메타데이터 정합성 오류 개선 • 오류 영향도 분석 • 개선 계획 수립 • 개선조치 및 결과 확인
데이터 송수신 관리	송수신 데이터	연계준비 • 연계계획 수립 • 규약에 의한 연계계획 협의 • 연계 테스트	생애주기 단계별 송수신 관리 • 수집단계 연계관리 • 가공·분석 단계 연계관리 • 활용단계 연계관리	송수신 모니터링 • 장애 및 오류 모니터링 • 장애 및 오류 원인분석	송수신 오류 개선 • 오류 영향도 분석 • 개선 계획 수립 • 조치 및 결과 확인
흐름 및 추적관리	출처 정보	출처정보 정의 • 규약에 의한 외부연계 및 내부연계 출처정보 정의 • 임의의 외부연계 출처정보 정의	흐름 및 추적정보 제공 • 활용처 정보 제공 • 출처 정보 제공	출처 정보 정합성 진단 • 누락 및 오류 진단 • 오류원인 분석	출처 정보 정합성 오류 개선 • 오류 영향도 분석 • 개선 계획 수립 • 개선조치 및 결과 확인

데이터 품질관리는 데이터 관리 원칙에 기반을 두어 “품질관리 기준 관리”, “품질 진단 및 개선 관리”의 2개 기능을 도출하고, PDCA 사이클을 기반을 두어 “데이터 품질이슈 조사”, “품질관리 기준 수립”, “품질관리 기준 점검”, “품질관리 기준 변경”, “품질관리 대상 정의”, “품질 진단 수행”, “품질 개선”, “품질통제” 등 8개 세부 기능을 도출하였다.

데이터 연계관리는 데이터 관리 원칙에 기반을 두어 “연계메타데이터 관리”, “흐름 및 추적 관리”, “데이터 송수신 관리”의 3개 기능을 도출하고, PDCA 사이클을 기반을 두어 “연계메타데이터 정의”, “연계메타데이터 변경 관리”, “연계메타데이터 정합성 진단”, “연계메타데이터 점검 및 개선”, “출처정보 정의”, “흐름 및 추적정보 제공”, “출처정보 정합성 진단”, “출처정보 점검 및 개선”, “연계준비”, “생애주기 단계별 송수신 관리”, “송수신 상황 모니터링”, “송수신 점검 및 개선”의 12개 세부 기능을 도출하였다.

데이터 정보보호 관리는 데이터 관리 원칙에 기반을 두어 “정보보호 운영관리”, “정보보호 기술관리”, “개인정보 보호관리”의 3개 기능을 도출하고, PDCA 사이클을 기반하여 “운영관리계획수립”, “운영관리 수행”, “감사”, “시정조치 및 개선”, “기술관리계획수립”, “암호화 및 접근제어 실행”, “취약점 점검”, “취약점 개선”, “개인정보보호 관리계획수립”, “비식별화 및 파기 실시”, “적정성 평가”, “적정성 개선”의 12개 세부 기능을 도출하였다.

PDCA 사이클을 통해서 데이터 품질관리, 데이터 연계관리, 데이터 정보보호 관리에서 총 8개 기능과 32개 세부 기능인 프로세스를 정의하였다.

그림 10 데이터 관리 기능 및 세부 기능(프로세스)



5.3 데이터 관리체계 조직

데이터 관리체계 조직은 빅데이터 플랫폼 및 센터에서 보유하고 있는 데이터에 대한 품질관리 활동을 수행하기 위한 담당부서 또는 담당자, 그리고 이와 관련된 이해관계자를 의미하며, 데이터 관리에 대한 명확한 역할과 책임을 제시하고 이를 수행하기 위한 조직 구성 및 담당에 대한 정의한다.

다양한 빅데이터 플랫폼 및 센터 특성을 볼 때, 기관 및 기업의 환경에 따라 다양한 조직 구성을 가지고 있으므로 다양한 형태의 데이터 관리 조직이 구성될 수 있다. 본 가이드에서는 데이터 관리 및 데이터 품질관리를 위해 구성되는 조직의 원칙과 역할을 제시하며 각 기관 및 기업은 본 가이드를 참조하여 환경에 맞게 구성해야 한다.

5.3.1 데이터 관리 조직 구성

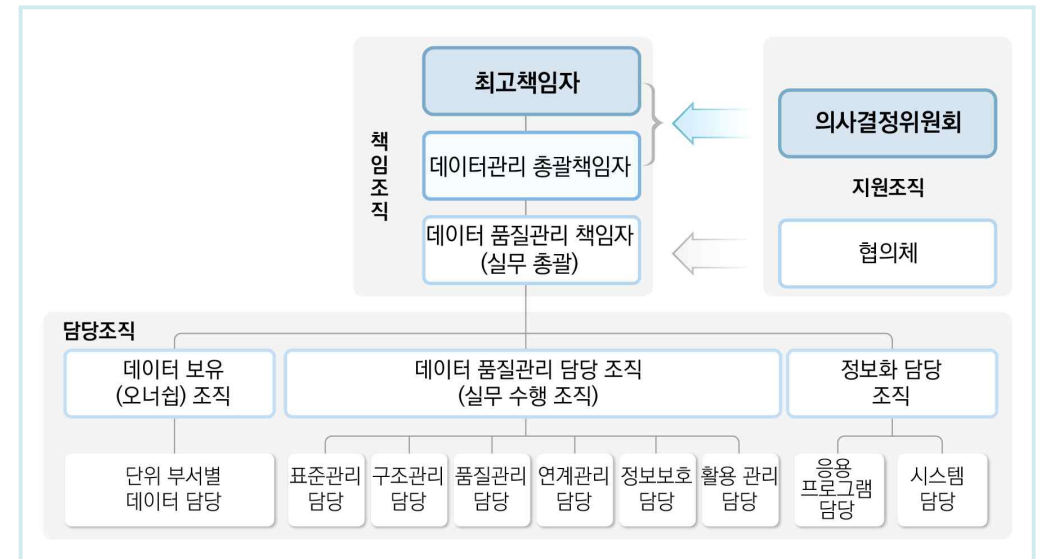
빅데이터 플랫폼 및 센터에서 데이터 관리 조직을 구성할 때 고려해야 할 사항이 있으며, 이를 기반으로 각 기관 및 기업의 환경에 맞게 데이터 관리 조직을 구성한다.

표 3 데이터 관리조직 구성 고려 사항

번호	고려사항	설명
1	데이터 관리 및 품질관리를 총괄하는 총괄 책임자를 선정한다.	<ul style="list-style-type: none"> 대외 및 대내적으로 데이터 관리 및 품질관리에 대한 결정 및 책임 수행 데이터 관리 정책 수립, 계획수립, 오류 개선 결정, 이해관계자 통보 및 협의 등 정책을 결정할 수준의 직급으로 선정
2	데이터 관리 및 품질관리 업무의 효율적 추진을 위해 품질관리 전담 조직 또는 품질관리 전담 담당자를 지정한다.	<ul style="list-style-type: none"> 빅데이터 플랫폼 및 센터 환경을 고려하여 전담 조직 및 전담 담당자로 지정 역할 및 책임 범위에 따라 분야별 전담 담당자를 지정
3	데이터 관리 및 품질관리 관련 주요 사항을 심의, 결정하고 플랫폼 및 센터 간 협의가 필요한 사항의 조정을 위한 조직을 설치한다.	<ul style="list-style-type: none"> 플랫폼 총괄책임자가 협의체 형태로 구성 데이터 관리 및 품질관리 계획 수립 심의 플랫폼 및 센터 간 협의, 조정 데이터 관리 및 품질 수준 평가 정책 또는 실무적 내용에 대한 의사소통

데이터 관리조직 구성을 위한 고려사항을 준수하며, 빅데이터 플랫폼 및 센터의 환경을 고려하여 효율적인 데이터 관리를 위해 관련 이해관계자들을 책임조직, 지원조직, 담당조직 등으로 분류하여 데이터 관리조직을 구성하였다.

그림 11 데이터 관리 조직 구성(안)



- ▶ 책임조직은 최고책임자, 데이터 관리 총괄책임자, 데이터 품질관리 책임자 등 데이터 관리 및 품질향상을 위한 책임과 권한을 가지며, 데이터 관리의 주요사항에 대한 의사결정 권한을 갖는다.
- ▶ 지원조직은 주요 정책과 의사결정을 지원하는 의사결정위원회와 빅데이터 플랫폼 및 센터 간 데이터 관리에 대한 협의, 조정을 수행하는 협의체 등 데이터 관리 및 품질향상을 위한 의사결정에 필요한 내용을 협의, 조정하는 역할을 수행한다.
- ▶ 담당 조직은 데이터를 직접적으로 수행하는 조직으로 크게, 데이터를 보유하고 관리하는 데이터 보유 조직, 각 데이터 관리 영역별 전담하여 데이터 품질관리를 수행하는 데이터 품질관리 담당조직, 데이터 품질관리 담당 조직과 협조하여 실제 DB 및 어플리케이션 등에 데이터 품질관리 업무를 수행하는 정보화 담당 조직으로 구성한다.

5.3.2 데이터 관리 조직 역할

다음은 빅데이터 플랫폼 및 센터의 데이터 관리 조직 역할을 정의한 내용으로 각 기관 및 기업에서는 환경에 따라 본 가이드를 참고하여 데이터 관리 조직을 구성한다.

표 4 데이터 관리 조직 역할 정의

조직 구성		역할 및 책임 정의	
		역할	내용
책임 조직	최고책임자	<ul style="list-style-type: none"> 데이터 관리 및 품질 향상을 위한 전사적, 전략적 최고 의사결정 권한 수행 	<ul style="list-style-type: none"> 데이터 관리 정책 수행의 주요 의사결정
	데이터 관리 총괄책임자	<ul style="list-style-type: none"> 데이터 관리 정책 및 외부 조직간 업무 조정 의사결정 의사결정위원회 소집 및 회의 주재 	<ul style="list-style-type: none"> 데이터 관리 정책 및 지침 수립 데이터 관리 조직 및 시스템 운영 데이터 관리 정책 수행의 주요 의사결정 데이터 관리 우선순위 및 목표 승인
	데이터 품질관리 책임자 (실무총괄)	<ul style="list-style-type: none"> 데이터 품질관리의 상세 추진계획 수립과 주요 의사결정 수행 품질관리 전반에 대한 실무적인 업무 수행 데이터 품질 지표 관리 실무 추진 및 내부 조직간 업무 조정을 위한 협의체 소집 및 회의 주재 	<ul style="list-style-type: none"> 품질관리 우선순위 및 목표 설정, 품질측정 범위 및 대상 선정 데이터 품질지표별 품질관리 현황 관리 데이터 표준/구조/값/관리체계 품질 개선을 위한 조직 역할 조율 품질 진단/개선의 구체적인 계획 및 결과 관리 데이터 개선방안 배포 및 담당자 교육 데이터 품질 정책, 지침, 가이드 개발 및 보완
	의사결정 위원회	<ul style="list-style-type: none"> 데이터 관리 정책 및 전략에 대한 주요 이슈 논의 최고 책임자 및 데이터 관리 총괄 책임자 의사결정 지원 	<ul style="list-style-type: none"> 주요 정책 심의(품질, 연계, 정보보호), 외부 조직간 업무 조정 협의체에서 상신된 사항 처리 회의 소집·주재는 최고책임자 또는 데이터 관리 총괄책임자 수행

조직 구성		역할 및 책임 정의		
		역할	내용	
	협의체	<ul style="list-style-type: none">플랫폼과 센터 간 협업 및 데이터 품질관리의 효율적 추진을 위한 실무 협의체	<ul style="list-style-type: none">데이터 품질 관련 이슈 조정 및 의사결정결정된 사항에 대한 실무 추진 조치데이터 관리에 필요한 실무 차원 지원 수행회의 소집·주재는 데이터 품질관리 책임자 수행	
담당 조직	데이터 보유 조직	<ul style="list-style-type: none">단위 부서별 데이터 담당오너십을 가지는 업무 영역의 데이터 품질관리 수행품질관리 담당자와 함께 데이터 품질관리 활동 수행	<ul style="list-style-type: none">사용자 뷰(메뉴)와 데이터 구조 설계핵심품질항목 선정, 업무규칙 도출/등록, 프로파일링, 업무규칙 진단 수행오류 데이터 확인과 원인파악, 데이터 정비	
	데이터 품질 관리 담당 조직	표준 관리 담당	<ul style="list-style-type: none">표준데이터 요소의 지속적인 품질 향상 활동 수행	<ul style="list-style-type: none">데이터 표준관리 및 메타데이터 관리작성된 데이터 산출물 표준준수 확인주기적인 표준준수 점검 및 조치 요구데이터 표준 관련 정책과 지침 개발, 교육 실시
		구조 관리 담당	<ul style="list-style-type: none">데이터 표준의 적용, 데이터 논리 / 물리 모델 생성 및 변경 관리 수행	<ul style="list-style-type: none">데이터 표준 적용 검토 및 현황 관리데이터 ERD 변경 관리데이터 구조 산출물 작성 및 관리
		품질 관리 담당	<ul style="list-style-type: none">품질관리 기준 관리, 품질진단, 품질 진단 분석 및 개선 활동 등 품질 관리의 실무 수행	<ul style="list-style-type: none">품질 진단 수행계획 수립품질 진단 수행 및 진단 현황 관리품질관리지표와 핵심품질항목, 업무규칙의 도출 및 현황 관리데이터 오류 원인분석 및 품질 진단결과 보고
		연계 관리 담당	<ul style="list-style-type: none">연계되는 데이터의 유형과 처리 현황의 전사적 관리중장기적으로 데이터 품질 관리자가 수행	<ul style="list-style-type: none">연계유형, 주기, 흐름·추적, 연계표준 관리외부 데이터 원천 기관 협업 관리연계 오류 분석 및 개선, 개선 결과 확인
		정보 보호 담당	<ul style="list-style-type: none">개인정보보호 및 데이터 외부 침입 보호, DB 접근 감시 수행	<ul style="list-style-type: none">DB 접근정책 수립, DB보안 관리, 보안사고 대응 및 조치, 사고 이력 관리개인정보보호 정책 수립 및 준수 관리데이터보안 교육 수행
		활용 관리 담당	<ul style="list-style-type: none">데이터 활용을 위한 가치 데이터 셋 발굴 및 정보 제공 관리	<ul style="list-style-type: none">수집데이터에 대한 데이터 융복합 수행 및 가치 데이터셋 수립수요조사 및 만족도 조사 수행정보 이용자 요구사항 관리
	정보화 담당 조직	응용 프로그램 담당	<ul style="list-style-type: none">데이터 관리 관련 프로그램 관리 및 검증 수행	<ul style="list-style-type: none">데이터 품질 관련 프로그램 개선작성된 SQL 실행계획 및 데이터 성능 검증SQL 튜닝 및 전파 교육
		시스템 담당	<ul style="list-style-type: none">DBMS 데이터 생성, 변경, 폐기 관리 수행	<ul style="list-style-type: none">데이터 표준 및 구조 변경에 따른 DB 변경DBMS 성능 모니터링, 개선 활동

5.4 데이터 관리체계 프로세스

데이터 관리체계 프로세스는 데이터 관리 구성 요소별 데이터 관리 기능을 데이터 관리 원칙 및 PDCA 사이클을 통해 도출하고 세부 기능별 절차를 프로세스로 수립한다.

데이터 관리 구성 요소별 프로세스는 각 데이터 구성 요소별 장에서 자세하게 설명하며 여기서는 데이터 관리 프로세스를 표현할 때 사용한 기호에 대하여 설명한다.

표 5 프로세스맵 작성 기호 표기법

기호	명칭	내용
●	시작	• 프로세스 시작 (Start)
종료	종료	• 프로세스 종료 (End)
단위업무	Activity(오프라인)	• 오프라인 단위업무를 표현하는 최소 단위
단위업무 (온라인)	Activity(온라인)	• 온라인 단위업무를 표현하는 최소 단위
판단	의사결정	• 해당 분기에서 나가는 Transition의 조건 표기
분기	조건분기	• 업무 수행에 있어 처리를 위한 조건에 대한 분기 표시
연결	연결자	• 연결된 업무 흐름이 페이지를 넘어갈 때 표현
프로세스	타 프로세스	• 선행, 후행 등 연관되는 프로세스 표현
데이터베이스	데이터 처리	• 정보 흐름상 데이터 처리활동 표현
문서	문서	• 생산, 연관 문서 정보 표현
→	업무흐름	• 단위업무와 단위업무를 연결하는 흐름
----->	정보흐름	• 단위업무와 관련된 데이터의 흐름

5.5 주요 데이터 관련 법제도 현황

빅데이터 플랫폼 및 센터는 주요 데이터 관련 법률의 영향을 받거나 활용이 가능함에 따라 데이터 관련 주요 법률 및 지침의 주요 내용과 빅데이터와의 연관관계를 소개한다. 빅데이터 플랫폼 및 센터에서는 주요 데이터 관련 법률 및 지침을 참고하여 데이터 생산, 수집, 저장, 가공, 분석, 관리, 결합, 거래, 제공, 유통, 활용 등을 수행해야 한다.

표 6 주요 데이터 관련 법률 및 지침 현황

분류	법률/지침명	주관부처	약칭	목적	제개정일/ 시행일
법률	데이터 기반행정 활성화에 관한 법률	행정 안전부	데이터 행정법	• 데이터를 기반으로 한 행정의 활성화에 필요한 사항을 정함으로써 객관적이고 과학적인 행정을 통하여 공공기관의 책임성, 대응성 및 신뢰성을 높이고 국민의 삶의 질을 향상	2020.6.9 2020.12.10
	공공 데이터의 제공 및 이용 활성화에 관한 법률	행정 안전부	공공 데이터법	• 공공기관이 보유·관리하는 데이터의 제공 및 그 이용 활성화에 관한 사항을 규정함으로써 국민의 공공데이터에 대한 이용권을 보장하고, 공공데이터의 민간 활용을 통한 삶의 질 향상과 국민경제 발전에 이바지함	2020.12.8 2020.12.10
	데이터 산업진흥 및 이용촉진에 관한 기본법	과학기술 정보 통신부	데이터 기본법	• 데이터의 생산, 거래 및 활용 촉진에 관하여 필요한 사항을 정함으로써 데이터로부터 경제적 가치를 창출하고 데이터산업 발전의 기반을 조성하여 국민생활의 향상과 국민경제의 발전에 이바지함	2021.10.19 2022.4.20
	국가지식정보 연계 및 활용 촉진에 관한 법률	과학기술 정보 통신부	국가지식 정보법	• 국민이 자유롭고 편리하게 국가지식정보를 이용할 수 있도록 국가지식정보의 연계 및 활용을 촉진하는 데에 필요한 사항을 정함으로써 국민의 지식재산 창출 및 활용 역량을 제고하고 국가경쟁력의 강화에 이바지함	2021.6.8. 2021.12.9
	공공데이터 관리지침	행정 안전부	-	• 「공공데이터의 제공 및 이용 활성화에 관한 법률」에 따라 효율적인 공공데이터 제공정책 시행을 위해 공공기관이 준수하여야 할 관리원칙과 기준을 정함을 목적으로 함	2021.10.26
지침	공공기관의 데이터베이스 표준화 지침	행정 안전부	-	• 「전자정부법」 및 같은 법 시행령 「공공데이터의 제공 및 이용활성화에 관한 법률」에 따라 공공기관이 생성 또는 취득하여 관리하는 데이터베이스의 표준화에 필요한 세부 사항을 정함을 목적으로 함	2021.6.7

5.5.1 데이터 기반행정 활성화에 관한 법

데이터 기반행정 활성화에 관한 법(데이터행정법)은 국민 생활전반에 걸쳐 데이터의 활용이 국가 경쟁력을 좌우하는 핵심자원이므로 그 중요성이 부각되고 있어 데이터를 기반으로 한 행정을 활성화하기

위하여 공공기관은 공동 활용 할 필요가 있는 데이터를 등록할 수 있도록 하고, 등록되지 아니한 데이터를 제공받으려는 경우 데이터 소관 공공기관의 장에게 데이터 제공을 요청할 수 있도록 그 절차 및 방법을 정하며, 공공기관의 데이터를 효율적으로 제공·연계 및 공동 활용하기 위하여 데이터통합관리 플랫폼을 구축하는 등 데이터기반행정을 활성화하기 위한 사항을 규정함으로써 행정의 책임성, 대응성 및 신뢰성을 높이고 국민의 삶의 질을 향상시키려는 것이다.

데이터행정부의 주요내용은 다음과 같다.

- 공공기관이 생성하거나 취득하여 관리하고 있는 데이터를 수집·저장·가공·분석·표현하는 등의 방법으로 정책수립 및 의사결정에 활용함으로써 객관적이고 과학적으로 수행하는 행정을 데이터기반행정으로 정의함(제2조제2호)
- 데이터기반행정 관련 정책, 제도 및 법령의 개선, 데이터 제공 거부에 대한 조정 등을 심의하기 위하여 데이터기반행정 활성화 위원회를 둠(제5조)
- 행정안전부장관은 관계 중앙행정기관의 장과 협의하여 3년마다 데이터기반행정 활성화를 위한 기본계획을 수립하도록 함(제6조)
- 공공기관의 장은 데이터기반행정과 관련하여 공동활용할 필요가 있는 데이터를 데이터통합관리 플랫폼에 등록할 수 있고, 등록되지 아니한 데이터의 제공을 요청하는 경우 데이터의 이용목적, 분석방법 등을 명시한 문서로 데이터 소관 공공기관에 요청하도록 함(제8조 및 제10조)
- 다른 법률 등에서 비밀로 규정하거나 국가안전보장 등에 관한 사항으로서 제공할 경우 국가의 중대한 이익을 크게 해칠 우려가 있는 경우 등에는 데이터 소관 공공기관은 데이터 제공을 거부할 수 있도록 함(제11조)
- 공공기관의 장은 업무협약 등을 통하여 민간 법인 등에 소관 데이터를 제공하여 줄 것을 요청할 수 있도록 함(제14조)
- 공공기관의 장은 생성하거나 취득하여 관리하는 데이터에 대한 메타데이터 및 데이터관계도를 관리하도록 하고, 행정안전부장관은 공공기관의 메타데이터 등을 종합하여 데이터관리체계를 구축·운영하도록 함(제16조)
- 행정안전부장관은 공공기관이 데이터를 효율적으로 제공·연계 및 공동 활용할 수 있도록 데이터 유형별 저장 체계에 관한 사항 등을 포함한 데이터통합관리 플랫폼을 구축·운영하도록 함(제18조)
- 행정안전부장관은 국가적 차원의 데이터 분석 등이 필요한 사항에 대해서 데이터분석 등을 통한 정책수립 및 의사결정에 활용할 수 있도록 정부 통합데이터분석센터를 설치·운영할 수 있음(제20조)
- 공공기관의 장은 데이터의 공동활용 성과 등 데이터기반행정의 실태를 자체 점검하고, 그 결과를 행정안전부장관에게 제출하며, 행정안전부장관은 점검 결과를 종합하여 공개하도록 함(제22조)

빅데이터 플랫폼에서는 데이터행정부법을 기반으로 공공기관이 구축하거나 제공·연계 및 공동 활용할 수 있는 빅데이터를 활용하여 다음과 같은 다양한 분야의 융합업무를 발굴할 수

있도록 유도 또는 추진할 수 있다.

- 환경, 산림, 소방안전, 스마트치안, 해양수산 등의 분야에서 공공안전 빅데이터를 활용거나 헬스케어, 라이프로그 등의 분야에서 질병 등의 빅데이터를 활용거나 연계하여 사전 위험예측 및 제거방법 등에 활용
- 금융, 문화, 교통, 유통소비, 통신, 중소기업, 지역경제, 농식품, 디지털산업혁신 등의 분야에서 경제사회 등 분야에서 미래 수요 충족을 위한 선제적 대응책 마련
- 금융, 지역경제, 소방안전, 농식품 등의 분야에서 공공데이터를 연계·활용하여 비용 절감이나 행정처리 절차 개선 기여
- 정부 및 공공기관에서는 빅데이터 플랫폼의 빅데이터를 연계·활용한 분석을 통하여 주요 정책 수립 등을 위한 신속하고 정확한 의견 수렴 및 빅데이터 플랫폼 데이터와 데이터행정부법 기반의 데이터의 비교, 분석을 통한 최적화된 대책마련 및 맞춤형 서비스 제공에 활용

데이터행정부 제정으로 공공기관 간 데이터 요청 및 제공이 활성화되고, 데이터 분석결과를 정책결정에 활용하여 과학적 행정을 구현함으로써 행정의 신뢰성을 제고하고 삶의 질을 향상시킬 수 있을 것으로 기대하고 있다. 또한 데이터행정부법을 기반으로 구축된 데이터를 관련 빅데이터 플랫폼에서 적극 활용하여 융합·분석 등을 통한 사회 현안에 대한 대책마련 및 맞춤형 서비스에 활용할 수 있을 것으로 기대하고 있다.

5.5.2 공공 데이터의 제공 및 이용 활성화에 관한 법률

공공 데이터의 제공 및 이용 활성화에 관한 법률(공공데이터법)은 공공기관이 보유·관리하는 데이터의 제공 및 그 이용 활성화에 관한 사항을 규정함으로써 국민의 공공데이터에 대한 이용권을 보장하고, 공공데이터의 민간 활용을 통한 삶의 질 향상과 국민경제 발전에 이바지함을 목적으로 제정된 법률이다.

공공데이터법의 주요내용은 다음과 같다.

- 행정안전부장관은 매년 공공기관을 대상으로 공공데이터의 제공기반조성, 제공현황 등 제공 운영실태를 대통령령으로 정하는 바에 따라 평가하여야 함(9조)
- 정부는 공공데이터의 제공 및 이용 활성화 정책을 효율적으로 수립·시행하기 위하여 국민, 민간기업 및 단체를 대상으로 공공데이터 이용수요, 이용현황, 애로사항 등을 조사할 수 있음(10조)
- 공공기관의 장은 해당 기관의 공공데이터 제공 및 이용 활성화에 관한 업무를 총괄하는 책임관 및 실무담당자를 임명하고 이를 이용자가 알기 쉽게 해당 기관의 인터넷 홈페이지를 통하여 공표하여야 함(12조)
- 공공데이터의 효율적인 제공 및 이용 활성화 지원을 위하여 공공데이터활용지원센터를 「지능정보화 기본법」에 따른 한국지능정보사회진흥원에 설치·운영해야 함(13조)
- 공공기관의 장은 공공데이터의 제공 및 이용 활성화를 위하여 개인 및 기업, 단체 등과 협력하여

관련 서비스를 제공할 수 있으며, 공공데이터를 활용하여 개인·기업 또는 단체 등이 제공하는 서비스와 중복되거나 유사한 서비스를 개발·제공하여서는 아니됨(15조, 16조)

- 공공기관의 장은 해당 공공기관이 보유·관리하는 공공데이터를 「공공기관의 정보공개에 관한 법률」에 따른 비공개대상정보 등을 제외하고 국민에게 제공하여야 함(17조)
- 행정안전부장관은 공공데이터의 효율적 제공을 위하여 통합제공시스템을 구축·관리하고 활용을 촉진하여야 함(21조)
- 공공기관의 장은 해당 기관이 생성 또는 취득하여 관리하는 공공데이터의 안정적 품질관리 및 적절한 품질수준의 확보를 위하여 필요한 조치를 취하여야 하며, 과학기술정보통신부장관과 행정안전부장관은 공공데이터의 적절한 품질수준의 확보와 제공 촉진을 위하여 품질 진단·평가, 개선지원 등 필요한 시책을 수립·추진하여야 함(22조)
- 공공데이터를 이용하고자 하는 자는 공표된 제공대상 공공데이터의 경우 소관 공공기관이나 공공데이터 포털 등에서 제공받을 수 있음(26조)

공공데이터법은 공공부문의 데이터 주도혁신을 이끌기 위한 법안으로 제정되었다. 기존의 경험이나 직관적 정책수립 의사결정을 탈피하여 데이터 분석을 통하여 데이터에 기반한 객관적으로 과학적인 행정을 수행하고 공공기관의 업무효율성, 책임성, 대응성 및 신뢰성을 높이고 궁극적으로 국민의 삶의 질을 향상시키는 것을 의미한다.

공공데이터법에서 중요한 부분중의 하나는 타 공공기관이 보유한 데이터를 요청하고 제공받아서 활용할 수 있다는 점이다. 개별 공공기관에서는 이미 방대한 량의 데이터를 보유하고 있지만 타 기관 간 연계 및 활용은 법적인 근거가 없어 쉽게 제공 및 활용할 수 없었기 때문에 공공데이터법을 통해 공공기관 간 데이터의 연계 및 공동 활용의 제도적 기반이 마련되었다.

각 빅데이터 플랫폼에서는 공공데이터법을 기반으로 관련되는 타 공공기관이 보유하고 있는 방대한 데이터를 연계 및 공동 활용할 수 있도록 추진할 수 있게 되었다.

5.5.3 데이터 산업진흥 및 이용촉진에 관한 기본법

데이터 산업진흥 및 이용촉진에 관한 기본법(데이터기본법)은 데이터의 가치에 대한 세계적인 관심이 집중되고 있고 경제·사회 전반에서 생산되는 데이터를 통해 새로운 산업과 서비스가 창출되면서 데이터 활용에 대한 요구가 증가하면서 안전한 데이터 관리를 위한 기반 마련도 중요해졌기 때문에 데이터 산업 발전의 제도적 기반을 마련하기 위해 추진된 법이다. 데이터기본법을 통하여 공공·민간의 데이터 활용도 더 빠르게 증가할 것으로 전망되고 있다.

데이터기본법에는 데이터의 생산, 거래 및 활용 촉진에 관한 필요한 사항을 규정함으로써 데이터로부터 경제적 가치를 창출하고 데이터 산업 발전의 기반을 조성하기 위한 목적으로 마련되었다. 데이터에 대한 개념 정의 및 정부의 역할, 데이터 보호 절차 등을 규정하고 데이터의 생산 및 거래, 활용 촉진에 필요한 사항들에 대한 내용을 담고 있다.

데이터기본법의 주요 내용은 다음과 같다.

- “데이터”란 다양한 부가가치 창출을 위하여 관찰, 실험, 조사, 수집 등으로 취득하거나 정보시스템 및 「소프트웨어 진흥법」 제2조제1호에 따른 소프트웨어 등을 통하여 생성된 것으로서 광(光) 또는 전자적 방식으로 처리될 수 있는 자료 또는 정보라고 정의함(2조)
- 정부의 역할로는 데이터 생산, 거래 및 활용을 촉진하고 데이터 산업의 기반을 조성하기 위하여 행정안전부장관이 과학기술정보통신부장관과 협의하여 매 3년마다 데이터 산업 진흥 기본계획을 수립하도록 하였고 다양한 분야와 형태의 데이터와 데이터 상품이 생산될 수 있는 환경을 조성하고 데이터 생산자의 전문성을 높이고 경쟁력을 강화하기 위한 시책을 마련할 것을 규정함(3조, 4조)
- 다양한 분야와 형태의 데이터와 데이터상품이 생산될 수 있는 환경을 조성하여야 하며, 데이터생산자의 전문성을 높이고 경쟁력을 강화하기 위한 시책을 마련하고 데이터생산자에게 데이터 생산에 필요한 재정적·기술적 지원을 하도록 함(9조)
- 과학기술정보통신부장관과 행정안전부장관은 데이터 간의 결합을 통해 새로운 데이터의 생산을 촉진하기 위하여 산업 간의 교류 및 다른 분야와의 융합기반 구축 등에 필요한 시책을 마련하여 추진도록 함(10조)
- 데이터를 안전하게 분석·활용할 수 있는 구역인 데이터안심구역을 지정하여 운영할 수 있으며 데이터안심구역 이용 지원을 위해 미개방데이터, 분석시스템 및 도구지원이 가능하며, 특히 미개방데이터 지원을 위하여 필요한 경우 정부, 지방자치단체, 공공기관, 민간법인 등에게 데이터 제공을 요청할 수 있음(11조)
- 데이터 보호를 위해 데이터 자산의 부정취득행위와 정당한 권한없이 데이터 생산자가 데이터 자산에 적용한 기술적 보호조치를 무력화하는 행위를 방지하고 데이터 기반의 정보 분석을 활성화하기 위해 데이터의 수집, 가공 등 정보분석에 필요한 사업을 정부가 지원할 수 있도록 하고 정보 분석을 위하여 데이터를 이용하는 경우에는 저작권법에 따라 공개된 데이터 등을 이용하고 보호하도록 함(12조, 13조)
- 과학기술정보통신부장관은 민간 데이터의 가치를 평가할 수 있는 기법 및 체계를 수립 후 공표하고, 데이터 가치평가 전문기관을 지정할 수 있음(14조)
- 정부에게 데이터의 생산, 거래 및 활용 촉진을 위하여 데이터를 컴퓨터 등 정보처리장치가 처리할 수 있는 형태로 본인 또는 제3자에게 원활하게 이동시킬 수 있는 제도적 기반 구축을 위해 노력하도록 하고, 데이터거래사업자 및 데이터분석제공사업자에 관한 사항들을 규정하는 한편 데이터 생산, 거래 및 활용에 관한 분쟁을 조정하기 위하여 데이터분쟁조정위원회를 두도록 함(15조, 34조)
- 과학기술정보통신부장관은 데이터 유통 및 거래를 활성화하기 위하여 데이터 유통 및 거래 체계를 구축하고, 데이터 유통 및 거래 기반 조성을 위하여 필요한 지원하고 데이터 유통과 거래를 촉진하기 위하여 데이터유통시스템을 구축·운영하도록 함(18조)
- 정부는 데이터의 수집·가공·분석·유통 및 데이터에 기반한 서비스를 제공하는 플랫폼을 지원하는 사업을 할 수 있음(19조)

- 과학기술정보통신부장관은 데이터의 품질향상을 위하여 행정안전부장관과 협의하여 품질인증 등 품질관리에 필요한 사업을 추진할 수 있으며 데이터 품질인증을 실시하기 위하여 인증기관을 지정할 수 있음(20조)
- 데이터의 합리적 유통 및 공정한 거래를 위하여 공정거래위원회와 협의를 거쳐 표준계약서를 마련하고, 데이터사업자에게 그 사용을 권고할 수 있음(21조)
- 과학기술정보통신부장관은 행정안전부장관과 협의하여 데이터의 호환성을 확보함으로써 각종 상품과 서비스에서의 데이터의 결합, 거래 및 활용을 촉진하기 위하여 데이터의 저장 형태 및 이전 방식, 데이터의 분류 체계, 데이터의 결합, 거래 및 활용을 위하여 필요한 사항에 대한 표준화 기준을 마련하여 고시하도록 함(28조)
- 정부는 데이터산업 전반의 기반 조성 및 관련 산업의 육성을 효율적으로 지원하기 위하여 필요한 때에는 그 업무를 전문적으로 수행할 전문기관을 지정할 수 있음(32조)

기존 데이터 3법이 개인정보보호 및 관련 산업 진흥을 위한 것이었다면 데이터기본법은 민간의 개인정보를 대상으로 한다는 점에서 의미가 있다. 데이터기본법의 통과로 데이터 산업의 생태계 변화가 이루어져 데이터를 이용한 다양한 산업이 활성화될 수 있는 동시에 개인정보를 해칠 우려가 있는 가운데 양질의 데이터 확보를 위한 체계적 접근과 데이터의 활용을 고려한 데이터 표준화와 개인정보의 보호가 선행되어야 한다.

데이터기본법에 포함되어 있는 빅데이터 관련 내용은 다음과 같다.

- 정부에서 다양한 분야와 형태의 빅데이터를 포함한 데이터와 데이터상품이 생산될 수 있는 환경을 조성
- 다양하고 방대한 빅데이터 간의 결합을 통해 새로운 데이터의 생산을 촉진하기 위해 산업 간의 교류 및 다른 분야와의 융합기반을 구축
- 공공데이터와 민간에서 보유하고 있는 다양한 빅데이터의 결합 촉진을 위한 교류 및 협력방안 마련
- 누구든지 데이터를 안전하게 분석·활용할 수 있는 데이터안심구역 지정 운영을 통하여 빅데이터의 활용 추진
- 데이터 기반의 정보분석을 활성화하기 위해 방대한 빅데이터의 수집, 가공 등 정보분석에 필요한 사업 지원
- 다양한 분야의 빅데이터의 유통 및 거래를 활성화하기 위한 체계 구축
- 빅데이터의 수집·가공·분석·유통 및 데이터에 기반한 서비스를 제공하는 빅데이터 플랫폼을 지원

5.5.4 국가지식정보 연계 및 활용 촉진에 관한 법률

국가지식정보 연계 및 활용 촉진에 관한 법률(국가지식정보법)은 국가지식의 상호 융합을 통해 새로운 지식을 창출하고 공유·확산·활용될 수 있는 장을 마련하여 국민 누구나 자신이 원하는

지식정보에 쉽게 접근하고 자유롭게 활용하여 지식역량을 제고하고 국가경쟁력을 강화하기 위하여 마련된 법률이다.

국가지식정보법의 주요 내용은 다음과 같다.

- “국가지식정보”란 국가기관, 지방자치단체 및 「지능정보화 기본법」에 따른 공공기관이 생산·보유·관리하고 있는 과학기술, 교육학술, 문화예술, 사회경제, 행정 등에 관한 정보 중 지식의 활용 및 교육을 목적으로 국가적 이용가치가 있는 디지털화된 정보나 디지털화의 필요성이 인정되는 정보로서 국가지식정보위원회의 지정을 받은 정보로 정의(2조)
- 과학기술정보통신부장관은 국가지식정보의 연계 및 활용 촉진에 관한 업무 지원 등의 사항을 종합적이고 효율적으로 추진하기 위하여 「지능정보화 기본법」에 따른 한국지능정보사회진흥원을 전담기관으로 지정할 수 있음(11조)
- 과학기술정보통신부장관은 국가지식정보의 효율적인 연계·활용을 위하여 관계 중앙행정기관의 장과 협의하여 국가지식정보 분야별 전문기관을 지식정보센터로 지정하고 해당 분야의 국가지식정보를 연계 및 활용하도록 할 수 있으며 지식정보센터는 국가지식정보를 통합플랫폼에서 활용할 수 있도록 필요한 조치를 하여야 함(12조)
- 과학기술정보통신부장관은 국가기관등의 장과 협의하여 통합플랫폼과의 연계대상이 되는 국가지식정보를 지정할 수 있으며 국가지식정보의 효율적인 연계 및 활용을 위하여 통합플랫폼을 구축·운영하여야 함(13조, 14조)
- 과학기술정보통신부장관은 국가지식정보의 제공 및 이용 활성화 정책의 효율적인 추진을 위하여 국가지식정보의 보유 현황 및 통합플랫폼과의 연계 현황 등에 대하여 조사할 수 있음(15조)
- 과학기술정보통신부장관은 국가지식정보의 이용을 활성화하기 위하여 민간사업자와 민간사업자단체 등을 통하여 국가지식정보를 제공할 수 있으며 민간의 지식정보 중 보존 및 이용가치가 있는 지식정보를 통합플랫폼과 연계하기 위하여 이를 보유한 민간사업자 및 민간사업자단체 등과 협력할 수 있음(18조)

국가지식정보법의 제정으로 국가지식정보의 효율적인 연계 및 활용을 위하여 통합플랫폼을 구축·운영해야 하며 이를 위하여 표준화를 추진하고 국가지식정보의 효율적 연계 및 활용을 위하여 필요한 행정적·기술적·재정적 지원을 해야 한다.

국가지식정보법 제정은 기존에 25개 국가기관 48개 개별사이트에서 분산되어 제공되고 있던 국가지식정보를 '디지털 집현전'이라는 하나의 통합플랫폼에서 연계하여 통합 검색 및 활용이 가능한 서비스를 제공할 수 있게 되었다.

5.5.5 데이터 관련 지침

정부에서 제정한 데이터 관련 지침은 행정안전부의 공공데이터 관리지침과 공공기관의 데이터베이스 표준화 지침 등이 있다.

공공데이터 관리지침(행정안전부고시 제2021-70호, 2021.10.26.)은 공공데이터 관리의

기본원칙, 공공데이터 제공·관리 단계별 기준, 공공데이터 품질관리, 공공데이터 제공 시 표준 준수, 공공기관의 민간 중복·유사 서비스 금지 등의 내용을 담고 있다. 2021년 10월 26일 개정된 지침에는 공공데이터 제공 관련 개선사항이 관리지침 새로 추가됐다. 지침에서는 공공데이터를 원천데이터로 제공하는 것과 공공기관이 정보시스템을 통합하거나 폐지하는 경우 공공데이터의 제공과 보존을 위해 노력할 것을 기본원칙으로 규정했다. 또한 기관에서 이미 공개된 데이터는 ‘공공데이터포털’에도 등록되도록 하고, 공공기관이 제공 중인 공공데이터의 일부 항목이 임의로 대체되거나 삭제되지 않도록 규정하는 조항도 포함되었다.

공공기관의 데이터베이스 표준화 지침(행정안전부고시 제 2021-32호, 2021.6.7 일부 개정)은 데이터베이스 구축·운영 시 공공기관이 표준 관련 제도를 쉽게 준수하게 하고, 제도의 관리·운영 편의성을 높이기 위해 표준 관련 제도를 통합하여 정비하고자 하였다.

지침에는 공통표준용어 관리원칙, 구성요소 및 관리항목, 공통표준용어 제·개정 등이 포함되어 있는데 세부항목·속성 등은 지침에 포함하지 않고 공공데이터 포털·메타데이터관리시스템을 통해 별도 관리·운영하도록 하였다.

공공데이터베이스의 구축부터 운영·관리 절차별 준수사항을 재정립하고 있는데, 기관의 표준화 관리체계 수립, 표준 수립 및 적용, 표준관련 필수 산출물 작성 및 관리, 사업수행 단계별 표준 적용방안 등이 포함되어 있고 공공데이터의 효율적 관리 및 활용을 위한 메타데이터 등록·관리, 메타관리시스템 운영 등이 포함되어 있다. 공공데이터베이스 표준화 관련 각종 서식 및 산출물 예시가 현행화되었다.

공공데이터 관리지침과 공공기관의 데이터베이스 표준화 지침 등 정부에서 제정한 지침에서는 누구든지 공공데이터를 편리하게 이용할 수 있도록 원천데이터를 제공하고 통계처리한 데이터의 경우에도 이용자의 활용성을 고려하여 최대한 상세한 형태로 제공하도록 하고 있다. 빅데이터 플랫폼에서 이와 같은 공공 원천데이터를 활용하여 연계·결합을 통한 활용이 가능하도록 하고 있다. 또한 정부 및 공공기관 공통의 데이터 관련 표준화를 수립하고 공공데이터에 일관되게 적용 및 지속적인 점검을 통하여 데이터의 품질을 향상시키도록 하고 있어서 고품질의 데이터를 활용할 수 있다.

01 개요



빅데이터 플랫폼 및 센터의 데이터 품질관리체계는 데이터 관리 원칙, 데이터 관리 조직, 데이터 관리 프로세스라는 데이터 관리 3요소 구성에 맞추어 품질관리 목적 및 전략, 과제를 선정하여 빅데이터 플랫폼 및 센터의 데이터 품질관리의 원칙과 방향을 설정하였다.

이를 기반으로 데이터 품질관리를 수행하는 기능을 도출하고 세부 기능으로 구분하여, 세부 기능별 프로세스를 수립하였으며, 세부 기능 수행을 위한 품질관리 역할 및 책임을 정의하여 데이터 품질관리 조직을 구성하였다.

1.1 목적 및 기능

빅데이터 플랫폼 및 센터의 데이터 품질관리 목적은 단기는 물론 중장기적인 품질관리를 통해 달성해야 하는 결과로, “높은 수준의 데이터 품질 확보”로 정의하였다. 이를 위해서 빅데이터 플랫폼 및 센터의 특성을 고려하여 범용적인 품질관리, 데이터 수생애주기 품질관리, 지속적인 품질관리의 전략을 수립하였다.

그림 12 데이터 품질관리 목적 및 기능



빅데이터 환경은 다양한 데이터를 다양한 원천에서 수집 및 가공, 활용해야 하므로 다양한 도메인의 데이터 품질관리를 위해서는 보편적인 기준을 바탕으로 품질관리를 수행하는 것이 효율적이며, 이에 “범용적인 품질관리” 전략을 수립하였다.

또한, 빅데이터는 수집된 데이터를 데이터 소비자의 요구에 맞게 가공하여 활용하고 있어 수집단계, 가공단계, 활용단계 등 데이터 생애주기 전반에 걸쳐 데이터 품질을 확보해야 하며, 이를 위한 “데이터 전생애주기 품질관리” 전략을 수립하였다.

데이터는 수집-가공-활용 활동으로 인한 데이터 변형, Garbage 데이터, 데이터 오류 발생 등으로 데이터 품질이 하락한다. 따라서 데이터 품질관리는 단기적인 관리보다는 지속적인 수행이 필요하며, 이를 위해 “지속적인 품질관리” 전략을 수립하였다.

빅데이터 플랫폼 및 센터 데이터 품질관리에 필요한 활동을 전략 기반으로 다음과 같이 “원천 및 활용 데이터 CTQ 적용”, “수집, 활용 영역별 품질 진단 수행”, “빅데이터 환경 고려 기준 및 체계 구성”, “생애주기 단계별 품질관리 요소 발굴”, “지속적인 품질관리 지원체계 구성” 등 5가지 과제를 도출하였다.

빅데이터 플랫폼 및 센터 데이터 품질관리 기능은 과제를 포괄할 수 있는 개념으로 “품질관리기준 관리”, “품질 진단 및 개선 관리” 등 2개로 정의하였다.

1.2 관리 범위

빅데이터 플랫폼 및 센터의 데이터 관리의 범위는 데이터 품질관리 전략 방향에 맞추어 데이터 수집, 데이터 가공·분석, 데이터 활용의 데이터 생애주기 전 단계에서 운영되는 데이터를 대상으로 하며 범용적이고 지속적인 체계를 구성하였다.

빅데이터 환경을 고려하여 데이터 생애주기 모든 단계의 데이터를 대상으로 품질관리를 수행하며, 원천데이터의 품질관리는 원천데이터를 보유하고 있는 기관의 품질관리기준을 준수하여 반영한다.

또한, 다양한 도메인의 데이터를 보유하고 있는 플랫폼 및 센터 및 빅데이터 환경을 고려하여 범용적인 품질관리 기준이 마련되어야 한다. 그리고 범용적인 품질관리 기준은 데이터 생애주기 모든 단계에 적용되어야 한다.

빅데이터 환경에는 핵심 데이터뿐 아니라 시점에 따라 필요하지 않은 데이터가 많이 존재하게 되며 시간이 흐름에 따라 이러한 데이터의 양이 폭발적으로 증가하게 된다. 따라서 RDB 환경에서와 같이 빅데이터 환경에서도 데이터 품질을 확보하기 위해서는 지속적으로 품질관리를 수행해야 하며, 데이터 품질기준을 수립한 후 품질 진단 및 개선 작업을 수행하고 다시 품질 점검을 하는 체계를 갖추어야 한다.

그림 13 데이터 품질관리 범위



02

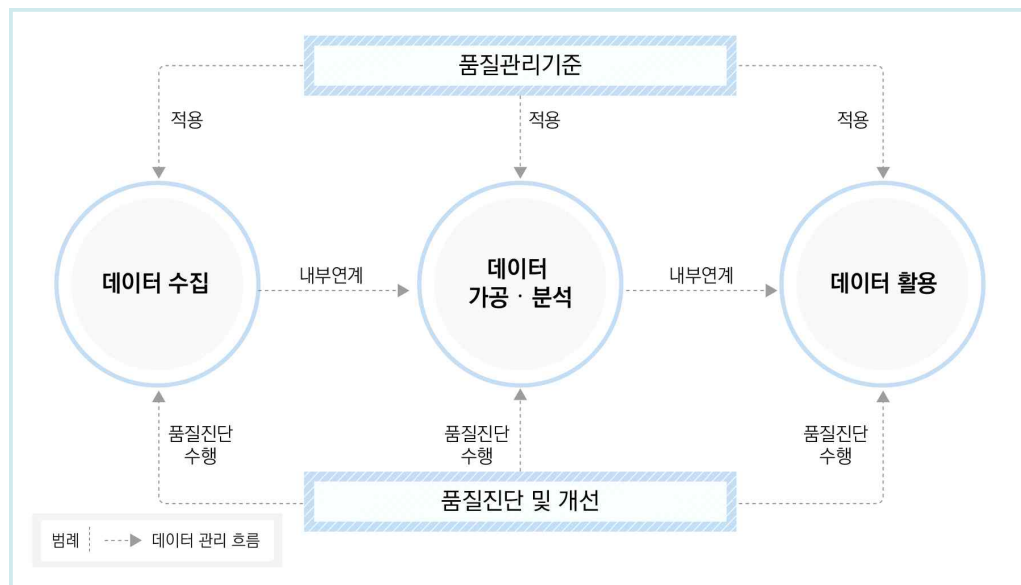
모델 수립



2.1 모델 개념

빅데이터 플랫폼 및 센터의 데이터 품질관리 모델을 데이터 품질관리 목적 및 전략을 기반으로 데이터 생애주기 전 단계를 거쳐 데이터 품질관리 기능을 수행되는 체계로 구성하였다.

그림 14 데이터 품질관리 모델



데이터 생애주기를 “데이터 수집”, “데이터 가공·분석”, “데이터 활용” 등으로 구분하고 “품질관리기준” 관리 및 “품질 진단 및 개선” 관리 등 데이터 품질관리 기능 간의 관계를 정의하였다.

품질관리기준은 데이터가 내부 연계를 통해 흘러가면서 변형되더라도 데이터 수집, 데이터 가공·분석, 데이터 활용 모든 단계에 적용된다. 그리고 원천데이터 보유 기관의 품질관리기준을

반영하여 적용한다.

품질 진단 및 개선은 데이터 생애주기 단계별 데이터 처리 후 오류를 점검하기 위해 모든 단계에 걸쳐 수행된다.

다음은 데이터 품질관리 모델 요소의 정의 내용이다.

표 7 품질관리 모델 요소 정의

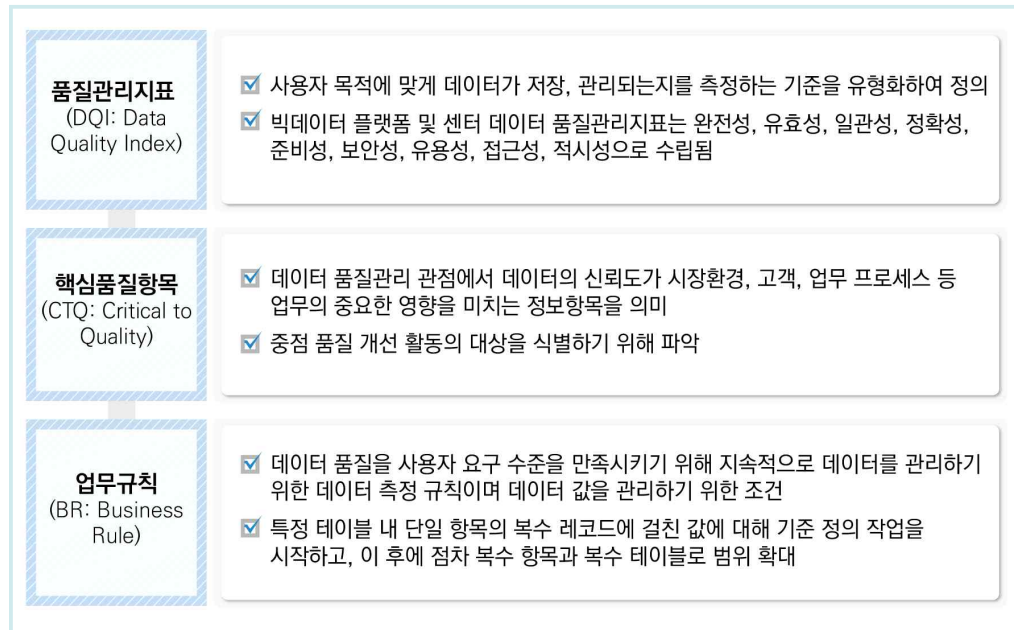
생애주기	데이터 관리 흐름	설명
데이터 수집	적용	• 데이터 수집 단계에 품질 진단 수행을 위한 데이터 품질관리기준을 적용
	품질 진단 수행	• 수집된 데이터에 대하여 적용된 데이터 품질관리기준을 바탕으로 품질 진단 수행
	내부 연계	• 데이터 가공 및 분석을 위해 품질진단 및 개선이 완료된 수집 데이터에 대하여 내부 연계 송신
데이터 가공·분석	적용	• 데이터 가공 및 분석단계에 품질 진단 수행을 위한 데이터 품질관리기준을 적용
	품질 진단 수행	• 가공된 데이터에 대하여 적용된 데이터 품질관리기준을 바탕으로 품질 진단 수행
	내부 연계	• 데이터 가공 및 분석을 위해 품질진단 및 개선이 완료된 수집 데이터에 대하여 내부 연계 수신 • 데이터 활용을 위해 품질 진단 및 개선이 완료된 가공 데이터에 대하여 내부 연계 송신
데이터 활용	적용	• 데이터 활용 단계에 품질 진단 수행을 위한 데이터 품질관리기준을 적용
	품질 진단 수행	• 활용을 위해 저장된 데이터에 대하여 적용된 데이터 품질관리기준을 바탕으로 품질 진단 수행
	내부 연계	• 데이터 활용을 위해 품질진단 및 개선이 완료된 가공 데이터에 대하여 내부 연계로 데이터 수신

2.2 관리 요소

빅데이터 플랫폼 및 센터의 데이터 품질관리 요소는 품질관리기준, 품질 진단 및 개선 등 2개의 기능별로 정의하였다.

품질관리 기준은 “품질관리지표(DQI: Data Quality Index)”, “핵심품질항목(CTQ: Critical to Quality)”, “업무규칙(BR: Business Rule)”으로 품질관리 요소를 정의하였다.

그림 15 품질관리기준의 데이터 품질관리 요소



품질 진단 및 개선은 “품질관리 대상 정의”, “품질 진단 수행”, “품질 진단결과 분석”, “품질 개선 수행”, “품질 통제”등 품질 진단 및 개선 절차를 기준으로 데이터 품질관리 요소를 정의하였다. 데이터 품질 진단 및 개선은 “IV. 데이터 품질 진단 방법”에서 상세하게 설명한다.

그림 16 품질진단 및 개선의 데이터 품질관리 요소



03 체계 구성

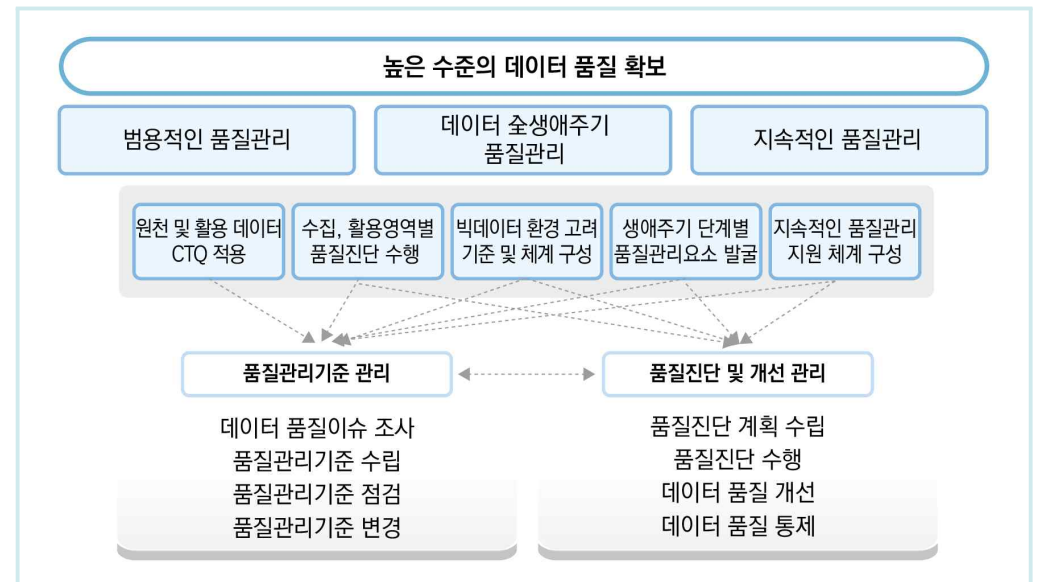
3.1 품질관리 원칙

3.1.1 원칙 도출

데이터 품질관리를 위한 원칙은 데이터 품질관리 목적 및 전략을 달성하기 위해 수행하는 과제에 대한 규칙으로 데이터 품질관리체계 안에서 목표 달성 및 개선을 수행할 수 있어야 한다.

개요에서 설명한 바와 같이 데이터 품질관리 목적은 중장기적으로 지속적인 관리를 통해 달성하는 결과로, 품질관리 목적을 달성하기 위한 3가지 전략과 전략을 수행하는 5가지 과제를 수립하였다. 데이터 품질관리 원칙은 이러한 목적과 전략, 과제를 수행하는데 필요한 수행 규칙을 정의하며, 빅데이터 플랫폼 및 센터 환경을 고려하였다.

그림 17 품질관리 원칙 도출



3.1.2 원칙 정의

데이터 품질관리 목적을 달성하기 위한 전략에 기반한 4개 품질관리 원칙을 수립하고, 10개 세부 원칙을 제시한다.

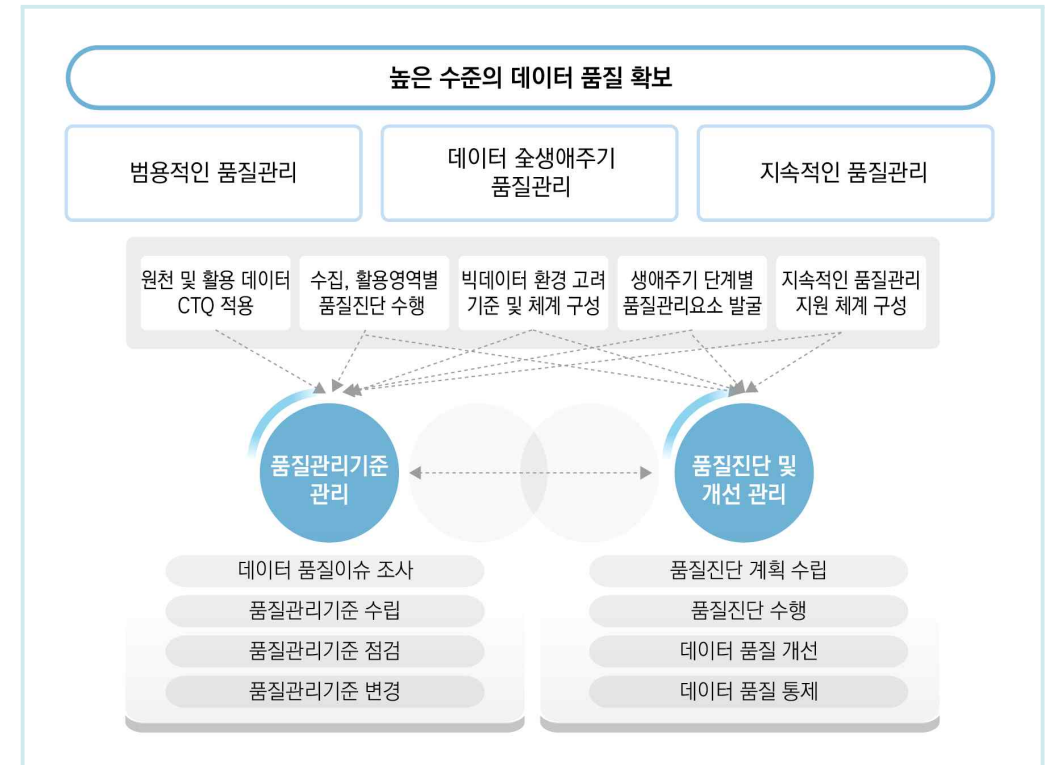
- ▶ 원칙1. 데이터를 핵심 자산으로 관리해야 한다.
 - 데이터 품질관리 담당자가 데이터 전방적인 관리를 수행한다.
 - 데이터 가치를 높이고 효율적 공유를 위해 데이터 품질을 높이는 정책 및 절차를 수립한다.
- ▶ 원칙2. 빅데이터 플랫폼 및 센터를 통합적으로 관리해야 한다.
 - 통합 관리를 통해 데이터 중복을 최소화하고 표준화된 기준에 의해 데이터를 관리한다.
 - 통합 관리를 위한 통제 역할의 의사결정 기구를 마련하며, 단위 기능별로 관리와 책임 소재를 명확히 한다.
- ▶ 원칙3. 지속적으로 품질관리를 수행해야 한다.
 - 데이터 변경은 관련 조직간 협조로 수행하며 데이터 품질관리 담당자가 통제한다.
 - 측정 가능한 지표를 설정하며 데이터 품질관리 담당자를 통해 지속적으로 관리를 수행한다.
 - 데이터 품질관리 담당자의 전문성 확보를 위해 지속적인 교육을 실시한다.
- ▶ 원칙4. 품질활동은 PDCA 사이클을 준수하고 오너십을 부여해야 한다.
 - 데이터 생애주기 전 단계의 데이터 품질을 관리해야 한다.
 - 품질관리 활동은 목적 달성과 지속적인 개선을 위한 접근체계인 PDCA 사이클 기반으로 수행한다.
 - 품질 주체들 간의 상호 협력을 강화하기 위해 역할 정의와 오너십을 명확히 해야 한다.

3.2 품질관리 기능

3.2.1 기능 정의

데이터 품질관리 목적 달성을 위한 전략과 과제에 기반한 데이터 품질관리 원칙을 준수하기 위한 기능을 “품질관리 기준 관리”, “품질 진단 및 개선 관리” 기능으로 정의한다.

그림 18 품질관리 기능



- ▶ 품질관리기준 관리
 - 품질관리지표, 핵심품질항목, 업무규칙 등을 관리
 - 데이터 품질이슈 조사를 수행하고 품질관리기준을 수립하며, 품질관리기준을 변경하는 프로세스를 수행
- ▶ 품질 진단 및 개선관리
 - 품질 진단 계획수립, 품질 진단 수행, 데이터 품질 개선, 데이터 품질통제 프로세스를 수행
 - 품질 진단 및 개선은 각 프로세스 내에 품질 진단 대상 정의(D), 품질 진단 수행(M), 진단결과 분석(A), 개선 수행(I), 품질 통제(C) 5단계 활동으로 진행

3.2.2 기능 세분화

품질관리 기능은 목표 달성과 개선 지향적인 접근체계인 PDCA (Plan ⇨ Do ⇨ Check ⇨ Act)에 따라 세분화하며, 세분화된 기능은 프로세스 정의의 기반이 된다.

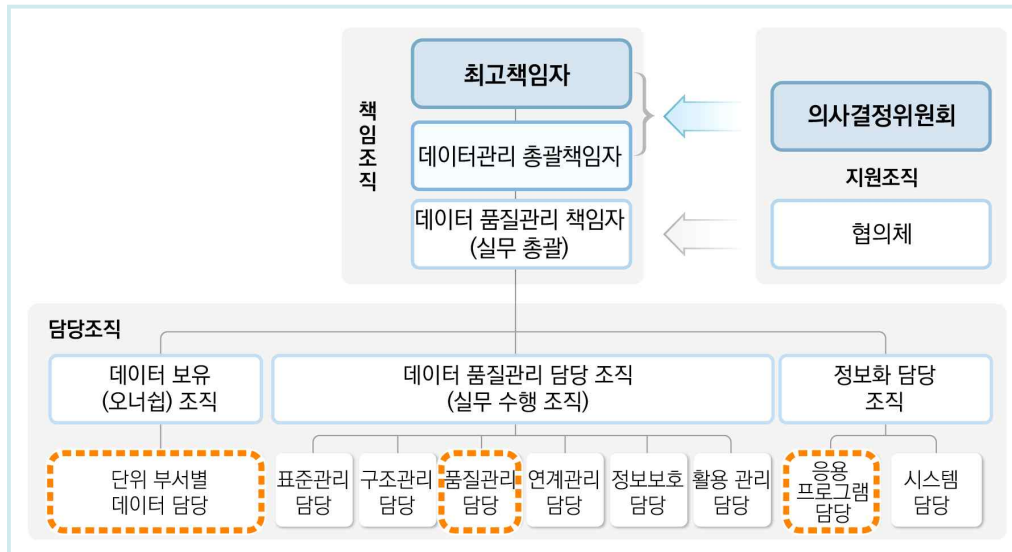
표 8 품질관리 기능 세분화

관리 기능	관리대상	세부 관리기능			
		Plan(계획)	Do(실행)	Check (모니터링 및 통제)	Act(개선)
품질관리 기준	품질관리지표, 핵심품질항목, 업무규칙	데이터 품질이슈 조사 <ul style="list-style-type: none"> 내·외부 품질이슈 파악 품질이슈 분석 품질관리기준 대상 파악 품질관리기준 도출 계획 수립 	품질관리기준수립 <ul style="list-style-type: none"> 품질관리지표 수립 핵심품질항목 도출 업무규칙 도출 품질관리기준 배포 	품질관리기준 점검 <ul style="list-style-type: none"> 점검 대상 데이터 추출 점검조건 도출 및 검증 점검 결과 분석 품질관리기준 개선 요소 도출 	품질관리기준변경 <ul style="list-style-type: none"> 품질관리기준 개선 요소 확인 품질관리기준 개선 계획 수립 품질관리기준 개선 수행 품질관리기준 개선 결과 확인 및 평가
품질진단 및 개선	수집 데이터, 활용 데이터	품질관리대상정의 <ul style="list-style-type: none"> 품질관리 계획 수립 이해관계자 협의 품질진단 계획 배포 	품질진단 수행 <ul style="list-style-type: none"> 품질진단계획 수립 품질진단 대상 정의 품질진단 실시 	품질 통제 <ul style="list-style-type: none"> 품질진단 결과 분석 품질 목표 관리 품질 목표 모니터링 	품질 개선 <ul style="list-style-type: none"> 품질개선 계획 수립 품질개선 준비 품질개선 수행 품질개선 결과 확인 및 평가

3.3 품질관리 조직

데이터 관리 조직 중에서 데이터 품질관리와 관련 높은 조직은 데이터 품질관리 담당, 단위 부서별 데이터 담당, 응용 프로그램 담당으로 정의한다.

그림 19 품질관리 조직



▶ 단위 부서별 데이터 담당

- 오너십을 가지는 업무 영역의 데이터 품질관리 수행
- 품질관리 담당자와 함께 데이터 품질관리 활동 수행

▶ 품질관리 담당

- 품질진단 수행계획 수립
- 품질진단 수행 및 진단 현황 관리
- 품질관리지표와 핵심품질항목, 업무규칙의 도출 및 현황 관리
- 품질진단결과, 오류 데이터 원인분석 및 데이터 정비 결과 보고

▶ 응용 프로그램 담당

- 데이터 품질 관련 응용 프로그램 개선 및 관리

데이터 품질관리 조직의 상세 역할 및 책임은 다음과 같이 정의하였다.

표 9 데이터 품질관리 조직 역할 정의

조직 구성			역할 정의
책임 조직	최고책임자		데이터 관리 및 품질향상을 위한 전사적, 전략적 최고 의사결정 권한 수행
	데이터 관리 총괄책임자		<ul style="list-style-type: none"> 데이터 관리 정책 및 외부 조직간 업무 조정 의사결정 의사결정위원회 소집 및 회의 주재
	데이터 품질관리 책임자 (실무총괄)		<ul style="list-style-type: none"> 데이터 품질관리의 상세 추진계획 수립과 주요 의사결정 수행 품질관리 전반에 대한 실무적인 업무 수행 데이터 품질지표 관리 실무 추진 및 내부 조직간 업무 조정을 위한 협의체 소집 및 회의 주재
지원 조직	의사결정위원회		<ul style="list-style-type: none"> 데이터 관리 정책 및 전략에 대한 주요 이슈 논의 최고책임자 및 데이터 관리 총괄책임자 의사결정 지원
	협의체		플랫폼과 센터간 협업 및 데이터 품질관리의 효율적 추진을 위한 실무 협의체
담당 조직	데이터 보유 조직	단위 부서별 데이터 담당	<ul style="list-style-type: none"> 오너십을 가지는 업무 영역의 데이터 품질관리 수행 품질관리 담당자와 함께 데이터 품질관리 활동 수행
	품질관리 담당		품질관리 기준 관리, 품질진단, 품질진단 분석 및 개선 활동 등 품질관리의 실무 수행
	정보화 담당조직	응용 프로그램 담당 시스템 담당	<ul style="list-style-type: none"> 데이터 품질관리 수행 관련 프로그램 관리 DBMS 데이터 생성, 변경, 폐기 관리 수행

3.4 품질관리 프로세스

3.4.1 데이터 품질관리 프로세스 구성

데이터 품질관리 프로세스는 데이터 품질관리 대 기능은 “품질관리기준 관리”, “품질진단 및 개선 관리”를 PDCA 사이클로 세분화하여 프로세스의 정의 기반인 세부 기능을 도출하였다. 세부기능명칭을 기준으로 데이터 품질관리 프로세스는 다음과 같이 정의한다.

그림 20 데이터 품질관리 프로세스



품질관리기준 관리 기능에서는 “데이터 품질이슈 조사”, “품질관리기준 수립”, “품질관리기준 점검”, “품질관리기준 변경” 프로세스로 구성하며, 품질진단 및 개선 관리 기능에서는 “품질관리대상 정의”, “품질진단 수행”, “품질개선”, “품질통제” 프로세스 등 총 8개 프로세스로 구성한다.

3.4.2 품질관리기준 관리

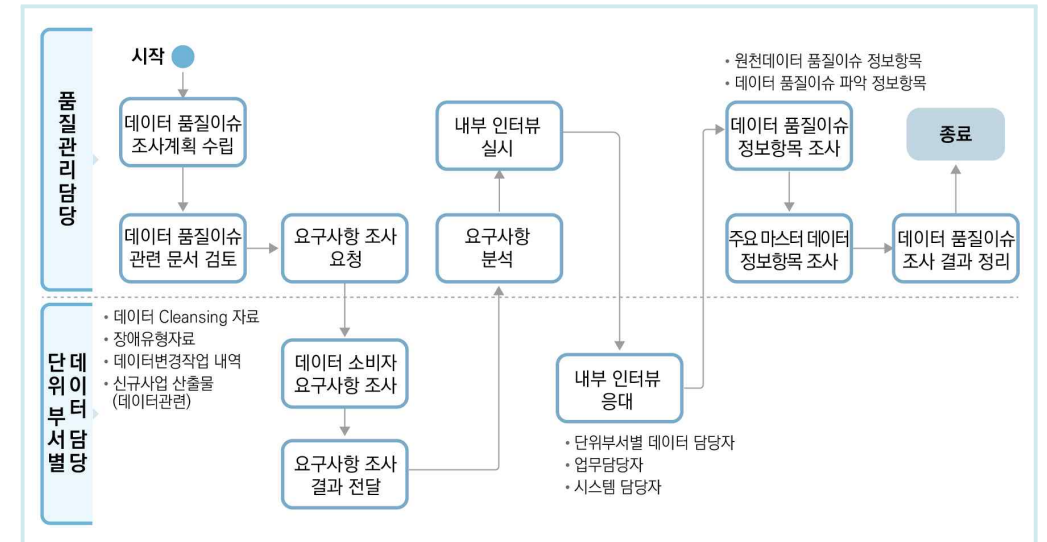
품질관리기준 관리 기능은 “데이터 품질이슈 조사”, “품질관리기준 수립”, “품질관리기준 점검”, “품질관리기준 변경” 프로세스로 구성된다.

▶ 데이터 품질이슈 조사 프로세스

- 데이터 품질관리 기준을 수립 및 변경 등을 위한 품질이슈를 조사하는 프로세스로 품질관리 담당자가 데이터 품질이슈 조사에 대한 계획을 수립하고 데이터 품질이슈 관련 문서 검토, 요구사항

조사, 내부 인터뷰를 수행하며, 데이터 품질이슈 정보항목을 조사하여 데이터 품질이슈 조사 결과를 정리하는 절차로 진행된다.

그림 21 데이터 품질이슈 조사 프로세스



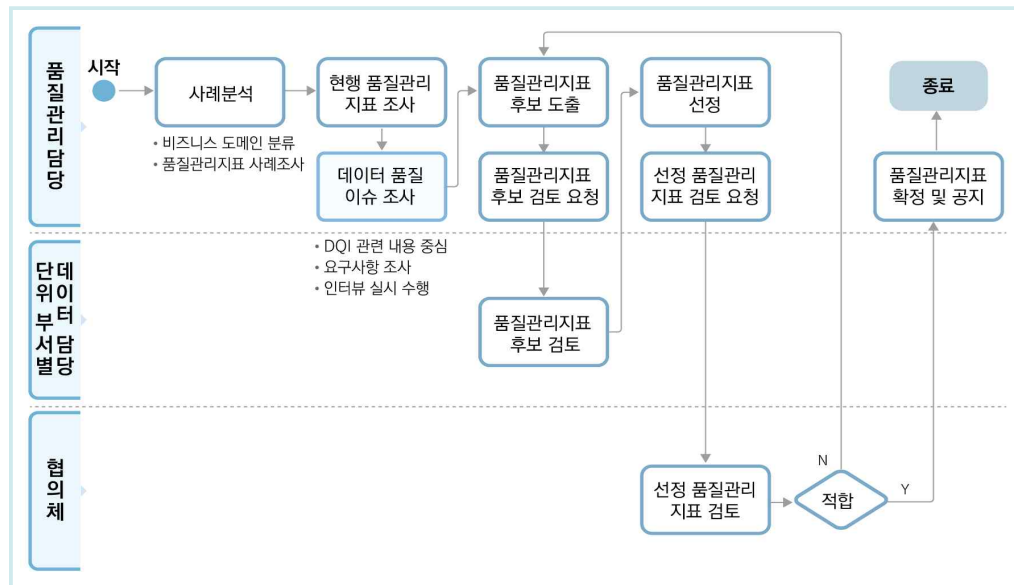
- [데이터 품질이슈 조사계획 수립] 품질관리 담당자는 데이터 품질이슈를 조사하기 위한 계획을 수립한다.
- [데이터 품질이슈 관련 문서 검토] 품질관리 담당자는 품질이슈를 확인할 수 있는 내·외부 관점의 요구사항, 중요데이터 목록, 데이터 Cleansing 자료, 장애유형 자료, 데이터 변경작업 내역, 신규 사업 산출물 등을 대상으로 검토한다.
- [요구사항 조사 요청] 품질관리 담당자는 단위 부서별 데이터 담당자에게 관련 이해관계자들의 데이터 품질이슈 및 요구사항 조사를 요청한다.
- [데이터 소비자 요구사항 조사] 단위 부서별 데이터 담당자는 해당 부서 업무와 관련된 이해관계자들에게 데이터 품질과 관련된 이슈 및 요구사항에 대하여 조사를 수행한다.
- [요구사항 조사 결과 전달] 단위 부서별 데이터 담당자는 품질관리 담당자에게 요구사항 조사 결과를 전달한다.
- [요구사항 분석] 품질관리 담당자는 수신한 요구사항 조사 결과 내용을 분석한다.
- [내부 인터뷰 실시] 품질관리 담당자는 단위 부서별 데이터 담당자 및 업무 담당자를 대상으로 내부 인터뷰를 실시한다.
- [내부 인터뷰 응대] 단위 부서별 데이터 담당자는 품질관리 담당자의 인터뷰에 대하여 적극적인 지원 및 참여를 지원한다.
- [데이터 품질이슈 정보항목 조사] 품질관리 담당자는 원천데이터 품질이슈 정보항목 및 데이터 품질이슈가 파악된 정보항목 등 데이터 품질이슈 정보항목을 조사한다.

- [주요 마스터 데이터 정보항목 조사] 품질관리 담당자는 표준화정보, 마스터데이터, 데이터 모델 등 주요 마스터 데이터의 정보항목을 조사한다.
- [데이터 품질이슈 조사 결과 정리] 품질관리 담당자는 데이터 품질이슈 조사 결과를 정리한다.

▶ 품질관리기준 수립(DQI) 프로세스

- 데이터 품질관리 기준 중 품질관리지표를 수립하는 프로세스로 품질관리 담당자가 사례분석과 현행 품질관리지표, 데이터 품질이슈 조사를 통해 품질관리지표 후보를 도출하고, 단위 부서별 데이터 담당자의 검토 결과를 반영하여 품질관리지표를 선정하고 협의체의 검토 후 품질관리지표를 확정하는 절차로 진행된다.

그림 22 품질관리기준 수립(DQI) 프로세스



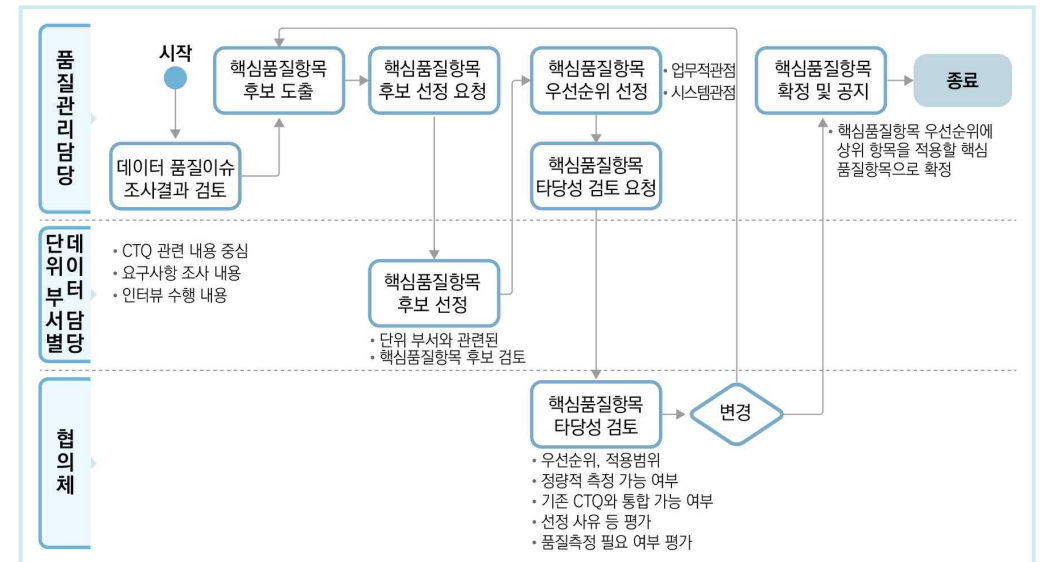
- [사례분석] 품질관리 담당자가 유사 비즈니스 도메인의 품질관리지표를 조사한다.
- [현행 품질관리지표 조사] 품질관리 담당자가 현행 품질관리지표를 조사한다.
- [데이터 품질이슈 조사] 품질관리 담당자가 데이터 품질이슈를 관련 문서, 인터뷰, 요구사항 분석 등을 통해 조사한다. (타 프로세스 수행)
- [품질관리지표 후보 도출] 품질관리 담당자가 사례조사 및 데이터 품질이슈 조사를 통해 품질관리 후보를 도출한다.
- [품질관리지표 후보 검토 요청] 품질관리 담당자가 단위 부서별 데이터 담당자에게 품질관리 후보에 대한 검토를 요청한다.
- [품질관리지표 후보 검토] 단위 부서별 데이터담당자는 요청받은 품질관리지표 후보에 대하여 수행 업무 관련 데이터 중심으로 검토한다.

- [품질관리지표 선정] 품질관리 담당자는 단위 부서별 데이터 담당자의 검토 내용을 참고하여 품질관리지표를 선정한다.
- [선정 품질관리지표 검토 요청] 품질관리 담당자는 선정된 품질관리지표를 협의체에 검토를 요청한다.
- [선정 품질관리지표 검토] 협의체는 품질관리 담당자가 요청한 품질관리지표 선정 내용에 대하여 검토를 수행한다.
- [적합] 협의체에서 검토한 선정 품질관리지표가 적합하지 않을 경우 품질관리지표 후보 도출을 다시 수행하며 적합할 경우 품질관리 담당자에게 의견을 전달한다.
- [품질관리지표 확정 및 공지] 품질관리 담당자는 협의체에서 적합 판정을 받은 선정된 품질관리지표를 확정 및 공지한다.

▶ 품질관리기준 수립(CTQ) 프로세스

- 데이터 품질관리 기준 중 핵심품질항목을 수립하는 프로세스로 품질관리 담당자가 데이터 품질이슈 조사결과를 분석하여 핵심품질항목 후보를 도출하고 단위 부서별 데이터 담당자의 검토를 통해 핵심품질 후보를 결정하며, 우선순위를 선정하여 협의체를 통해 타당성을 검토 후 핵심품질항목을 확정하는 절차로 진행된다.

그림 23 품질관리기준 수립(CTQ) 프로세스



- [데이터 품질이슈 조사결과 검토] 품질관리 담당자가 관련 문서, 인터뷰, 요구사항 분석 등을 통해 데이터 품질이슈 조사 결과에서 핵심품질항목 관련 내용에 대하여 검토한다.
- [핵심품질항목 후보 도출] 품질관리 담당자가 데이터 품질이슈 조사 결과를 바탕으로 핵심품질항목 후보를 도출한다.

- [핵심품질항목 후보 선정 요청] 품질관리 담당자가 단위 부서별 데이터 담당자에게 핵심품질항목 후보 선정을 요청한다.
- [핵심품질항목 후보 선정] 단위 부서별 데이터 담당자는 핵심품질항목 후보를 해당 부서 업무와 관련된 데이터 중심으로 후보를 선정한다.
- [핵심품질항목 우선순위 선정] 품질관리 담당자는 단위 부서별 데이터 담당자가 선정한 핵심품질항목 후보 간 우선순위를 선정하며 업무적 관점(업무영향도, 대외영향도), 시스템 관점(테이블영향도, AP연관도, 도메인중요도) 기준으로 평가한다.
- [핵심품질항목 타당성 검토 요청] 품질관리 담당자는 핵심품질항목들의 우선순위 결과 및 핵심품질항목 타당성에 대하여 협의체에 검토를 요청한다.
- [핵심품질항목 타당성 검토] 협의체는 품질관리 담당자가 요청한 핵심품질항목의 타당성에 대하여 검토하고 우선순위 및 적용 범위 타당성, 정량적 측정 가능 여부, 기존 CTQ와 통합 가능 여부, 선정 사유, 품질측정 필요 여부 등을 평가한다.
- [변경] 협의체에서 타당성 검토를 요청한 핵심품질항목의 변경이 필요하다고 결정되면 핵심품질항목 후보 도출부터 다시 수행하며, 변경이 필요하지 않다고 결정되면 결과를 품질관리 담당자에게 전달한다.
- [핵심품질항목 확정 및 공지] 품질관리 담당자는 협의체에서 타당하다고 결정된 핵심품질항목을 확정 및 공지한다.

▶ 품질관리기준 수립(BR) 프로세스

- 데이터 품질관리 기준 중 핵심품질항목을 수립하는 프로세스로 품질관리 담당자가 데이터 품질이슈 조사 결과를 검토하여, Inside-Out, Outside-In, 근거규정분석에 따라 업무규칙 대상 후보를 도출하며, 단위 부서별 데이터 담당자와 협의하고 협의체 검토를 통해 최종 업무규칙을 선정하는 절차로 진행된다.

그림 24 품질관리기준 수립(BR) 프로세스 (1/2)

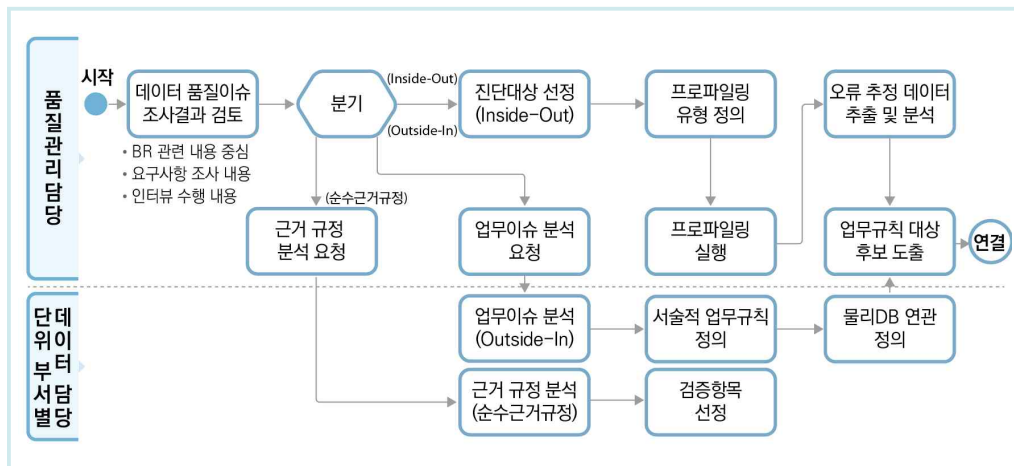
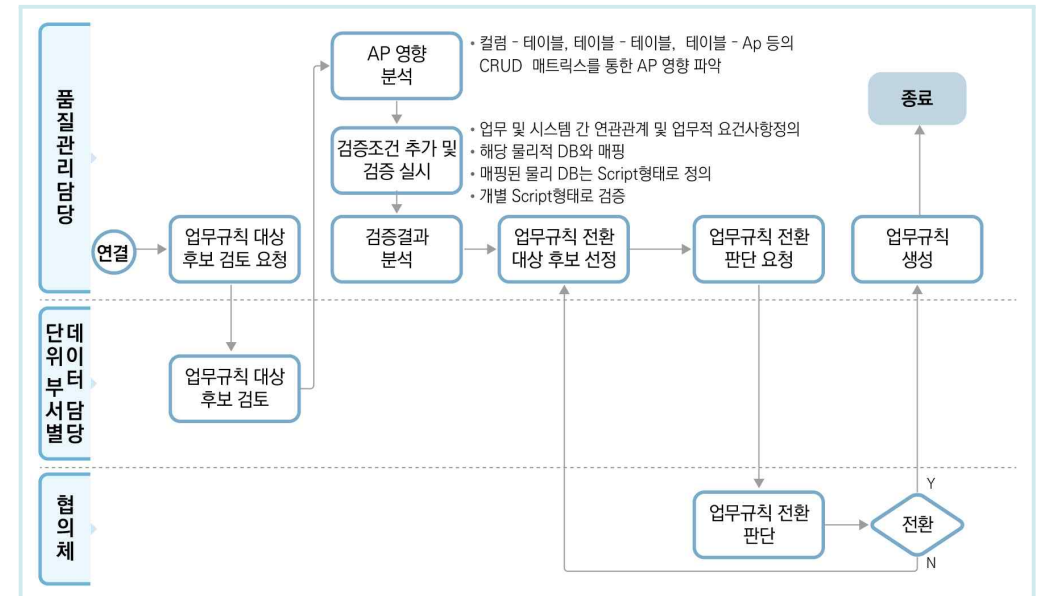


그림 25 품질관리기준 수립(BR) 프로세스 (2/2)



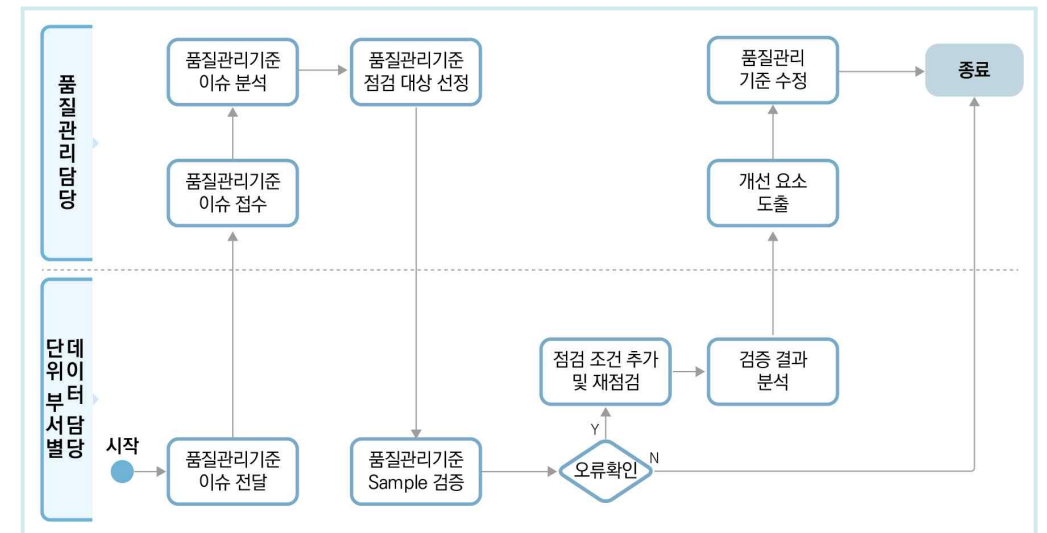
- [데이터 품질이슈 조사결과 검토] 품질관리 담당자가 관련 문서, 인터뷰, 요구사항 분석 등을 통해 데이터 품질이슈 조사 결과에서 핵심품질항목 관련 내용에 대하여 검토한다.
- [분기] 품질관리 담당자는 업무규칙 도출을 위해 Inside-Out 도출방법, Outside-In 도출방법, 순수 근거 규정 도출 방법 중 선택하여 수행하면 중복 수행 가능하다.
- [진단대상 선정 (Inside-Out)] 품질관리 담당자는 프로파일링 분석을 위한 실제 물리적인 DB에 대한 진단 대상을 선정하고 정보시스템의 업무 분류 체계에 따라 진단 대상을 선정한다.
- [업무이슈 분석 요청] 품질관리 담당자는 단위 부서별 데이터담당자에게 업무이슈 분석을 요청한다.
- [근거 규정 분석 요청] 품질관리 담당자는 단위 부서별 데이터담당자에게 근거규정 분석을 요청한다.
- [업무이슈 분석 (Outside-In)] 단위 부서별 데이터담당자는 데이터 품질이슈 조사에서 도출된 업무이슈를 분석하고 업무분석, 현업분석, 업무프로그램 및 산출물 분석, 핵심품질항목 분석을 중요도에 따라 우선순위 지정 후 수행한다.
- [근거 규정 분석 (순수근거규정)] 단위 부서별 데이터담당자는 업무와 관련된 근거 규정에 대하여 근거 규정의 세부 조항, 해당 근거 규정의 내용 파악, 근거 규정에 적용된 용어를 파악하며 근거 규정을 세분화하고 관련 규정을 연계하여 분석한다.
- [검증항목 선정] 단위 부서별 데이터담당자는 상세 근거 규정의 조 또는 항 단위로 항목을 추출하며 명사나 어절을 추출 대상으로 검증항목은 “조”제목 또는 “조항” 의미를 기준으로 선정하고, 조건 항목은 “아니된다”, “아니다”, “아니한” 등 이유를 설명한 문장으로 선정한다.

- [서술적 업무규칙 정의] 단위 부서별 데이터담당자는 업무규칙을 구성하기 위해 서술적 형태 및 쉽게 이해할 수 있도록 업무규칙을 정의하고 검증항목, 조건항목 등 실제 검증에 필요한 가능한 모든 조건을 포함한다.
- [물리DB 연관 정의] 단위 부서별 데이터담당자는 정의된 서술적 업무규칙과 실제 검증에 필요한 물리 DB와 연관관계를 정의하고 시스템-주제영역-테이블-컬럼 순으로 가능한 구체화 및 검증항목, 조건 항목과 물리 DB 속성을 매핑한다.
- [프로파일링 유형 정의] 품질관리 담당자는 진단 대상을 프로파일링 유형(코드, 여부, 번호, 날짜, 금액, 수량, 율)으로 정의한다.
- [프로파일링 실행] 품질관리 담당자는 검증환경 및 시스템 영향도 등을 고려하여 프로파일링을 실행한다.
- [오류 추정 데이터 추출 및 분석] 품질관리 담당자는 프로파일링 수행 결과에 따른 오류 추정 데이터를 추출하고 통계적 기법을 통해 분석하며 업무적, 법률적 요건사항을 고려하여 추가 분석 조건을 설정하여 재검증 할 수 있다.
- [업무규칙 대상 후보 도출] 품질관리 담당자는 오류 데이터에서 업무규칙 대상 후보를 도출하며 기술적, 업무적 중요도에 따라 도출 가능하고 업무규칙 대상 후보는 최대한 측정 가능한 Script 형태로 구성한다.
- [업무규칙 대상 후보 검토 요청] 품질관리 담당자는 업무규칙 대상 후보에 대하여 단위 부서별 데이터 담당자에게 검토를 요청한다.
- [업무규칙 대상 후보 검토] 단위 부서별 데이터 담당자는 업무규칙 대상 후보에 대하여 해당 업무 관련 데이터 중심으로 검토한다.
- [AP영향 분석] 품질관리 담당자는 컬럼-테이블, 테이블-테이블, 테이블-Ap 등의 CRUD 매트릭스를 통한 AP 영향도를 파악한다.
- [검증조건 추가 및 검증 실시] 품질관리 담당자는 업무 및 시스템간 연관관계 및 업무적 요건사항을 정의하고 Script 형태로 물리 DB 매핑(검증 조건을 정의)하고 검증을 바로 실시한다.
- [검증결과 분석] 품질관리 담당자는 검증 조건을 통해 업무규칙 대상 후보의 검증 결과를 분석한다.
- [업무규칙 전환 대상 후보 선정] 업무규칙 대상 후보의 검증 결과에 따라 업무규칙 전환 대상 후보를 선정한다.
- [업무규칙 전환 판단 요청] 품질관리 담당자는 도출한 업무규칙 전환 대상 후보에 대하여 협의체에 전환 여부 검토 및 판단을 요청한다.
- [업무규칙 전환 판단] 협의체에서 품질관리 담당자가 요청한 업무규칙 전환 대상 여부를 결정한다.
- [전환] 협의체에서 전환 결정이 되지 않은 업무규칙은 대상 후보선정부터 재 수행하며, 전환 결정 대상은 품질관리 담당자에게 결정 내용을 전달한다.
- [업무규칙 생성] 품질관리 담당자는 확정된 업무규칙을 업무규칙명, 설명, DQI, CTQ, 검증시스템/테이블/컬럼, 조회 테이블/컬럼, 관련법규 등으로 상세하게 정의한다.

▶ 품질관리기준 점검 프로세스

- 수립된 데이터 품질관리기준을 점검하는 프로세스로 단위 부서별 데이터 담당자가 품질관리기준 이슈를 전달하면 품질관리 담당자는 이슈를 접수 및 분석하여, 품질관리기준 점검 대상을 선정하고 Sample 검증을 통해 오류를 확인하여 오류 발견 시 재검증 수행으로 품질관리기준을 수정하는 절차로 진행된다.

그림 26 품질관리기준 점검 프로세스



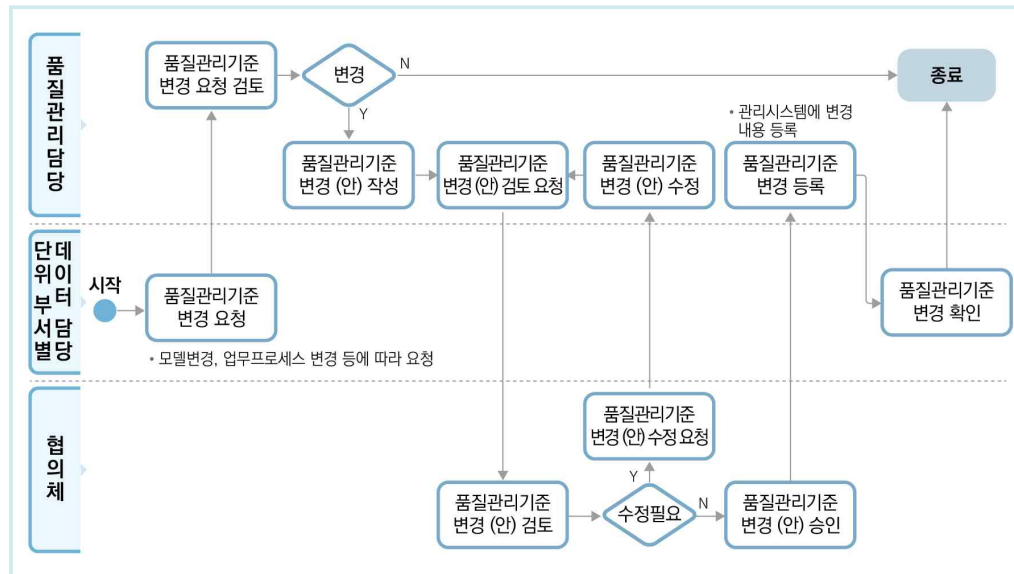
- [품질관리기준 이슈 전달] 단위 부서별 데이터 담당자가 품질관리기준 이슈를 발견하면 품질관리 담당자에게 전달한다.
- [품질관리기준 이슈 접수] 품질관리 담당자는 단위 부서별로 전달되는 품질관리기준 이슈를 접수한다.
- [품질관리기준 이슈 분석] 품질관리 담당자는 단위 부서별로 전달된 품질관리기준 이슈를 통합적으로 분석한다.
- [품질관리기준 점검 대상 선정] 품질관리 담당자는 분석 결과를 토대로 품질관리 기준의 오류를 확인하기 위해 점검 대상을 선정한다.
- [품질관리기준 Sample 검증] 단위 부서별 데이터 담당자는 품질관리 담당자가 선정한 품질관리기준 점검 대상 데이터의 Sample 검증을 수행한다.
- [오류확인] 단위 부서별 데이터 담당자는 Sample 검증을 통해 오류를 확인하면 점검 조건을 추가하여 재점검을 수행하며 오류를 발견하지 않으면 종료한다.
- [점검 조건 추가 및 재점검] 단위 부서별 데이터 담당자는 오류를 확인하고 점검 조건을 추가하여 재점검을 수행하여 오류를 정밀하게 탐색한다.
- [점검 결과 분석] 단위 부서별 데이터 담당자는 재점검 이후 점검 결과를 분석한다.

- [개선 요소 도출] 품질관리 담당자는 단위 부서별 데이터 담당자가 점검한 결과 분석 내용을 바탕으로 품질관리기준 개선 요소를 도출한다.
- [품질관리기준 수정] 품질관리 담당자는 품질관리기준 개선 요소를 바탕으로 품질관리기준을 수정한다.

▶ 품질관리기준 변경 프로세스

- 데이터 품질관리기준을 변경하는 프로세스로 모델변경, 업무프로세스 변경 등에 따라 품질관리기준에 변경이 필요하면, 단위 부서별 데이터 담당자가 품질관리 담당자에게 요청을 하고 품질관리 담당자는 품질관리기준 변경(안)을 작성 후 협의체 검토를 거쳐 품질관리기준을 변경 등록하는 절차로 진행된다.

그림 27 품질관리기준 변경 프로세스



- [품질관리기준 변경 요청] 단위 부서별 데이터 담당자가 데이터 모델 변경, 업무 프로세스 변경 등에 따라 품질관리기준 변경을 품질관리 담당자에게 요청한다.
- [품질관리기준 변경 요청 검토] 품질관리 담당자는 변경 요청이 온 품질관리기준에 대하여 변경 여부를 검토한다.
- [변경] 품질관리 담당자는 변경이 결정되면 품질관리기준 변경(안)을 작성하며, 변경이 결정되지 않으면 해당프로세스는 종료한다.
- [품질관리기준 변경(안) 작성] 품질관리 담당자는 변경이 결정된 품질관리기준에 대하여 변경안을 작성한다.
- [품질관리기준 변경(안) 검토 요청] 품질관리 담당자는 변경한 품질관리기준에 대하여 협의체에 검토를 요청한다.

- [품질관리기준 변경(안) 검토] 협의체는 품질관리 담당자가 요청한 변경된 품질관리기준(안)에 대하여 검토를 수행한다.
- [수정필요] 협의체가 검토 요청이 온 변경된 품질관리기준(안)에 대하여 수정 필요를 결정하면 품질관리 담당자는 해당 품질관리기준(안)을 수정한다.
- [품질관리기준 변경(안) 수정 요청] 협의체에서 수정이 필요한 품질관리기준(안)에 대하여 품질관리 담당자에게 품질관리기준(안) 수정을 요청한다.
- [품질관리기준 변경(안) 수정] 품질관리 담당자는 협의체의 요청에 따라 품질관리기준(안)을 수정한다.
- [품질관리기준 변경(안) 승인] 협의체에서 검토 요청이 온 변경된 품질관리기준(안)에 대하여 변경 승인한다.
- [품질관리기준 변경 등록] 품질관리 담당자는 변경 승인이 떨어진 품질관리기준을 데이터 품질을 관리하는 시스템에 변경 등록한다.
- [품질관리기준 변경 확인] 단위 부서별 데이터 담당자는 요청한 품질관리기준이 변경되었음을 확인한다.

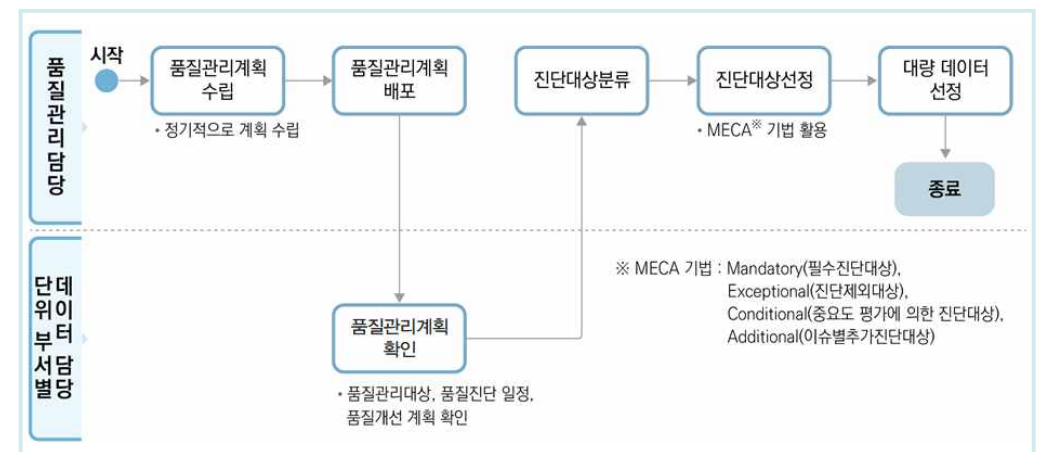
3.4.3 품질진단 및 개선 관리

품질진단 및 개선 관리 기능은 “품질관리 대상 정의”, “품질진단 수행”, “품질개선”, “품질통제”프로세스로 구성된다.

▶ 품질관리 대상 정의 프로세스

- 품질관리를 위해 대상을 정의하는 프로세스로 품질관리 담당자가 정기적으로 품질관리계획을 수립하고, 단위부서별 데이터담당(업무담당자)에게 배포하여 확인하며, 데이터 분류기준에 의해 진단 대상을 분류, MECA 기법 및 통계기법으로 진단 대상을 선정하는 절차로 진행된다.

그림 28 품질관리 대상 정의 프로세스

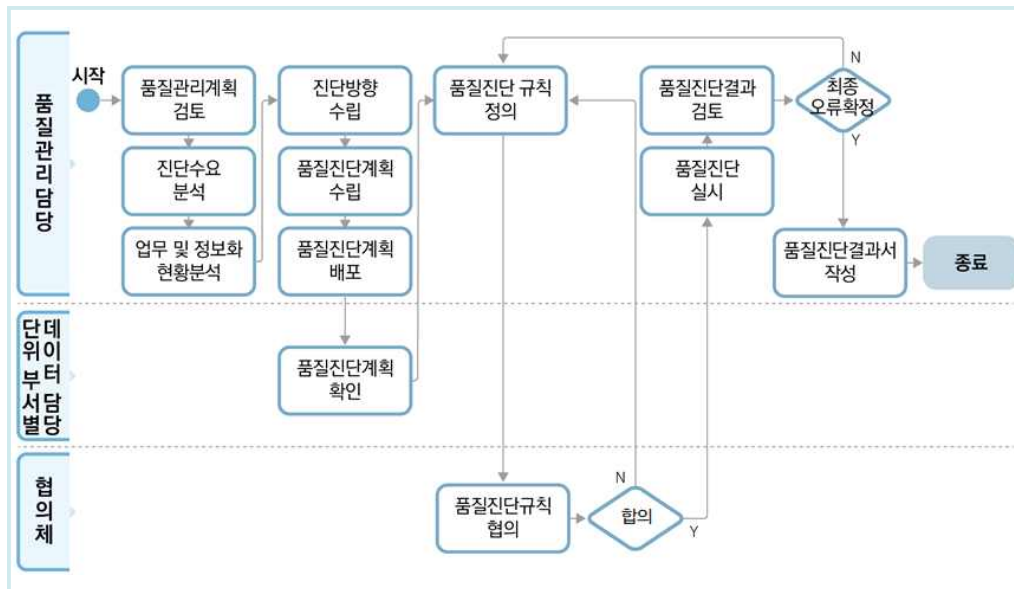


- [품질관리계획 수립] 품질관리 담당자가 정기적으로(예:1년 주기) 데이터 품질진단과 개선에 대한 계획을 수립하고 품질진단 및 개선 목적, DB 현황, 품질관리조직, 품질진단 및 개선 계획 등을 정의한다.
- [품질관리계획 배포] 품질관리 담당자가 수립한 품질관리계획을 단위부서별 데이터담당자 및 업무 담당자에게 배포한다.
- [품질관리계획 확인] 단위 부서의 데이터 담당자는 품질관리 담당자가 배포한 품질관리계획을 수신하고 내용을 확인하며 품질진단 일정, 품질진단 대상, 품질진단 조치 절차 등을 확인한다.
- [진단대상분류] 품질관리 담당자는 빅데이터 플랫폼 및 센터 데이터 관리 방법에 따른 데이터분류체계를 통해 진단대상을 분류한다.
- [진단대상선정] 품질관리 담당자는 빅데이터 플랫폼 및 센터 데이터 관리방법의 MECA 기법에 따라 필수진단대상, 진단제외대상, 중요도 평가에 따른 진단 대상, 이슈별 추가진단 대상을 선정하여 진단대상을 선정한다.
- [대량 데이터 선정] 품질관리 담당자는 통계기법을 적용한 표본 추출을 통하여 대량데이터의 진단 대상을 선정한다.

▶ 품질진단 수행 프로세스

- 수집된 데이터에 대하여 품질기준에 의해서 품질진단을 수행하는 프로세스로 품질관리 담당자가 품질관리계획 검토 후 진단수요 및 업무/정보화 현황을 분석하여 진단방향을 수립하고 진단 계획을 수립 및 배포하며, 품질진단규칙을 정의하여 협의체를 통해 품질진단규칙을 선정하고 품질진단을 수행하는 절차로 진행된다.

그림 29 품질진단 수행 프로세스

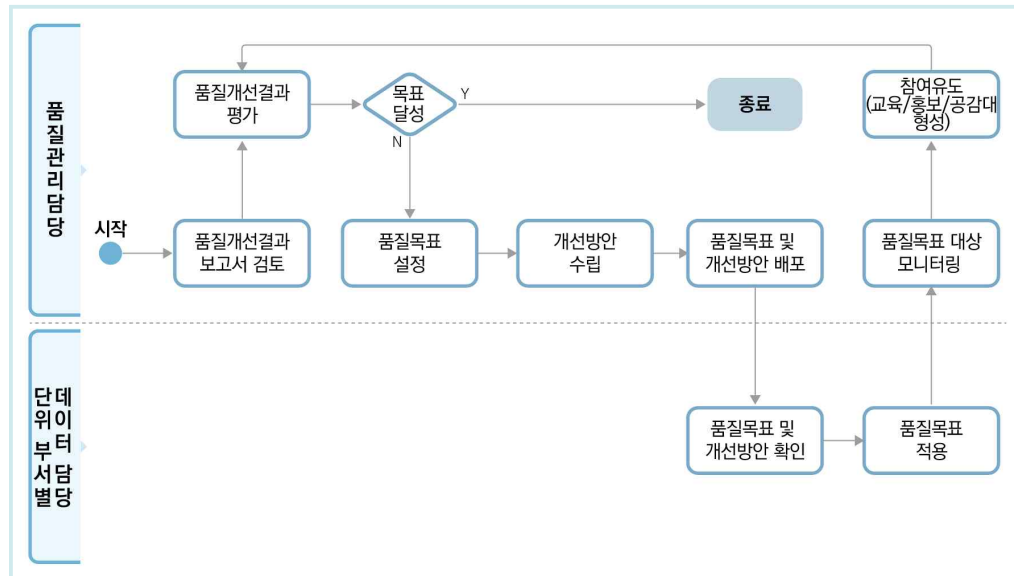


- [품질관리계획 검토] 품질관리 담당자는 기 수립된 품질관리계획을 검토하여, 품질관리대상, 품질관리일정 등을 확인한다.
- [진단수요분석] 품질관리 담당자는 품질진단에 대한 진단수요 등을 분석한다.
- [업무 및 정보화 현황분석] 품질관리 담당자는 진단수요 내용에 대하여 데이터 오류, 이슈 등과 관련된 업무 및 정보화 현황 및 데이터에 대한 조사 분석을 수행한다.
- [진단방향 수립] 주요 품질이슈 도출 후 관리수준, DB 내용, 지원도구, 표준화 등의 관점에서 중요도를 산정하고 이에 따라 진단 대상, 진단범위, 품질목표 수준 등 진단방향을 수립한다.
- [품질진단계획 수립] 품질관리 담당자는 품질진단 목적, 품질진단 범위 및 일정, 품질진단 절차, 품질진단 환경 구성, 이해당사자의 역할 및 책임, 소요예산 등을 고려하여 품질진단 계획을 수립한다.
- [품질진단계획 배포] 품질관리 담당자는 수립한 품질진단계획을 이해관계자(단위 부서별 데이터담당자 = 업무담당자)에게 배포한다.
- [품질진단계획 확인] 단위 부서별 데이터 담당자는 품질관리 담당자가 배포한 품질진단계획 및 품질진단 대상, 품질진단 일정, 품질진단 절차 등을 확인한다.
- [품질진단 규칙 정의] 품질관리 담당자는 품질관리지표, 업무규칙, 핵심품질항목 등 품질진단 규칙을 정의하고 진단대상 항목의 유효값, 유효범위를 통계적으로 분석하여 오류 예상 항목에 대한 진단 규칙을 도출한다.
- [품질진단규칙 협의] 협의체(플랫폼 및 센터)에서는 품질관리 담당자가 요청한 품질진단 규칙 정의에 대하여 협의를 수행한다.
- [합의] 협의체(플랫폼 및 센터)에서 품질진단 규칙 정의에 대한 협의 결과로 품질진단 규칙에 대하여 합의한다.
- [품질진단실시] 품질관리 담당자는 품질진단 규칙을 기반으로 품질진단 대상에 대하여 품질진단을 실시한다.
- [품질진단결과 검토] 품질관리 담당자는 품질진단 결과를 검토하여 오류데이터 확정 및 품질진단 규칙 재정의 여부를 판단한다.
- [최종 오류확정] 최종 오류 확정이 되지 않으면 진단규칙정의/품질진단실시/품질 진단결과 검토 절차를 반복 수행한다.
- [품질진단결과서 작성] 최종 오류가 확정되면 오류확정 내역(데이터셋-테이블, 항목(컬럼)수, 전체건수, 오류건수, 오류율 등)을 값, 구조, 표준, 관리체계 관점에서 품질진단결과서를 작성한다.

▶ 품질 통제 프로세스

- 품질기준이 적용되어 품질진단이 수행되는 환경을 지속적으로 확인하는 프로세스로 품질관리 담당자가 품질개선결과를 검토, 평가하여 품질목표를 설정하고 개선방안을 수립하여 단위부서별 데이터 담당자에게 배포하고, 품질 목표가 적용된 대상에 대하여 모니터링을 실시하며, 참여유도를 위한 다양한 활동을 수행하는 절차로 진행된다.

그림 30 품질 통제 프로세스

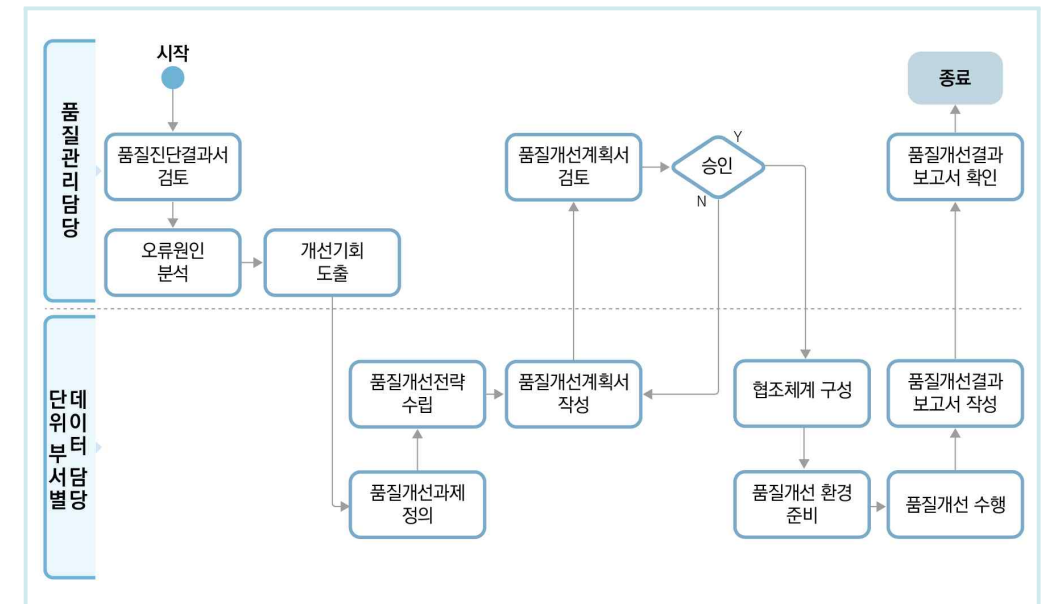


- [품질개선결과보고서 검토] 품질관리 담당자가 단위 부서별 품질개선결과보고서를 검토한다.
- [품질개선결과 평가] 품질관리 담당자가 단위 부서별 품질개선 영역에 대한 효과를 정량적으로 평가하고 목표한 과업 범위에 대한 달성 비율과 개선 전후를 비교한다.
- [목표 달성] 품질관리 담당자가 단위 부서별 품질개선 영역에 대하여 품질개선결과와 평가 후 목표 달성 여부를 확인한다.
- [품질목표 설정] 품질통제 대상에 대한 품질 목표를 설정하고 개선적용 영역의 품질유지 및 개선 미적용 영역의 개선 적용 활동이 지속적으로 실행되도록 품질목표를 관리한다.
- [개선방안 수립] 품질목표를 달성, 유지하기 위한 개선 방안을 수립한다.
- [품질목표 및 개선방안 배포] 품질관리 담당자는 수립된 품질목표 및 개선방안을 단위 부서별 데이터 담당자에게 배포한다.
- [품질목표 및 개선방안 확인] 단위 부서별 데이터 담당자는 배포된 품질목표 및 개선방안을 확인하여 수행한다.
- [품질목표 적용] 단위 부서별 데이터 담당자는 품질목표 관리대상에 품질목표를 적용하여 관리한다.
- [품질목표 대상 모니터링] 품질관리 담당자는 품질목표 관리대상에 대하여 수시 모니터링을 수행한다.
- [목표 달성] 품질관리 담당자가 품질개선결과를 평가하여 품질목표에 달성하지 못하면 품질목표를 다시 설정하고 품질개선방안을 수립하는 절차를 진행한다.
- [참여유도] 품질관리 담당자는 지속적인 품질개선 활동을 수행하도록 이해관계자의 품질개선 참여를 유도하는 교육, 홍보, 공감대형성 등 다양한 활동을 수행한다.

▶ 품질 개선 프로세스

- 품질개선은 품질 오류를 개선하는 프로세스로 품질관리 담당자가 품질진단결과를 검토하여 오류원인 분석과 개선기회를 도출하고, 단위부서별 데이터 담당자가 개선과제 정의 및 개선전략 수립 후 품질개선계획서 작성하고 협조체계를 구성하여 품질개선 활동을 수행하는 절차로 진행된다.

그림 31 품질 개선 프로세스



- [품질진단결과서 검토] 품질관리 담당자가 품질진단결과서 내용을 검토한다.
- [오류원인 분석] 품질관리 담당자는 품질진단결과서의 오류원인을 바탕으로 최종 오류의 근본 원인을 발굴한다.
- [개선기회 도출] 품질관리 담당자는 근본적인 오류 원인을 해소할 수 있는 구체적인 개선기회를 값, 구조, 표준, 관리체계, 응용 프로그램 등을 대상으로 도출하여 이해관계자들에게 배포한다.
- [품질개선과제 정의] 단위부서별 데이터 담당자가 오너십을 가지고, 해당부서의 품질진단결과를 바탕으로 품질개선과제를 정의한다.
- [품질개선전략 수립] 단위부서별 데이터 담당자는 정의한 품질개선과제를 시급성, 중요도 등을 고려하여 우선순위를 결정하고 추진일정을 수립하는 품질개선전략을 수립한다.
- [품질개선 계획서 작성] 단위부서별 데이터 담당자는 품질개선과제 및 품질개선전략을 바탕으로 품질개선 계획을 수립하고 추진체계, 일정, 예산, 수행방식 등 구체적인 개선 계획서를 작성한다.
- [품질개선 계획서 검토] 품질관리 담당자는 단위부서별 데이터 담당자가 작성한 품질개선 계획서를 검토한다.

- [승인] 품질관리 담당자는 단위부서별 데이터 담당자가 작성한 품질개선 계획서를 검토하여 적절여부를 판단하여 승인 또는 반려한다.
- [협조체계 구성] 품질관리 담당자가 품질개선계획서를 승인한 후 품질개선을 위한 관련 이해관계자들에게 협조를 요청한다.
- [품질개선 환경 준비] 단위 부서별 데이터 담당자는 품질개선을 효과적으로 수행하기 위해 사전에 품질개선 환경 준비를 수행한다.
- [품질개선 수행] 단위 부서별 데이터 담당자는 품질개선계획에 따라 오류 원인별 품질 개선을 품질관리체계, 표준화, 값, 구조, 응용프로그램 등 영역별로 수행한다.
- [품질개선결과보고서 작성] 단위 부서별 데이터 담당자는 품질개선결과를 품질관리체계, 표준화, 값, 구조, 응용프로그램 등 개선 영역별로 구분하여 보고서에 작성한다.
- [품질개선결과보고서 확인] 품질관리 담당자는 단위 부서별 데이터 담당자가 작성한 품질개선 결과보고서를 확인한다.

01

개요



1.1 목적 및 기능

- ▶ 연계 정의 : 빅데이터 플랫폼 및 센터 데이터 품질관리에서 연계란 기관 간 데이터 송수신 뿐 아니라 흐름 및 추적관리, 플랫폼간의 데이터 활용까지를 포함하는 광의의 연계를 대상으로 한다.
- ▶ 연계관리 정의 : 연계관리는 연계를 효율적이고 체계적으로 운영하여 연계를 통해 달성하고자하는 목적을 달성하게 하는 것으로 정의하며, 이러한 연계 및 연계관리는 다음의 표와 같이 정리된다.

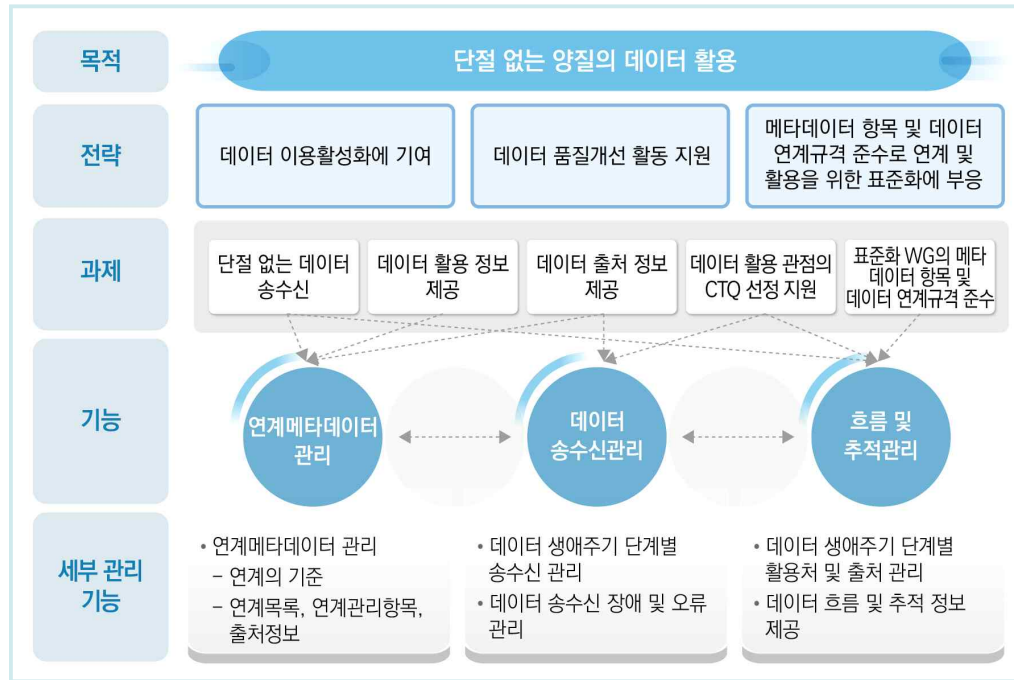
표 10 연계관리 정의

구분	연계 정의	연계관리 정의
협의의 정의	<ul style="list-style-type: none"> • 데이터 송수신 • 빅데이터 플랫폼 및 센터에서 데이터 생애주기 단계별로 데이터를 제공하는 송신과 데이터를 활용하는 수신 간에 데이터를 주고받는 것 	<ul style="list-style-type: none"> • 단절 없는 데이터 송수신 • 송신과 수신 간에 단절 없이 데이터 교환이 이루어지도록 하는 것 • 데이터 송수신 상황을 모니터링하고 장애 및 오류가 발생할 경우 조치 및 개선을 하는 것
확장된 정의	<ul style="list-style-type: none"> • 활용처 정보와 출처정보 생성 및 활용 • 데이터 생애주기 단계별로 송수신 데이터의 활용처와 출처 정보를 생성하고 활용정보와 출처정보를 제공하는 것 	<ul style="list-style-type: none"> • 이용활성화와 품질개선활동 지원 등 데이터 관리 • 활용처 정보 및 출처 정보를 이용한 데이터 활용도·오류원인·영향도 분석을 통하여 이용활성화와 데이터 품질개선 활동을 지원할 수 있도록 하는 것 • 활용처와 출처 정보의 정확성을 진단하고 오류 발생 시 개선하도록 하는 것
광의의 정의	<ul style="list-style-type: none"> • 기관 간 데이터 활용 • 범국가, 플랫폼 및 센터가 데이터 송수신을 표준화 하여 모든 참여 기관이 단절 없이 데이터를 주고받는 것 	<ul style="list-style-type: none"> • 기관 간 데이터 활용을 효율적으로 수행할 수 있도록 하는 것 • 표준화된 연계메타데이터를 이용하여 효율적으로 데이터를 활용할 수 있도록 하는 것 • 연계메타데이터 표준 준수여부를 점검하고 오류 발생 시 조치 및 개선하도록 하는 것

1.1.1 목적

연계관리는 연계메타데이터를 기반으로 데이터 제공기관과 활용기관 간의 데이터 송수신, 흐름 및 추적 관리를 통해 단절 없이 양질의 데이터를 효율적으로 활용할 수 있도록 하는 것을 목적으로 하며, 다음의 그림처럼 도식화할 수 있다.

그림 32 연계관리 목적



▶ 목적 : 단절 없는 양질의 데이터 활용

- 연계는 단절 없이 데이터 송수신이 이루어지게 관리를 해야 한다.
- 연계는 양질의 데이터 즉, 활용 가치가 높은 고품질의 데이터 송수신을 도모해야 한다. 즉, 데이터 이용 활성화 등의 데이터 관리를 지원해야 한다.

1.1.2 기능

▶ 전략

- 연계는 데이터를 제공하는 주체와 이용하는 주체 간의 데이터 이동, 즉 안정적이고 효율적인 송수신이 이루어지도록 표준을 준수하도록 관리해야 한다.
- 데이터를 이용하는 활용처(Target) 정보를 이용하여 데이터의 활용빈도에 따른 데이터 가치를 측정하고, 특정 데이터의 변경이나 이상이 미치는 영향도를 파악할 수 있는 정보를 제공하여 효율적인 데이터 관리를 지원해야 한다.

- 데이터를 제공하는 출처(Source) 정보를 이용하여 문제가 있는 데이터의 근원을 찾을 수 있도록 지원하고, CTQ로 선정된 데이터의 출처를 식별하여 CTQ와 관련된 추가적인 CTQ 대상 선정을 지원하는 등 데이터 품질관리 활동을 지원해야 한다.

▶ 과제

- 단절 없는 데이터 송수신 : 연계의 기본적인 과제로서 출처와 활용처간 단절 없이 데이터 송수신이 이루어지도록 모니터링하고, 문제가 발생할 경우 원인 및 영향도를 분석하여 개선이 이루어지도록 한다.
- 데이터 활용 정보 제공 : 출처별 활용처 정보(Source-to-Target) 정보를 제공하여 효율적인 데이터 관리 활동을 지원할 수 있도록 한다.
- 데이터 출처 정보 제공 : 활용처별 출처 정보(Target-to-Source) 정보를 제공하여 데이터 품질 관리 활동을 지원할 수 있도록 한다.

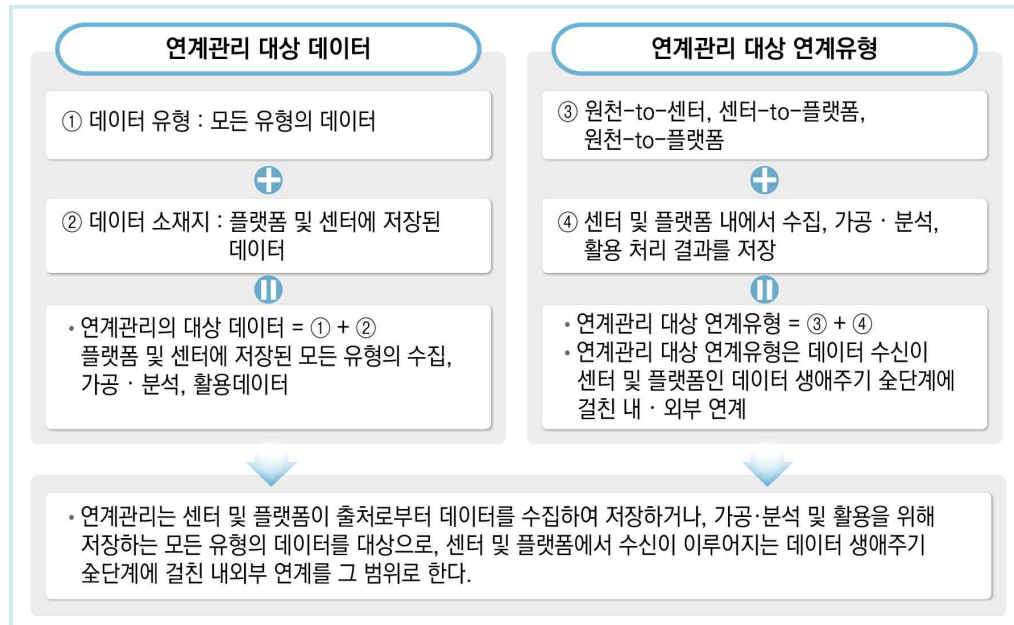
▶ 기능

- 연계메타데이터 관리 : 데이터 송수신 등 연계관리의 기준이 되는 연계메타데이터 (연계목록, 연계관리항목, 출처정보)에 대한 표준을 정하고 이를 준수할 수 있도록 관리하는 기능
- 데이터 송수신 관리 : 빅데이터 생애주기(수집 ⇨ 가공·분석 ⇨ 활용) 및 연계유형 (협약에 의한 외부연계, 소셜 데이터 수집 등의 임의의 외부연계, 내부연계)에 따른 데이터 송수신을 관리하고 장애 및 오류 발생 시 개선하도록 관리하는 기능
- 흐름 및 추적 관리 : 데이터 생애주기 단계별 활용처 및 출처를 관리하여 흐름 및 추적정보를 제공할 수 있도록 관리하는 기능

1.2 관리 범위

연계관리는 출처로부터 수집하여 저장하거나, 가공·분석 및 활용을 위해 저장하는 모든 유형의 데이터를 대상으로, 데이터 생애주기 초단계에 걸쳐서 센터 및 플랫폼에서 수신하는 내·외부 연계를 그 범위로 한다.

그림 33 연계관리 범위



▶ 연계관리 대상 데이터 측면의 범위

- 데이터 유형 : 연계관리 데이터는 정형, 반정형, 비정형 등 모든 유형의 데이터를 대상으로 한다.
- 데이터 소재지 : 연계관리 대상 데이터는 센터 및 플랫폼에 저장된 데이터를 대상으로 한다. 저장된 데이터는 수집데이터, 가공·분석데이터, 활용데이터로 구분한다.
- 종합 : 데이터 유형 측면의 관리 범위 + 데이터 소재지 측면의 관리 범위
연계관리 대상이 되는 데이터는 센터 및 플랫폼에 저장된 모든 유형의 수집, 가공·분석, 활용데이터로 한다.

▶ 연계관리 대상 연계유형 측면의 범위

- 외부연계 : 원천-to-센터, 센터-to-플랫폼, 원천-to-플랫폼 등 데이터의 수신이 센터 및 플랫폼인 연계를 대상으로 한다.
- 내부연계 : 센터 및 플랫폼 내의 수집, 가공·분석, 활용을 위한 연계를 대상으로 한다.
- 종합 : 외부연계 측면의 범위 + 내부연계 측면의 범위
연계관리 대상 연계유형은 데이터 수신이 센터 및 플랫폼인 모든 내·외부 연계로 한다.

▶ 연계관리 범위 종합

- 연계관리는 센터 및 플랫폼이 출처로부터 데이터를 수집·저장하거나, 가공·분석 및 활용을 위해 저장하는 모든 유형의 데이터를 대상으로, 센터 및 플랫폼에서 수신이 이루어지는 데이터 생애주기 쉼단계에 걸친 내·외부 연계를 그 범위로 한다.

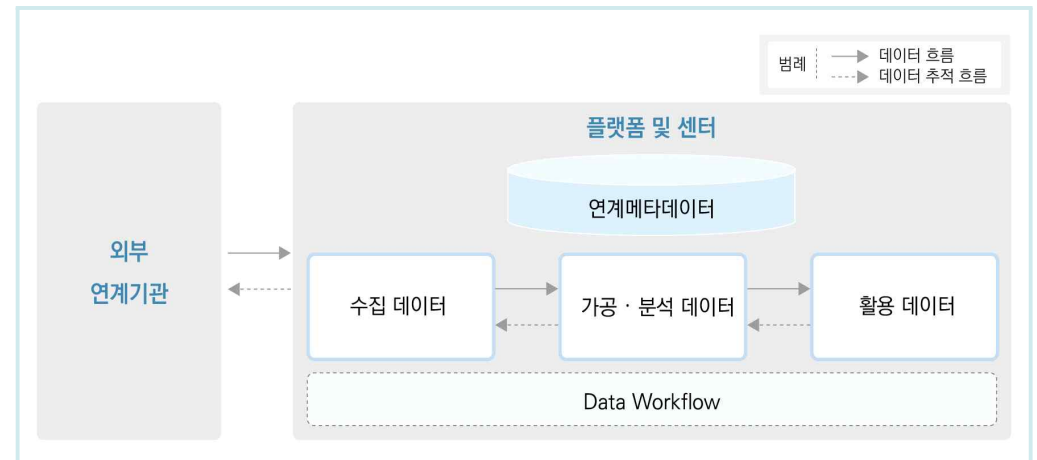
02 모델 수립

2.1 모델 개념

연계관리 모델은 연계관리체계의 대상으로 데이터 송수신 관리 등의 연계관리를 위한 요소를 정의하며, 연계메타데이터 관리모델, 송수신 관리모델, 흐름 및 추적관리 모델로 구성된다.

▶ 연계메타데이터 관리모델

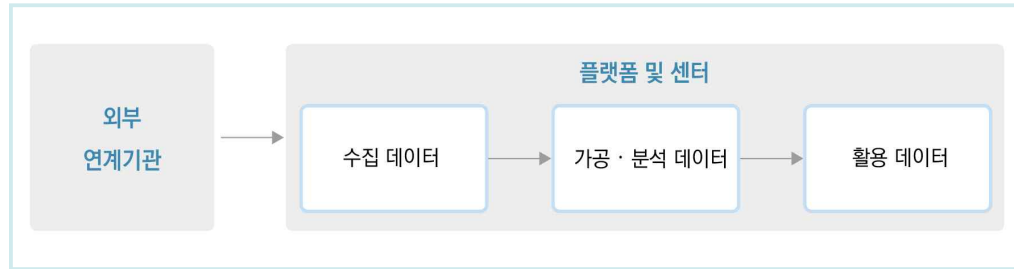
그림 34 품질진단 수행 프로세스



- 송수신, 흐름 및 추적관리 기준인 연계메타데이터(연계목록, 연계관리항목, 출처정보)의 생성, 변경, 오류점검 및 개선 관리
- 표준화 WG(Working Group)에서 정의한 '메타데이터 항목 및 데이터 연계규격'을 준수하도록 관리
- 오픈 포맷 데이터 제공을 위한 활용데이터 생성
- Data Workflow 관리 도구를 이용한 효율적인 연계표준 관리

▶ 데이터 송수신 관리모델

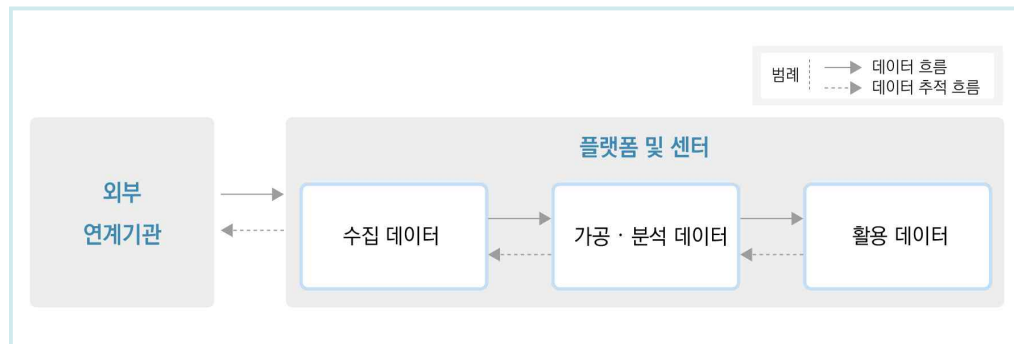
그림 35 데이터 송수신 관리모델



- 데이터 생애주기 쏘단계의 데이터 송수신을 관리
- 데이터 송수신 상황을 모니터링하여 단절 없는 송수신이 이루어지도록 관리
- 모니터링 결과 오류 또는 장애가 발생할 경우 원인과 영향도를 분석하여 개선조치 하도록 관리

▶ 흐름 및 추적 관리모델

그림 36 흐름 및 추적 관리모델



- 데이터 생애주기 단계별로 송수신이 발생할 때 연계유형에 맞는 출처와 활용처 정보를 관리하여 데이터 흐름 및 추적이 가능하도록 관리
- 출처 및 활용처 정보를 이용하여 활용도 분석, 각종 원인 및 영향도 분석 정보를 제공하여 데이터 관리를 지원 할 수 있도록 관리
- 흐름 및 추적 관리를 모니터링하여 오류 여부를 확인하고 개선되도록 관리

2.1.1 연계메타데이터 관리모델

효율적인 데이터 송수신 관리와 흐름 및 추적관리, 범국가적인 데이터 이용 활성화를 위해서는 연계의 기준이 되는 연계메타데이터의 표준화와 표준적인 연계기술을 선정해서 관리해야 한다. 연계메타데이터 관리는 아래의 표와 같이 구성된다.

표 11 연계메타데이터 관리모델 구성

구 분		설 명	연계관리체계 포함 여부
연계 메타 데이터 관리	연계메타데이터 정의	• 연계의 기준이 되는 연계메타데이터 (연계항목, 연계관리항목, 출처정보) 정의	대상
	생애주기별 연계메타데이터 관리	• 데이터 생애주기 단계 및 연계유형별로 적합한 연계메타데이터 생성 및 관리	대상
	연계메타데이터 표준	• 연계메타데이터 구성 표준안	대상
연계기술 관리	데이터 송수신 기술	• 데이터 송수신에 적용되는 기술로 DCAT, CKAN 등 오픈 기술 적용을 권장	• 제외(빅데이터 특성상 다양한 기술 구성이 가능하고, 다양하고 이질적인 민관협력 사업)
	연계관리 기술	• 데이터 워크플로우 관리 도구를 이용한 효율적인 연계메타데이터 관리와 모니터링	• 제외(빅데이터 특성상 다양한 기술 구성이 가능하고, 다양하고 이질적인 민관협력 사업)

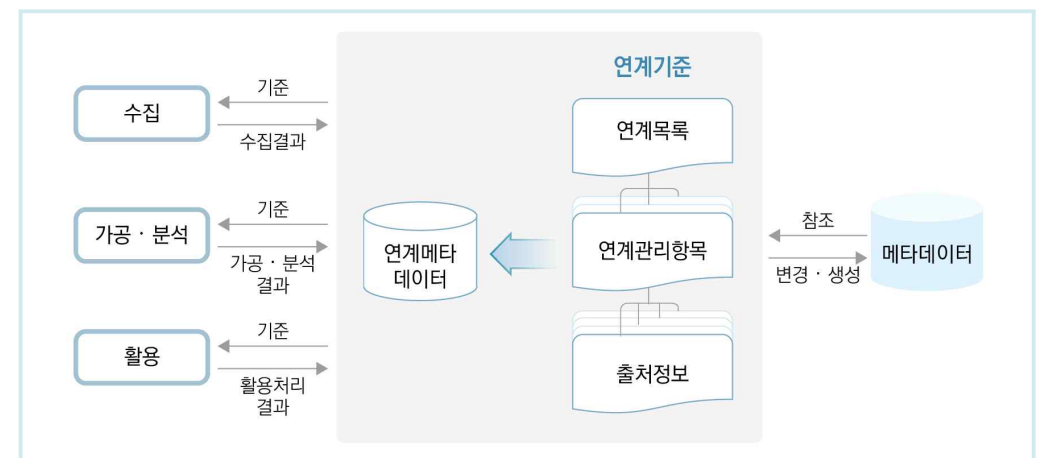
▶ 연계메타데이터 개요

- 데이터 송수신, 즉 연계의 기준 데이터
- 연계목록, 연계관리항목 및 출처정보로 구성

▶ 생애주기별 연계메타데이터 생성 및 관리

- 연계메타데이터는 데이터 생애주기 단계별로 적합한 생성 및 관리 방식이 필요
- 연계메타데이터는 메타데이터를 참조하여 정하는 것이 원칙이지만, 임의의 연계 및 내부연계의 경우 메타데이터를 생성 및 관리

그림 37 데이터 생애주기별 연계메타데이터 관리



▶ 연계메타데이터의 기능

- 데이터 송수신 기준
- 연계메타데이터를 이용하여 데이터 송수신 자동화
- 출처정보 및 연계관리항목의 활용처 정보를 이용한 데이터 흐름 및 추적 관리

▶ 연계메타데이터 관리 자동화

- 파서 및 데이터 워크플로우 관리 도구를 이용하여 연계메타데이터 생성 등을 자동화하여 효율적인 연계메타데이터를 관리

▶ 연계메타데이터 표준

표 12 연계메타데이터 표준(안)

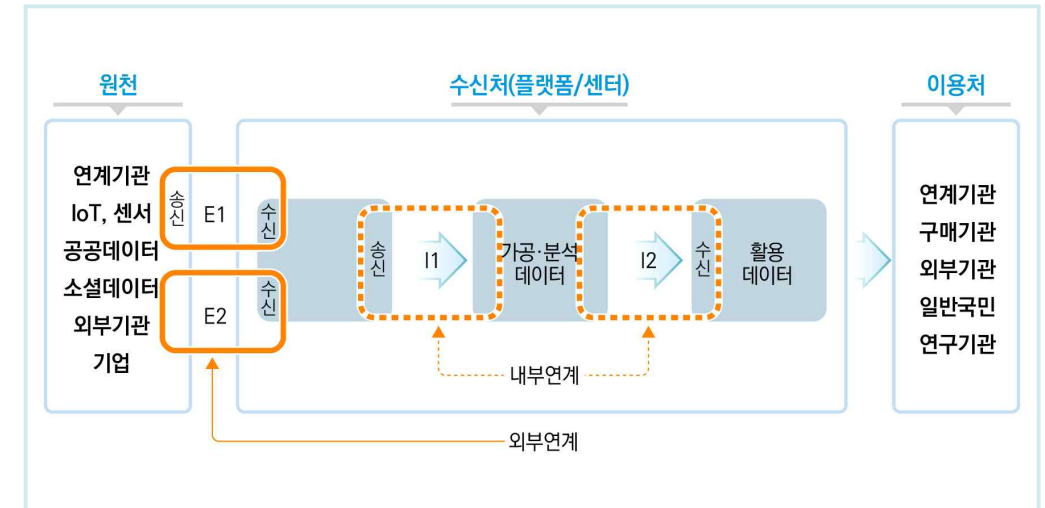
연계메타데이터 항목			설명
연계 목록	연계 정보 정의	연계정보구분	• 연계정보 제공 시 “제공”, 활용 시 “활용”
		● 연계정보분류	• 연계정보의 데이터분류체계에서 해당되는 명칭
		● 연계항목명	• 연계정보가 무엇을 의미하는지 정의
		연계주기	• 연계정보의 연계주기 정의(예, 실시간, 매주, 매월 등)
		연계적용기술	• 연계기능 구현을 위해 적용한 기술(예, ESB, FTP, 크롤링 등)
	연계 주체	제공기관	• 연계정보를 제공하는 기관정보(기관명, 담당자, 연락처 등)
		활용기관	• 연계정보를 활용하는 기관정보(기관명, 담당자, 연락처 등)
연계 관리 항목	연계 항목	● 연계항목명	• 연계정보를 구성하는 세부연계항목들에 대한 명칭
		연계항목설명	• 연계정보를 구성하는 세부연계항목들에 대한 정의
		● 데이터타입	• 연계항목의 데이터 타입을 정의
		● 데이터길이	• 정의된 연계항목별 데이터의 길이를 정의
	활용처 정보	● 데이터베이스명	• 연계항목을 활용하는 데이터베이스명
		● 테이블명	• 연계항목을 활용하는 테이블명
		● 컬럼명	• 연계항목을 활용하는 컬럼명
출처정보	● 데이터베이스명	• 연계항목을 제공하는 데이터베이스명	
	● 테이블명	• 연계항목을 제공하는 테이블명	
	● 컬럼명	• 연계항목을 제공하는 컬럼명	
기타		• 연계 관련 행정사항 등 (예, 연계시작일, 종료일, 변경사유, 변경일, 변경전후 내역)	

● 표준 및 메타데이터와 정합성 확보를 권장하는 항목

2.1.2 데이터 송수신 관리모델

데이터 송수신은 연계유형 및 데이터 생애주기 단계별로 관리하며, 연계관리 대상은 데이터는 수집, 가공·분석, 활용 데이터로 구분하고, 수집 즉, 외부연계의 경우 합의에 의한 수집과 소셜 데이터 등의 임의의 수집으로 구분하여 송수신을 관리한다.

그림 38 데이터 송수신 관리



▶ 생애주기 단계별 데이터 송수신

- 외부연계에 대한 수신처 활동은 수집 활동이다. 외부연계는 외부기관과 협의에 의한 수집(E1)과 소셜 데이터 수집과 같은 임의의 수집(E2)로 구분한다. 연계관리에서는 수집활동의 결과로 저장하는 수집 데이터를 관리한다.
- 내부연계는 수집 등을 통해 저장된 데이터를 이용하여 가공·분석 활동과 외부에 제공하기 위한 처리인 활용 활동으로 구분하며, 각각의 활동의 결과로 저장하는 가공·분석데이터와 활용데이터를 관리한다.
- 내부연계를 관리하지 않을 경우 외부에 제공하는 데이터에 문제가 발생할 경우 문제의 원인이 되는 데이터를 추적할 수 없으므로 내부연계 역시 외부연계와 동일하게 관리해야 한다.

▶ 데이터 송수신 관리는 연계유형 및 데이터 생애주기에 따라 다음의 표와 같이 요약할 수 있다.

표 13 데이터 송수신 관리모델 구성

연계 구분	구분	생애주기 단계	단계설명	송신	수신	비고
외부 연계	E1	합의수집	<ul style="list-style-type: none"> 외부연계기관에서 데이터를 수집하여 저장 외부연계기관과 연계기준을 합의하여 결정 	원천	수신처	
	E2	임의수집	<ul style="list-style-type: none"> 외부연계기관에서 데이터를 수집하여 저장 소셜 데이터 수집 등 외부연계기관과 합의 없는 임의의 수집 	원천	수신처	<ul style="list-style-type: none"> 소셜 데이터의 경우 파싱을 통한 데이터 유형 분류 필요
내부 연계	I1	가공분석	<ul style="list-style-type: none"> 내부 데이터를 이용하여 가공 및 분석 처리하고, 그 결과를 저장하는 내부연계 	원천	수신처	
	I2	활용	<ul style="list-style-type: none"> 내부의 수집, 가공·분석 데이터를 외부에 제공하기 위해 처리하여 저장하는 내부연계 	원천	수신처	<ul style="list-style-type: none"> 오픈 포맷의 데이터 제공

데이터 송수신 관리모델은 공통적인 관리사항과 연계유형 및 생애주기에 따라 상이한 관리 사항으로 구분된다.

▶ 공통적인 관리사항

- 연계계획 수립 및 필요시 연계기능 개발
- 장애 및 오류 모니터
- 장애 및 오류 발생 시 원인분석
- 오류 원인에 따른 영향도 분석
- 개선계획 수립
- 개선조치 및 조치결과 확인

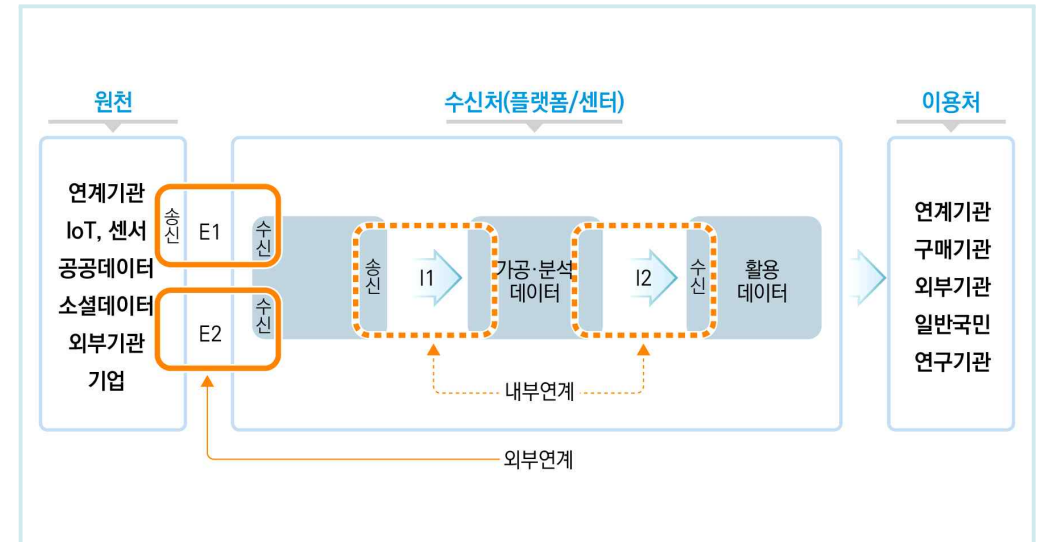
▶ 연계유형 및 생애주기 단계별 상이한 관리사항

- 협의에 의한 수집 : 외부연계기관과 합의에 의한 연계기준(연계메타데이터) 결정과 합의에 의한 연계기준 기반의 연계계획 수립
- 임의의 수집 : 파싱을 통한 연계목록 및 출처정보 추출
- 내부연계(가공·분석, 활용) : 메타데이터를 이용하여 연계기준(연계메타데이터) 자동 생성
- 내부연계(활용) : 오픈 포맷 데이터 제공을 위한 처리

2.1.3 데이터 흐름 및 추적관리

데이터 흐름 및 추적관리를 위해서는 연계메타데이터의 연계관리항목 중 활용처 및 출처정보를 관리해야 하며, 연계유형 및 생애주기에 따라 상호합의에 의한 생성관리, 수신처 임의의 생성관리로 구분한다.

그림 39 데이터 흐름 및 추적관리



▶ 데이터 흐름 및 추적관리는 연계메타데이터의 활용처 및 출처정보 생성 및 관리를 통해 이루어지며 다음 표와 같이 정리할 수 있다.

표 14 데이터 송수신 관리모델 구성

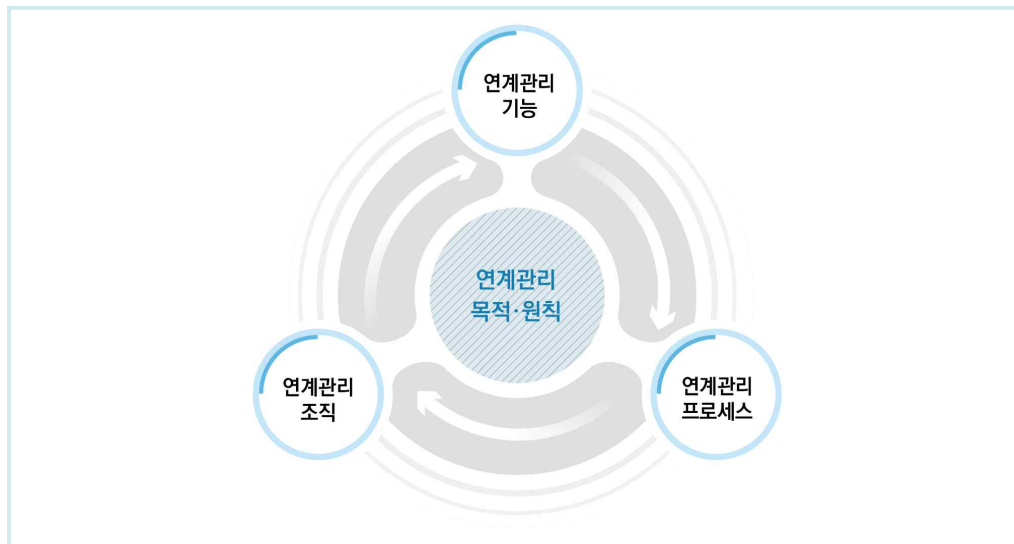
연계 구분	구분	생애주기 단계	규약형태	활용처 및 출처정보 관리
외부연계	E1	합의수집	기관 간 합의 (협약에 의한 수집)	<ul style="list-style-type: none"> 메타데이터를 참조하여 협의 기관 간 합의에 의해 결정
	E2	임의수집	센터 및 플랫폼 표준 (임의의 수집)	<ul style="list-style-type: none"> 메타데이터를 참조하여 협의 기관 간 합의에 의해 결정
내부연계	I1	가공분석	센터 및 플랫폼 표준	<ul style="list-style-type: none"> 메타데이터를 참조하여 생성관리 데이터 워크플로우 관리도구의 지원으로 자동화
	I2	활용	센터 및 플랫폼 표준	<ul style="list-style-type: none"> 메타데이터를 참조하여 생성관리 데이터 워크플로우 관리도구의 지원으로 자동화

03 체계 구성

▶ 개요

연계관리체계는 단절 없는 양질의 데이터 활용이라는 연계목적을 효율적으로 달성하기 위해 연계관리 원칙, 기능, 전담 인력 및 관련 조직과 프로세스를 체계화하는 것이다. 또한 연계수행 뿐 아니라 연계관리를 통해 문제가 있는지 평가하고 문제가 있는 경우 해결을 통해 스스로 개선해 나아갈 수 있는 체계이기도 하다.

그림 40 연계관리체계 개요



- ▶ 연계관리 목적 및 원칙 : 연계 목적(단절 없는 양질의 데이터 활용) 달성을 위한 연계 관리의 기본방향 및 원칙으로 요구사항 및 환경 분석을 통해 도출
- ▶ 연계관리 기능 : 연계메타데이터 관리, 데이터 송수신 관리, 흐름 및 추적관리
- ▶ 연계관리 조직 : 연계관리 조직체계 및 연계관리를 위한 조직별 역할과 책임
- ▶ 연계관리 프로세스 : 목표달성과 개선을 위한 접근체계인 PDCA (Plan ⇨ Do ⇨ Check ⇨ Act)로 연계관리 기능을 세분화하며 세부 기능에 연계관리 조직체계에 맞춰 오너십 부여

3.1 연계 관리 원칙

3.1.1 원칙 도출

연계관리 원칙은 아래와 같이 연계관리 목적 및 전략을 달성하기 위해 수행하는 과제에 대한 수행 규범이어야 하며, 관리체계는 스스로 목표 달성과 개선을 수행할 수 있도록 설계되어야 한다.

▶ 목적

- 연계관리는 단절 없는 양질의 데이터 활용이라는 목적을 달성할 수 있어야 한다.

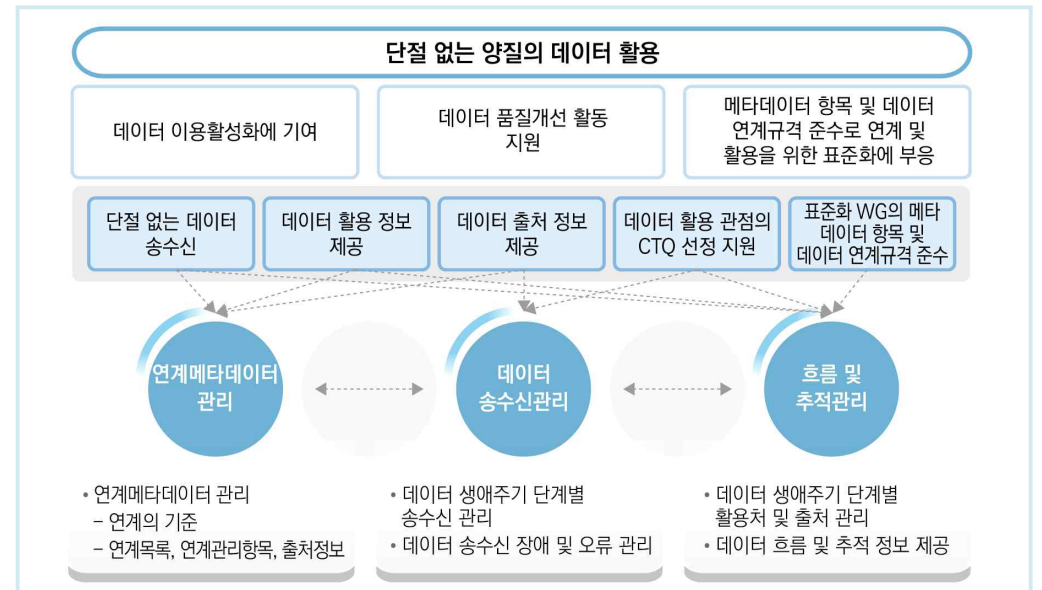
▶ 전략

- 연계관리는 데이터 이용 활성화에 기여해야 한다.
- 데이터 품질개선 활동을 지원할 수 있도록 흐름 및 추적정보를 제공해야 한다.
- 표준화를 통하여 연계 및 데이터 활용을 용이하게 해야 한다.

▶ 과제

- 목적과 전략을 달성하기 위해서 연계는 기본적으로 단절 없는 데이터 송수신이 이루어져야 한다.
- 이용 활성화 정도를 평가하기 위해 데이터 활용 정보를 제공하고, 데이터 관련 문제 발생 시 근원을 찾기 위한 출처정보를 제공하며, 데이터 품질 개선 활동을 지원하기 위한 CTQ 선정 정보를 제공한다.
- 데이터 연계 및 이용을 용이하게 하기 위해서는 플랫폼 표준화 WG에서 정한 표준을 준수해야 한다.

그림 41 연계관리 원칙 도출



3.1.2 원칙 정의

연계관리 목적, 전략 달성을 위한 과제수행과 목적 및 개선 지향적인 연계관리기능을 수행하기 위한 연계관리 원칙은 다음과 같다.

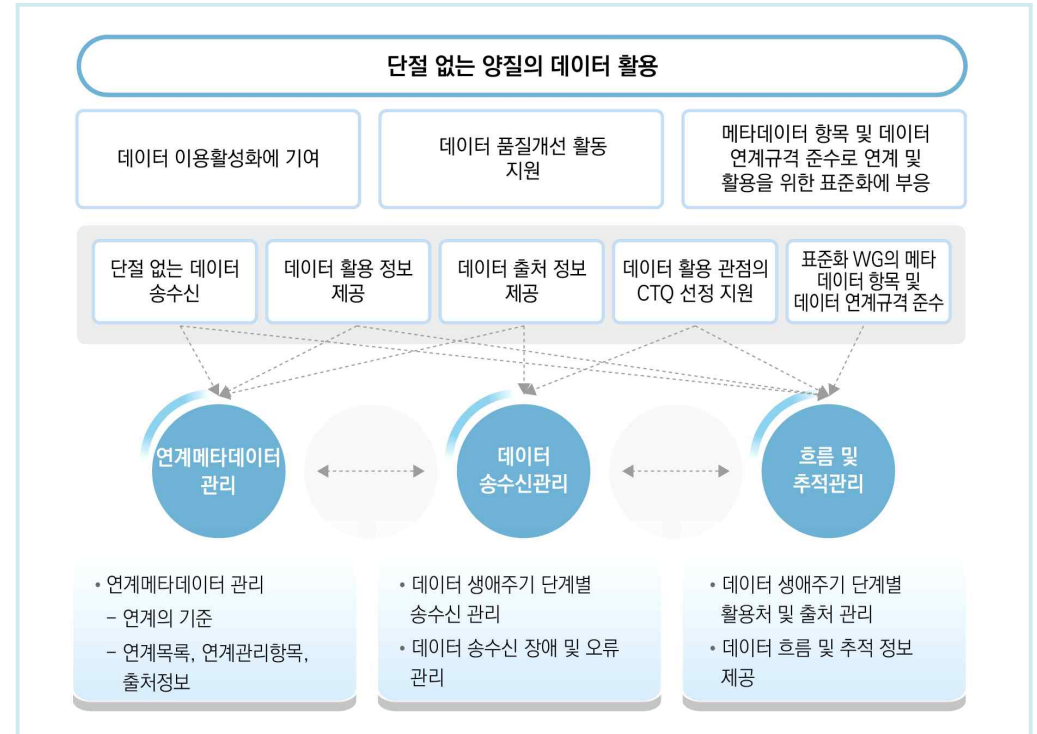
- ▶ 원칙1. 데이터 생애주기 소단계에 걸쳐 송수신 단절이 없어야 한다.
 - 데이터 생애주기 소단계에서 데이터 송수신 단절이 없도록 관리해야 한다.
 - 장애 또는 오류 발생 여부를 모니터링하여 조치 및 개선이 가능해야 한다.
- ▶ 원칙2. 데이터 활용처 정보를 제공해야 한다.
 - 데이터 이용 실적 및 집계 등을 위해 흐름정보를 제공할 수 있어야 한다.
 - 데이터의 생애주기 소단계에서 활용처 정보를 관리해야 한다.
- ▶ 원칙3. 데이터 출처정보를 제공해야 한다.
 - 품질 이상 발생 시 원인파악과 영향도 분석 등을 위해 출처 추적정보를 제공할 수 있어야 한다.
 - 데이터의 생애주기 소단계에서 출처정보를 관리해야 한다.
- ▶ 원칙4. 연계 메타데이터를 관리해야 한다.
 - 데이터 생애주기 소단계에서 연계메타데이터(연계목록, 연계관리항목, 출처정보)를 관리해야 한다.
 - 연계메타데이터와 메타데이터는 정합성을 유지해야 한다.
- ▶ 원칙5. 연계활동은 PDCA 사이클을 준수하고 오너십을 부여해야 한다.
 - 연계활동은 목적달성과 지속적인 개선을 위한 접근체계인 PDCA 사이클을 기반으로 수행한다.
 - 연계주체들 간의 상호협력을 강화하기 위해 역할 정의와 오너십을 명확히 해야 한다.
- ▶ 원칙6. 표준화 WG의 표준을 준수하여 데이터 이용 활성화에 기여한다.
 - 표준 WG의 메타데이터 항목 및 데이터 연계규격을 준수하여 연계 및 데이터 이용 활성화를 제고한다.
 - 오픈 포맷의 데이터 제공을 지향한다.

3.2 연계 관리 기능

3.2.1 기능 정의

연계관리 목적을 달성하고, 연계관리 원칙을 준수하기 위한 연계관리 기능은 연계메타데이터 관리, 데이터 송수신관리 기능, 흐름 및 추적관리 기능으로 구성된다.

그림 42 연계관리 기능



▶ 연계메타데이터 관리 기능

- 데이터 송수신, 흐름 및 추적관리의 기준이 되는 연계메타데이터 생성 및 변경 관리
- 연계메타데이터와 메타데이터 간의 정합성을 진단하여, 오류가 있는 경우 개선할 수 있도록 관리
- 표준화 WG의 메타데이터 항목 및 데이터 연계규격을 준수하여 데이터 이용 활성화를 제고 할 수 있도록 관리

▶ 데이터 송수신 관리 기능

- 단절 없는 데이터 송수신이 이루어지도록 관리
- 빅데이터 특성을 감안한 생애주기 단계별(수집 ⇨ 가공·분석 ⇨ 활용)로 데이터 송수신을 관리

- 송수신 상황을 모니터링하여 장애 및 오류 발생 여부를 확인하고 장애 및 오류가 있을 경우 개선하여 단절 없는 송수신이 이루어지도록 관리

▶ 흐름 및 추적관리 기능

- 데이터 송수신이 이루어지는 생애주기별 출처 및 활용처 정보를 관리
- 출처 및 활용처 정보를 이용한 다양한 분석 정보 제공
- 출처 및 활용처 정보의 정합성을 진단하여, 오류가 있을 경우 개선할 수 있도록 관리

3.2.2 기능 세분화

연계관리 기능은 목표 달성과 개선 지향적인 접근체계인 PDCA (Plan ⇔ Do ⇔ Check ⇔ Act)에 따라 세분화하여 정리하면 아래의 표와 같으며, 세분화된 기능은 프로세스 정의의 기반이 된다.

표 15 연계관리 기능 세분화

관리기능	관리대상	세부 관리기능			
		Plan (계획)	Do (실행)	Check (모니터링 및 통제)	Act (개선)
연계 메타데이터 관리	연계 메타 데이터*	연계메타데이터 정의 <ul style="list-style-type: none"> 연계메타데이터 표준정의 메타데이터와 정합성 유지항목 정의 표준화 WG의 표준 참조 	연계메타데이터 변경 관리 <ul style="list-style-type: none"> 사유별 연계 메타데이터 변경 관리 변경 영향도 분석 	연계메타데이터 정합성 진단 <ul style="list-style-type: none"> 연계메타데이터 및 메타데이터와의 정합성진단 표준화 WG의 표준 준수여부 진단 오류 원인 분석 	연계메타데이터 정합성 오류 개선 <ul style="list-style-type: none"> 오류 영향도 분석 개선 계획 수립 개선조치 및 결과 확인
데이터 송수신 관리	송수신 데이터	연계준비 <ul style="list-style-type: none"> 연계계획 수립 규약에 의한 연계 계획 협의 연계 테스트 	생애주기 단계별 송수신 관리 <ul style="list-style-type: none"> 수집단계 연계관리 가공·분석 단계 연계관리 활용단계 연계관리 	송수신 모니터링 <ul style="list-style-type: none"> 장애 및 오류 모니터링 장애 및 오류 원인 분석 	송수신 오류 개선 <ul style="list-style-type: none"> 오류 영향도 분석 개선 계획 수립 조치 및 결과 확인
흐름 및 추적관리	출처정보	출처정보 정의 <ul style="list-style-type: none"> 규약에 의한 외부 연계 및 내부연계 출처정보 정의 임의의 외부 연계 출처정보 정의 	흐름 및 추적정보 제공 <ul style="list-style-type: none"> 활용처 정보 제공 출처 정보 제공 	출처 정보 정합성 진단 <ul style="list-style-type: none"> 누락 및 오류 진단 오류원인 분석 	출처 정보 정합성 오류 개선 <ul style="list-style-type: none"> 오류 영향도 분석 개선 계획 수립 개선조치 및 결과 확인

*연계메타데이터 : 플랫폼의 표준화 WG(Working Group)에서 정의한 '메타데이터 항목 및 데이터 연계규격' 준수

3.3 연계관리 조직

데이터 관리 조직 중 연계관리 담당 및 관련 의사결정조직, 지원조직, 데이터 관리 담당조직별 역할을 다음의 표와 같이 정의한다.

표 16 연계관리 조직별 역할 정의

조직구분	연계관련 역할
의사결정조직	최고책임자 <ul style="list-style-type: none"> 연계관련 최고 의사결정권자 필요시 의사결정위원회를 통해 의사결정
	데이터관리 총괄 책임자 <ul style="list-style-type: none"> 연계관리 총괄 최고책임자를 보좌 연계관련 주요 정책 및 외부 조직간 업무 조정을 위한 의사결정위원회 소집 및 회의 주재
	데이터품질관리 책임자 (실무 총괄) <ul style="list-style-type: none"> 연계관리 실무 총괄 데이터 관리 총괄 책임자 보좌 의사결정위원회 회의 지원 연계 실무 추진 및 내부 조직간 업무 조정을 위한 협의체 소집 및 회의 주재
지원조직	의사결정위원회 <ul style="list-style-type: none"> 연계관련 주요 정책 심의, 외부 조직간 업무 조정 최고책임자 및 데이터관리 총괄 책임자의 의사결정을 지원 회의 소집 및 주재는 최고책임자 또는 데이터관리 총괄 책임자가 수행
	협의체 <ul style="list-style-type: none"> 연계데이터 품질검증, 내부조직 간의 연계관리 업무공유 내부 연계관련 실행과제 및 이슈 등의 조치 연계 업무 수행에 필요한 실무 차원의 지원 수행 내부 조직 간의 업무분담 및 조정 회의 소집 및 주재는 데이터품질관리 책임자가 수행

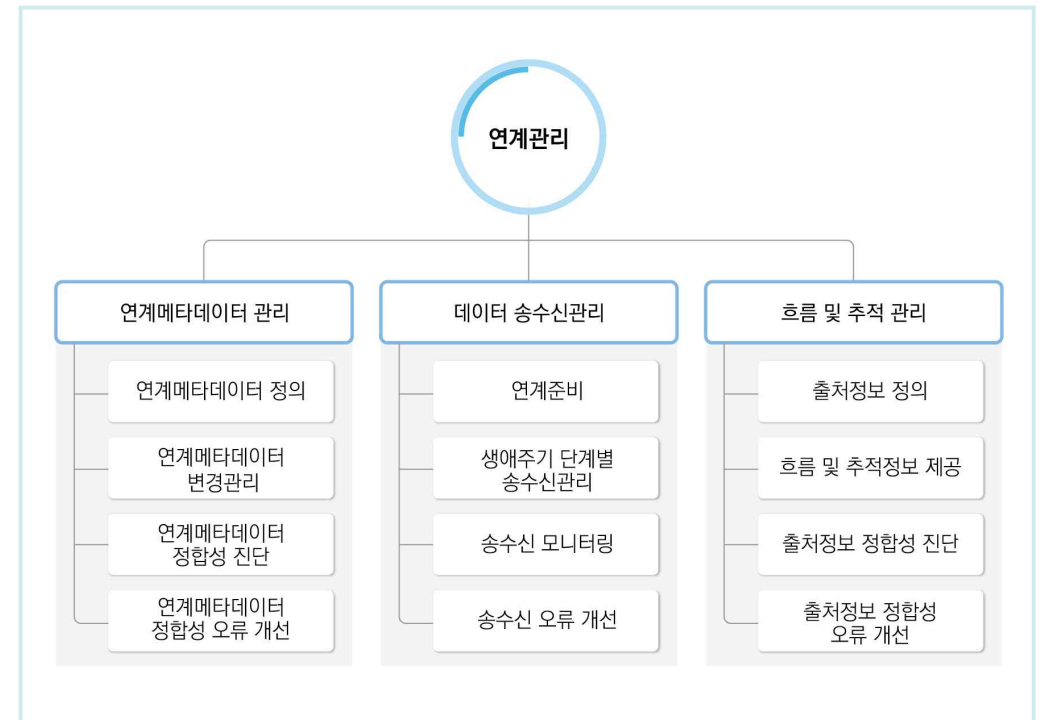
조직구분		연계관련 역할
담당 조직	데이터보유(오너십) 조직	<ul style="list-style-type: none"> 보유 데이터에 대한 연계 타당성 및 제공여부 평가 보유 데이터에 관한 연계관련 오류정제 등 개선활동 수행
	데이터 품질관리 담당 조직 (실무수행 조직)	품질관리 담당
		구조관리 담당
		표준관리 담당
		연계관리 담당
	정보화 담당조직	<div>응용 프로그램 담당</div> <div>시스템 담당</div>

3.4 연계 관리 프로세스

3.4.1 데이터 연계관리 프로세스 구성

연계관리 프로세스는 주요 기능인 연계메타데이터 관리기능, 데이터 송수신 관리기능, 흐름 및 추적관리 기능을 목표 및 개선 지향적인 관리 사이클인 PDCA (Plan ⇨ Do ⇨ Check ⇨ Act) 관점에서 세분화하여 아래의 그림과 같이 구성한다.

그림 43 연계관리 프로세스 구성도



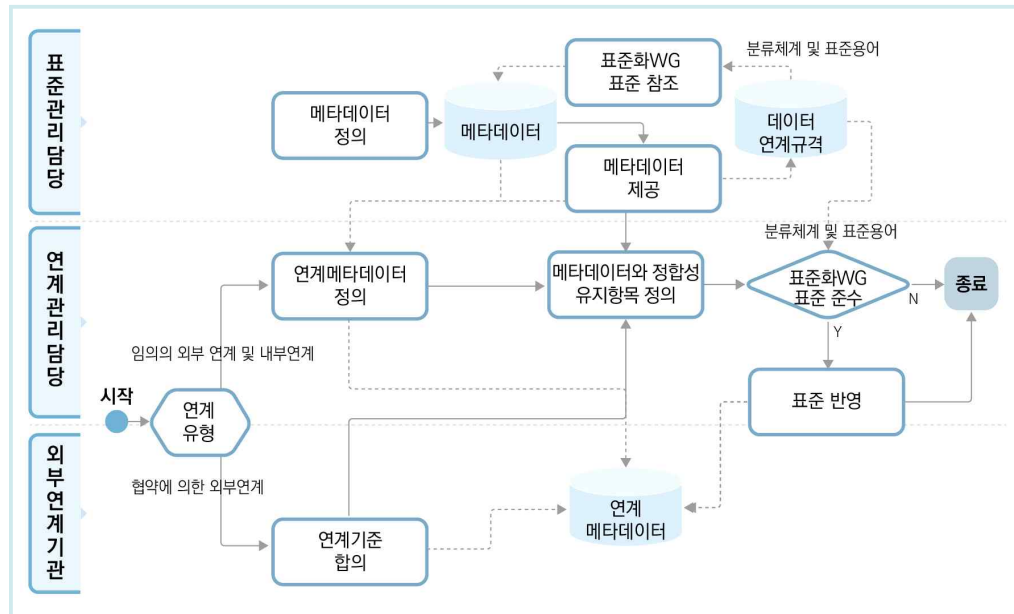
3.4.2 연계메타데이터 관리

연계메타데이터 관리는 연계메타데이터 정의, 연계메타데이터 변경관리, 연계메타데이터 정합성 진단, 연계메타데이터 정합성 오류 개선 프로세스로 구성된다.

▶ 연계메타데이터 정의 프로세스

연계메타데이터를 정의하는 프로세스로 주요 활동은 연계메타데이터 표준을 정의하고 메타데이터와 정합성 유지 항목 정의 및 표준화 WG의 표준준수 여부를 확인하고 오류 시 개선하는 활동으로 구성된다.

그림 44 연계메타데이터 정의



- [연계유형] 연계메타데이터는 빅데이터 생애주기 및 연계유형에 따라, 협약에 의해 연계메타데이터가 결정되는 외부연계와, 파싱에 의해 출처정보를 추출하는 소셜 데이터 수집 등의 임의의 외부연계, 메타데이터에 의해 자동으로 연계메타 데이터를 생성 할 수 있는 가공분석, 활용 등의 내부연계로 분류된다.
- [표준화 WG 표준 참조] 플랫폼 표준관리담당은 메타데이터 정의를 위해 한국지능정보사회진흥원의 메타표준 및 표준용어를 참조한다. 센터 표준관리담당은 메타데이터 정의 시 플랫폼 메타표준 및 표준용어를 참조한다.
- [메타데이터 제공] 플랫폼 표준관리담당은 한국지능정보사회진흥원의 메타표준 및 표준용어 관리를 지원하기 위해 메타데이터를 제공한다. 센터 표준관리담당은 플랫폼의 메타표준 및 표준용어 정의를 지원하기 위해 메타데이터를 제공한다.
- [메타데이터 정의] 플랫폼 표준관리 담당은 표준화 WG의 메타데이터 항목 및 데이터 연계 규격을 참조하여 메타데이터를 정의한다. 센터 표준관리 담당은 플랫폼 메타데이터를 참조하여 메타데이터를 정의한다.
- [연계메타데이터 정의] 소셜 데이터 등의 임의의 수집 및 가공·분석, 활용을 위한 내부연계는 연계관리담당이 연계메타데이터를 정의하고, 데이터 워크플로우 관리도구 등을 이용한 자동화 생성이 가능할 경우 최대한 자동화를 통해 효율적으로 연계메타데이터를 생성할 수 있도록 한다. 임의의 외부연계는 메타데이터를 참조하여 수집하고, 수집된 데이터에서 파싱 및 데이터 워크플로우 관리 도구를 이용하여 출처정보를 자동 생성할 수 있도록 한다. 내부연계인 가공·분석, 활용은 메타데이터를 참조하여 연계메타데이터를 정의하고 데이터 워크플로우 관리 도구를

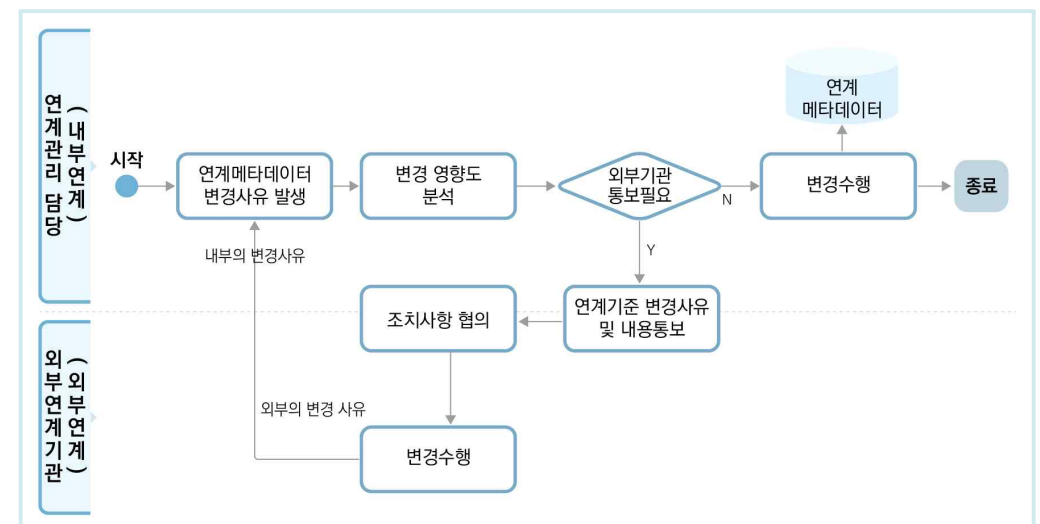
이용하여 연계메타데이터를 자동으로 생성할 수 있도록 한다.

- [연계기준합의] 협약에 의한 외부연계의 경우 메타데이터를 참조하여 외부연계기관과 협의를 통하여 연계메타데이터를 정의한다. 외부연계기관과 상호 합의한 연계메타데이터를 등록한다.
- [메타데이터와 정합성 유지항목 정의] 연계관리 담당은 연계메타데이터 정의 시 메타데이터를 참조하여 정의해야 하고, 활용처 정보 및 출처정보 등과 같이 연계메타데이터와 메타데이터 간에 정합성을 유지해야 하는 항목을 정의한다.
- [표준화 WG 표준준수] 메타데이터는 표준화 WG의 표준을 참조하고, 연계메타데이터가 메타데이터를 참조하여 전체적인 정합성을 유지하고 있지만, 플랫폼 간 연계 효율성 제고를 위해 한국지능정보사회진흥원에서 관리하는 분류체계 및 표준용어를 반영하고 있는지 여부를 재확인한다.
- [표준반영] 한국지능정보사회진흥원에서 관리하는 표준을 반영할 필요가 있는 경우 한국지능정보사회진흥원의 표준용어와 데이터분류체계를 참조하여 연계메타데이터를 정의한다.

연계메타데이터 변경관리 프로세스

연계메타데이터의 변경을 관리하는 프로세스로 변경사유별 변경에 따른 영향도를 분석하여 변경계획을 수립하고 변경조치 실시 후 그 결과를 확인하는 프로세스이다.

그림 45 연계메타데이터 변경관리



- [연계메타데이터 변경사유 발생] 연계메타데이터는 최초 정의 이후 상황에 따라 변경해야 하는 사유가 발생한다. 변경 사유로는 메타데이터 변경, 조직의 변경 및 역할변경, 데이터 및 시스템 구조 변경 등이 있고, 외부연계기관의 변경으로 인한 연계메타데이터 변경이 발생하기도 한다.
- [변경 영향도 분석] 연계메타데이터 변경사유가 발생하면 변경으로 인한 영향도를 분석하여 업무량과 관련 조직 등을 파악한다. 영향도 분석은 활용처 정보(Source-to-Target)를 이용하여

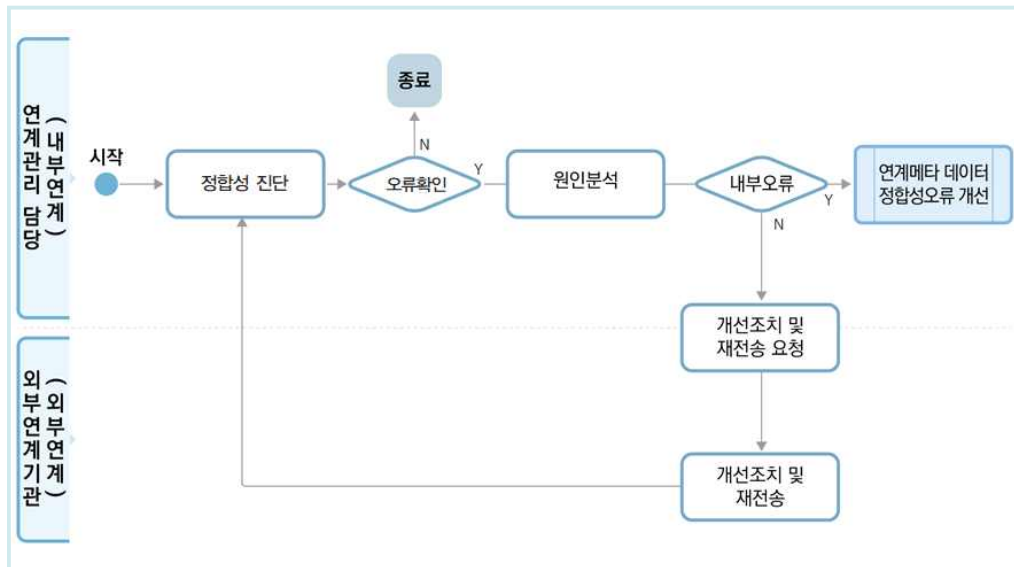
효율적으로 파악할 수 있도록 한다.

- [외부기관 통보필요] 연계메타데이터의 변경으로 외부연계기관도 조치를 취해야 하는지를 판단한다.
- [연계기준 변경사유 및 내용통보] 연계메타데이터 변경으로 외부연계기관도 조치를 취해야 할 경우 외부연계기관에게 변경사유 및 변경 내용을 통보하여 상호 협의하여 변경할 수 있도록 한다.
- [조치사항 협의] 연계메타데이터 변경에 따라 단절 없는 송수신이 이루어질 수 있도록 메타데이터를 참조하여 외부연계기관과 협의한다.
- [변경수행] 연계메타데이터 변경을 수행하고 필요시 테스트 및 데이터 재전송 등의 후속조치를 수행한다.

▶ 연계메타데이터 정합성 진단 프로세스

연계메타데이터의 정합성을 진단하여 오류를 발견할 때 원인을 분석하는 프로세스로 메타데이터와 정합성 진단 및 표준화 WG의 연계규격 준수여부 점검하고 오류 발생 시 원인 분석을 통해 개선을 수행한다.

그림 46 연계메타데이터 정합성 진단



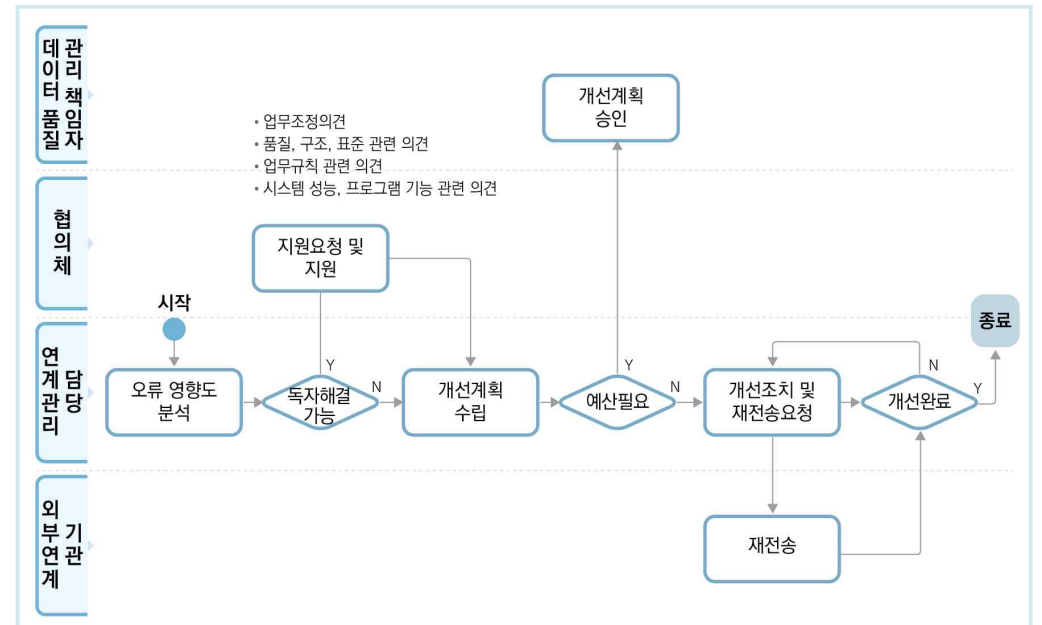
- [정합성 진단] 연계메타데이터와 메타데이터의 정합성 및 연계메타데이터와 한국지능정보사회진흥원의 데이터분류체계 및 표준용어 준수여부를 진단한다.
- [오류확인] 메타데이터와 정합성 및 표준WG 표준과 한국지능정보사회진흥원의 데이터분류체계 및 표준용어 준수여부에 오류가 있는지 여부를 확인한다.
- [원인분석] 정합성 및 표준 준수 오류가 발생할 경우 그 원인을 파악한다. 연계메타데이터의 출처정보(Target-to-Source)를 이용하여 효율적인 원인분석이 가능하도록 한다.
- [내부오류] 오류 원인 분석 결과 센터 및 플랫폼 내부의 문제인지 외부연계기관의 문제인지를

확인한다.

- [개선 조치 및 재전송 요청] 외부요인에 의한 오류인 경우 외부연계기관에게 정보를 전달하고 개선 조치 및 조치 후 필요시 재전송을 요청한다.
- [개선 조치 및 재전송] 오류 문제 해결을 위한 조치를 수행하고 필요시 연계 데이터를 재전송한다.
- [정합성 오류 개선] 내부요인에 의한 정합성 오류의 경우 연계메타 데이터 정합성 오류 개선 프로세스를 수행한다.

▶ 연계메타데이터 정합성 오류 개선 프로세스

그림 47 연계메타데이터 정합성 오류 개선



연계메타데이터의 정합성 문제를 개선하고 조치 결과를 확인하는 프로세스로서, 오류원인 분석결과에 따른 영향도를 분석하고, 분석 결과에 따라 개선계획 수립 및 개선 조치를 수행하며, 개선 조치에 따른 개선 여부를 확인하는 활동으로 구성된다.

- [오류 영향도 분석] 연계메타데이터의 정합성 오류가 발생할 경우 원인에 따라 오류가 미치는 영향을 분석하여 문제해결, 개선 등의 조치를 위한 업무량을 파악한다. 연계메타데이터의 활용처 정보(Source-to-Target)를 이용하여 오류로 인해 영향을 받는 데이터를 효율적으로 파악한다.
- [독자해결가능] 오류 및 표준 미준수를 연계관리담당이 독자적으로 해결할 수 있는지를 확인하여, 독자해결이 불가능 할 경우 협의체 및 데이터품질관리 책임자를 통하여 관련자에게 개선방안 지원을 요청한다. 개선방안 요청 시 연계메타데이터, 오류 및 미준수 현상, 원인 및

영향도 분석정보를 제공한다.

- [지원요청 및 지원] 오류 해결을 위한 개선방안수립 지원 요청을 받은 분야별 담당자는, 담당별 의견을 제시한다. 필요시 협의체는 조직별 역할을 조정하여 의견을 제시한다.
- [개선계획] 연계관리담당은 독자적 해결이 가능할 경우 독자적으로, 관련자의 조력을 받아야 할 경우 관련자의 의견을 수렴하여 오류 문제를 해결하기 위한 계획을 수립한다.
- [예산필요] 개선계획에 시스템 성능 개선, 프로그램 개발, 정보자원 추가 도입 등의 사유로 예산이 필요한지 여부를 확인하여 예산이 필요할 경우에는 데이터품질관리 책임자에게 개선계획 승인을 요청한다.
- [개선계획승인] 개선계획 실행에 예산이 필요할 경우 센터 및 플랫폼 의사결정 절차 및 규정에 따라 최고책임자, 데이터관리 총괄 책임자, 의사결정위원회, 데이터품질관리 책임자의 승인을 얻는다.
- [개선조치 및 재전송 요청] 개선계획에 따라 개선조치를 수행하고, 연계데이터 등의 재전송이 필요할 경우 외부연계기관에 재전송을 요청한다.
- [재전송] 개선조치 이외에 연계데이터 재전송이 필요할 경우 재전송을 수행한다. 협약에 의한 외부연계의 경우 외부연계기관에 요청하고, 임의의 외부연계 및 내부연계의 경우 동일 작업을 다시 수행한다.
- [개선완료] 개선조치 및 재전송 결과 이상이 없는지 확인한다. 이상이 있는 경우 문제가 해결될 때까지 조치 및 재전송을 반복 수행한다.

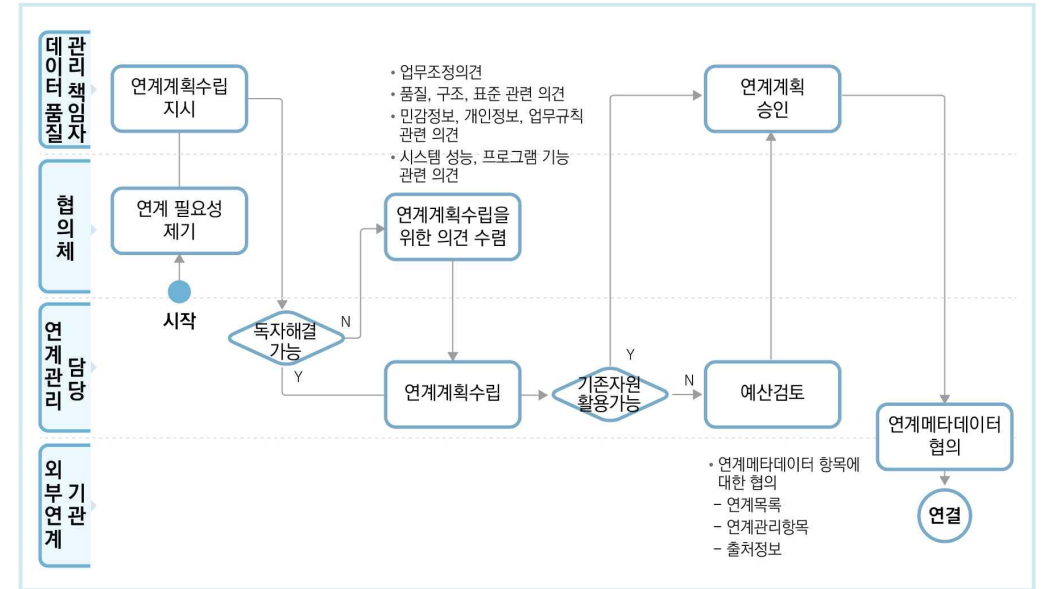
3.4.3 데이터 송수신 관리

데이터 송수신 관리 기능은 데이터 송수신을 위한 연계준비, 생애주기 단계별 송수신관리, 송수신 모니터링 및 송수신 오류 개선 프로세스로 구성된다.

▶ 연계준비 프로세스

연계를 위한 준비 프로세스로, 연계계획을 수립하고, 규약에 의한 연계의 경우 연계기관들 간의 협의를 통해 연계메타정보를 교환하며, 연계메타데이터를 정의하고 필요한 데이터를 생성한다. 필요시 시스템 및 기능을 구현한 후 테스트를 거쳐서 연계를 위한 검증을 수행하는 활동을 수행한다.

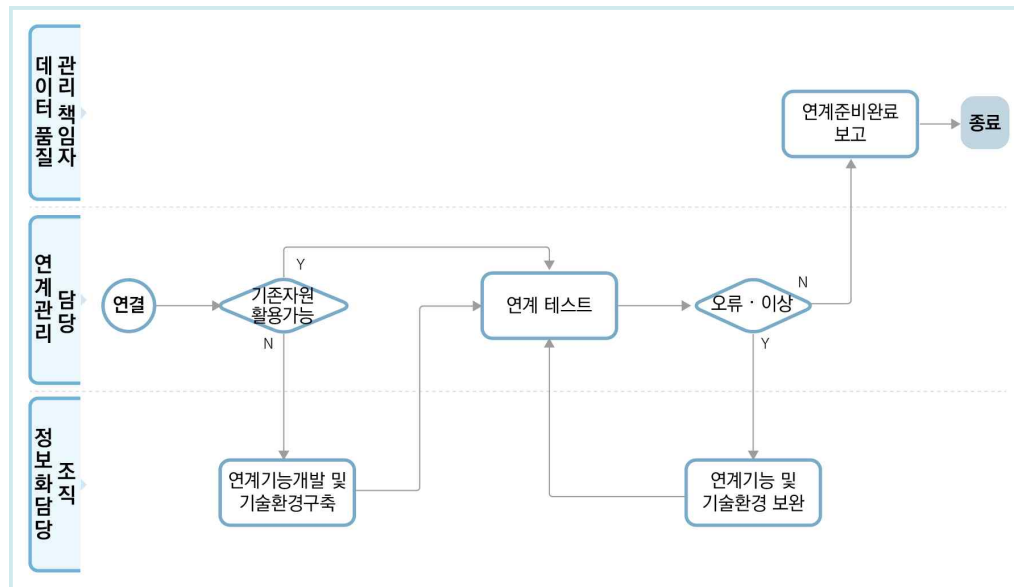
그림 48 연계준비(1/2)



- [연계 필요성 제기] 센터 또는 플랫폼에서 보유하고 있지 않은 외부 데이터를 필요로 하는 경우 데이터 연계가 필요하다는 요구를 제기한다. 연계 필요성은 외부기관의 요청에 의하거나 내부의 다양한 조직에서 요청할 수 있으며 협의체를 통하여 필요성을 제기하기도 한다. 분석 및 가공 결과를 저장하는 내부연계의 경우에는 연계 필요성 제기 없이 자동으로 데이터 송수신과 연계관리가 이루어지도록 한다.
- [연계계획수립 지시] 연계 필요성이 제기되면 센터 및 플랫폼은 의사결정 절차나 규정에 따라 데이터관리총괄책임자, 의사결정위원회 또는 데이터품질관리 책임자가 연계 필요성 여부를 승인하고 연계관리담당에게 연계계획을 수립할 것을 지시한다. 센터 및 플랫폼에 의사결정 절차나 규정이 없는 경우 데이터품질관리 책임자가 연계 필요성 여부를 승인하고 연계계획수립을 지시한다. 외부기관이 요청한 연계 필요성에 대해 승인을 얻지 못할 경우 결정권자는 연계관리담당에게 연계거절을 통보하도록 지시한다.
- [독자해결가능] 연계관리담당이 독자적으로 연계계획을 수립할 수 없을 경우 관련 담당자 및 협의체에 지원을 요청한다. 지원요청 시 연계 필요성 및 연계계획수립 지시를 참조하여 연계데이터, 연계목적, 외부연계기관, 연계적용기술 및 연계주기, 연계시작일 등과 같은 기본적인 사항을 전달한다.
- [연계계획 수립을 위한 의견수렴] 연계계획 수립을 위한 지원 요청을 받은 분야별 담당자는 담당별 의견을 제시한다. 필요시 협의체는 조직별 역할을 조정하여 의견을 제시한다.
- [연계계획수립] 연계계획은 연계관련 의사결정과 실무 추진 모두 가능하도록 연계대상 및 범위, 연계추진 수행조직, 연계타당성 검토결과, 연계 방식, 연계추진 일정 및 소요예산에 대해 상세하게 작성한다.

- [기존자원 활용가능] 연계를 위해 현재 사용하는 기술 환경에 추가, 변경 등이 필요한지를 확인한다. 새로운 프로그램 개발 및 기존 프로그램 수정, 새로운 기기 도입 및 증설 등이 필요한지를 확인한다.
- [예산검토] 기존의 기술 환경에 변화가 필요할 경우, 연계를 위한 예산 타당성을 검토한다.
- [연계계획 승인] 센터 및 플랫폼 의사결정 규정 및 절차에 따라 데이터품질관리 책임자, 의사결정위원회, 데이터품질관리 책임자, 데이터관리 총괄 책임자 및 최고책임자는 연계 관련 예산을 포함한 연계계획을 승인한다.
- [연계메타데이터 협의] 연계계획 승인을 얻으면 연계관리담당은 연계메타데이터를 결정한다. 합의에 의한 외부연계의 경우 외부연계기관 담당자와 조직도 및 메타데이터를 참조하여 연계목록, 연계관리항목, 출처정보를 합의하여 결정한다. 임의의 수집 및 내부연계의 경우 연계관리담당이 메타데이터를 참조하여 연계메타데이터를 정의한다.

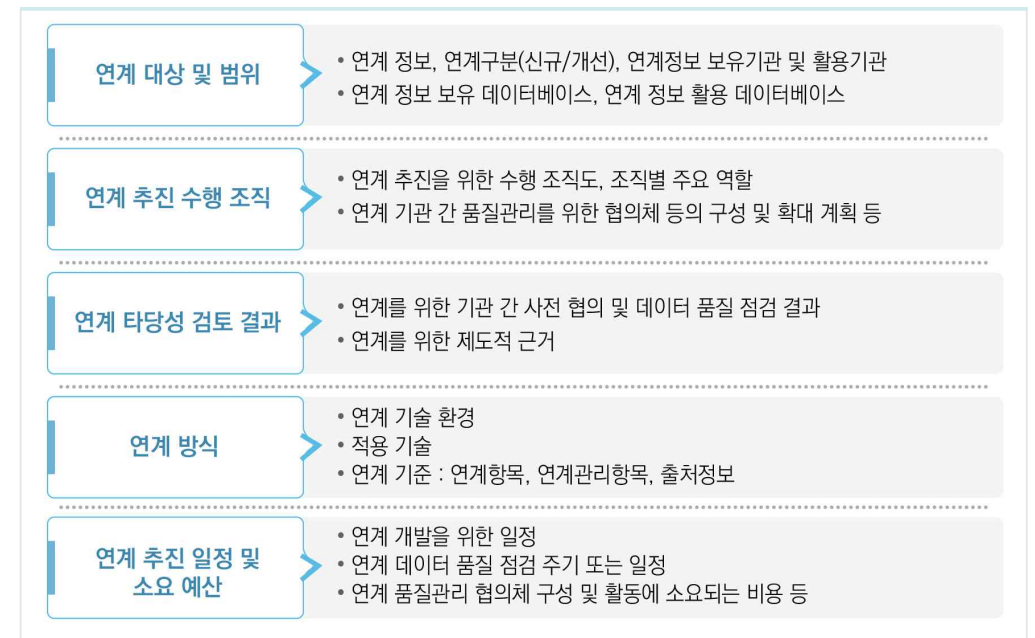
그림 49 연계준비(2/2)



- [기존자원 활용가능] 새로운 연계를 위해 현재 사용하는 기술 환경에 추가, 변경 등이 필요한지를 확인한다.
- [연계기능개발 및 기술 환경 구축] 기술 환경에 추가 및 변경이 필요할 경우 연계프로그램의 추가 및 수정, 기술 환경 구축 및 변경 등을 통하여 연계를 수행할 수 있도록 기능을 개발하고 기술 환경을 구축한다.
- [연계테스트] 연계기능 개발 및 기술 환경 구축이 완료되면 연계 테스트를 수행한다. 소셜 데이터 수집 등 임의의 외부연계, 가공분석, 활용 등 내부연계의 경우 메타데이터 및 연계메타데이터 생성 및 정합성 여부까지 확인한다.

- [오류·이상] 연계테스트 결과 오류 또는 이상여부를 확인하여 이상이 있는 경우 보완조치를 실시한다.
- [연계기능 및 기술 환경 보완] 연계테스트 결과 이상 및 오류가 있는 경우 그 원인을 파악하여 프로그램 기능 및 기술 환경을 개선한다.
- [연계준비 완료보고] 연계테스트 결과 이상이 없으면 연계관리 담당은 센터 및 플랫폼 절차 및 규정에 따라 연계준비가 되었음을 보고한다. 별다른 규정 및 절차가 없는 경우 데이터품질관리 책임자에게 연계준비가 되었음을 보고한다. 내부연계의 경우 완료보고를 생략할 수 있다. 협약에 의한 외부연계의 경우 연계 착수일을 포함하여 보고하고, 협의체 등을 통해 관련자들에게 전파한다.

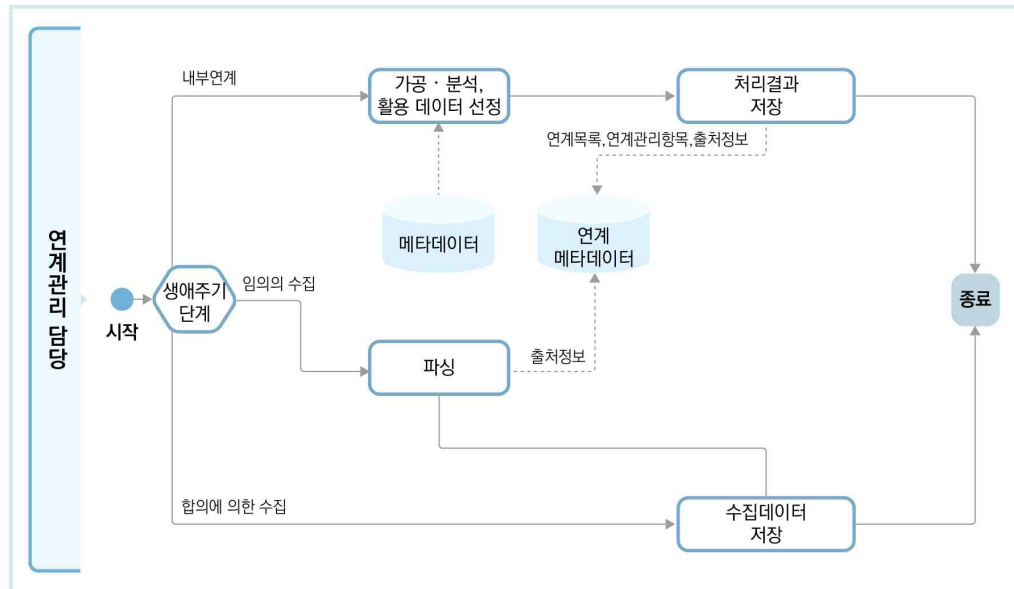
그림 50 연계계획에 포함할 내용



▶ 생애주기 단계별 송수신 관리 프로세스

데이터 생애주기 단계별 연계 데이터 송수신을 관리하는 프로세스로 외부연계를 통한 데이터 수집단계의 데이터 송수신 관리, 수집 시 수행되는 실시간 및 일괄 분석 단계의 데이터 송수신 관리와 활용 단계의 내부연계 데이터 송수신 관리로 구분된다.

그림 51 생애주기 단계별 송수신 관리

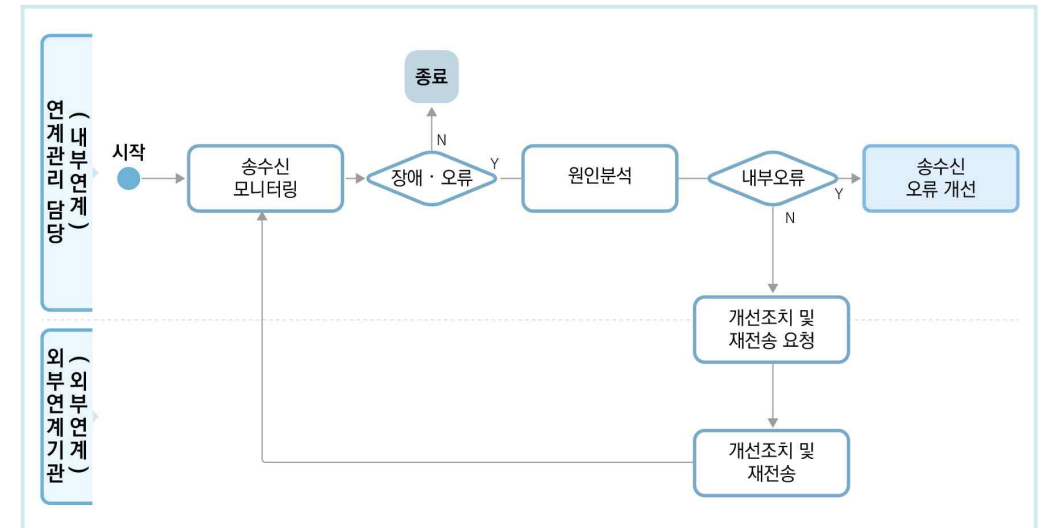


- [생애주기 단계] 빅데이터 생애주기에 따라 협약에 의한 수집, SNS 데이터 크롤링 등의 임의의 수집, 합의에 의한 수집 등 내부연계로 송수신관리 유형별로 분기한다.
- [가공·분석, 활용 데이터 선정] 메타데이터 등을 참조하여 필요 데이터를 선정하고, 데이터 워크플로우 도구를 이용하여 가공·분석, 활용에 이용된 데이터를 출처정보로 자동 생성한다.
- [가공·분석, 활용을 위한 처리결과 저장] 가공·분석, 활용 처리가 종료 후 결과를 저장하면 데이터 워크플로우 관리 도구를 통한 가공·분석, 활용데이터와 출처 및 활용처 정보를 포함한 연계메타데이터를 자동 생성한다.
- [파싱] 소셜 데이터 등 임의의 수집의 경우 파싱을 통해 출처정보를 추출한다.
- [수집데이터 저장] 협약에 의한 데이터 수집은 연계 착수 전 생성된 연계메타 데이터를 통해 수집·저장된다. 소셜 데이터 등 임의의 수집 데이터는 파싱으로 추출한 정보를 함께 저장한다.

▶ 송수신 모니터링 프로세스

송수신 상태 및 데이터의 정합성을 모니터링하여 이상여부를 진단하는 프로세스로서, 송수신 데이터 품질 이상 여부와 장애 발생 여부를 점검 및 모니터링하고, 오류 및 장애 발생 시 원인을 파악하는 프로세스이다.

그림 52 송수신 모니터링

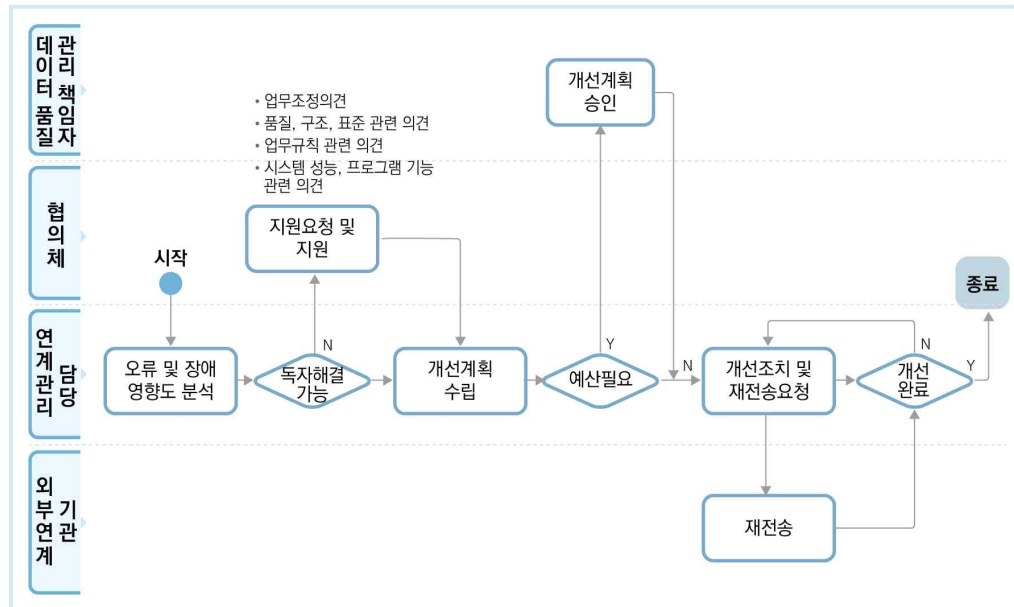


- [송수신 모니터링] 단절 없는 연계를 위해 데이터 송수신 상황을 모니터링 또는 정기 점검을 수행하며 데이터 워크플로우 관리 도구를 이용한 송수신 상황 모니터링을 자동화하여 효율적인 모니터링을 수행한다. 사전에 오류현상-to-오류원인-to-개선방안 관련 자료를 작성하여 신속한 대응을 할 수 있도록 한다.
- [장애·오류] 데이터 송수신 장애 또는 오류 발생여부를 확인하고 데이터 워크플로우 관리 도구를 이용한 송수신 상황 모니터링 자동화 도구에 장애 및 오류 발생 시 경고 기능을 추가하여 장애 및 오류 발생 여부를 신속하게 파악한다.
- [원인분석] 장애 또는 오류가 발생할 경우 그 원인을 파악한다. 연계메타데이터의 출처정보(Target-to-Source)를 이용하여 효율적인 원인분석이 가능하도록 한다.
- [내부오류] 장애 또는 오류 원인 분석 결과 센터 및 플랫폼 내부의 문제인지 외부연계기관의 문제인지를 확인한다.
- [개선조치 및 재전송 요청] 외부요인에 의한 장애 및 오류인 경우 외부연계기관에게 장애 및 오류 관련 정보를 전달하고 개선조치 및 조치 후 재전송을 요청한다.
- [개선조치 및 재전송] 외부기관은 장애 및 오류 문제 해결을 위한 조치를 수행하고 필요시 연계 데이터를 재전송한다.
- [송수신 오류 개선] 내부요인에 의한 오류 및 장애의 경우 송수신 오류 개선 프로세스를 수행한다.

▶ 송수신 오류 개선 프로세스

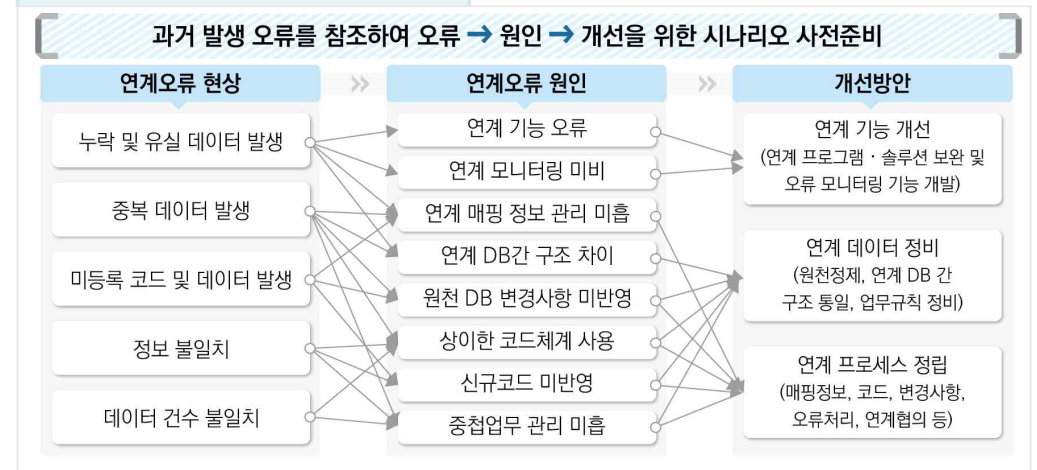
송수신 장애 및 오류 발생시 데이터 정합성 관련 개선조치를 실시하고 개선 결과를 확인하는 프로세스로서 장애 및 오류 원인분석 결과에 따라 영향도 분석과 개선계획 수립 및 개선을 수행하며, 개선조치가 계획에 따라 수행되었는지 확인하는 활동을 실시한다.

그림 53 송수신 오류 개선



- [오류 및 장애 영향도 분석] 내부 원인일 경우 오류 및 장애가 미치는 영향을 분석하여 문제해결, 개선 등의 조치를 위한 업무량을 파악한다. 연계메타데이터의 활용처 정보(Source-to-Target)를 이용하여 오류로 인해 영향을 받는 데이터를 효율적으로 파악 가능하도록 한다.
- [독자해결가능] 연계관리담당이 독자적으로 오류 및 장애를 해결할 수 없는 경우 관련 담당자 및 협의체에 지원을 요청한다.
- [지원요청 및 지원] 연계 오류 및 장애 개선계획 수립을 위한 지원 요청을 받은 분야별 담당자는, 담당별 의견을 제시한다. 필요시 협의체는 조직별 역할을 조정하여 의견을 제시한다.
- [개선계획 수립] 연계관리담당자는 장애 및 오류 문제 해결 계획을 수립한다.
- [예산필요] 개선계획에 시스템성능개선, 프로그램개발, 정보자원추가도입 등 예산 필요 여부를 확인하고 필요한 경우 데이터품질관리 책임자에게 개선계획 승인을 요청한다.
- [개선계획승인] 개선계획에 예산이 필요할 경우 규정에 따라 최고책임자, 데이터 관리총괄책임자, 의사결정위원회, 데이터품질관리 책임자 승인을 받는다.
- [개선조치 및 재전송 요청] 개선계획에 따라 개선조치를 수행하고, 연계데이터 재전송이 필요할 경우 외부연계기관에 재전송을 요청한다.
- [재전송] 협약에 의한 외부연계의 경우 외부연계기관에 요청하고, 임의의 외부연계 및 내부연계의 경우 동일 작업을 다시 수행한다.
- [개선완료] 개선조치 및 재전송 결과 이상이 없는지 확인한다. 이상이 있는 경우 문제가 해결될 때까지 반복해서 개선조치 및 재전송을 수행한다.

오류 개선을 위한 사전 시나리오 준비



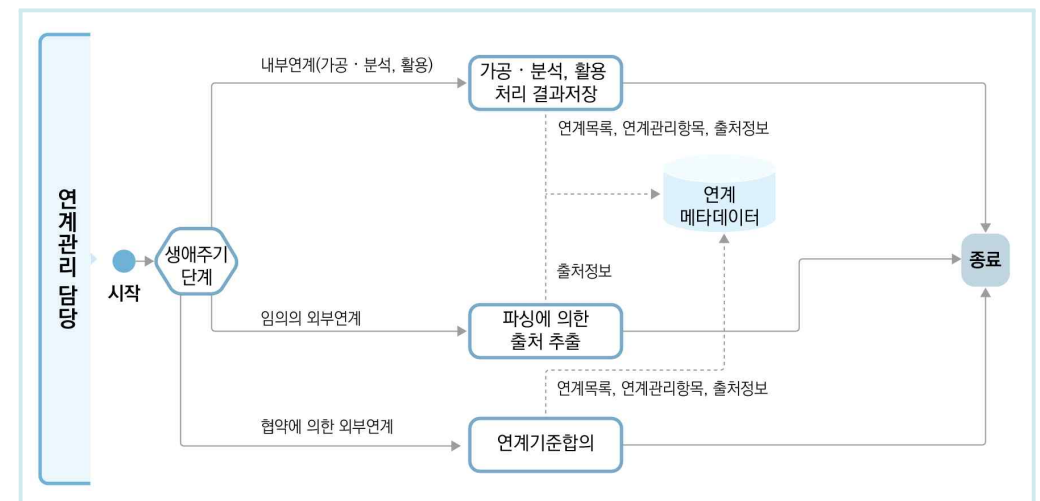
3.4.4 흐름 및 추적관리

흐름 및 추적관리 기능은 활용처 정보와 출처정보를 이용하여 데이터 생애주기 전 단계에 걸쳐 빠짐없이 데이터의 흐름과 출처를 관리 할 수 있도록 출처정보 정의, 흐름 및 추적정보 제공, 출처정보 정합성 진단, 출처정보 정합성 오류를 개선하는 프로세스로 구성된다.

출처정보 정의 프로세스

연계유형에 따라 활용처 및 출처정보를 정의하는 프로세스로서, 데이터 생애주기에 따라 규약에 의한 외부연계는 연계기관 간의 합의로 출처정보를 정의하고, 임의의 외부연계 및 내부연계는 메타데이터를 참조하여 센터 및 플랫폼에서 자체적으로 결정한다.

그림 54 출처정보 정의

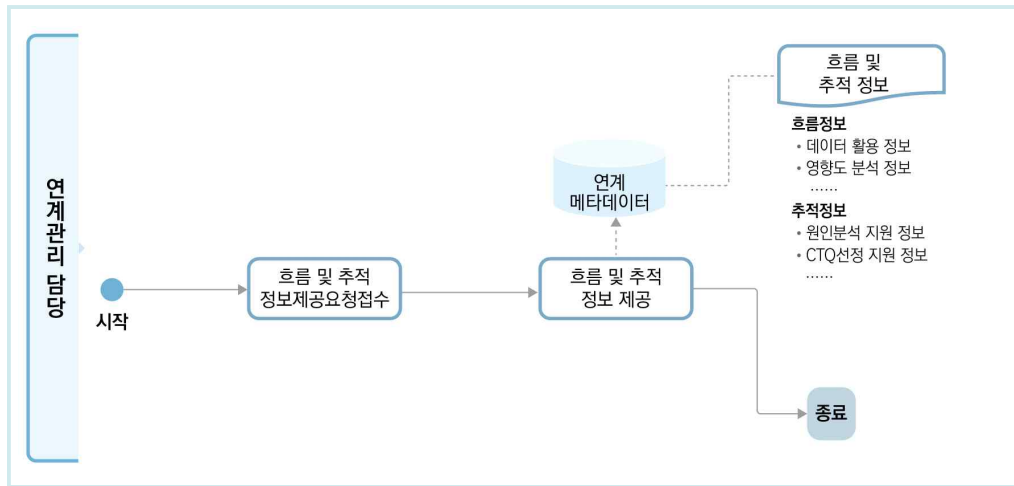


- [연계유형] 빅데이터 생애주기에 따라 협약에 의한 수집, 소셜 데이터 등의 임의의 수집, 가공·분석, 활용 등의 내부연계로 송수신 관리 유형을 분류하여 분기한다.
- [가공·분석, 활용 처리 결과 저장] 내부연계인 가공·분석, 활용 처리 결과를 저장할 때에, 저장 데이터에 대한 메타데이터를 생성하고 활용처 정보를 생성한다. 동시에 처리를 위해 사용한 입력 데이터를 출처정보로 생성한다. 데이터 워크플로우 관리 도구를 이용하여 출처정보 및 활용처 정보 생성을 자동화한다.
- [파싱에 의한 출처 추출] 크롤러 등의 소셜 데이터 수집기는 메타데이터를 참조하여 외부데이터를 수집한다. 수집된 데이터는 파싱을 통해 수집데이터명과 수집주소를 추출하여 출처정보로 관리한다. 수집한 데이터를 저장 할 때에 추출한 출처정보를 생성한다. 데이터 워크플로우 관리 도구를 이용하여 출처정보 생성을 자동화한다.
- [연계기준합의] 협약에 의한 외부연계의 경우 메타데이터를 참조하여 외부연계 기관과 협의를 통하여 활용처 정보 및 출처정보를 정의하고 연계메타데이터로 등록한다.

▶ 흐름 및 추적정보 제공 프로세스

출처 및 활용처 정보를 이용하여 데이터 관리를 위한 다양한 정보를 제공하는 프로세스로서 제공하는 정보는 데이터 활용, 영향도 분석, 원인분석 지원, CTQ선정지원 정보 등이 있다.

그림 55 흐름 및 추적정보 제공



- [흐름 및 추적 정보제공 요청접수] 데이터 흐름 및 추적정보가 필요한 조직으로부터 정보제공 요청을 접수한다.
- [흐름 및 추적 정보 제공] 연계관리담당자는 흐름 및 추적정보 요청 목적 및 사유를 판단하여 제공 할 정보의 형태 및 정보 추출 방식을 결정하여 정보를 제공한다. 제공 가능한 정보는 활용처 정보(Source-to-Target) 및 출처정보(Target-to-Source)로서 다양한 형태의 데이터 관리에 활용할 수 있다. 흐름 및 추적정보의 활용 예는 다음의 표와 같다.

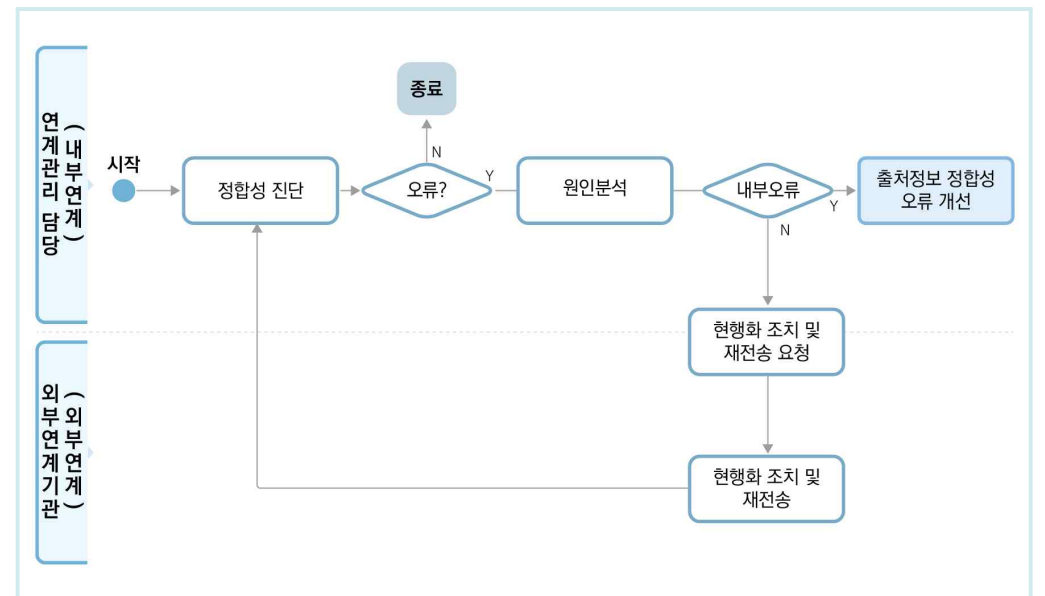
표 17 흐름 및 추적정보 활용 예

구분		내 용
흐름정보 (Source-to-Target)	데이터 활용 정보	<ul style="list-style-type: none"> 데이터의 활용처 정보를 제공 데이터별로 이용도가 높은지 낮은지를 판단할 수 있다.
	영향도 분석 정보	<ul style="list-style-type: none"> 데이터의 활용처 정보를 제공 특정 데이터에 품질 이상 등의 문제가 있는 경우 영향을 받는 데이터를 파악할 수 있다.
추적정보 (Target-to-Source)	원인분석 지원 정보	<ul style="list-style-type: none"> 데이터 출처정보를 제공 특정 데이터에 문제가 있는 경우 문제의 원인이 되는 데이터를 추적하여 찾을 수 있도록 지원한다.
	CTQ선정 지원 정보	<ul style="list-style-type: none"> 활용 관점에서 CTQ로 선정된 데이터의 출처정보 파악이 가능 CTQ로 선정된 데이터의 출처 데이터를 찾아서 추가적인 CTQ 선정을 위한 후보 데이터 정보를 제공한다.

▶ 출처정보 정합성 진단 프로세스

연계메타데이터의 출처 및 활용처 정보의 정합성을 진단하고 문제 발생시 원인을 분석하는 프로세스로서, 출처정보의 누락, 출처(제공기관) 및 활용처(활용기관) 변경사항 미반영 등의 출처정보의 부정합 여부를 진단하고 오류 등 부정합 발생 원인을 분석하는 활동을 수행한다.

그림 56 출처정보 정합성 진단

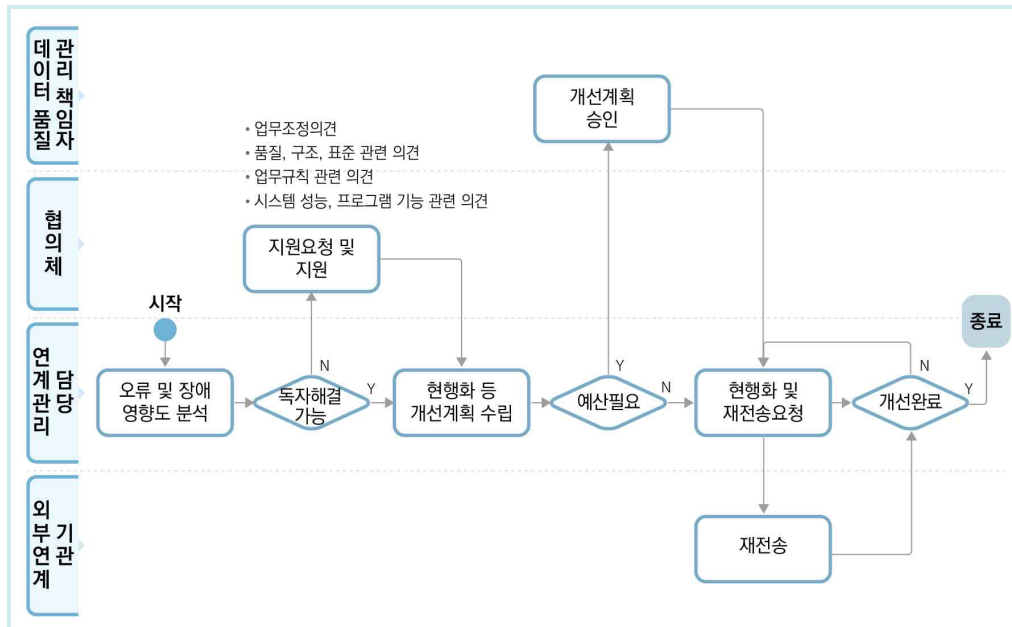


- [정합성진단] 데이터 흐름 및 추적을 위해 활용처 및 출처정보가 정합성을 유지하고 있는지 여부를 진단한다. 누락, 조직변경 및 역할 변경으로 인한 변동 사항이 반영되어 있는지 확인한다.
- [오류] 활용처 및 출처 정보의 정합성 오류 발생여부를 확인한다.
- [원인분석] 정합성 오류가 발생할 경우 그 원인을 파악한다. 연계메타데이터의 출처정보를 이용하여 효율적인 원인분석이 가능하도록 한다.
- [내부오류] 장애 또는 오류 원인 분석 결과 센터 및 플랫폼 내부의 문제인지 외부연계기관의 문제인지를 확인한다.
- [현행화 조치 및 재전송 요청] 외부요인에 의한 오류인 경우 외부연계기관에게 정보를 전달하고 현행화 조치 및 필요시 재전송을 요청한다.
- [현행화 조치 및 재전송] 오류 문제 해결을 위한 현행화 등의 조치를 수행하고 필요시 연계 데이터를 재전송한다.
- [출처정보 정합성 오류 개선] 내부요인에 의한 정합성 오류의 경우 출처정보 정합성 오류 개선 프로세스를 수행한다.

▶ 출처정보 정합성 오류 개선 프로세스

출처정보 오류 및 정합성 문제를 개선하고 조치 결과를 확인하는 프로세스로서, 오류 및 부정합 원인에 따른 영향도 분석 및 개선계획 수립하고, 개선계획에 따라 개선조치를 수행하고, 개선 조치에 따른 개선 여부를 확인하는 활동으로 구성된다.

그림 57 출처정보 정합성 오류 개선



- [오류 및 장애 영향도 분석] 활용처 및 출처정보의 정합성 오류 원인이 내부적인 문제일 경우 오류가 미치는 영향을 분석하여 문제해결, 개선 등의 조치를 위한 업무량을 파악한다. 연계메타데이터의 활용처 정보(Source-to-Target)를 이용하여 오류로 인해 영향을 받는 데이터를 효율적으로 파악 가능하도록 한다.
- [독자해결가능] 연계관리담당이 정합성 오류 개선 계획을 독자적으로 수립할 수 없는 경우 협의체 및 데이터품질관리 책임자를 통하여 관련자에게 개선방안 수립 지원을 요청한다. 지원 요청 시 출처 및 활용처 정보 정합성 오류현상, 오류원인, 영향도 분석 결과, 관련 메타데이터 및 연계메타데이터를 제공한다.
- [지원요청 및 지원] 연계 오류 및 장애 개선계획 수립을 위한 지원 요청을 받은 분야별 담당자는, 담당별 의견을 제시한다. 필요시 협의체는 조직별 역할을 조정하여 의견을 제시한다.
- [현행화 등 개선계획 수립] 연계관리담당은 독자적 해결이 가능할 경우 독자적으로, 관련자의 조력을 받아야 할 경우 관련자의 의견을 수렴하여 정합성 오류 문제를 해결하기 위한 계획을 수립한다.
- [예산필요] 현행화 등의 개선계획에 시스템 성능 개선, 프로그램 개발, 정보자원 추가 도입 등의 사유로 예산이 필요한지 여부를 확인하여 예산이 필요할 경우에는 데이터품질관리 책임자에게 개선계획 승인을 요청한다.
- [개선계획승인] 개선계획 실행에 예산이 필요할 경우 센터 및 플랫폼 의사결정 절차 및 규정에 따라 최고책임자, 데이터관리 총괄 책임자, 의사결정위원회, 데이터품질관리 책임자의 승인을 얻는다.
- [현행화 및 재전송 요청] 개선계획에 따라 현행화 조치 등의 개선조치와 연계데이터 등의 재전송이 필요할 경우 외부연계기관에 재전송을 요청한다.
- [재전송] 현행화 등의 개선조치 이외에 연계데이터 재전송이 필요할 경우 재전송을 수행한다. 협약에 의한 외부연계의 경우 외부연계기관에 요청하고, 임의의 외부연계 및 내부연계의 경우 동일 작업을 다시 수행한다.
- [개선완료] 현행화 등의 개선조치 및 재전송 결과 이상이 없는지 확인한다. 이상이 있는 경우 문제가 해결될 때까지 조치 및 재전송을 반복 수행한다.

01

개요



1.1 목적 및 기능

1.1.1 배경

최근 빅데이터 플랫폼 구축 및 활용, 데이터 경제 활성화 등 데이터 활용을 강조하고 있으나 개인정보보호도 중요한 가치로, 데이터의 안전성과 유용성 모두 보장이 필요하다.

이에 데이터 플랫폼의 구축과 데이터 개방, 공공·민간 데이터 저장과 유통, 개인 맞춤형 서비스 등 다양한 분야에 활용 가능한 데이터 정보보호 관리체계를 수립한다.

▶ 정보보호관리 특징

빅데이터 플랫폼 및 센터의 데이터 수집과 활용은 정보보호법 및 개인정보보호법에 따라 빅데이터 생애주기(수집, 가공·분석, 활용, 파기) 소단계에서 정보보호 관리를 수행해야 한다.

그림 58 빅데이터 생애주기



표 18 빅데이터 생애주기 단계

항 목	내 용
수 집	• 빅데이터 처리를 목적으로 다양한 경로를 통해 필요한 데이터를 모으는 과정
가공·분석	• 수집한 데이터를 장치에 저장하고, 저장된 데이터를 다양한 방법을 통해 가공하여 새로운 정보를 생성하는 과정
활 용	• 수집 데이터와 분석 정보를 빅데이터 처리자가 직접 사용 또는 제3자가 사용하게 하는 과정
파 기	• 수집한 데이터 또는 이를 분석한 정보를 삭제하는 과정

(*정보보호관리는 빅데이터 전 사이클에서 파기의 단계가 포함되어 관리됨)

▶ 정보보호 동향

빅데이터 환경에서 데이터 활용에 어려움에 처해 있는 문제를 해결하기 위하여 데이터3법이 시행되고 있다. 데이터 3법은 빅데이터 산업 육성을 위해 데이터 이용에 따른 규제를 푸는 법으로써 개인정보보호법, 정보통신망 이용촉진 및 정보보호 등에 관한 법률(정보통신망법), 신용정보의 이용 및 보호에 관한 법률(신용정보보호법)으로 이뤄져 있다.

개인정보보호법은 가명 정보 데이터를 제품·서비스 개발에 활용하고 개인정보 관리감독 기능을 하는 개인정보보호위원회를 일원화하는데 초점을 맞춘다. 정보통신망법은 온라인상 개인정보보호 규제·감독 권한을 개인정보보호위원회로 변경하는 법이다. 신용정보보호법은 금융 분야 가명 정보를 빅데이터 분석·이용에 활용할 수 있도록 하고 가명 정보 주체의 동의 없이 정보의 이용·제공을 허용하는 내용이다.

데이터 3법은 새로운 개념 도입과 개인정보 관련 감독 기관의 일원화, 가명처리 정보 제3자 제공 및 활용 등의 내용을 포함하고 있다. 데이터3법은 추가 정보의 결합 없이는 개인을 식별할 수 없도록 안정하게 처리된 가명정보의 개념을 도입하는 것이 핵심이다. 가명정보를 이용하면 개인정보를 활용해 마이데이터 사업과 같은 새로운 서비스나 기술, 제품 등을 개발할 수 있어 기업들이 신사업을 전개할 수 있다.

▶ 정보보호관리 정의

빅데이터 플랫폼 및 센터 데이터 품질관리에서 정보보호관리란 특정한 목적을 위해서 가공한 정보를 안전한 상태를 유지시키기 위하여 정보를 보호하기 위해 기술적·관리적으로 보호하는 행위로 정의한다.

표 19 정보보호관리 정의

구분	정보보호관	정보보호관리란
협의의 정의	<ul style="list-style-type: none"> 빅데이터 플랫폼 및 센터에서 데이터 생애주기 단계별(수집, 가공분석, 활용) 데이터를 안전하게 보호하는 것 	<ul style="list-style-type: none"> 빅데이터 플랫폼 및 센터에서 데이터 생애주기 단계별(수집, 가공분석, 활용) 데이터를 안전하게 보호할 수 있도록 하는 것 정보보호 조치들에 대해 진단하고 문제가 있는 경우 개선하도록 하는 것
광의의 정의	<ul style="list-style-type: none"> 정보보호와 이용활성화의 균형을 위한 법개정 동향에 맞추어 데이터 생애주기 단계별(수집, 가공분석, 활용) 데이터를 안전하게 보호한 상태에서 데이터를 이용활성화 하는 것 	<ul style="list-style-type: none"> 정보보호와 이용활성화의 균형을 위한 법개정 동향에 맞추어 데이터 생애주기 단계별(수집, 가공분석, 활용) 데이터를 안전하게 보호한 상태에서 데이터 이용활성화를 할 수 있도록 하는 것 가명 및 익명정보 등 비식별 조치 등에 대해 진단하고 문제가 있는 경우 개선하도록 하는 것

1.1.2 정보보호관리 목적

빅데이터 플랫폼 및 센터에서 데이터를 운용·관리·활용함에 있어 정보보호를 통해 데이터를 안전하게 활용될 수 있도록 하는데 있다.

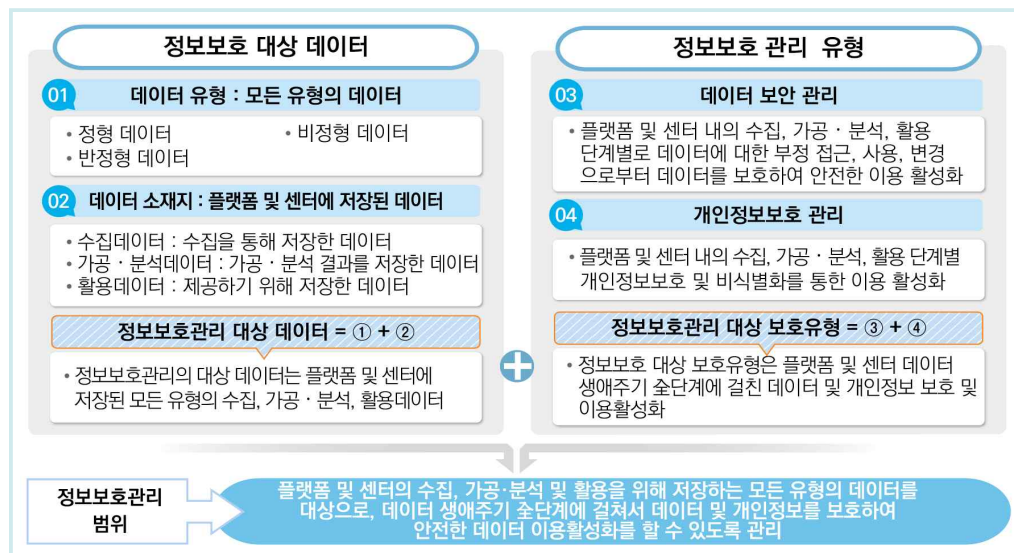
그림 59 정보보호관리 목적



▶ 정보보호관리 범위

정보보호관리는 빅데이터 플랫폼 및 센터에서 이루어지는 수집, 가공·분석, 활용되는 데이터를 보호하고, 정보 훼손 등을 방지 및 조치 관리하기 위한 관리적, 기술적, 개인정보보호 관리를 범위로 정의한다.

그림 60 정보보호관리 범위

02
모델 수립

2.1 정보보호관리 모델 개요

정보보호관리 모델은 정보보호 운영관리 모델, 정보보호 기술관리 모델 및 개인정보보호관리 모델을 빅데이터 생애주기인 수집, 가공, 활용 단계별로 관리한다.

▶ 정보보호 운영관리 모델

- 정보보호 관련 법규 등을 참고하여 정보보호 운영계획을 수립하고 실행
- 정보보호가 정상적으로 이루어지는지를 확인 및 점검
- 점검 결과 문제가 발견될 경우 개선계획을 수립하고 개선조치를 수행

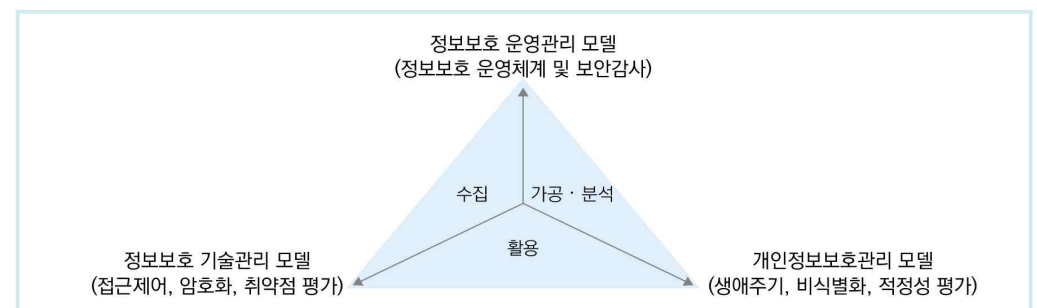
▶ 정보보호 기술관리 모델

- 데이터 보호를 위한 접근제어, 암호화 등의 계획을 수립
- 접근제어/암호화 실시 및 취약점 분석, 취약점 발견 시 개선조치 수행

▶ 개인정보 보호관리 모델

- 개인정보보호 관련 법규 및 가이드를 참조하여 개인정보보호 계획을 수립
- 계획에 따라 비식별조치 및 파기 등 실시
- 비식별조치 등의 적정성을 평가 및 점검, 점검결과 문제 시 개선조치 수행

그림 61 정보보호 관리 모델



2.2 정보보호 운영관리 모델

2.2.1 개요

정보보호 운영관리는 정보보호 관련 법규를 바탕으로 플랫폼 및 센터의 정보보호를 위한 운영계획을 수립하고, 정보보호에 문제가 있는지를 감사하며, 감사결과에 문제가 있는 경우 개선조치를 수행한다.

▶ 정보보호 운영계획 수립

- 정보보호 관련 법규를 바탕으로 기본계획과 분야별 계획을 수립
- 정보보호에 관련된 통제절차 및 행동을 포함
- 분야별 계획에는 정보보안감사 계획, 정보보호 기술계획을 포함

▶ 정보보호 운영

- 정보보호 운영계획을 실행
- 정보보호는 데이터 생애주기 쉐 단계(수집, 가공·분석, 활용)에 걸쳐서 수행

▶ 보안감사

- 주기적 또는 수시로 정보보안 감사를 실시하여 문제점 및 원인을 파악
- 독립적인 별도의 임시조직을 구성하여 감사를 실시
- 정보보호 관련 통제절차에 대한 기록 등을 조사 관찰하고 증거를 수집하여 분석

▶ 개선 조치

- 감사 등을 통해 발견된 정보보호 운영관련 문제를 해결하기 위한 개선조치를 실시
- 문제 원인별 영향도를 분석하여 개선계획을 수립하여 개선조치를 수행

2.2.2 보안감사

정보보호 운영관리가 적절하고 효율적으로 관리되고 있는지는 보안감사를 통해 확인 할 수 있으며, 객관적인 감사가 되기 위해서는 독립적인 임시 조직을 구성을 해야 한다.

▶ 감사주기

- 보안감사는 년 1회 실시를 원칙으로 하고, 필요 시 수시 감사를 실시

▶ 감사 조직

- 정보보호 및 정보관련 비전문가를 포함한 임시조직을 구성
- 통제절차에 대한 감사는 정보 관련 비전문가 활용
- 로그 정보 분석은 전문가 활용

▶ 감사 범위

- 플랫폼 및 센터의 정보보호 관리체계에서 규정한 사항 확인을 위한 모든 활동
- 관리적 보안, 기술적 보안(필요시 개인정보보호 관련 감사를 통합하여 실시)

▶ 감사 대상

- 정보보호를 위한 모든 통제절차
- 통제절차 수행을 위한 근거문서 및 산출문서

▶ 감사 증거

- 사람이 생산한 문서
- 시스템이 생산한 로그 정보

▶ 정보보안 감사 관리기준

- 데이터 접근제어, 데이터 암호화, 데이터 취약점 점검의 기술적 보안 관리와 개인정보보호 관리에서 발생하는 모든 수행 및 변경·관리 사항에 대하여 점검 확인

정보보안 감사

- 감사실시
 - 기본 감사 : 정보보안담당관이 정기적으로 진행하는 정기 감사로 일반적인 상황에서 로그 정보를 활용하여 감사 진행
 - 특수 감사 : 정보보안담당관이 비정기적으로 진행하는 불시 보안감사로서 침해가 의심되는 경우 이를 추적하는 보안감사 및 개인정보보호에 의거 보호를 받아야 하는 개인식별정보 및 민감정보에 대한 감사 진행
 - 상시 감사 : 보안관리담당자가 보안강화를 위해 상시 감사를 진행 가능
- 보안관리 대응
 - 보안관리담당자는 상시감사에 의해 드러난 보안 위배 사항은 즉각 조치
 - 침해사고는 침해사고목록 및 침해사고 대응 결과 보고서를 작성 필수
- 감사 시 제출 자료
 - 정보보안시스템에서 발생된 로그 자료는 기본 제출 자료로써 제공 필요
 - 정책 및 규칙 정의서 : DB 접근제어 규칙 정의서, 암호화 적용 대상 및 사용 알고리즘 정의서 등의 기술적 규칙을 정의한 기준서 제출
 - 로그관리 항목 : DB 접근제어, DB 암호화, DB 취약점 평가 등 기술적 보안관리에서 발생하는 모든 수행 및 변경사항 기록
 - 취약점 점검 결과서 : 취약점 점검·평가 결과서 제출

2.3 정보보호 기술관리 모델

2.3.1 개요

정보보호 기술을 위해 관리해야 하는 요소는 정보보호 데이터의 접근 권한을 관리하는 접근제어, 인가된 정보주체 이외에는 정보를 볼 수 없는 암호화, 취약점을 파악하여 원인을 분석하고 취약점을 개선하는 취약점 점검으로 구성된다.

▶ 접근제어

- 정보보호를 데이터에 대한 접근 권한은 직제 규정 등에 따라 허용된 사람에 한하여 업무수행에 필요한 최소한의 범위로 부여
- 인사이동에 따라 직제 변경 시 현행화 필수 수행
- 접근제어 관리 기준, 접근제어 관리 요소, 접근제어를 위한 보안설정 등 관리

▶ 암호화

- 암호화는 민감정보의 정보보호를 강화하기 위해 인가된 정보주체 이외에는 정보를 볼 수 없도록 데이터를 변경하는 행위
- 정보보호 기술관리에서는 암호화 관리기준 및 적용기준을 관리

▶ 취약점 점검

- 데이터 보호 기술에 대한 취약점을 파악하여 원인을 분석하고 취약점을 개선을 통해 정보보호를 강화하는 과정
- 취약점 점검 관리기준을 관리
- 취약점 관리 활동으로는 단계별 취약점 점검, 취약점 완화 처리, 취약점 관리 고려사항 점검 등을 수행

2.3.2 접근제어

접근제어는 사용자가 DBMS에 로그인하거나 SQL을 수행할 때, 기 정의된 보안규칙에 따라 권한 여부를 판단하여 통제하고, 로그 저장이 필요한 SQL에 대하여 SQL 수행(CRUD)과 관련된 정보를 저장(로그)하여 부당한 접근을 제약하는 보안 통제 관리방법이다.

정보보호를 데이터에 대한 접근 권한은 직제 규정 등에 따라 허용된 사람에 한하여 업무수행에 필요한 최소한의 범위로 부여하고, 인사이동에 따라 직제 변경 시 현행화되어야 한다.

▶ 접근제어 관리기준

- 데이터에 대한 접근 권한은 직제규정 등에 따라 허용된 자에 한하여 업무수행에 필요한

최소한의 범위로 부여

- 데이터관리자와 정보보호 담당자의 역할 분리
- 데이터 사용자 계정과 소유자 계정의 분리
- 접근권한은 양도·대여 불가(정보보호 담당자의 수시 확인, 지도 및 감독 의무)
- 접근권한 목록을 미리 준비하고 이에 접근권한 부여

▶ 접근제어 관리요소

- 접근제어시스템 : 데이터 접근제어를 위한 시스템으로 정보보안시스템이라 함
- 사용자 계정 : 업무적으로 데이터 사용이 인가된 자의 계정(관리자 계정, 데이터 테이블 소유자 계정, 어플리케이션 계정 등이 있음)
- 제어영역 : 사용자 인증, 접근권한, SQL통제, 객체 통제 등
- 다단계 보안 : 물리적 보안부터 데이터보안까지 단계적 보안 시스템의 흐름
- 마스킹 : 중요한 데이터 항목 보호를 위해 '*' 등으로 변경처리

▶ 접근제어를 위한 보안설정 단계

- DB설치파일 접근제한 단계 : DB 파일, 리두 로그(redo log) 파일, 컨트롤 파일 등 DB 설치 파일에 대한 접근제어
- 애플리케이션 사용자 접근 제한 단계 : 애플리케이션 사용자는 DB서버에 바로 접속할 수 없으며 반드시 DB 접근 제어 시스템을 경유하도록 접근제어
- DB 사용자 계정 접근 제한 단계 : DB 테이블 소유자 계정과 어플리케이션 사용자 계정을 분리하여 애플리케이션 사용자 계정은 DB 테이블 소유자 계정으로부터 업무에 필요한 최소 접근 권한 보유
- DB 프로그램 접근제한 단계 : DB 프로그램은 DB 테이블 소유자 계정만 사용하도록 접근제어
- DB 테이블 접근제한 단계 : 애플리케이션은 DB 테이블을 직접 참조할 수 없으며 동의어나 뷰를 통해 참조하도록 접근제어
- DB SQL 접근제한 단계 : DB 접근제어시스템은 등록된 SQL만 DB에 접근 가능

2.3.3 암호화

암호화는 민감정보 보호를 강화하기 위해 인가된 정보주체 이외에는 정보를 볼 수 없도록 데이터를 변경하는 활동으로 접근 권한의 오·남용을 통한 개인정보의 유·노출을 방지하고 정보를 보호하는데 필요한 사항을 규정하여 정보주체의 권리를 보호한다.

▶ 암호화 관리기준

- 데이터 암호화 적용 및 암호키 관리시스템 구축 필요

- 암호키는 암호화시스템에 저장
- 암호화 기술은 정부가 권장하는 암호 알고리즘을 사용
- 필수 암호화 항목 : 주민등록번호, 운전면허번호, 여권번호, 외국인등록번호, 개인계좌번호, 개인카드번호, 바이오 정보(지문, 홍채, 음성, 필적) 및 비밀번호
- 정보주체가 아니면 민감정보 항목을 볼 수 없도록 암호화 처리

▶ 암호화 적용 기준

표 20 암호화 적용기준

구 분			필수여부
정보통신망 보조저장매체를 통한 송신 시	비밀번호, 바이오 정보, 고유 식별정보		암호화 송신
개인정보처리시스템 에 저장 시	비밀번호		일방향(해쉬 함수)암호화 저장
	바이오정보		암호화 저장
	고유 식별정보	주민등록번호	암호화 저장 (2017.12.31.까지 암호화 저장 : 100만명 이상 정보주체)
		여권번호 외국인번호 운전면허번호	인터넷 구간 DMZ 구간 암호화 저장 내부망에 저장 암호화 저장 또는 암호화 적용여부 및 적용범위를 정하여 시행
업무용 컴퓨터 모바일 기기에 저장 시	비밀번호, 바이오 정보, 고유 식별정보		암호화 저장

2.3.4 취약점 점검

취약점 점검은 데이터 보호 기술에 대한 취약점을 파악하여 원인을 분석하고 취약점 개선을 통해 정보보호를 강화하는 과정으로 정보에 대한 위협 대응 및 개인정보 침해 예방을 위한 활동으로 정보보호 관리활동에 대하여 점검 관리한다.

▶ 취약점 점검 관리기준

- 공개하여 알려진 취약점으로부터 보호하기 위해 데이터베이스는 최신 버전으로 유지하고 보안관련 패치의 최신성을 유지
- 이미 드러나 알려진 취약점의 보완·조치로는 취약점 분석 솔루션을 활용
- 취약점을 최소화하기 위해 데이터 서버 설치 시 최소화 관리기준에 의거 설치(운영에 필요한 최소한의 환경변수, 최소한의 계정, 최소한의 권한부여)

▶ 취약점 관리 활동

표 21 취약점 관리 활동

항 목	내 용
단계별 취약점 점검	<ul style="list-style-type: none"> • 기본 계정의 패스워드 정책이 적용되어 운영 및 관리 되는지 진단 • DB서버, DBA 이외의 인가되지 않은 사용자의 접근 가능여부 진단 • 암호화 기준과 적용 대상에 암호화가 되었는지 진단
취약점 완화 처리	<ul style="list-style-type: none"> • DB 환경 설정 취약점인 경우 적절한 다른 값으로 변경 • 업무에 사용되지 않는 계정은 설치하지 않음(제거가 어려운 경우 무효화) • 비밀번호 취약성이 드러난 계정은 변경 • 사용자 계정에 부적절한 권한이 발견되면 해당 권한의 제거 • DB가 자체적으로 제공한 내부 프로그램에 보안적 문제가 발생하면 보안에 문제가 발생된 기능의 사용을 중지(패치 버전이 있을 경우 적용)
취약점 관리 고려사항	<ul style="list-style-type: none"> • 취약점 조사, 완화, 평가 관련 업무량이 많을 경우 솔루션을 도입 • DB접근제어 및 암호화 서버의 일반적인 취약점은 서버 취약점으로 분류 • 접근제어 및 암호화 로그로부터 발견되는 취약점은 DB 취약점으로 분류

2.4 개인정보 보호관리 모델

2.4.1 개요

개인정보 보호 관리는 개인정보의 침해, 훼손 등으로부터 보호하는데 필요한 사항에 대한 규정과 개인정보에 대하여 수집, 가공, 활용, 폐기까지의 과정에서 침해·훼손 등으로 부터 보호하기 위한 관리를 말한다.

정보보호의 3대 요소는 기밀성(Confidentiality), 무결성(Integrity), 가용성(Availability)으로 구분되며 주요 내용은 다음과 같다.

▶ 기밀성

- 인가되거나 허가받은 사람만 시스템이나 정보에 액세스할 수 있어야 하는 것
- 송신자와 수신자만 내용을 봐야 하는데 중간에 제3자가 가로채지 못하도록 예방하고 조치하는 것이 중요하며 본인과 본인이 허락한 유저만 접근하도록 통제
- 기밀성을 가장 많이 위협하는 것은 훔쳐보거나 해킹으로 이를 예방하기 위해서는 주기적인 비밀번호 변경은 필수

▶ 무결성

- 데이터나 네트워크 보안에서 보호되어야 할 정보가 허가받은 사람만이 서버나 정보에 접근하지

나 정보를 수정할 수 있어야 하는 것

- 네트워크를 통해 소통되는 정보가 누군가에 의해 변조되고 변경되거나 추가되지 않도록 보호대책을 필수적으로 강구
- 무결성을 위협하는 공격은 정보의 변경, 신분을 위장하거나 정보 데이터를 위장하여 사실인 것처럼 하는 것이 있음

▶ 가용성

- 정보를 액세스할 수 있는 사람이 언제 어디서든 원할 때 정보에 액세스할 수 있어야 하는 것
- 악의적인 공격자에 의해 좀비PC를 이용해 서버가 감당할 수 없는 양의 트래픽을 유발시켜 서버를 마비시키는 DDoS(서비스 거부공격) 공격을 가해 서비스 제공이 방해되는 경우도 있어 가용성을 침해당한 것
- 데이터 3법 개정 이후 : 플랫폼에서 통일된 기준으로 임시 대체키 부여하여 비식별 데이터 생성

데이터 3법이 시행되면서 데이터를 활용한 다양한 사업이 확장되고 있다. 대량의 데이터를 저장, 분류, 유통하는 빅데이터 플랫폼에서는 기밀성, 무결성, 가용성 등 보안 3대 요소는 물론 다양한 사이버 범죄 행위 시도가 불가능하도록 모든 시스템과 제도를 개선해야 한다.

2.4.2 관련 법 동향

빅데이터 이용을 활성화하기 위해 규제를 완화하는 데이터 경제 3법의 주요 내용은 다음과 같다.

표 22 데이터 경제 3법의 주요 내용

데이터 3법		
해당부처	법률명	주요내용
행정안전부	개인정보보호법	<ul style="list-style-type: none"> 개인정보보호위원회를 국무총리 소속의 중앙행정기관화 하여 행정안전부와 방송통신위원회의 개인정보 관련 사무를 개인정보보호위원회로 이관하여 개인정보 보호 컨트롤 타워로서의 기능 강화 개인정보 범위의 판단기준 제시 개인정보처리자는 가명정보를 통계작성, 과학적 연구, 공익적 기록보존 등의 목적으로 정보주체의 동의없이 처리할 수 있게 되었고, 개인정보처리자 간에 지정된 가명정보 결합전문기관을 통해 가명정보를 결합하여 이용할 수 있어 가명정보를 제도화 함 개인정보와 관련된 개인정보, 가명정보, 익명정보, 총 3가지 개념을 규정

데이터 3법		
해당부처	법률명	주요내용
금융위원회	신용정보보호법	<ul style="list-style-type: none"> 개인신용정보의 처리, 그 업무의 위탁, 유통 및 관리와 신용정보주체의 보호에 관해 일반법인 개인정보보호법의 일부 규정을 금융 분야에 맞게 규정 추가정보를 사용하지 아니하고는 특정 개인을 알아볼 수 없도록 처리(가명처리)한 개인신용정보로서 가명정보의 개념을 도입 고지사항의 중요한 사항만을 발제한 요약정보를 신용정보주체에게 알리고 정보활용 동의를 받을 수 있도록 하되, 신용정보주체가 요청할 경우 고지사항 전부를 알리도록 하는 등 신용정보주체에게 요약정보를 고지한 후에 동의를 얻는 가능성을 허용 개인인 신용정보주체가 금융회사, 정부·공공기관 등에 대하여 본인에 관한 개인신용정보를 본인이나 본인신용정보관리회사, 다른 금융회사 등에게 전송하여 줄 것을 요구할 수 있는 개인신용정보의 전송 요구권을 도입
과학기술 정보통신부, 방송통신위원회	정보통신망법	<ul style="list-style-type: none"> 개인정보 정의 등 개인정보보호법과 유사, 중복 규정 삭제 삭제된 일부 규정을 개인정보보호법 내에 특례 규정으로 이관 단말기 접근권한에 대한 동의, 주민등록번호 처리 관련 본인확인기관의 지정 등 규정은 존치함

2.5 가명처리

2.5.1 개인정보의 가명처리 단계

개인정보의 가명처리는 가명처리에 필요한 사전준비, 위험성을 검토하여 가명처리 수행, 사전준비와 가명처리 결과에 대한 적정성 검토 및 추가처리, 가명정보를 안전하게 관리하는 사후관리 단계로 구성된다.

표 23 개인정보의 가명처리 단계

단계	설명
1. 사전준비	• 설정 및 가명처리 대상 정보 확보
2. 가명처리	• 처리 상황에 따른 수준정의 및 처리
3. 적정성 검토 및 추가처리	• 가명처리가 적절한 수준으로 이루어 졌는지에 대한 최종적인 판단절차를 수행
4. 사후 관리	• 적정성 검토 결과 가명처리가 적정하다고 판단되면 가명정보를 본래 처리 목적을 위해 활용할 수 있으며, 법에 따라 기술적·관리적·물리적 안전조치 등 사후관리를 이행

그림 62 개인정보의 가명처리 단계별 절차도



* 가명정보 처리 가이드라인. 개인정보보호위원회, 2021

2.5.2 사전 준비

가명정보 처리 목적을 명확히 하고 가명처리를 위한 적합성 검토 및 계약서, 개인정보 처리방침, 내부 관리계획 등 필요한 서류를 작성한다.

▶ 사전 준비 사항

- 가명처리 개인정보파일 대상 선정: 처리목적 달성에 필요한 정보의 종류, 범위를 명확히 하여 가명처리 대상을 선정
- 가명처리 여부 검토(개인정보 보유부서 또는 가명정보 활용 관련 전담부서 등): 개인정보의 수집 목적 및 성격, 가명정보 활용 목적 등을 고려하여 가명처리 여부를 결정
- 가명정보 처리 상황 정의: 가명처리는 활용 형태, 처리 장소, 방법 등 처리 상황을 고려하여 수행해야하므로 가명처리 전 해당 사항을 미리 확인
- 가명정보 처리를 위한 안전조치: 개인정보 처리방침 수립·공개, 내부 관리계획 수립·시행 등 가명정보 처리에 앞서 이행하여야 할 사항 등 준비
- 필요서류 작성: 가명정보의 처리 또는 가명처리를 위탁(보호법 제26조에 따라 수행)하거나 가명 정보를 제3자에게 제공하는 경우 필요에 따라 재식별 금지에 관한 사항, 기타 처리에 있어 유의해야 할 사항 등을 포함한 계약서를 작성할 수 있음
- 기타: 그 밖에 개인정보 활용 및 가명처리 등에 대해 내부 승인 절차를 별도로 두고 있는 개인정보처리자는 이 단계에서 해당 절차를 진행하여야 함
- 가명정보 처리에 관한 내부 관리계획이 없는 경우, 계획 수립 필요

2.5.3 가명처리

가명처리 단계는 세부적으로 대상선정, 위험성 검토, 가명처리 방법 및 수준정의, 가명처리를 하는 4가지 단계로 구성되어 있다.

가명처리 시에는 가명정보 그 자체만으로 특정 개인을 알아볼 수 있는 지와 가명정보를 처리할 자가

보유하고 있거나 접근·입수 가능한 정보와의 사용·결합을 통해 식별가능한 지를 고려해야 한다.

▶ 대상선정

- [1단계, 사전준비]에서 설정한 목적을 달성하기 위해 필요한 항목을 개인정보파일에서 선정
- 가명처리 대상항목 선정 시 가명정보 처리 목적 달성에 필요한 최소 항목으로 해야 함

▶ 위험성 검토

- 위험성 검토는 처리하고자 하는 데이터의 식별 위험성을 가명처리 방법 및 수준에 반영하기 위한 절차이며, 식별 위험성은 1) 데이터 자체의 위험성과 2) 처리 환경의 위험성으로 구분하여 검토

그림 63 위험성 검토

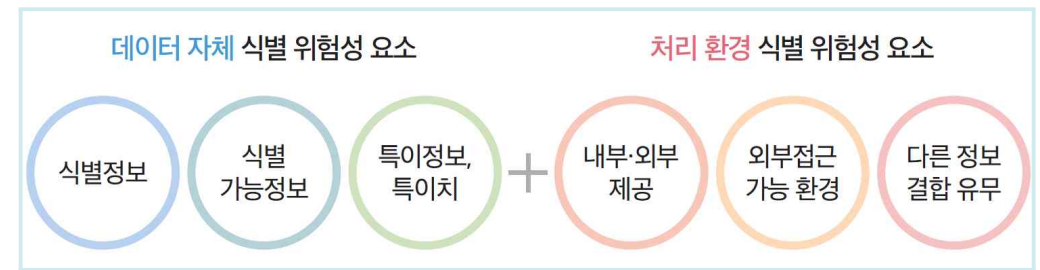


표 24 위험성 검토 내용

검토	검토내용
데이터 자체의 위험성 검토	<ul style="list-style-type: none"> • 데이터 자체 위험성 검토는 가명처리의 대상이 되는 정보에 식별 가능한 요소가 있는지를 파악하는 것으로, 그 자체로 식별될 위험이 있는 항목, 다른 항목과 결합을 통해 식별될 가능성이 있는 항목, 그 밖에 특이정보, 특이치 등이 있는지 검토 <ul style="list-style-type: none"> - (이용 항목별 식별 위험) 다른 사람과 구분하기 위해 부여된 식별 정보는 특정 개인과 고유하게 연결되어 있으므로, 해당 정보가 포함되어 있을 경우 특정 개인을 알아볼 가능성이 높음 - (다른 이용 항목과의 결합 유무) 단일 이용 항목으로는 식별 가능성이 없으나, 가명처리 대상이 되는 다른 이용 항목과 결합하여 식별 가능성이 높아지는 이용 항목이 있는지 검토 - (특이정보, 특이치 유무) 가명처리 대상 전체 데이터에 식별 가능성을 가지는 고유(희소)한 값이 있는지, 편중된 분포를 가지는 단일·다중 이용 항목이 있는지 검토

검토	검토내용
처리 환경의 위험성 검토	<ul style="list-style-type: none"> 개인정보처리자는 가명정보 활용 형태, 처리 장소, 방법 등 가명정보 처리 상황에 따라 발생할 수 있는 식별 위험성 검토 <ul style="list-style-type: none"> (내부이용 및 외부제공 여부) 처리 주체가 보유하고 있는 정보 또는 접근·입수 가능한 정보와 이용 범위 및 유형을 고려하여 식별가능한 항목이 있는지 검토 (처리 장소) 가명정보가 해당 가명정보 외에 다른 정보의 접근·입수가 제한된 장소에서 처리되는지 검토 (다른 정보 결합 유무) 가명정보를 다른 정보와 연계 분석할 예정인 경우 다른 정보와 결합하여 식별가능한 항목이 있는지 검토 가명정보를 다른 정보와 내부 결합 할 예정인 경우 다른 정보와 결합하여 식별가능한 항목이 있는지 검토 가명정보를 반복 제공할 예정인 경우 반복 제공을 통해 식별 위험이 높아지는 항목이 있는지 검토

▶ 가명처리 방법 및 수준정의

- 개인정보처리자는 '식별 위험성 검토 결과보고서'를 기반으로 가명정보의 활용 목적 달성에 필요한 수준을 고려하여 가명처리 방법 및 수준 정의를 하여야 한다.

표 25 가명처리 방법 및 수준 정의표 예시

순번	항목명	처리수준	비고
1	소유자명	가명처리	소유자명과 연락처는 추후 시계열 분석을 위해 가명처리 수행
2	연락처	(암호화: SHA2+Salt)	
3	지번	가명처리(삭제)	세부 지번의 정보는 분석목적에 필요하지 않음
4	전세		
5	보증금	기타기술 (라운딩: 만원 단위)	만 원 단위의 금액만 분석목적에 필요
6	월세		
7	주택구분		
8	시도		
9	시군구	처리하지 않음	
10	읍면동	※ 항목이 다수여서 작성이 어려운 경우 '별지'를 활용하여 목록만 제시	처리하지 않는 항목을 작성
11	전용면적		
12	공급면적		

▶ 가명처리

- 개인정보처리자는 '가명처리 방법 및 수준 정의표'를 기반으로 가명처리를 수행하여야 함
- 가명처리 단계에서 생성되는 추가정보는 원칙적으로 파기하고 필요한 경우 가명정보와 분리하여 별도로 저장하여야 함

2.5.4 적정성 검토 및 추가 가명처리

가명처리가 적절한 수준으로 이루어 졌는지에 대한 최종적인 판단절차를 수행하여야 한다. 가명처리에 대한 적정성 검토는 개인정보처리자의 판단에 따라 내부 인원을 활용하여 자체적으로 검토할 수 있으며, 필요시 외부전문가를 통하여 검토할 수 있다.

▶ 적정성 검토

- (필요서류 및 위험성 검토) 필요서류 내용, 데이터 자체 위험도, 처리환경 등 위험성 판단 항목을 누락 없이 검토하였는지 확인
- (가명처리 방법 및 수준의 적정성) 가명처리 단계에서 위험성 검토 결과를 반영하여 가명처리 방법 및 수준을 적정하게 정의하였는지 확인
- (가명처리의 적정성) 정의한 가명처리 방법 및 수준에 따라 실제 가명처리를 수행하였는지 확인
- (처리 목적 달성 가능성) 가명처리를 수행한 정보가 당초 가명정보 처리 목적을 달성할 수 있는지 여부 검토

▶ 추가 가명처리

- 적정성 검토 결과 가명처리가 적정하지 않다고 판단되면 가명처리를 다시 수행하거나 부분적으로 추가적인 가명처리를 수행할 수 있음

2.5.5 사후관리

적정성 검토 결과 가명처리가 적정하다고 판단되면 가명정보를 본래 처리 목적을 위해 활용할 수 있으며, 법에 따라 기술적·관리적·물리적 안전조치 등 사후관리를 이행하여야 한다.

▶ 재식별 금지 및 모니터링

- 개인정보처리자는 누구든지 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 아니되며, 가명정보를 처리하는 중 우연히 특정 개인이 식별되는 경우 즉시 처리중지, 회수, 파기 등 위와 같은 위험을 제거하기 위해 적절한 조치를 수행하여야 함
- 또한, 개인정보처리자는 가명정보 처리 과정에서 특정 개인이 식별될 위험이 있는지 여부를 지속적으로 모니터링 하여 안전하게 처리하여야 함

▶ 안전조치 시행

- 개인정보처리자는 사전준비 단계에서 수립한 내부 관리계획에 따라 가명정보를 안전하게 관리하여야 함

▶ 개인정보 처리방침 수립 및 공개

- 개인정보처리자는 가명정보 처리와 관련하여 처리 목적, 처리하는 개인정보의 항목 등을 개인정보 처리방침에 공개하여야 함

▶ 가명정보 처리 관련 기록 작성 및 보관

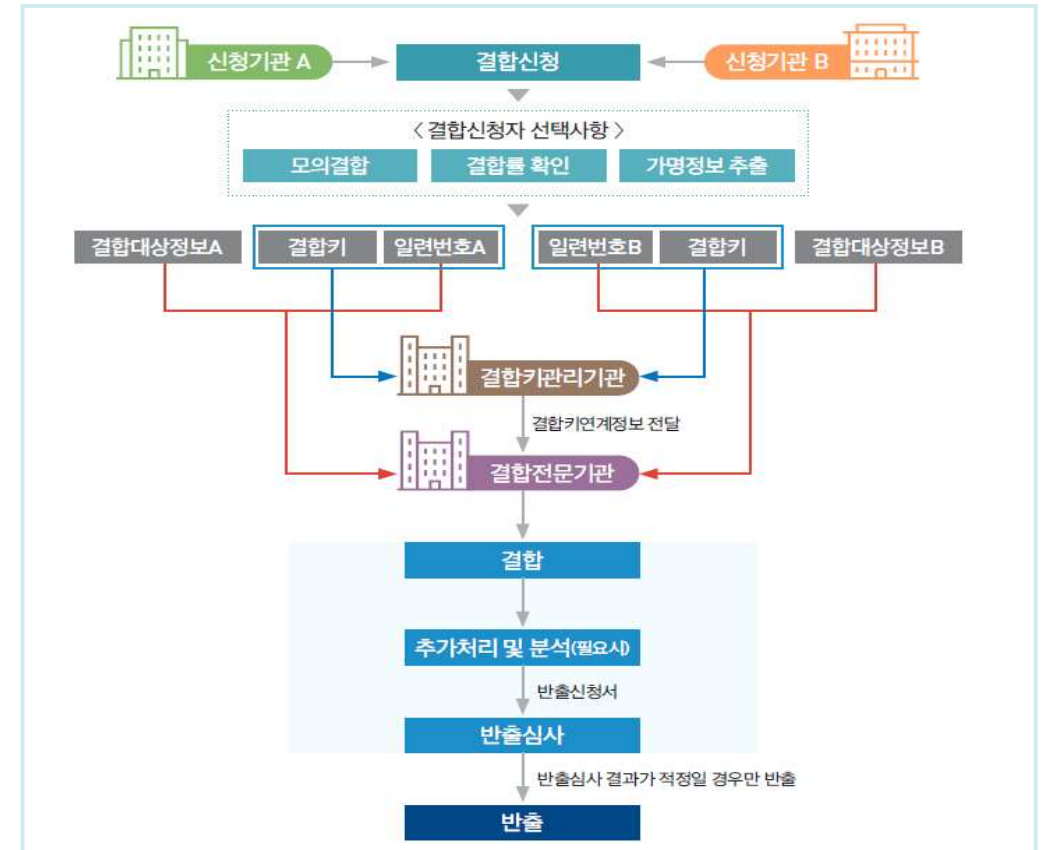
- 개인정보처리자는 가명정보의 처리 목적, 개인정보 항목, 이용내역, 제3자 제공 시 제공받는 자를 작성하여 보관하여야 함

2.6 가명정보 결합 및 반출

2.6.1 가명정보 결합·반출 개요

가명정보 결합 및 반출절차는 다음과 같다.

그림 64 가명정보 결합·반출 절차



* 가명정보 처리 가이드라인. 개인정보보호위원회. 2021

▶ 가명정보의 결합

- 개인정보처리자는 결합전문기관을 통해 통계작성, 과학적 연구, 공익적 기록보존 등을 위한 가명정보 결합 가능

▶ 가명정보의 결합 방식

- 가명정보의 결합은 보통 1회로 종료되지만, 처리목적을 달성하기 위해 필요한 경우 동일한 서로 다른 개인정보처리자 간의 가명정보를 반복적으로 결합할 수도 있음

▶ 가명정보 결합·반출 절차

- 1단계: 결합신청자 간의 결합신청에 필요한 사항을 협의하여 결합신청서를 작성하는 등 가명정보 결합에 필요한 사전 준비사항을 수행하고 그 결과를 반영하여 결합전문기관에 결합 신청
- 2단계: 가명정보를 제공하고자 하는 결합신청자는 결합기관리기관으로부터 결합키 생성에 이용되는 정보(Salt값)를 수신하여 결합키를 생성하고 필요시 모의결합, 결합물 확인, 가명정보 추출 등을 수행한 후 결합에 필요한 정보를 각 기관에 전송
- 3단계: 결합정보를 이용하고자 하는 결합신청자는 결합정보를 반출하기 전 결합전문기관 내에 설치된 별도의 공간에서 추가 가명·익명처리를 하거나, 결합전문기관이 분석기능을 지원하는 경우 분석을 수행할 수 있으며, 결합정보 또는 분석결과 등을 반출하고자 하는 경우 결합전문기관에 반출신청서를 제출하여 반출 신청
- 4단계: 결합정보를 이용하고자 하는 결합신청자는 반출정보를 당초 결합신청서 및 반출신청서에 기재한 목적에 따라 처리할 수 있으며, 가명정보 처리 시 안전조치를 준수

그림 65 가명정보 결합·반출 절차

절차	결합신청자	결합기관리기관	결합전문기관
1 사전준비 및 결합 신청	① 결합 신청	-	② 결합신청서 검토 및 접수 ③ 결합 일정·절차 등 협의(결합신청자)
2 결합 수행	1. 결합키 생성 ① 결합키 생성 협의 ② 결합키 및 일련번호 생성	② 결합키 생성 협의(Salt값 전송)	-
	2. 모의결합 (선택) ① 결합키 전송 ④ 모의결합 대상 가명처리 ⑤ 가명처리된 모의결합대상정보, 가명처리내역 및 결합키 전송 ⑧ 모의결합된 정보 분석 (결합전문기관 내) * 반출 제한	-	② 모의결합 가능성 검토 및 통지 ③ 모의결합 대상 결합키 선정 및 전송 ④ 가명처리 수준 검토 (필요시 추가처리 요청) ⑦ 모의결합 수행 ⑨ 모의결합 관련 정보 파기
	3. 결합물 확인(선택) ① 결합키 및 일련번호 전송 ④ 결합물 확인	② 결합키연계정보 생성 ③ 결합물 측정 및 통보	-
	4. 가명정보 추출(선택) ① 결합키 및 일련번호 전송 ② 추출 요청	③ 추출 가능 여부 검토 및 통지 ④ 추출에 필요한 일련번호 선정 및 전송	-
	5. 가명처리 및 검토 ① 결합대상정보 확정 ② 가명처리 ③ 가명처리된 결합대상정보, 가명처리 내역 및 일련번호 전송	-	★ 가명처리 지원(가능한 경우) ④ 가명처리 수준 검토 (필요시 추가처리 요청)
	6. 결합 ① 결합키 및 일련번호 전송	② 결합키연계정보 생성 및 전송 * 반복결합의 경우 반복결합키 포함	③ 결합키연계정보 수신 및 가명정보 결합 * 반복결합의 경우 반복결합키 포함
3 추가처리 및 반출	1. 추가처리 (필요시) ① 결합된 정보의 추가처리 및 분석(결합전문기관 내) * 결합전문기관에 지원 요청 가능	-	★ 추가처리 및 분석 지원 (가능한 경우)
	2. 반출 ① 반출신청	④ 결합키 및 결합키연계정보 파기(필요시 유지) * 반복결합의 경우, 결합키 생성방법(Salt)-반복결합키 생성방법(Salt) 보관	② 반출심사위원회 구성·운영 ③ 반출 승인 및 결합정보 반출 * 반복결합의 경우 반복결합키 포함 ④ 결합키연계정보 파기
4 활용 및 사후관리	✓ 안전성 확보 조치 이행 ✓ 가명정보 처리 내역 기록·보관	-	★ 반출한 정보 분석 지원 (가능한 경우) ★ 개인정보 보호 교육 제공 (가능한 경우)

2.6.2 사전준비 및 결합 신청

▶ 사전준비

- 결합신청자는 서로 다른 결합신청자 간의 협의를 통해 가명정보 결합에 대한 사전준비를 수행 하여야 함
- 결합신청자는 필요시 가명정보 결합에 관한 별도의 내부승인절차 등을 이행할 수 있으며, 결합 전에 대한 계약 체결 등 필요한 조치를 할 수 있음

▶ 결합 신청

- (결합 신청) 결합신청자는 가명정보를 보유하고 있는 개인정보처리자 뿐만 아니라, 현재 가명정보를 보유하고 있지 않으나 결합된 가명정보를 처리할 예정인 개인정보처리자도 결합신청자가 됨
- (신청 방법) 결합신청자는 결합전문기관 선택 후 가명정보 결합종합지원시스템(ink.privacy.go.kr)을 이용하여 결합 신청
- (신청 서류) 결합신청자는 '가명정보 결합 및 반출 등에 관한 고시의 [별지 제3호] 결합신청서와 첨부 서류를 결합 신청 시 제출하여야 함
- (신청서 검토 및 접수) 결합전문기관이 신청서 작성내용(결합 목적 적합성 등) 및 첨부 서류에 대한 보완을 요청한 경우 결합신청자는 해당사항을 보완하여 다시 제출하여야 함
- (결합 일정 및 절차 등 협의) 결합신청자는 결합 신청 내역에 따라 결합 절차 및 필요한 정보 등을 결합전문기관 및 결합기관리기관과 협의하여야 함

2.6.3 결합 수행

▶ 결합키 생성

- 가명정보를 제공하고자 하는 결합신청자는 결합기관리기관과 결합키 생성에 관한 사항을 협의하고 결합기관리기관으로부터 결합키 생성에 필요한 Salt값을 전송받아야 함
- 반복결합을 신청하는 경우 추후 반출되는 정보와의 연계·분석을 위하여 결합키에 사용된 결합키 생성항목, 인코딩 방식, 알고리즘(Salt값 제외)을 보관하여야 함

▶ 모의결합(선택사항)

- 모의결합을 신청한 결합신청자는 결합기관리기관과의 협의에 따라 생성한 결합키를 결합전문기관에 전송하여야 함
- 결합전문기관은 모의결합대상정보의 특성, 결합물 등을 고려하여 모의결합이 가능한 경우 모의결합대상정보를 선정하여 해당 결합키를 결합신청자에게 전송하여야 함
- 결합신청자는 결합전문기관으로부터 결합키를 제공받아 해당 모의결합대상정보를 가명처리하여 가명처리 내역과 함께 결합전문기관에 전송하여야 함
- 결합전문기관은 가명처리 내역을 확인한 후 보완 사항이 없으면 결합키를 사용하여 모의결합대상정보의 결합을 수행함

- 결합신청자는 결합전문기관 내에서 모의결합된 정보를 분석할 수 있음
- 결합전문기관은 결합신청자의 분석이 완료된 후에는 모의결합에 사용된 정보를 파기하여야 함

▶ 결합률 확인(선택사항)

- 결합률 확인을 신청한 결합신청자는 결합키와 일련번호를 결합키관리기관에 전송하여야 함
- 결합키관리기관은 결합신청자로부터 결합키와 일련번호를 제공받아 결합률을 확인한 후 통지하여야 함

▶ 가명정보 추출(선택사항)

- 가명정보 추출을 신청한 결합신청자는 결합키와 일련번호를 결합키관리기관에 전송하여야 함
- 결합키관리기관은 추출 여부를 판단하는데 필요한 정보(결합 목적 등)를 결합신청자로부터 제공받아 추출 가능 여부를 검토하고, 추출이 가능한 경우 추출에 필요한 일련번호를 결합신청자에게 전송

▶ 가명처리 및 검토

- (가명처리) 결합신청자는 모의결합, 결합률 확인, 가명정보 추출 등의 선택절차가 모두 완료되고 결합절차를 진행하기로 결정하면 결합대상정보를 가명처리하여 결합전문기관에 전송하여야 함
- (가명처리 검토) 결합전문기관은 결합신청자가 제출한 결합대상정보 및 가명처리 내역을 검토하여야 함

▶ 결합

- (결합키 및 일련번호 전송) 결합신청자는 결합키와 일련번호를 결합키관리기관에 전송
- (결합키연계정보 생성) 결합키관리기관은 결합키와 일련번호를 사용하여 결합키연계정보를 생성하고 결합전문기관에 결합키연계정보 전송
- (결합) 결합전문기관은 결합키연계정보와 일련번호, 결합대상정보를 사용하여 결합

2.6.4 추가처리 및 반출

▶ 추가처리 및 분석

- (추가처리) 결합정보를 이용하고자 하는 결합신청자는 결합전문기관 내에서 결합정보가 특정 개인을 알아 볼 수 있는지 여부를 확인하고, 개인식별 가능성이 확인된 경우 해당 부분에 대한 추가처리를 수행하여야 함
- (분석) 결합신청자는 결합전문기관 내에 마련된 분석에 필요한 시설, 장비를 갖춘 공간에서 결합정보를 분석할 수 있음

▶ 반출신청 및 심사

- (반출신청) 결합정보를 반출하려는 결합신청자는 결합 고시 [별지 제4호] 반출신청서와 첨부 서류를 제출하여야 함
- (신청서 검토 및 접수) 반출신청서를 제출받은 결합전문기관은 신청서 및 첨부 서류에 누락이 없는지 확인하고 보완 사항이 없으면 해당 반출신청서를 접수하여야 함
- (반출심사) 결합신청자가 반출을 요청하면 결합전문기관은 접수일로부터 영업일 기준 5일 이내 반출심사위원회 구성 등에 관한 사항을 결합신청자에게 통지하고, 반출심사위원회를 개최하여 반출심사를 진행하여야 함
- (추가설명 등) 결합신청자는 반출심사위원회의 요청에 따라 추가 서류를 제출하거나 직접 출석하여 설명할 수 있음

▶ 반출

- (반출정보) 결합전문기관이 반출 승인을 하면 결합신청자는 결합전문기관 내에서 결합정보를 분석한 결과물만을 반출하거나, 결합정보(데이터셋)를 반출할 수 있음

2.6.5 활용 및 사후관리

▶ 활용 및 사후관리

- (반출정보의 이용 범위) 반출정보는 결합신청자가 반출심사 시 제출한 환경(가명정보 활용 형태, 처리 장소, 방법)과 목적범위 내에서 활용하는 것이 원칙임
- (반복결합) 반복결합을 신청하여 반출한 경우에는 반출정보에 반복결합키가 포함되어 있으므로, 이를 이용하여 내부에서 연계하여 활용
- (재식별 금지) 결합신청자는 반출정보를 특정 개인을 알아보기 위한 목적으로 처리하여서는 아니되며, 재식별되지 않도록 지속적으로 모니터링 하여야 함
- (안전조치) 결합신청자는 반출정보를 활용하는 경우 안전성 확보에 필요한 관리적·기술적·물리적 조치를 수행하여야 함

표 26 가명정보의 안전한 관리

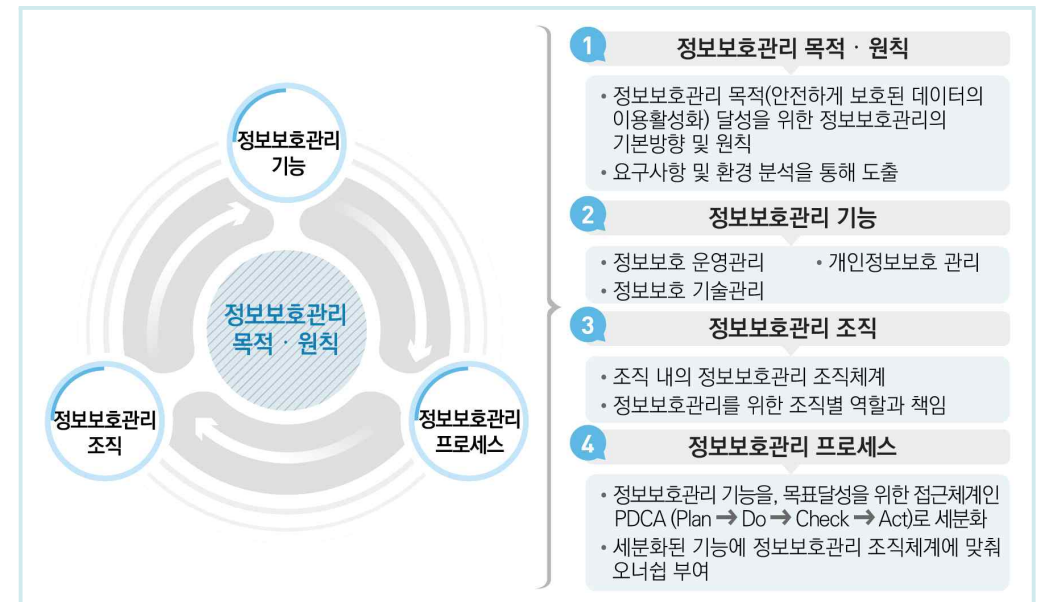
구분	내용	관련법규
관리적 보호조치	<ul style="list-style-type: none"> 개인정보처리자는 가명정보 및 추가정보를 안전하게 관리하기 위한 내부 관리계획을 수립·시행하고, 관리책임자는 연1회 이상 내부 관리계획의 이행 실태를 점검 관리하여야 함 위탁업무 수행 목적 이외의 가명정보 처리 금지 사항을 규정하고, 분실·도난·유출·위변조·훼손 또는 재식별 되지 않도록 교육하고 감독하여야 함 개인정보처리자는 가명정보 처리와 관련하여 개인정보 처리방침을 수립하여야 함 	<ul style="list-style-type: none"> 보호법 시행령 제29조의5 제1항 제1호 보호법 제26조 수탁자 관리·감독의 의무 보호법 제30조 개인정보 처리방침 수립 및 공개
기술적 보호조치	<ul style="list-style-type: none"> 개인정보처리자는 추가정보와 가명정보를 분리하고 불법적인 결합 및 재식별이 되지 않도록 접근통제를 강화하는 등의 필요한 조치를 적용하여야 함 가명정보 처리 자의 비밀번호 규격 및 설정, 사용자 계정, 접근권한 및 이력정보 3년 보관, 사용자 계정 유효성 점검 등의 접근권한 관리하고 인력을 최소한으로 통제하여야 함 개인정보처리자는 가명정보의 처리목적, 가명처리한 개인정보처리 항목, 가명저의 이용내역, 제3자 제공 시 제공받는 자를 작성하여 보관하여야 함 	<ul style="list-style-type: none"> 보호법 시행령 제29조의5 제1항 제2호 추가정보의 분리 보관 보호법 시행령 제29조의5 제1항 제3호 접근권한의 분리 보호법 시행령 제29조의5 제2항 가명정보 처리 관련 기록 작성·보관
물리적 보호조치	<ul style="list-style-type: none"> 개인정보처리자는 가명정보 또는 추가정보의 안전한 관리를 위하여 물리적 안전조치를 취하여야 함 	
정보주체 권리보장	<ul style="list-style-type: none"> 개인정보처리자는 정보주체가 자신의 개인정보에 대한 가명 처리 정지를 요구하는 경우 이를 보장하여야 함 	<ul style="list-style-type: none"> 보호법 제37조

03 체계 구성



정보보호관리는 정보보호를 통한 안전한 데이터 활용을 목적으로 정보보호 기능, 수행조직과 정보 보호관리 프로세스를 체계화하여 수행한다.

그림 66 정보보호관리 체계 개요



3.1 정보보호관리 원칙

정보보호의 목적을 달성하기 위해 3개의 정보보호관리 원칙과 7개의 세부 원칙을 제시한다.

▶ 원칙1. 정보보호 관리는 관련 법규를 준수해야 한다.

- 정보보호는 개인정보보호 및 정보보호 관련 법규, 가이드라인에 따라 정보보호기준을 만들고, 이의 교육·전파, 적용, 감사 및 개선조치를 수행해야 한다.

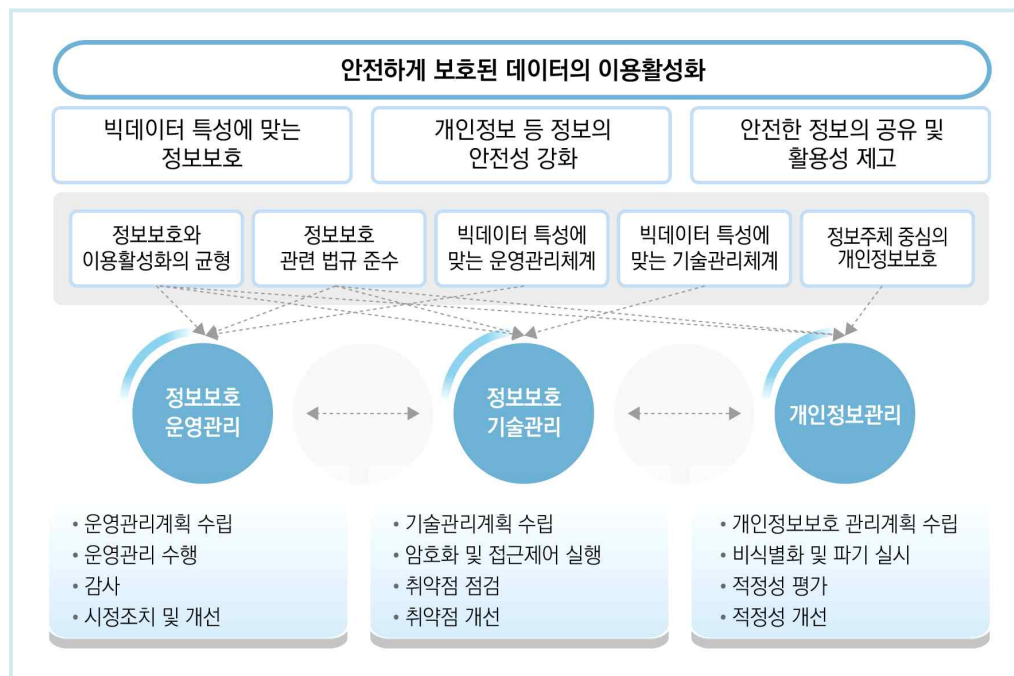
- 개인정보보호 및 정보보호 관련 법규와 가이드라인의 개정사항을 확인하고, 영향도를 분석하는 등 지속적으로 법규 준수를 위한 노력을 해야 한다.
- ▶ 원칙2. 정보보호 관리는 데이터 생애주기 쉼단계에 걸쳐 데이터가 안전하게 보호되도록 관리해야 한다.
- 정보보호관리는 기밀성, 무결성, 가용성이 보장되어야 한다.
 - 정보를 수집하고, 수집 목적 범위 내에서 가공·분석하며, 제3자에게 제공하기기 위한 처리를 하는 활용 단계 모두 정보보호를 해야 한다.
 - 보유기간이 경과하거나 사용목적을 달성한 개인정보는 복구가 불가능하도록 파기해야 한다.
- ▶ 원칙3. 정보보호관리 활동은 PDCA 사이클을 준수하고 오너십을 부여해야 한다.
- 목적달성과 지속적인 개선을 위한 접근체계인 PDCA 사이클 기반 정보보호관리 활동을 수행한다.
 - 정보보호 관련 주체들 간의 상호협력을 강화하기 위해 역할 정의와 오너십을 명확히 해야 한다.

3.2 정보보호관리 기능

3.2.1 개요

정보보호관리 원칙을 준수하며 정보보호의 목적을 달성하기 위한 3개의 정보보호 관리 기능을 도출한다.

그림 67 정보보호관리 기능 개요



▶ 기능1. 정보보호 운영관리

- 정보보호 관련 법규 등을 참고하여 정보보호를 위한 운영계획을 수립하고 실행하도록 관리
- 정보보호가 정상적으로 이루어지는 지를 감사 등을 통해 확인하고 점검하도록 관리
- 점검 및 확인 결과 문제가 발생될 경우 원인분석, 영향도 분석 등을 통해 개선계획을 수립하고 개선조치를 수행하도록 관리

▶ 기능2. 정보보호 기술관리

- 데이터 보호를 위한 접근제어, 암호화 등의 계획을 수립하여 실행하도록 관리
- 접근제어 및 암호화 관련 기술적인 취약점이 있는지를 확인하고 점검하도록 관리
- 취약점 발견 시 원인 및 영향도를 분석하여 개선계획을 수립하고 개선조치를 수행하도록 관리

▶ 기능3. 개인정보관리

- 개인정보보호 관련 법규 및 가이드를 참조하여 개인정보보호 계획을 수립하고 비식별조치 및 파기 등을 실시하도록 관리
- 비식별조치 등의 적정성을 평가하고 점검하도록 관리
- 평가결과 문제가 있는 경우 개선조치를 수행하도록 관리

3.2.2 기능 세분화

정보보호관리를 정보보호 운영관리, 정보보호 기술관리, 개인정보보호 관리 기능을 목표달성을 위한 접근체계인 PDCA 사이클에 따라 세분화한다.

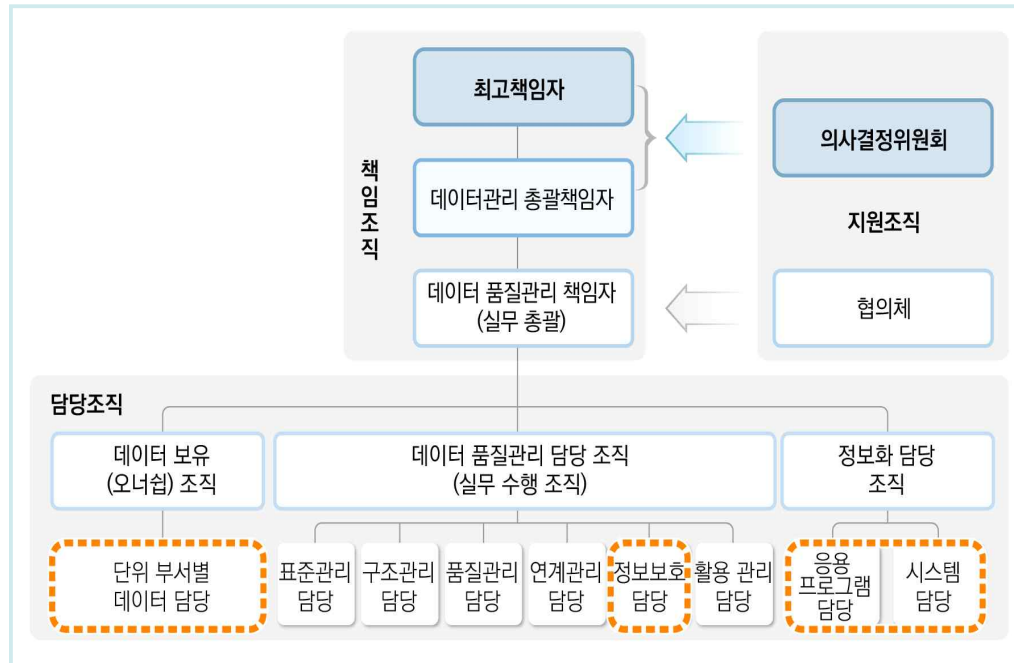
표 27 정보보호관리 기능 세분화

기능	관리 대상	세부 관리기능			
		Plan(계획)	Do(실행)	Check(모니터링 및 통제)	Act(개선)
정보 보호 운영 관리	수집 데이터	운영관리계획 수립 • 관련법규 분석 • 정보보호 운영 절차, 운영 현황에 대한 점검 수행을 위한 계획을 수립	운영관리 수행 • 운영계획에 따른 정보보호 조치 실시	감사 • 운영계획 및 법규 준수여부를 확인 • 미준수 확인 시 원인을 분석	시정조치 및 개선 • 원인에 따라 영향도를 분석하여 개선계획을 수립 • 개선계획에 따른 개선 조치 실시 및 확인
정보 보호 기술 관리	수집 데이터, 보안툴	기술관리계획 수립 • 운영계획에 따라 암호화 및 접근제어 관련 기준 정의 등 실행계획 수립	암호화 및 접근제어 실행 • 기술계획에 따라 접근제어 및 암호화 실시	취약점 점검 • 기술계획 준수여부를 확인 • 미준수 확인 시 원인을 분석	취약점 개선 • 원인에 따라 영향도를 분석하여 개선계획을 수립 • 개선계획에 따른 개선 조치 실시 및 확인
개인 정보 보호 관리	수집 데이터	개인정보보호 관리계획 수립 • 운영계획에 따라 개인정보 식별 등 개인정보보호계획 수립	비식별화 및 파기 실시 • 개인정보보호 계획에 따라 개인정보 비식별화 및 기간이 경과된 개인정보 파기	적정성 평가 • 개인정보보호 계획 및 법규 준수여부를 확인 • 미준수 확인 시 원인을 분석	적정성 개선 • 원인에 따라 영향도를 분석하여 개선계획을 수립 • 개선계획에 따른 개선 조치 실시 및 확인

3.3 정보보호관리 조직 및 역할

- ▶ 전체 데이터 관리조직에서 정보보호 관리와 관련된 조직은 정보보호 담당, 정보화 담당 조직(응용프로그램 담당, 시스템 담당), 협의체로 정의한다.

그림 68 정보보호관리 조직 및 역할



- ▶ 의사결정조직, 지원조직 및 담당조직별 정보보호 관련 역할을 정의한다.

표 28 정보보호관리 역할 정의

조직구분		정보보호 관련 역할
의사결정조직	데이터관리 총괄 책임자	<ul style="list-style-type: none"> 정보보호 관리 총괄 최고책임자를 보좌 정보보호 관련 주요 정책 및 외부 조직간 업무 조정을 위한 의사결정위원회 소집 및 회의 주재
	데이터 품질관리 책임자 (실무 총괄)	<ul style="list-style-type: none"> 정보보호 관리 실무 총괄 의사결정위원회 회의 지원 데이터 관리 총괄 책임자 보좌

조직구분		정보보호 관련 역할
지원조직	의사결정위원회	<ul style="list-style-type: none"> 정보보호 및 개인정보관리 정책 심의, 외부 조직간 업무 조정 조직의 주요 정보 보호 활동과 업무 수행에 관한 심의 책임과 최고 의사결정 기구 역할을 함 정보 사용 전반에 걸친 중요한 정보보호 관련 심의 및 승인 기능을 수행하며 필요한 자원 할당을 통해 정보보호관리 활동이 적절히 수행될 수 있도록 하여야 한다.
	협의체	<ul style="list-style-type: none"> 내부 조직 간의 정보보호관리 업무 공유 정보보호 및 개인정보보호 업무 수행에 필요한 실무 차원의 지원 수행 회의 소집 및 주재는 데이터품질관리 책임자가 수행
담당조직	데이터 보유 (오너십) 조직	<ul style="list-style-type: none"> 개인정보 취급조직으로서 개인정보 처리 관련업무 수행 정보주체의 이전 수렴 및 불만사항 접수 개인정보를 보호하고 개인정보와 관련한 이용자의 고충을 처리하기 위하여 개인정보 보호책임자 보유데이터에 대한 정보관리 및 위험에 대한 대응.개선 활동 수행
	데이터 품질관리 담당 조직 (실무수행 조직)	<ul style="list-style-type: none"> 정보보호 실무 전반을 수행 정보보호 관련 업무를 기획하고 시행하기 위한 세부계획을 마련하며, 정보보호관련 기술적 보안통제 사항을 관리 정보보호 예방 및 위험에 대한 주기적인 모니터링을 통해 정보를 보호 관리 개인정보 보유현황 및 업무를 파악하고 개인정보의 전체 수명주기에 걸쳐 법적 요건을 만족하기 위한 취급방침을 수립하고 관리 데이터관리자의 업무 보조
	정보화 담당 조직	<ul style="list-style-type: none"> 정보시스템 구축, 운영에 있어 개인정보를 수집할 수 있는 기반에 대한 설계, 개발, 운영 및 보완 작업 지원 응용프로그램, 데이터베이스의 정보보호 위험관리에 있어 원인분석 및 개선 지원 침해사고 등 비상상황 시 대응 지원 기타 정보보호 관리자의 업무 지원

3.4 정보보호관리 프로세스

3.4.1 구성도

PDCA 관점으로 세분화된 정보보호관리 기능을 정보보호관리 프로세스로 구성한다.

그림 69 정보보호관리 프로세스 구성도



정보보호 관리체계의 효율적인 운영을 위하여 3개의 기능과 12개의 정보보호 프로세스를 정의한다.

표 29 정보보호관리 프로세스 정의

구분	프로세스 정의
정보 보호 운영 관리	운영관리 계획 수립 <ul style="list-style-type: none"> 정보보호 관련 법규 등을 바탕으로 정보보호를 위한 계획을 감사계획을 포함하여 수립 운영계획에는 정보보호 기술관리 및 개인정보보호 관리 기본 방향을 포함한 총괄 운영계획을 수립
	운영관리 수행 <ul style="list-style-type: none"> 운영계획에 따라 정보보호 조치를 실시
	감사 <ul style="list-style-type: none"> 수지 또는 정기적으로 감사를 위한 임시조직을 구성하여 운영계획 및 법규 준수 여부를 확인 운영계획 및 법규를 준수하지 않는 사항이 있는 경우 원인을 분석
	시정조치 및 개선 <ul style="list-style-type: none"> 원인에 따라 영향도를 분석하여 개선계획을 수립 개선계획에 따라 시정조치 및 개선조치를 해결될 때까지 반복적으로 수행
정보 보호 기술 관리	기술관리 계획 수립 <ul style="list-style-type: none"> 운영계획의 정보보호 기술관리 방향을 참조하여 접근제어 및 암호화 관련 기준을 정의하고 실행계획을 수립
	접근제어 및 암호화 실행 <ul style="list-style-type: none"> 정보보호 기술계획에 따라 접근제어 및 암호화를 수행

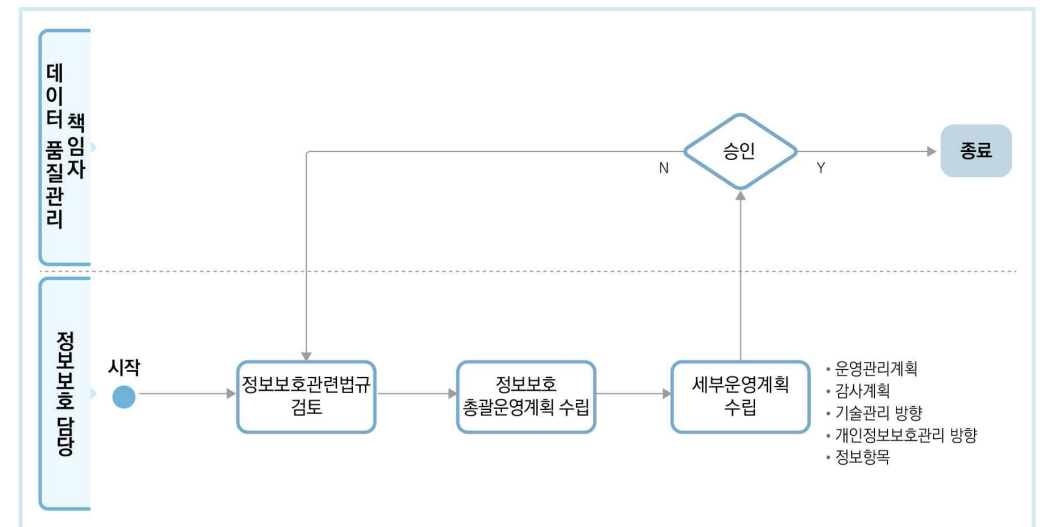
구분	프로세스 정의
취약점 점검	<ul style="list-style-type: none"> 기술계획 준수 여부 확인 등의 기술관리 취약점을 파악 취약점 발견 시 원인을 분석
	<ul style="list-style-type: none"> 원인에 따라 영향도를 분석하여 취약점 완화 등의 개선계획을 수립 개선계획에 따라 개선조치를 해결될 때까지 반복적으로 수행
개인 정보 보호 관리	개인정보 보호 관리 계획 수립 <ul style="list-style-type: none"> 운영계획의 개인정보보호 관리 방향을 참조하여 개인정보 식별 등의 개인정보보호 실행계획을 수립
	비식별화 및 파기 실시 <ul style="list-style-type: none"> 개인정보보호 계획에 따라 개인정보 비식별화 및 기간경과 등의 개인정보 파기를 수행
	적정성 평가 <ul style="list-style-type: none"> 개인정보보호계획 및 법규 준수 여부 확인 등의 개인정보보호 적정성을 평가 문제 발견 시 원인을 분석
	적정성 개선 <ul style="list-style-type: none"> 원인에 따라 영향도를 분석하여 개인정보보호 적정성을 달성할 수 있는 개선계획을 수립 개선계획에 따라 개선조치를 해결될 때까지 반복적으로 수행

3.4.2 정보보호 운영관리

▶ 운영관리계획 수립

- 정보보호 관련 법규 등을 바탕으로 정보보호를 위한 계획을 감사계획을 포함하여 수립한다.

그림 70 운영관리계획 수립



- [정보보호관련 법규 검토] 정보보호 운영계획 수립을 위해 개인정보보호 법 등의 정보보호 관련

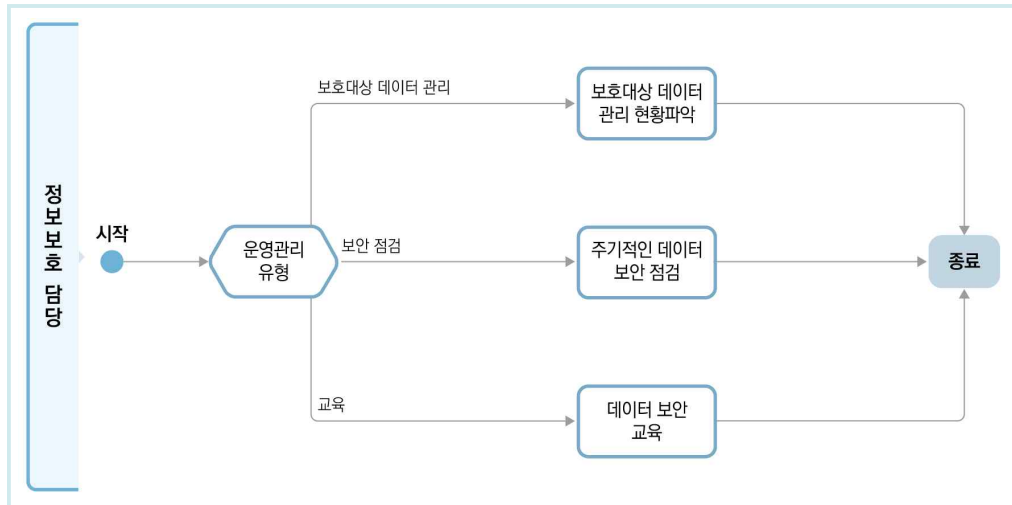
법규 및 개정 동향 검토가 필요하며, 공공기관의 경우 국가정보화기본법 등, 민간기관의 경우 해당 산업에 맞는 법규를 검토해야 한다.(정보통신망법, 신용정보법, 의료법, 생명윤리법 등)

- [정보보호 총괄운영계획 수립] 정보보호 관련법규 검토 결과를 바탕으로, 정보보호 관리체계를 포함하여 정보보호를 위한 총괄운영계획을 수립한다.
- [세부운영계획 수립] 총괄운영계획에 따라 세부운영계획을 수립하며, 세부운영계획에는 운영관리계획 외에 감사계획, 정보보호 기술관리 방향, 개인정보 보호관리 방향을 포함하여 수립함으로써, 통합된 정보보호를 수행할 수 있도록 한다.
- [승인] 정보보호 운영관리계획이 수립되면, 플랫폼 및 센터 의사결정체계 및 결재와 관련된 규칙에 따라, 데이터 품질관리 책임자의 승인을 얻는다.

▶ 운영관리 수행

- 정보보호 운영관리를 위한 운영계획을 수행한다.

그림 71 운영관리 수행

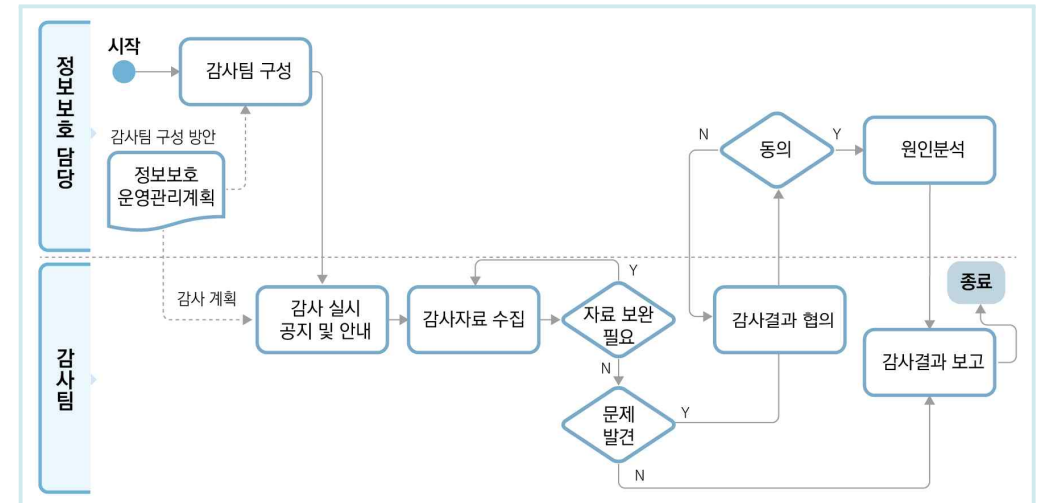


- [운영관리 유형] 정보보호를 위한 운영관리 유형에 따라 정보보호 상이한 조치를 실시한다.
- [보호대상 데이터 관리] 시스템 및 수집데이터 추가, 가공·분석데이터의 증가 및 활용데이터 변경 등에 따라 보호대상 데이터의 변경 및 현황을 관리한다.
- [주기적인 데이터 보안 점검] 일일 점검 목록(Check-List) 등을 작성하여 효율적으로 주기적인 데이터 보안 상황 점검이 이루어지도록 한다.
- [데이터 보안 교육] 데이터 보안의 중요성을 강조하기 위해 주기적으로 보안 교육을 실시하며, 보안 교육 시 보호대상 데이터의 변경상황 및 일일점검 결과 반복되는 데이터 보안 문제 등에 대한 전파를 실시한다.

▶ 감사

- 정보보호 운영계획 및 관련법규에 부합하는지 여부를 확인하기 위해 감사를 위한 조직을 구성하고, 감사를 실시하며, 감사 결과 부합하지 않은 문제가 발견될 경우 원인을 분석한다.

그림 72 감사



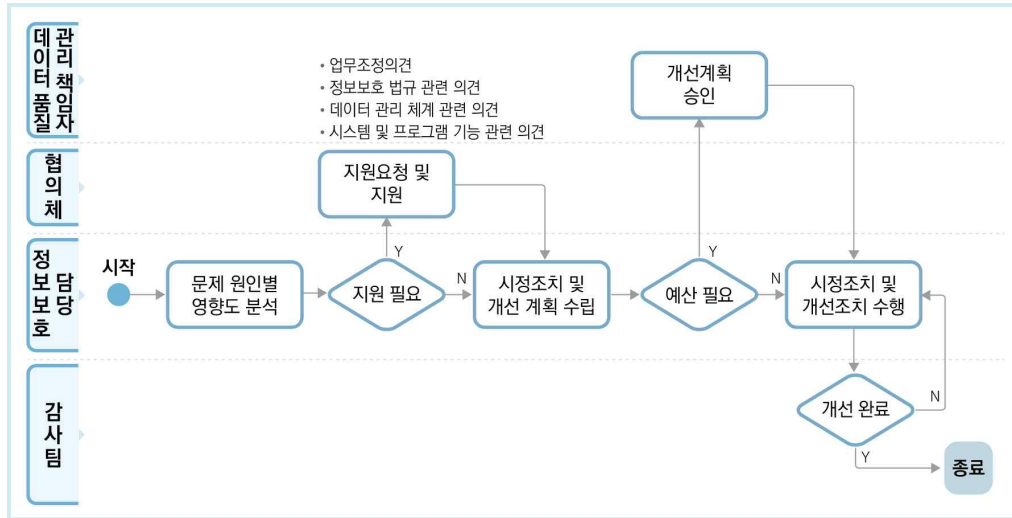
- [감사팀 구성] 감사를 위한 팀을 구성한다. 감사팀 구성 방안 및 감사계획은 및 정보보호 운영관리계획에 포함시켜 수립한다.
- [감사 실시 공지 및 안내] 감사팀이 구성되면 세부 감사 일정을 수립한 후 이를 관련 조직에 안내하며, 관련 감사계획은 정보보호 운영관리계획에 포함시켜 감사팀의 활동을 효율적으로 수행할 수 있도록 한다.
- [감사자료 수집] 감사팀은 관련 조직으로부터 감사에 필요한 자료를 수집하며, 필요시 시스템 로그 데이터를 수집한다.
- [자료보완 필요] 수집한 자료가 감사목적 달성에 충분한지를 검토한 후, 미흡할 경우 추가로 자료를 수집하며, 감사계획에 포함되지 않은 자료라 하더라도 감사목적 달성을 위해 필요한 자료는 관련 조직에 요청하여 수집한다.
- [문제 발견] 자료를 바탕으로 감사를 수행한 결과 문제가 발견될 경우 정보보호 담당 및 관련조직과 감사결과를 협의하며, 문제가 없는 경우 감사결과 보고를 한다.
- [감사결과 협의] 감사결과에 대해 감사팀과 정보보호 담당 및 관련 조직과 협의를 한다. 감사팀은 문제와 근거 자료를 함께 제시한다.
- [동의] 감사팀의 문제제기에 대해 정보보호 담당 및 관련조직이 동의하는지 확인하고, 동의하지 않을 경우 추가자료 등을 바탕으로 감사결과에 대해 추가 협의를 한다.
- [원인분석] 감사결과 문제가 있다고 결정된 경우, 감사팀이 함께 참여하여 원인파악 및 분석을 효율적으로 수행한다.
- [감사결과보고] 감사결과를 플랫폼 및 센터 의사결정 체계에 따라 보고하며, 보고 시 감사

지적사항이 있는 경우 원인분석 결과와 함께 보고한다.

▶ 시정조치 및 개선

- 감사 결과 문제점이 있는 경우 영향도를 분석하여, 필요 시 관련 조직의 도움을 받아 개선계획을 수립하고, 예산이 필요할 경우 의사결정권자의 승인을 얻어서, 문제가 해결될 때까지 시정조치 및 개선 활동을 반복적으로 수행한다.

그림 73 시정조치 및 개선



- [문제 원인별 영향도 분석] 감사결과 발견된 문제에 따른, 영향 및 범위는 시정조치 및 개선조치에 대한 업무량을 결정하므로, 문제가 미치는 영향 및 해결을 위해 조치 범위 등을 분석을 필수적으로 수행한다.
- [지원필요] 정보화담당은 문제 원인별 영향도 분석 결과 해결을 위한 조치를 독자적으로 수행할 수 있는지를 판단하며, 독자적 해결 또는 해결방안 도출이 어려울 경우 관련조직의 지원을 요청 한다.
- [지원요청 및 지원] 정보화담당은 독자적으로 해결 또는 해결방안 도출이 어려울 경우 관련조직에 지원을 요청하며, 지원요청 시 관련 조직은 업무조정 의견, 법규관련 의견, 데이터관리체계관련 의견, 시스템 및 프로그램 기능 관련 의견을 제시한다.
- [시정조치 및 개선계획 수립] 정보화 담당은 문제 원인별 영향도 분석결과를 바탕으로 개선계획을 수립하며, 필요시 관련조직의 의견을 반영한다.
- [예산 필요] 시정 및 개선에 예산이 필요한지를 파악한 후, 새로운 시스템 및 프로그램 기능 추가 등의 사유로 예산이 필요할 경우 개선계획에 반영하고 이의 승인을 요청한다.
- [개선계획 승인] 예산이 필요한 시정 및 개선 조치에 대하여는, 플랫폼 및 센터의 의사결정체계 승인 절차에 따라, 데이터 품질관리 책임자의 승인을 얻는다.
- [시정조치 및 개선조치 수행] 개선계획에 따라 시정 및 개선조치를 수행하며, 정보화 담당이

독자적으로 수행할 수 없는 경우 관련 조직과 협력하여 시정 및 개선 조치를 수행한다.

- [개선완료] 감사팀은 시정조치 및 개선조치 결과를 확인하며, 시정 및 개선 조치가 미흡할 경우 반복적으로 시정 및 개선 조치를 수행한다.

3.4.3 정보보호 기술관리

▶ 기술관리계획 수립

- 정보보호 총괄계획인 운영계획의 기술관리 방향을 참조하여 기술관리계획을 수립하고 필요시 기술환경을 구축하고 접근제어 및 암호화 테스트를 수행한다.

그림 74 기술관리계획 수립(1/2)

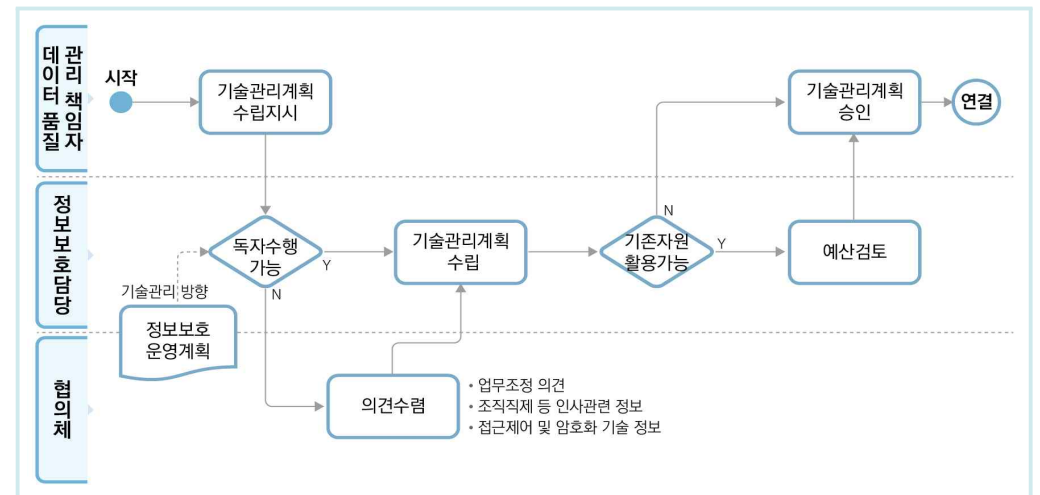
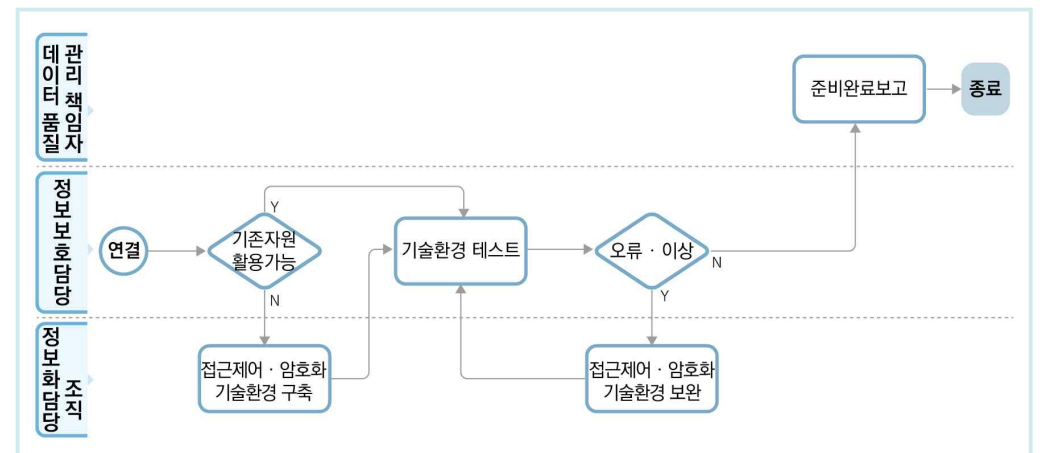


그림 75 기술관리계획 수립(2/2)



- [기술관리계획 수립 지시] 플랫폼 및 센터는 의사결정 절차나 규정에 따라 데이터관리총괄책임자,

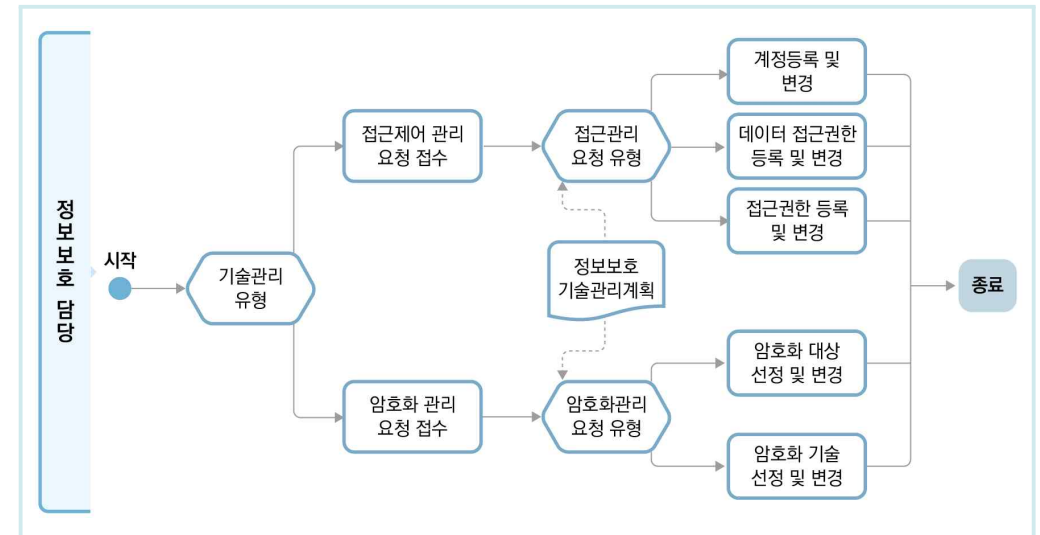
의사결정위원회 또는 데이터 품질관리 책임자는 정보보호담당에게 정보보호 기술관리계획을 수립할 것을 지시하며, 플랫폼 및 센터에 관련한 의사결정 절차나 규정이 없는 경우 데이터 품질관리 책임자가 계획수립을 지시한다.

- [독자수행가능] 정보보호담당이 독자적으로 기술관리계획을 수립할 수 없을 경우 관련 담당자 및 협의체에 지원을 요청하며, 지원요청 시 정보보호 기술관리 필요성을 참조하여 관리대상 기술, 목적, 등과 같은 기본적인 사항을 전달한다.
- [의견수렴] 정보보호담당의 요청을 받은 관련조직은 업무조정 의견, 조직직제, 인사 관련 정보, 접근제어 및 암호화 관련 기술정보 등을 제공한다.
- [기술관리계획수립] 기술관리계획은 접근제어 및 암호화 대상 및 범위, 기술관리 수행조직, 기술적용 타당성 검토결과, 추진 일정 및 소요예산에 대해 상세히 작성한다.
- [기존자원활용가능] 접근제어 및 암호화를 위해 현재 사용하는 기술환경에 추가, 변경 등이 필요한지를 확인하며, 새로운 프로그램 개발 및 기존 프로그램 수정, 새로운 기기 도입 및 증설 등이 필요한지를 검토한다.
- [예산검토] 기존의 기술환경에 변화가 필요할 경우, 연계를 위한 예산 타당성을 검토한다.
- [기술관리계획 승인] 플랫폼 및 센터 의사결정 규정 및 절차에 따라 데이터품질관리 책임자, 의사결정위원회, 데이터 품질관리 책임자, 데이터관리 총괄 책임자 및 최고책임자는 관련 예산을 포함한 기술관리계획을 승인한다.
- [접근제어·암호화 기술환경 구축] 기술환경에 추가 및 변경이 필요할 경우 연계프로그램의 추가 및 수정, 기술환경 구축 및 변경 등을 통하여 연계를 수행할 수 있도록 기능을 개발하고 기술환경을 구축한다.
- [기술환경 테스트] 접근제어 및 암호화 기능 개발 및 기술환경 구축이 완료되면 테스트를 수행한다.
- [오류·이상] 테스트 결과 오류 또는 이상여부를 확인하여 이상이 있는 경우 보완조치를 실시한다.
- [접근제어·암호화 기술환경 보안] 테스트 결과 이상 및 오류가 있는 경우 그 원인을 파악하여 프로그램 기능 및 기술환경을 개선한다.
- [준비완료보고] 플랫폼 및 센터 의사결정 규정 및 절차에 따라 데이터품질관리 책임자, 의사결정위원회, 데이터품질관리 책임자, 데이터관리 총괄 책임자 및 최고책임자는 연계 관련 예산을 포함한 연계계획을 승인한다.

▶ 접근제어 및 암호화 실행

- 접근제어 및 암호화 관련 관리 요청이 접수되면 정보보호 기술관리계획에 의해 접근관리 및 암호화 관리 활동을 수행한다.

그림 76 접근제어 및 암호화 실행

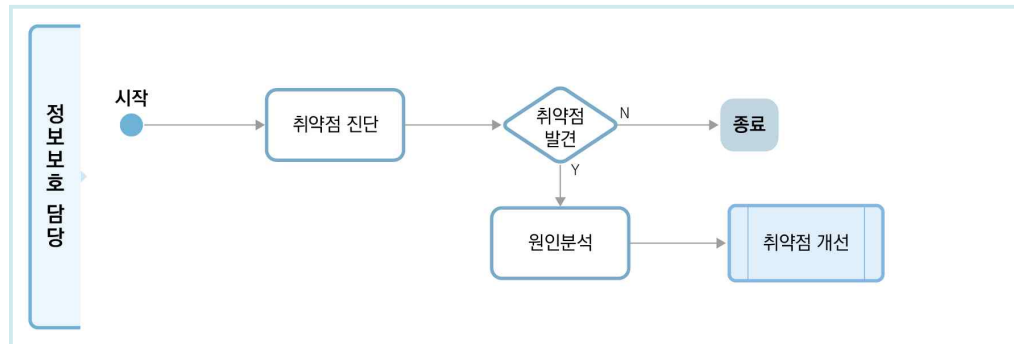


- [기술관리 유형] 기술관리 유형인 접근제어와 암호화에 따라 분기한다.
- [접근제어 관리 요청 접수] 데이터 접근제어에 대한 등록 및 변경 요청을 접수한다.
- [접근관리 요청 유형] 데이터 접근관리 요청 유형에 따라 분기한다.
- [계정등록 및 변경] 계정 등록 및 변경 요청 시 기술관리계획을 참조하여 이를 수행한다.
- [데이터 접근권한 등록 및 변경] 데이터의 특성에 따라 접근정책을 정의한다. 접근정책은 정보보호 기술관리계획에 따라 정의한다.
- [접근권한 등록 및 변경] 계정별로 데이터 접근 권한을 설정하고, 인사이동 및 데이터 보호 정책 변경 시 변경 등록한다.
- [암호화 관리 요청 접수] 암호화 대상 데이터에 대한 적용 및 변경 요청을 접수하며, 암호화 대상은 데이터 보호정책에 따라 결정한다.
- [암호화 관리 요청 유형] 암호화 관리 요청 유형에 따라 분기한다.
- [암호화 대상 선정 및 변경] 암호화 대상 데이터를 결정하여 적용하고, 변경사유 발생시 변경하여 적용한다.
- [암호화 기술 선정 및 변경] 최적의 암호화 대상 기술 및 프로그램을 선정하여 적용하고, 필요시 변경하여 적용한다.

▶ 취약점 점검

- 접근제어 및 암호화 취약점을 점검하고 취약점 발견 시 그 원인을 분석한다.

그림 77 취약점 점검

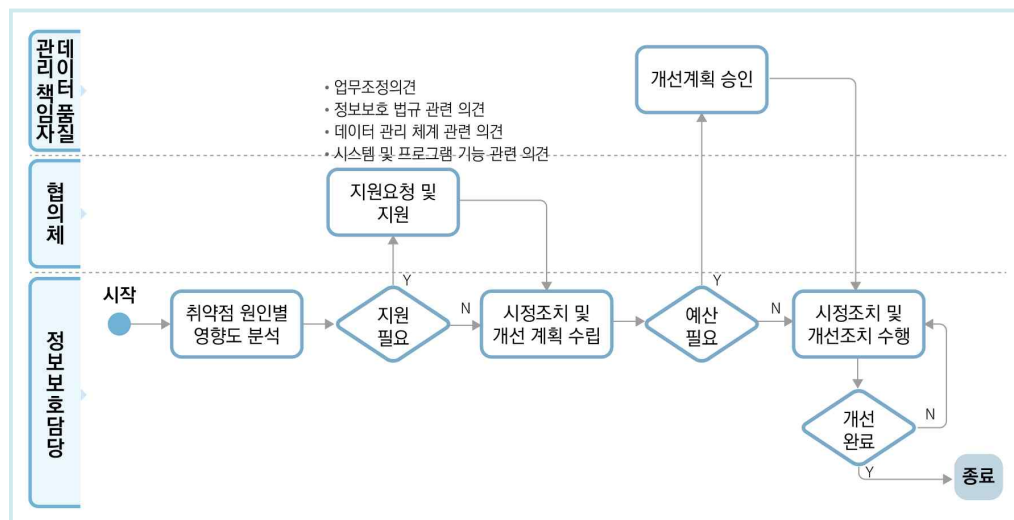


- [취약점 진단] 접근제어 및 암호화 등의 기술관리 측면의 취약점을 진단하며, 관련 취약점 진단을 위한 계획은 정보보호 기술관리 계획에 포함시켜 수립함으로써 효율적인 진단이 되도록 한다.
- [취약점 발견] 접근제어 및 암호화 취약점 존재 여부를 확인하며, 취약점이 발견되면 원인 분석을 실시한다.
- [원인분석] 접근제어 및 암호화에 취약점이 발견될 경우 그 원인을 분석하며, 원인분석에는 필요시 정보화담당 조직이 함께 참여하여 원인파악 및 분석을 효율적으로 수행한다.
- [취약점 개선] 취약점 발견과 원인분석이 완료되면 취약점 개선 프로세스를 수행한다.

▶ 취약점 개선

- 취약점 점검결과 문제점이 있는 경우 영향도를 분석하여, 필요시 관련 조직의 도움을 받아 개선계획을 수립하고, 예산이 필요할 경우 의사결정권자의 승인을 얻어서, 문제가 해결될 때까지 시정조치 및 개선 활동을 반복적으로 수행한다.

그림 78 취약점 개선



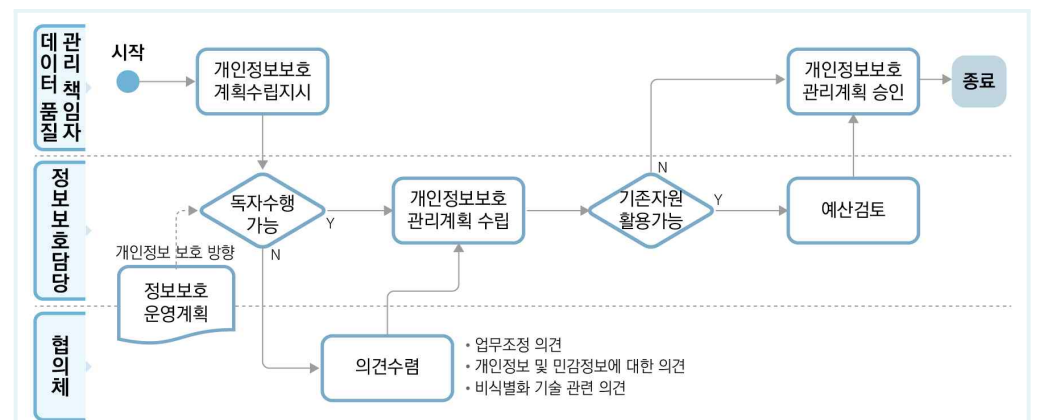
- [취약점 원인별 영향도 분석] 취약점 점검 결과 발견된 문제에 따른, 영향 및 범위는 개선조치에 대한 업무량을 결정하므로, 문제가 미치는 영향 및 해결을 위해 조치 범위 등을 분석을 필수적으로 수행한다.
- [지원필요] 정보화담당은 해결을 위한 조치를 독자적으로 수행할 수 있는지를 판단하며, 독자적 해결 또는 해결방안 도출이 어려울 경우 관련조직의 지원을 요청 한다.
- [지원요청 및 지원] 정보화담당은 독자적으로 해결 또는 해결방안 도출이 어려울 경우 관련조직에 지원을 요청하며, 관련조직은 업무조정 의견, 법규관련 의견, 데이터관리체계관련 의견, 시스템 및 프로그램 기능 관련 의견을 제시한다.
- [시정조치 및 개선계획 수립] 정보화 담당은 문제 원인별 영향도 분석결과를 바탕으로 개선계획을 수립하며, 필요시 관련조직의 의견을 반영한다.
- [예산필요] 시정 및 개선에 예산이 필요한지를 파악하며, 새로운 시스템 및 프로그램 기능 추가 등의 사유로 예산이 필요할 경우 개선계획에 반영하고 이의 승인을 요청한다.
- [개선계획 승인] 예산이 필요한 시정 및 개선 조치에 대하여는, 플랫폼 및 센터의 의사결정체계 승인 절차에 따라, 데이터 품질관리 책임자의 승인을 얻는다.
- [시정조치 및 개선조치 수행] 개선계획에 따라 시정 및 개선조치를 수행하며, 정보화 담당이 독자적으로 수행할 수 없는 경우 관련 조직과 협력하여 시정 및 개선 조치를 수행한다.
- [개선완료] 시정조치 및 개선조치 결과를 확인하며, 시정 및 개선 조치가 미흡할 경우 취약점이 해결될 때까지 반복적으로 시정 및 개선 조치를 수행한다.

3.4.4 개인정보보호 관리

▶ 개인정보보호 관리계획 수립

- 정보보호 총괄계획인 운영계획의 개인정보보호 관리 방향을 참조하여 개인정보보호계획을 수립하고 필요시 예산을 확보하는 등 계획을 승인받는다.

그림 79 개인정보보호 관리계획 수립

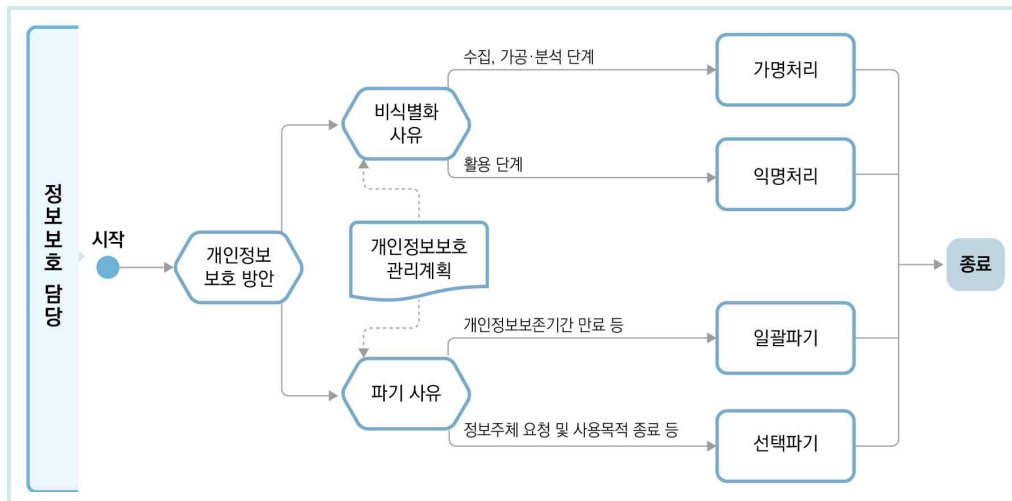


- [개인정보보호 계획 수립 지시] 플랫폼 및 센터는 의사결정 절차나 규정에 따라 데이터관리총괄책임자, 의사결정위원회 또는 데이터 품질관리 책임자는 정보보호담당에게 개인정보보호 관리계획을 수립할 것을 지시하며, 플랫폼 및 센터에 관련한 의사결정 절차나 규정이 없는 경우 데이터 품질관리 책임자가 계획수립을 지시한다.
- [독자수행가능] 정보보호담당이 독자적으로 개인정보보호 관리계획을 수립할 수 없을 경우 관련 담당자 및 협의체에 지원을 요청하며, 지원요청 시 관련 담당자 및 협의체는 개인정보 관련법규, 비식별화 조치 목적, 등과 같은 기본적인 사항을 전달한다.
- [의견수렴] 정보보호담당의 요청을 받은 관련조직은 업무조정의견, 개인정보 및 민감정보에 대한 의견, 비식별화 조치 관련 기술정보 등을 제공한다.
- [개인정보보호 관리계획 수립] 개인정보보호 관리계획은 비식별화 조치 대상 및 범위, 수행조직, 적용기술 및 타당성, 추진 일정 및 소요예산에 대해 상세히 작성한다.
- [기존자원 활용가능] 비식별화 및 파기를 위해 현재 사용하는 기술환경에 추가, 변경 등이 필요한지를 확인하며, 새로운 프로그램 개발 및 기존 프로그램 수정, 새로운 기기 도입 및 증설 등이 필요한지를 검토한다.
- [예산검토] 기존의 기술환경에 변화가 필요할 경우, 연계를 위한 예산 타당성을 검토한다.
- [개인정보 관리계획 승인] 플랫폼 및 센터 의사결정 규정 및 절차에 따라 데이터품질관리 책임자, 의사결정위원회, 데이터품질관리 책임자, 데이터관리 총괄 책임자 및 최고책임자는 관련 예산을 포함한 개인정보보호 관리계획을 승인한다.

▶ 비식별화 및 파기 실시

- 개인정보보호를 위해 데이터 생애주기 단계별로 가명처리 및 익명처리를 실시하고, 개인정보보존기간 만료 등의 사유가 발생할 경우 개인정보보호를 위해 파기를 수행한다.

그림 80 비식별화 및 파기 실시

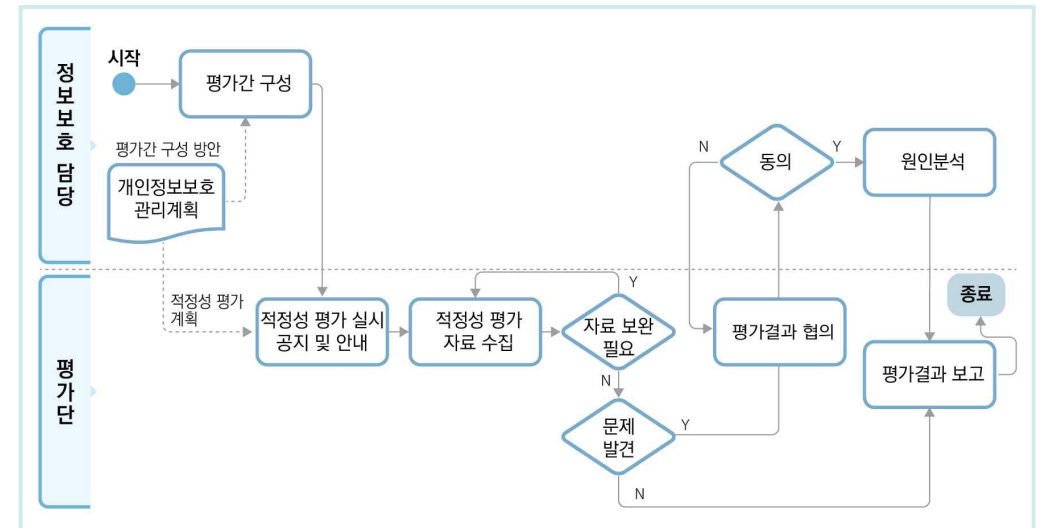


- [개인정보보호 방안] 개인정보 보호 방안인 비식별화와 개인정보 파기로 분기한다.
- [비식별화 사유] 데이터 생애주기 단계별로 비식별화 처리를 분기한다.
- [가명처리] 수집, 가공·분석 단계는 개인정보 비식별화 가이드라인에 따라 가명처리 한다.
- [익명처리] 활용 단계는 개인정보 비식별화 가이드라인에 따라 재식별이 불가능하도록 익명처리 한다.
- [파기사유] 파기 사유별로 파기처리를 분기한다.
- [일괄파기] 개인정보 보존기간 만료 등의 사유가 발생할 경우, 개인정보를 일괄 파기 한다.
- [선택파기] 개인정보 비식별화 오류, 개인정보 주체의 요청, 개인정보 사용목적 종료 등의 사유가 발생할 경우, 해당 개인정보를 선별하여 파기한다.

▶ 적정성 평가

- 개인정보보호 관리계획 및 관련법규에 부합하는지 여부를 확인하기 위해 적정성 평가를 위한 조직을 구성하고, 평가를 실시하며, 평가 결과 부합하지 않은 문제가 발견될 경우 원인을 분석한다.

그림 81 적정성 평가



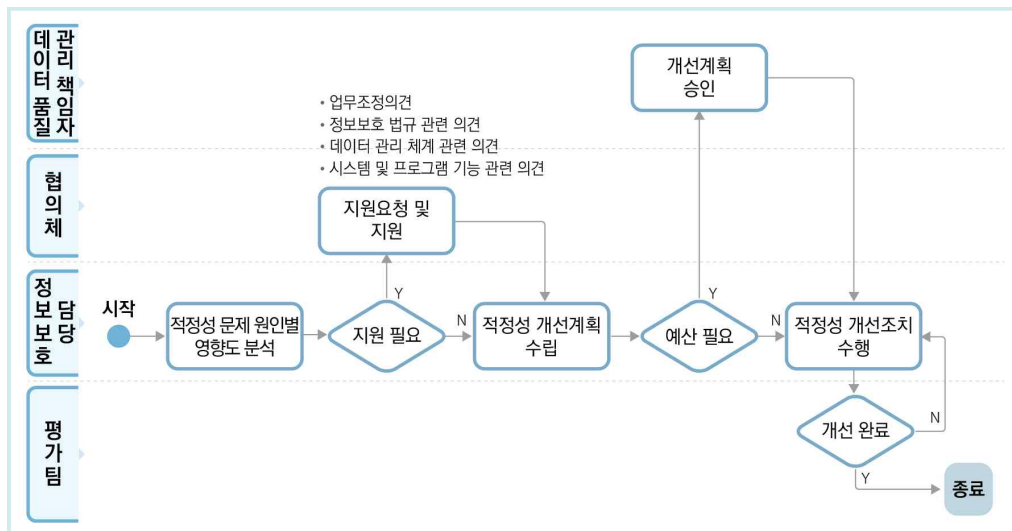
- [평가단 구성] 적정성 평가를 위한 팀을 구성하며, 평가단 구성 방안 및 적정성 평가계획은 및 개인정보보호 관리계획에 포함시켜 수립한다.
- [적정성 평가 실시 공지 및 안내] 평가팀이 구성되면 세부 평가 일정을 수립한 후 이를 관련 조직에 안내하며, 관련 적정성 평가계획은 개인정보보호 관리계획에 포함시켜 적정성 평가팀의 활동을 효율적으로 수행할 수 있도록 한다.
- [적정성 평가 자료 수집] 평가팀은 관련 조직으로부터 적정성 평가에 필요한 자료를 수집한다.

- [자료보완 필요] 수집한 자료가 감사목적 달성에 충분한지를 검토한 후, 미흡할 경우 추가로 자료를 수집하며, 감사계획에 포함되지 않은 자료라 하더라도 감사목적 달성을 위해 필요한 자료는 관련 조직에 요청하여 수집한다.
- [문제 발견] 자료를 바탕으로 적정성 평가를 수행한 결과 문제가 발견될 경우 정보보호 담당 및 관련조직과 평가결과를 협의하며, 문제가 없는 경우 적정성 평가 결과 보고를 한다.
- [평가결과 협의] 적정성 평가 결과에 대해 평가팀과 정보보호 담당 및 관련 조직과 협의를 하며, 적정성 평가팀은 문제와 근거 자료를 함께 제시한다.
- [동의] 평가팀의 문제제기에 대해 정보보호 담당 및 관련조직이 동의하는지 확인한 후, 동의하지 않을 경우 추가자료 등을 바탕으로 적정성 평가 결과에 대해 추가 협의를 한다.
- [원인분석] 평가결과 문제가 있다고 결정된 경우 그 원인을 분석하는데, 원인분석 수행 시 평가팀이 함께 참여하여 원인파악 및 분석을 효율적으로 수행한다.
- [평가결과보고] 평가결과를 플랫폼 및 센터 의사결정 체계에 따라 보고하며, 보고 시 적정성에 문제가 있는 경우 원인분석 결과와 함께 보고한다.

▶ 적정성 개선

- 개인정보보호 관리계획 및 관련법규에 부합하는지 여부를 확인하기 위해 적정성 평가를 위한 조직을 구성하고, 평가를 실시하며, 평가 결과 부합하지 않은 문제가 발견될 경우 원인을 분석한다.

그림 82 적정성 개선



- [문제 원인별 영향도 분석] 적정성 평가 결과 발견된 문제에 따른, 영향 및 범위는 개선조치에 대한 업무량을 결정하므로, 문제가 미치는 영향 및 해결을 위해 조치 범위 등을 분석을 필수적으로 수행한다.

- [지원필요] 정보화담당은 문제 원인별 영향도 분석 결과 해결을 위한 조치를 독자적으로 수행할 수 있는지를 판단하며, 독자적 해결 또는 해결방안 도출이 어려울 경우 관련조직의 지원을 요청한다.
- [지원요청 및 지원] 정보화담당은 독자적으로 해결 또는 해결방안 도출이 어려울 경우 관련조직에 지원을 요청하며, 지원요청 시 관련조직은 업무조정 의견, 법규관련 의견, 데이터관리체계관련 의견, 시스템 및 프로그램 기능 관련 의견을 제시한다.
- [적정성 개선계획 수립] 정보화 담당은 문제 원인별 영향도 분석결과를 바탕으로 개인정보보호 관련 적정성 개선계획을 수립하며, 필요시 관련조직의 의견을 반영한다.
- [예산필요] 시정 및 개선에 예산이 필요한지를 파악하며, 새로운 시스템 및 프로그램 기능 추가 등의 사유로 예산이 필요할 경우 개선계획에 반영하고 이의 승인을 요청한다.
- [개선계획 승인] 예산이 필요한 시정 및 개선 조치에 대하여는, 플랫폼 및 센터의 의사결정체계 승인 절차에 따라, 데이터 품질관리 책임자의 승인을 얻는다.
- [시정조치 및 개선조치 수행] 개선계획에 따라 시정 및 개선조치를 수행하며, 정보화 담당이 독자적으로 수행할 수 없는 경우 관련 조직과 협력하여 시정 및 개선 조치를 수행한다.
- [개선완료] 평가팀은 시정조치 및 개선조치 결과를 확인하며, 시정 및 개선 조치가 미흡할 경우 반복적으로 시정 및 개선 조치를 수행한다.

빅데이터 플랫폼 및 센터 데이터 품질관리 가이드

발 행 2021년 12월

발 행 인 문용식

기획/작성 한국지능정보사회진흥원 빅데이터플랫폼팀
이용진 단장, 손기문 팀장, 유시형 수석, 김은영 연구원, ㈜씨에이에스

문 의 ysh@nia.or.kr [한국지능정보사회진흥원 빅데이터플랫폼팀 유시형 수석]

주 소 [41068] 대구광역시 동구 첨단로 53

홈페이지 www.nia.or.kr

- * 본 가이드의 내용은 한국지능정보사회진흥원의 공식견해와 다를 수 있습니다.
- * 본 가이드 내용에 대한 무단 전제를 금하여, 가공 및 인용할 때에는 반드시 출처를 명기하여 주시기 바랍니다.
- * 본 책자는 정보통신기금으로 제작되었습니다.(비매품)