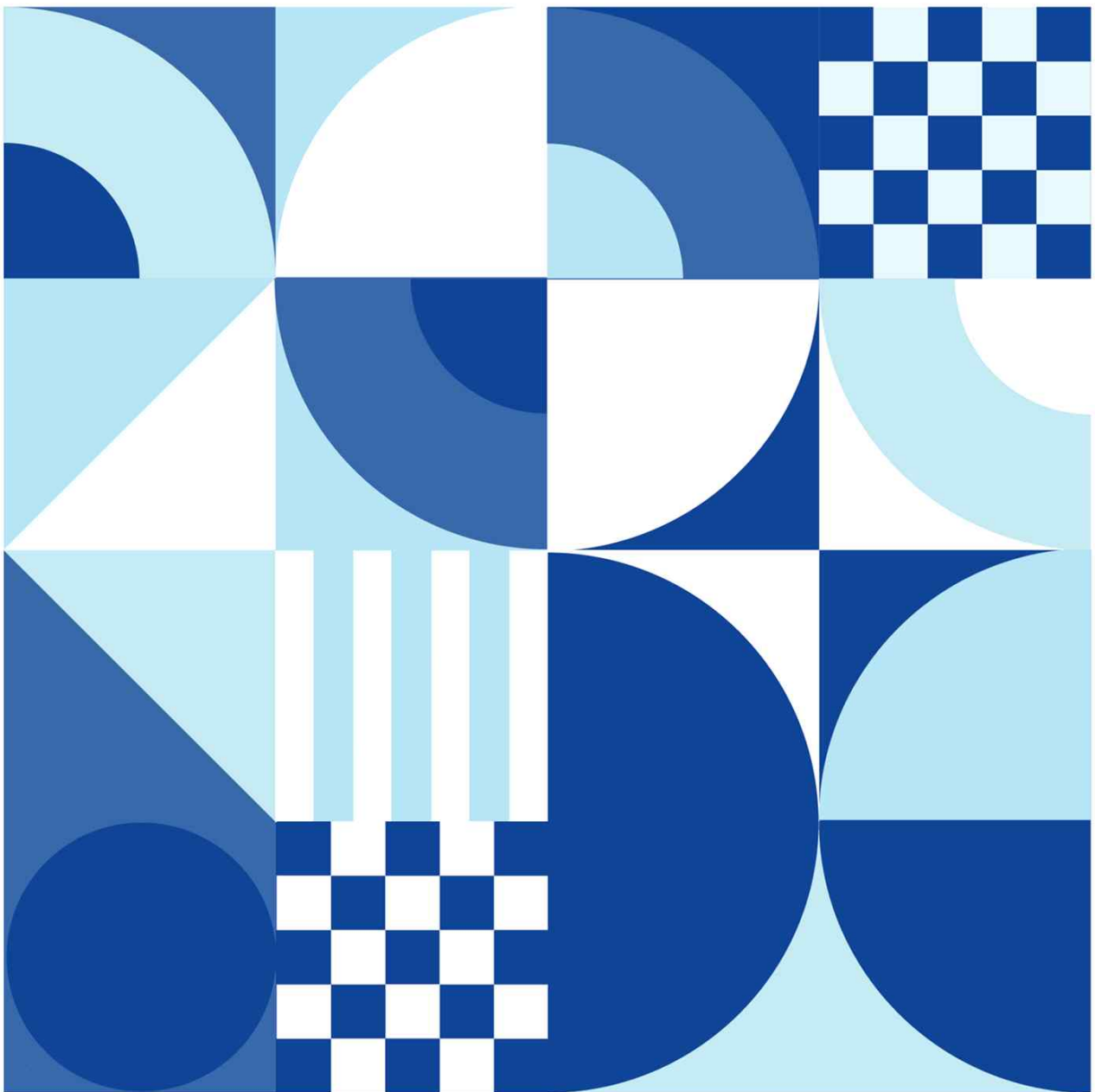




신뢰 기반 데이터 유통 거래 활성화를 위한

한국형 데이터 스페이스

참조모델 v.1.0



과학기술정보통신부

NIA 한국지능정보사회진흥원

신뢰 기반 데이터 유통 거래 활성화를 위한 한국형 데이터 스페이스 참조모델 v.1.0

『한국형 데이터 스페이스 참조모델』은

데이터 제공자와 이용자가 신뢰 기반의 규칙과 표준 하에서
데이터를 안전하게 공유·활용할 수 있도록 거버넌스, 기술, 서비스
요소를 구조화한 개념적 기준 모델입니다.

- ① 본 참조모델에서 언급하는 공공데이터 및 공공저작물의 활용은 「공공저작물 자유이용 허락표시 기준(공공누리)」에 따라 허용된 범위 내에서 이루어져야 하며, 공공누리 제0유형 및 AI 유형이 표시된 저작물에 한하여 인공지능 학습 목적의 자유로운 활용이 가능함

※ 실제 데이터 제공 및 활용 시에는 각 데이터의 개별 이용 조건, 관련 법령 및 소관 기관의 세부 지침 준수

- ② 보고서 내용의 무단전재를 금하며, 가공·인용할 때는 반드시 출처를 「한국지능정보원(NIA)」이라고 밝혀 주시기 바랍니다.
- ③ 본 보고서의 내용은 한국지능정보원(NIA)의 공식 견해와 다를 수 있습니다.



개정 이력

개정 이력			
연번	버전	발간일	주요 변경 내용
1	v1.0	'26.3	o 최초 발간 - 기준 제시

목 차

I. 개요

1. 추진 배경

1.1 국내 데이터 유통·활용 생태계 주요 병목 현상	1
1.2 국내 데이터 유통·활용 생태계 정책 및 법제도 주요 동향	3
1.3 데이터 스페이스 조성을 위한 국제 사회 움직임	6
1.4 데이터 스페이스 도입 필요성	9
2. 한국형 데이터 스페이스 추진 전략	10

II. 데이터 스페이스 주요 개념

III. 한국형 데이터 스페이스 참조모델

1. 참조모델 수립 목적	12
2. 참조모델 세부 구성	
2.1 한국형 데이터 스페이스 구성 요소	13
2.2 한국형 데이터 스페이스 설계 핵심 원칙	15
2.3 한국형 데이터 스페이스 거버넌스	16
2.3.1 참여 주체 및 역할 체계	16
2.3.2 공통 정책·규칙 체계	20
2.3.3 지속 가능한 운영 체계	23
2.4 한국형 데이터 아키텍처 모델	26
2.4.1 한국형 데이터 스페이스 아키텍처	26
2.4.2 데이터 스페이스 표준 흐름	27
2.4.3 데이터 스페이스 공통 서비스	44

I 개 요

1. 추진 배경

1.1 국내 데이터 유통·활용 생태계 주요 병목 현상

■ 기존 데이터 유통·거래 제도의 체감 효과 부족

- ▶ 국내는 「데이터산업진흥 및 이용촉진에 관한 기본법」 제정과 함께 표준계약서, 데이터 가치평가, 라이선스 유형 등 제도적 기반을 갖췄으나 실제 데이터 거래·활용 현장에서 충분히 확산되지 못하고 있음

- ◆ 현재 국내 데이터 거래 마켓플레이스·플랫폼에서는 거래 단계에서 실제 활용량이나 방식에 따라 정산·가격이 책정되는 체계가 부재하여, 공급자에게는 데이터 제공에 대한 명확한 보상 구조를, 수요자에게는 합리적인 비용 부담 구조를 제공하지 못하는 한계가 있음
- ◆ 표준계약서 제도는 존재하나, 계약 이행 과정에서 이를 모니터링·추적하거나 목적 외 사용을 예방·조정할 수 있는 운영체계나 합의기구가 부재하여, 실제 분쟁 발생 시에는 사후적인 법적 대응이나 행정적 제재에 의존할 수밖에 없음

■ 중앙집중형 플랫폼의 구조적 제약

- ▶ 현재 21개 분야 빅데이터 플랫폼은 데이터를 단일 저장소에 데이터를 집적하는 중앙집중형 방식으로 운영되어, 데이터 제공자의 통제권 상실 및 수요자의 자율적 참여 제약
- ▶ 산업/분야별 참여 주체 간 자율적인 교차 산업 분석이나 데이터 융합에 제약이 발생하며 공급자 중심의 데이터 제공 구조에 따라 수요가 높은 고품질 데이터 생산·유통이 어려움

- ◆ 국내 데이터 플랫폼들은 공통표준용어가 제정·공표되었음에도 불구하고 현장 활용이 미흡하여 기관별 데이터 형식과 용어가 제각각이고, 품질관리 수준도 달라 정합성 확보가 어려움¹⁾
- ◆ 제공자 입장에서는 한번 데이터를 제공하면 이후 데이터 활용과정에서 권리와 통제권을 상실할 우려로 소극적인 입장(NIA, 2024)

1) https://online.kofst.or.kr/news/316787?category=COM045_7DaOeSe

■ 데이터 협업·비즈니스 모델 창출의 한계

- ▶▶ 국내 데이터 마켓플레이스와 플랫폼은 데이터 유통·거래 시 수요자-공급자 협업 메커니즘이 부재하여 공급자가 직접 수요자를 찾아 참여시켜야 하는 구조로, 지속적인 데이터 기반 비즈니스 모델 창출로 이어지기 어려움
- ▶▶ 특히 고품질의 AI 학습용 데이터와 같은 고부가가치 데이터는 저작권·개인정보보호법 등 현행 법·제도의 불명확성과 규제 부담으로 인해 제공·활용이 제한되면서, 실제 수요가 가장 큰 영역임에도 데이터 시장에 원활하게 유통되지 못하고 있음

◆ 다수의 데이터 플랫폼·마켓플레이스는 원천 데이터를 제공하거나 이를 가공·결합한 결과물(리포트) 등 데이터의 가공·판매에 초점을 맞춘 기능을 중심으로 운영되고 있음
◆ 대부분의 플랫폼은 제공자→이용자의 단방향적 구조로 운영되고 있으며, 새로운 서비스나 비즈니스 모델을 창출하는 프로젝트형 협업 사례는 거의 부재

[국내 데이터 플랫폼 개선 방향]

구분	AS-IS	TO-BE
	빅데이터 플랫폼 (금융, 유통 등 분야별 플랫폼)	데이터 스페이스
주요 기능	공공·민간 데이터를 수집·집적하여 가공·결합·분석을 통한 서비스 제공	신뢰 기반 데이터 공유, 협업을 통한 혁신 서비스 창출
데이터 저장 및 관리 방식	각 기관·기업이 데이터를 중앙 플랫폼으로 이전, 집적화	기관별 자체 보유·관리, 분산·연합형 접근 제어
운영 방식	플랫폼 운영기관 주도 계약 및 거래 관리	공급자·수요자·운영기관 공동의 합의 기반 운영
수요 창출	공급자가 수요자 발굴	공급자·수요자 공동 참여, 비즈니스 모델 및 협업 기획 발굴
수익 분배·정산	정액제·일회성 계약 중심, 거래 후 활용량·조건 반영 어려움	스마트 컨트랙트·클리어링하우스 기반 자동 정산 및 수익 분배

1.2 국내 데이터 유통·활용 생태계 정책 및 법제도 주요 동향

■ 신뢰 기반 데이터 생태계 기반을 위한 법제도 수립

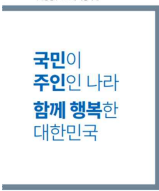
- » 한국은 디지털 전환과 데이터 경제 활성화를 국가 전략 과제로 설정하고, 2022년 4월 데이터산업법을 제정하여 과학기술정보통신부를 중심으로 데이터의 생산·유통·거래·활용 전반을 체계적으로 육성하기 위한 법적 기반을 마련함

 - ※ 기존의 데이터 이용 촉진 중심 정책에서 나아가 데이터 거래와 산업 육성을 명시적으로 포괄하며, 데이터 제공자와 이용자 간 권리·책임 구조, 신뢰 확보, 보호·관리 체계를 핵심 원칙으로 제시
- » 이를 토대로 정부는 「데이터 산업진흥 기본계획(’23~’25)」을 수립해 데이터 산업을 전주기 관점에서 정의하고, 데이터 유통·거래 과정에서의 투명성, 이력 관리, 안전한 거래 환경 조성 등 신뢰 기반 데이터 유통·거래를 위한 정책 과제 구체화
- » 아울러 데이터 거래 지원 체계, 제도 정비, 기술·인프라 확충 등을 통해 데이터가 신뢰 가능한 방식으로 유통·활용되어 경제적 가치로 이어질 수 있는 기반 마련

■ 新 정부 국정과제와 주요 정책 발표

- » 2025년 출범한 新 정부는 AI를 국가 성장 전략의 핵심 동력으로 설정하고, 이를 뒷받침하기 위한 국가 차원의 공통 실행 기반으로 ‘AI 고속도로’를 국정과제로 제시함. AI 고속도로는 개별 기술이나 단일 사업이 아니라, 데이터·컴퓨팅·서비스가 대규모로 흐를 수 있도록 국가 차원에서 AI 생태계 전반을 공공 인프라 관점에서 지원하겠다는 정책적 비전을 담고 있음

[데이터 스페이스 국정과제 개요]

	<p>[국정목표 2] 세계를 이끄는 혁신경제 [전략-1] AI 3대 강국 도약</p> <p>▶ (20) AI 3대 강국 도약을 위한 AI 고속도로 구축</p> <ul style="list-style-type: none"> • 데이터 구축·개방 국가 전체 공공·민간 AI 학습용 데이터 통합 플랫폼(~’26) 및 분야별(의료·제조 등) 데이터 스페이스 구축, AI 가독성 확보를 위한 통계 메타데이터 구축 등 - 데이터 활용 촉진을 위해 데이터 안심구역 간 연계 및 「데이터산업법」 정비
---	---

» 이와 함께 AI 활용 확산 과정에서 데이터가 기관·기업·산업별로 분산되어 있고, 데이터의 출처·이용 조건·책임 구조가 체계적으로 관리되지 않을 경우 AI 고속도로의 실효성에 한계가 발생할 수 있다는 문제의식에 따라 데이터의 연계·공유·활용체계를 보완하기 위한 대안으로 데이터 스페이스 조성이 국정과제에 포함됨

■ 데이터 스페이스 초기 조성 및 확산 지원을 위한 한국 정부의 움직임

① 과학기술정보통신부

	<p>» 과기정통부는 데이터 자산 보호 중심의 환경으로 발생한 유통 분질을 극복하고자, 분야별 AX 확산을 위한 데이터 스페이스 체계 추진</p> <p>※ 의료 분야를 시작으로 원천 데이터는 각 주체가 보유한 채 필요 시 활용이 가능하도록 하는 데이터 스페이스 모델을 단계적으로 확산 추진</p>
	<p>» 과기정통부는 공공·민간 빅데이터 플랫폼과 데이터 스페이스 간 운영 분질을 해소하기 위해, 범국가 차원의 『국가 데이터 인프라』 구축('24~)</p> <p>» 이를 통해 공공·민간 데이터를 단일 창구에서 탐색·연계하고, 연합 데이터 카탈로그와 공통 컴플라이언스를 기반으로 데이터 스페이스 간 상호운용성과 확산 기반 강화 목표</p>

② 산업통상부

	<p>» 산업통상부는 제조 산업의 AI 전환과 글로벌 규제 대응을 위해 산업 데이터 스페이스를 핵심 수단으로 활용하고, 이를 공급망 전 주기의 데이터 유통·공유·검증·거래를 지원하는 신뢰 기반 인프라로 정의</p> <p>» 아울러 M.AX 얼라이언스를 중심으로 산업 데이터 스페이스 표준모델 개발 및 실증 추진</p>
--	---



- » 산업통상부는 산업 간 상호운용성과 독일 Manufacturing-X와의 연계를 포함한 국제 협력을 통해 신뢰 기반 데이터 교환 구조를 구축하고,
- » EU 디지털 제품 여권(DPP, Digital Product Passport) 등 글로벌 공급망 규제 확산에 따른 데이터 기반 규제 대응

- » 과학기술정보통신부와 산업통상부의 정책 방향은 AI 확산과 글로벌 규제 대응을 개별 사업이나 단일 플랫폼 차원이 아닌, 데이터 활용을 뒷받침하는 국가·산업 단위의 구조적 기반을 선제적으로 구축하려는 공통된 인식을 보여주고 있음
- » 과학기술정보통신부가 분야별 AX 확산을 위한 데이터 스페이스 체계와 국가 데이터 인프라 구축을 통해 공공·민간 데이터의 연계와 활용 기반을 강화하고자 하는 한편, 산업통상부는 산업 데이터 스페이스를 공급망 전 주기의 데이터 유통·검증·거래를 지원하는 신뢰 인프라로 정의하고 국제 협력까지 병행 추진하고 있다. 이는 AI 경쟁력이 개별 기관이나 기업의 역량을 넘어, 데이터의 연계·활용 구조와 신뢰 체계에 의해 좌우된다는 정책적 공감대가 형성되고 있음을 시사함

1.3 「데이터 스페이스」 조성을 위한 국제 사회 움직임

① European Union (EU)

» EU는 「European Data Strategy(2020)」를 통해 공통 유럽 데이터 스페이스(Common European Data Spaces)를 제시하며, 데이터를 중앙에 집적하지 않고 각 주체가 데이터 주권을 유지한 채 신뢰 기반으로 연결·공유하는 연합형 구조를 정책 비전으로 설정하였으며, 이후 보건·제조·에너지·모빌리티 등 산업·도메인별 데이터 스페이스로 구체화하면서 데이터 주권, 상호운용성, 신뢰 확보를 핵심 원칙으로 삼아 데이터 스페이스를 유럽 디지털 전략의 구조적 축으로 발전



[EU Data Union Strategy]

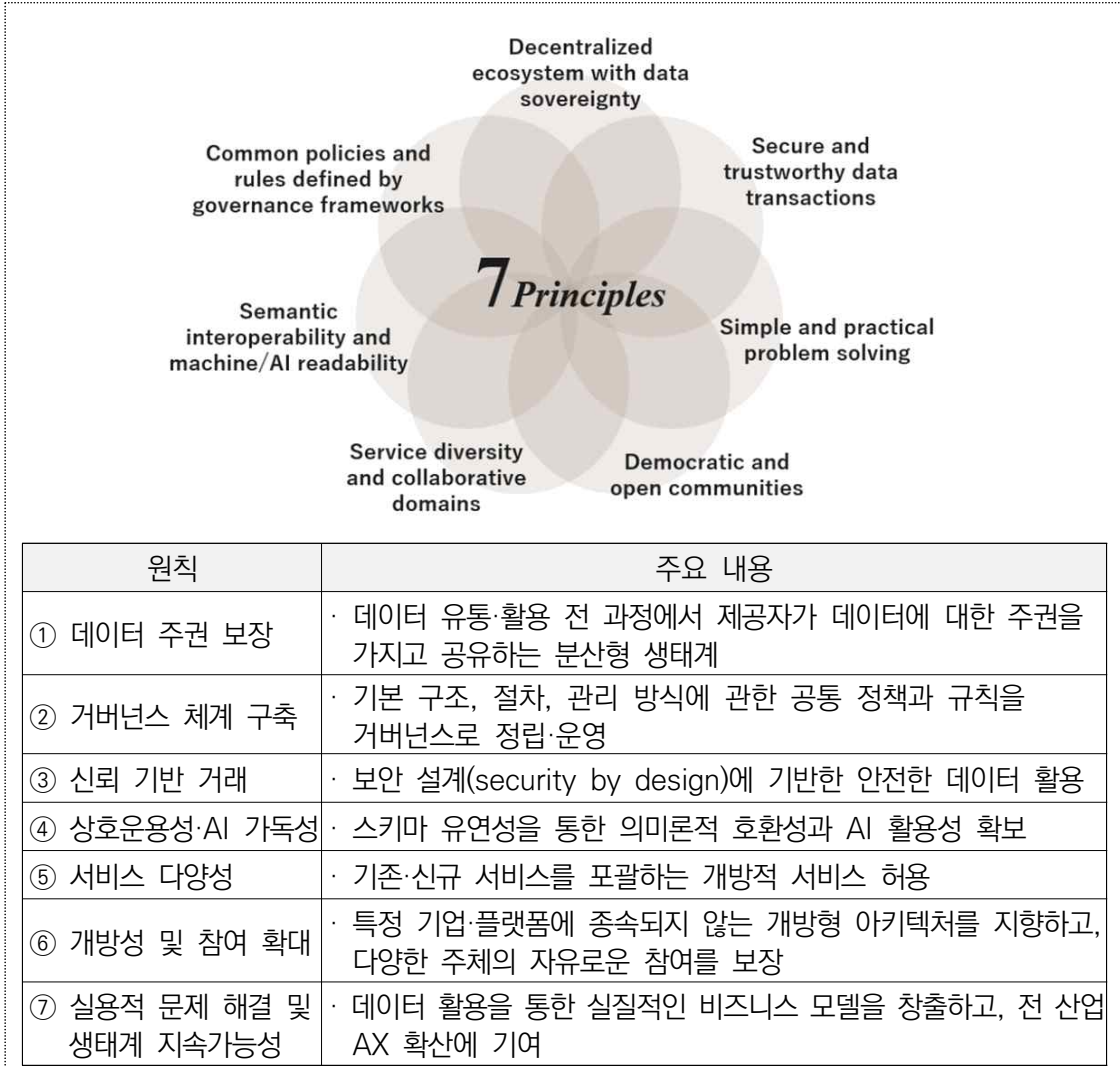
- » 다만 이러한 접근만으로는 시장 확산과 경제적 동력이 충분히 형성되지 않는 한계가 드러나면서, EU는 Data Union 전략²⁾(‘25.11월 발표)을 통해 데이터 스페이스를 EU 단일 데이터 시장 질서로 통합
- » Data Labs 개념을 통해 실제 데이터 활용·AI 실증·서비스 실험이 이루어지는 단계로 정책을 확장하고 있음. 이는 데이터 스페이스 정책이 구조를 만드는 단계에서, 실제로 활용되고 가치가 창출·환류되는 단계로 진화하고 있음

② 일본

- » 일본은 데이터의 자유로운 신뢰 흐름(Data Free Flow with Trust, DFFT)를 국가 차원의 디지털 질서로 구현하기 위해, 경제산업성(METI) 주도로 산업계·학계·정부가 참여하는 『우라노스 생태계 이니셔티브』를 추진하고 있음 (’23년~)
- » 일본 제조업은 기업 단위의 디지털 전환은 진전되었으나, 데이터가 기업·산업·플랫폼별로 분절되어 산업 전반의 최적화와 공급망 전주기 협업에는 구조적 한계가 지적되어 있음
- » 이에 우라노스 생태계는 데이터의 소유권과 통제권을 각 주체가 유지한 채 공통 규칙과 신뢰 기반으로 데이터를 연계·활용하는 연합형 상호운용 구조를 통해, 산업·기업·국가 간 데이터 상호운용성을 제고하고 공급망 경쟁력 강화와 사회적 과제 대응을 동시에 추진

2) <https://digital-strategy.ec.europa.eu/en/policies/data-union>

[우라노스 생태계 데이터 스페이스의 운영 원칙³⁾]



* (출처) METI, Whitepaper: Ouranos Ecosystem Dataspaces Reference Architecture Model

③ 중국

▶▶ 중국은 2019년 당 중앙위원회 회의를 통해 데이터를 토지·노동·자본·기술과 함께 새로운 생산요소로 규정한 이후, 데이터의 자원화·시장화를 추진하기 위해 「데이터 요소 X 3개년 행동계획(2024~2026)」을 발표하고 데이터 스페이스를 데이터 유통 환경 개선의 핵심 개념으로 제시하였음

▶▶ 이어 국가데이터국은 「신뢰할 수 있는 데이터 스페이스 개발 실행 계획(2024~2028)」을 통해 데이터 스페이스를 국가 차원의 새로운 데이터 인프라로 공식화하고, 신뢰 기반 데이터 유통·거래 체계를 단계적으로 구축하겠다는 전략을 명확히 하고 있음

3) <https://www.ipa.go.jp/en/digital/architecture-guidelines/ouranos-ecosystem-dataspaces-ram-white-paper.html>

[중국 「신뢰할 수 있는 데이터 스페이스 개발 실행 계획(‘24-’28)」 주요 목표]

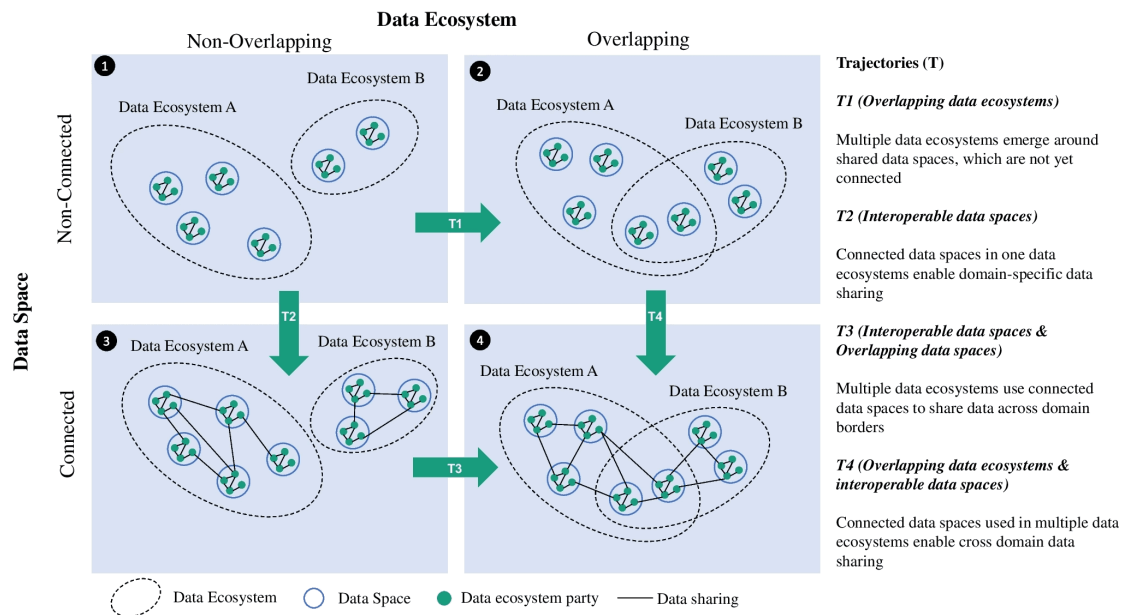
- ① 100개 이상의 신뢰할 수 있는 데이터 공간 구축 및 데이터 공간 모범사례 확립
- ② 데이터 요소의 원활한 흐름과 데이터 자원의 효율적인 배분
- ③ 신뢰할 수 있고 통제가능·상호연결·가치창출이 가능한 데이터 공간 구축
- ④ 기업, 산업, 도시, 개인, 국경 간 신뢰할 수 있는 데이터 공간을 구분하여 구축

* (출처) KISDI, 중국과 일본의 데이터 공간(Data Space) 전략 비교

④ 호주

호주는 데이터 스페이스를 신뢰 기반 데이터 공유 인프라로 도입하고 있으며, 이를 통해 정부·산업·연구기관 간 안전하고 통제 가능한 데이터 활용 환경을 구축하고자 함. 호주는 특히 Australian Research Data Commons(ARDC)를 중심으로, IDSA (국제데이터스페이스연합) 참여 및 글로벌 사례 분석을 통해 국제 표준을 자국 환경에 맞게 적용하는 방향으로 정책을 추진하고 있음

현재 환경·바이오 등 특정 분야를 중심으로 파일럿 프로젝트를 추진하며 데이터 스페이스의 실효성을 검증하는 단계에 있으며, 대표적으로 Planet Research Data Commons 내 기존 프로젝트와 연계하여 테스트베드 및 시범 데이터 스페이스를 구축하고, 이를 통해 기술·법·거버넌스 측면의 구현 가능성을 점검하고 있음⁴⁾



[호주 데이터 스페이스 생태계 구현도]

4) <https://ardc.edu.au/program/australian-dataspaces/>

1.4 국내 데이터 스페이스 도입 필요성

■ 데이터 유통·활용 생태계 패러다임 전환 필요

- » 표준 계약 부재, 데이터 가치평가 기준 미정립, 데이터 주권 및 권리관계 관리 미흡, 활용 이력 추적 어려움 등으로 인해 신뢰 기반 데이터 거래 활성화에 구조적 한계가 존재하며, 특히 기관 간 데이터 결합·활용 시 법·제도적 불확실성과 책임소재 문제로 인해 실제 활용까지 이어지지 못하는 사례 다수 발생
- » 데이터 스페이스는 분산·연합형 구조를 기반으로 데이터의 저장 위치를 유지하면서도, 표준화된 계약·정책·인증 체계를 통해 안전한 데이터 공유를 가능하게 하며, 이를 통해 참여 주체 간 신뢰 기반 협력 생태계를 조성하고, 데이터 등록-발견-계약-활용-정산까지 전주기 관리가 가능한 구조로 전환함으로써 데이터 유통·활용 패러다임의 근본적 혁신 제시

■ AX 확산에 따른 도메인 특화 데이터 중요성 증대

- » AI 전환(AI) 가속화로 범용 데이터 대비 산업·도메인 특화 데이터의 경쟁력이 핵심 요소로 부상하였으며 제조, 물류, 에너지 등 각 산업에서는 고품질·고신뢰 데이터 확보 여부가 AI 성과와 직결됨. 하지만 개별 기업 또는 기관 단독으로는 특화 데이터 확보에 한계가 있으며, 특히 민감정보·산업 데이터는 외부 공유가 제한적임
- » 데이터 스페이스는 데이터 주권을 유지하기 위해 정책 기반 접근제어, 사용조건, 이력 추적 등을 통해 안전한 데이터 공유를 지원하고, 다기관·다기업 간 협력 활용을 가능하게 하여 버티컬 AI 데이터 유통 촉진 및 산업별 AI 고도화 등 AX 확산에 기여

■ 글로벌 규제 대응을 위한 데이터 스페이스 활용 필요

- » EU의 디지털 제품 여권(DPP), 탄소국경제도(CBAM) 등 글로벌 규제는 제품 전주기 데이터의 투명성·추적성 확보를 요구하며, 생태계 단위의 데이터 공유·관리 체계 구축이 필수화되는 추세이며 유럽 등은 데이터 스페이스를 활용해 규제 대응 및 데이터 경제 활성화하고 있으나, 국내는 개별 대응 중심으로 한계 존재
- » 이에 국가 차원에서 데이터 스페이스 기반 표준화된 데이터 관리·공유 체계를 구축하여 글로벌 규제 대응 역량 확보와 동시에 데이터 주권 및 산업 경쟁력 강화 필요

2. 한국형 데이터 스페이스 추진 전략

① 단계적 확산을 통한 데이터 스페이스 구축

- ◆ 초기에는 법·제도적 요구사항과 기술적 복잡도를 고려하여 **산업·서비스 단위의 소규모 데이터 스페이스를 다수 구축**하고, 다양한 실증을 통해 **표준, 계약·정책, 기술 구성요소 간 상호운용성 검증**
- ◆ 실증 결과를 기반으로 참여 주체, 데이터 범위, 서비스 기능을 점진적으로 확대하고, **데이터 스페이스 간 연계(Federation)를 통해 확장 가능한 데이터 생태계로 발전 유도**

② 규제 대응을 넘어 버티컬 AI 중심 활용 확산

- ◆ 유럽 등은 규제 대응 중심으로 데이터 스페이스를 도입하고 있으나, **산업적 활용성과 비즈니스 창출 측면에서는 한계 존재**
- ◆ 국내는 단순 규제 대응을 넘어 산업별 버티컬 AI 활용을 중심으로 데이터 스페이스 적용 **Use Case를 적극 발굴·확산**하고 **데이터 제공·활용-AI 서비스 창출로 이어지는 선순환 구조를 구축**하여 데이터 스페이스 기반 신규 비즈니스 창출 및 산업별 AI 경쟁력 강화

③ 공공 주도 실증 후 민간 중심 자생 생태계 전환

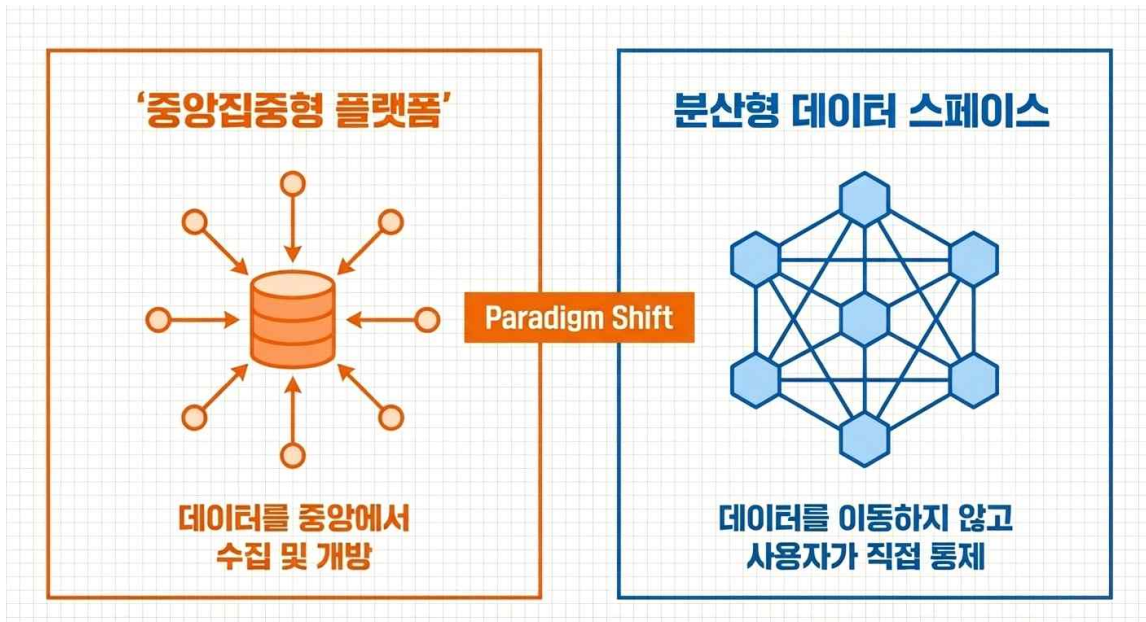
- ◆ 초기에는 **공공이 데이터 스페이스 구축·운영을 주도**하여 표준, 기술 아키텍처, 계약·정산 모델 등 기반 요소를 정립하고, 다양한 실증을 통해 확산 가능한 레퍼런스 확보
- ◆ 이후에는 구축된 기반을 토대로 **민간이 데이터 제공·활용·서비스 개발을 자율적으로 수행**하는 구조로 전환을 유도하고, **시장 중심의 데이터 생태계로 점진적 안착**

④ DaaS(Data Space as a Service) 기반 참여 접근성 강화

- ◆ 궁극적으로는 커넥터, 데이터 등록·검색, 계약·정책 관리 등 핵심 기능을 서비스형으로 제공하여 기업이 **별도의 기술 구축 없이도 데이터 스페이스에 쉽게 참여할 수 있도록 지원**
- ◆ **중소·중견기업 및 다양한 참여 주체의 진입 장벽을 완화**하고, 데이터 활용 기반 신규 서비스 및 **부가가치 창출을 촉진**하여 생태계 전반의 참여 확대 유도

II 데이터 스페이스 주요 개념

- » 데이터 스페이스는 데이터 주권을 전제로 한 분산형 데이터 공유 구조를 통해, 특정 조직이나 플랫폼에 데이터를 집적하지 않고도 다양한 참여자 간에 데이터를 안전하게 유통·거래할 수 있도록 설계된 체계
- » 중앙 집중형 플랫폼과 달리, 연합(Federation) 기반 구조를 채택하여 각 참여 주체가 자신의 데이터를 직접 보유·통제한 상태에서, 사전에 합의된 정책과 규칙에 따라 데이터 접근, 이용 목적, 결합·재사용 여부, 공유 이력 등을 정책 기반으로 자동 관리할 수 있음
- » 이러한 구조는 데이터 주권 보장, 신뢰 기반 교환, 상호운용성을 핵심 원칙으로 하며, 데이터 활용 과정 전반에서 투명성과 책임성을 확보하는 것을 목표로 함. 이와 같은 개념은 International Data Spaces Association (IDSA), Gaia-X 등 국제 논의를 통해 구체화 되어 왔으며, 데이터 스페이스는 단순한 기술 플랫폼이 아니라 정책·제도·기술·운영 요소가 결합된 신뢰 기반 데이터 유통·활용 인프라로 인식되고 있음



III 한국형 데이터 스페이스 참조모델

1. 참조모델 수립 목적

■ 신뢰 기반 데이터 유통·거래 전 주기를 포괄하는 국가 기준 구조 마련

- » 한국형 데이터 스페이스 참조모델은 데이터가 단순히 개방·공유되는 단계를 넘어, 신뢰를 전제로 유통·거래되고 지속적으로 활용되는 환경을 국가 차원에서 체계적으로 구축하기 위한 기준 구조를 마련하는 데 목적이 있음
- » 본 참조모델은 데이터 주권과 책임성, 데이터 출처·이력·권리관계에 대한 신뢰 확보, 계약·정산·분쟁 대응 등 데이터 유통·거래 과정에서 요구되는 요소들을 데이터의 등록·식별·이력 관리·인증·교환·거래·정산·감사에 이르는 전 주기를 포괄하는 하나의 구조로 제시함

■ 공공·민간 데이터 스페이스 확산을 위한 공통 구조와 자율성의 균형 확보

- » 본 참조모델은 공공과 민간, 산업별 데이터 스페이스가 각자의 목적과 특성을 유지하면서도 최소한의 공통 규칙과 신뢰 구조를 공유할 수 있도록 설계
- » 공공 영역에서는 데이터의 공공성, 법·제도 준수, 책임성 확보를 전제로 하면서 민간과의 연계를 가능하게 하고, 민간 및 산업 영역에서는 다양한 비즈니스 모델과 기술 구현 방식이 공존할 수 있는 유연성을 보장함으로써, 데이터 스페이스가 특정 플랫폼이나 중앙 시스템에 종속되지 않고 확산될 수 있는 구조적 기반 마련

■ 단계적 실증과 연계를 통한 국가 데이터 스페이스 생태계 기반 구축

- » 본 참조모델은 특정 산업·도메인을 중심으로 데이터 연계·활용을 단계적으로 실증하고, 이를 통해 기술 요소, 거버넌스 체계 및 정책 적용의 적정성을 종합적으로 검증하며, 그 결과를 기반으로 표준화 및 공통 구조를 정립하고, 참여 주체와 적용 범위를 점진적으로 확대하여 데이터 스페이스 생태계 전반으로 확산을 유도함
- » 또한 데이터 스페이스의 정책 실증(PoC)과 시범 사업을 통한 단계적 적용·확산을 전제로 설계되었으며, 이를 통해 실행 가능성을 검증하고 제도적·기술적 한계를 보완하고자 함

2. 참조모델 세부 구성

2.1 한국형 데이터 스페이스 구성 요소

- ▶▶ 해외 주요 데이터 스페이스 참조모델(IDSA, DSSC, Ouranos RAM 등)은 데이터 카탈로그, 인증·접근관리, 커넥터 기반 연계 등 참여주체 간 데이터 교환과 활용을 지원하는 공통 기반 중심으로 구성 요소를 제시하고 있음. 다만, 데이터 스페이스를 통한 AI 학습·분석 및 서비스화, 데이터 거래·정산 등 활용 및 비즈니스 확장 기능은 일부 제한적으로만 반영되어 있음
- ▶▶ 기존 해외 참조모델을 기반으로 데이터 스페이스 구성 요소를 거버넌스와 아키텍처로 재구성하고, 이를 공통 요소(필수 기능)와 확장 요소(선택적 기능)로 체계화
- ▶▶ 특히, 기존 모델에서 제한적으로 다루어진 AI 활용 영역을 보완하기 위해 AI 인프라 요소를 확장 구성 요소로 포함하여, 데이터 스페이스가 데이터 교환을 넘어 AI 기반 분석 및 서비스 창출까지 지원하도록 설계

[한국형 데이터 스페이스 공통·확장 거버넌스 구성 요소]

구분		구성 요소	핵심 내용
공통	참여주체 및 역할 체계	운영·거버넌스 주체	· 운영 총괄, 정책 적용, 모니터링 및 분쟁 조정 수행
		데이터 스페이스 이용 주체	· 데이터 제공·활용, AI 학습 및 서비스 구현 수행
		데이터 스페이스 공통 기반 제공 주체	· 데이터 연계, 신뢰·인증, 정산 및 공통 기술·서비스 제공
	공통 정책·규칙 체계	참여 승인·관리 및 인센티브·제재 체계	· 참여 관리 및 인센티브·제재 기준 운영
		데이터 표준·연계 정책	· 데이터 표준 및 메타데이터 기준(DCAT 등) 설정
		데이터 접근·이용 기본 정책	· 데이터 제공자의 통제권 기반 조건부 데이터 접근·이용 정책
		신뢰 기반 데이터 교환 정책	· 인증·접근통제·이력 관리 및 교환 기술 기준 적용
		데이터 거래·계약 및 정산	· 데이터 거래 계약 표준 및 절차 정의, 참여자 간 직접 계약 구조 지원
		데이터 결합·AI 활용·서비스 확산 실행 규칙	· 분산 데이터 환경에서의 AI 학습·분석·서비스 활용 절차 기준 마련
		분쟁 해결 및 책임	· 데이터 이용·거래 관련 분쟁 조정 및 책임 규정 체계 마련

	지속 가능한 운영 체계	거버넌스 협의체 및 운영 조직 구성	· 분과형 협의체를 구성하고, 참여 주체 간 협의 기반의 운영·의사결정 체계 마련
		Use Case 공동 발굴 및 파이프라인 적용·관리 기준	· 활용 수요 기반 Use Case 발굴, 데이터 연계-AI 학습-서비스 파이프라인 정의
		지속 가능한 운영을 위한 멤버십·수수료 정책	· 멤버십·이용료 기반 운영 자원 체계 마련
확장	공통 정책·규칙 체계	데이터 가치평가 기준	· 데이터 품질, 희소성, 활용도 기반 가치 산정 기준 마련
		AI 결과물 권리 및 책임 규정	· AI 모델, 분석 결과물의 소유권·책임·재사용 기준 정의
	지속 가능한 운영 체계	애플리케이션·서비스 생태계 활성화 정책	· 컴플라이언스 검증된 앱·서비스 중심 활용·연계 체계 구축

[한국형 데이터 스페이스 공통·확장 아키텍처 구성요소]

구분		구성요소	핵심내용
공통	등록	온보딩	참여기관·사용자 등록 및 초기 설정
		데이터 카탈로그	메타데이터 기반 데이터 탐색·검색
	인증	인증 및 접근 관리	참여기업, 사용자, 데이터, 커넥터 등 인증
	교환	커넥터	기관 간 데이터 연계 및 교환
		Compute to Data	데이터 이동 없이 원천 위치에서 분석·처리가 가능한 분산형 처리 방식 적용
		상호운용성 인터페이스	API 및 표준 기반 연계
	활용	정책 집행 엔진	데이터 이용 조건을 기술적으로 집행·통제하는 정책 실행 메커니즘 적용
		데이터 전처리 파이프라인	정제·표준화·가명처리 등 전처리 자동화
		계약 및 정산 시스템	데이터 이용 조건 기반 계약 및 이용량·기여도 기반 정산 체계 운영
		로깅 및 감사	데이터 이용 이력 추적
확장	유통	워크스페이스(AI Lab)	데이터 연계 기반 AI 모델 개발·학습·검증 환경
		데이터 마켓플레이스	데이터 검색·거래·구매 기능 통합 플랫폼
	활용	앱·서비스 마켓플레이스	SaaS·API 형태의 서비스 제공 구조
		연합학습, 폐쇄 분석	데이터 이동 없이 다기관 간 공동 학습·분석이 가능한 분산형 분석 구조
	보안	디지털 트윈 관리	물리 객체·시스템의 디지털 트윈 모델 등록 및 메타데이터 관리
		보안 강화 기술	PETs(차등프라이버시, MPC 등) 기반 보호

2.2 한국형 데이터 스페이스 설계 핵심 원칙

① 신뢰성

» 참여 주체 간 역할과 책임, 데이터 사용 규칙을 사전에 명확히 정의하고 일관되게 적용함으로써 신뢰를 확보하고, 데이터 이용 조건의 명확성, 이력 추적 가능성, 책임 규정을 통해 불확실성과 분쟁 가능성을 최소화해야 함

② 주권 및 통제

» 각 참여 주체가 자신의 데이터와 시스템에 대한 주권과 통제권을 유지한 채 연결되는 분산형 구조를 전제로 하며, 데이터 제공자는 사용 목적·범위·기간 등에 대한 통제 권한을 지속적으로 보유함으로써, 데이터 공유에 대한 부담을 낮추고 참여 확대를 유도해야 함

③ 상호운용성

» 특정 시스템이나 기술에 종속되지 않고 이기종 참여자와 시스템이 유연하게 연계될 수 있도록 상호운용성을 제고하고 표준화된 데이터 구조와 개방형 인터페이스를 기반으로 향후 산업·국가 간 연계를 포함한 확장성과 지속성을 확보해야 함

④ 보안 및 프라이버시

» 보안과 프라이버시는 설계 단계부터 전주기에 걸쳐 내재화되어야 할 기본 전제로 참여자 인증과 권한 관리, 접근 통제, 이력 추적, 개인정보·민감정보 보호를 통해 안전한 데이터 활용과 책임성 있는 사후 관리가 가능하도록 해야 함

⑤ 확장성 및 모듈성

» 초기 구축 이후에도 참여자, 데이터 유형, 활용 시나리오가 단계적으로 확장될 수 있도록 모듈형 구조로 설계되어야 하며, 이러한 확장성과 모듈성은 민간 주도 운영 전환과 다양한 산업·데이터 스페이스 간 연계를 가능하게 하는 핵심 조건임

2.3 한국형 데이터 스페이스 거버넌스

- ◆ 한국형 데이터 스페이스 거버넌스는 기술이나 플랫폼을 중심에 두기보다, 데이터와 AI 활용에 대한 공통의 합의와 신뢰를 중심으로 작동하는 연합형(Federated) 협력 생태계를 지향함
- ◆ 이 모델에서 데이터 스페이스는 중앙 집중적 통제 구조가 아니라, 데이터 제공자와 이용자가 각자의 데이터 주권을 유지한 채 참여하고, 공통 정책·규칙과 운영 절차에 대한 합의를 통해 협력이 성립되는 구조로 설계됨
- ◆ 거버넌스는 단순한 운영 규칙의 집합이 아니라, 데이터 접근과 결합, AI 학습과 서비스 확산 과정 전반에서 무엇이 허용되고, 어떻게 책임이 귀속되며, 신뢰가 어떻게 집행되는지를 지속적으로 조정·관리하는 핵심 작동 원리로 기능함
- ◆ 이를 통해 한국형 데이터 스페이스는 데이터 교환을 넘어, 합의된 규칙 아래에서 AI 학습과 서비스가 안전하게 실험·확산되고, 참여자 간 가치가 공정하게 배분되는 자율적이면서도 책임 있는 데이터·AI 협력 생태계로 운영될 수 있음

2.3.1 참여 주체 및 역할 체계

- ◆ 데이터 스페이스는 단일 기관이 모든 기능을 수행하는 구조가 아니라, 운영, 이용, 공통 기반 제공의 역할이 분리된 다중 참여 구조를 전제로 함
- ◆ 각 주체는 데이터의 제공과 활용, 정책 운영, 기술 지원, 신뢰 확보 등 서로 다른 기능을 담당하며, 이러한 역할 분담을 통해 데이터 주권을 유지하면서도 안전하고 지속 가능한 협력 추진

■ 데이터 스페이스 운영·총괄 주체

- » 데이터 스페이스 전반의 운영 원칙과 절차를 수립하고, 참여 주체 간 협력 구조를 조정하는 총괄 관리 주체로써 특정 데이터나 서비스의 직접적 소유자라기보다, 데이터 스페이스가 공정하고 안정적으로 운영될 수 있도록 관리·조정하는 중립적 운영자 역할 수행

- » 주요 역할은 참여기관의 가입 및 탈퇴 관리, 공통 정책·규칙의 제정 및 적용, 운영 현황 점검, 참여자 간 이해관계 조정, 운영 성과 모니터링 등을 수행하며 데이터 이용 과정에서 정책 위반, 책임 귀속, 수익 배분 갈등 등 다양한 운영 이슈가 발생할 수 있으므로, 운영 주체는 이러한 문제를 사전에 예방하고 필요시 조정할 수 있는 관리 체계를 함께 갖추어야 함

■ 데이터 스페이스 이용 주체

① 데이터 제공 주체 (Data Provider)

- » 데이터 제공 주체는 자신이 보유한 데이터를 데이터 스페이스 내에서 등록·제공하는 주체로써, 데이터 스페이스의 가장 기본적인 참여자이자, 데이터 공급의 출발점이 되는 핵심 주체
- » 데이터 제공 주체의 가장 중요한 특징은 데이터를 외부에 완전히 이전하거나 통제권을 포기하는 것이 아니라, 자신의 시스템과 관리 체계 안에서 데이터 주권과 통제권을 유지한 채 참여하는 것이며 데이터 제공자는 사용 목적, 이용 범위, 이용 기간, 재이용 가능 여부, 정산 조건 등 데이터 활용과 관련된 조건을 직접 설정할 수 있음

② 데이터 활용 주체 (Data Consumer)

- » 데이터 활용 주체는 데이터 제공자가 설정한 조건에 따라 데이터를 탐색하고, 분석, 연구, AI 모델 학습, 추론, 의사결정 지원, 서비스 기획 등에 활용하는 주체
- » 전통적인 데이터 유통 구조에서 활용 주체가 단순한 데이터 소비자로 인식되었다면, 데이터 스페이스에서는 활용 주체가 데이터를 기반으로 실질적인 부가가치를 창출하는 핵심 행위자로 간주됨
- » 특히 한국형 데이터 스페이스에서는 산업·분야별 AX를 촉진하는 것이 중요한 목표이므로, 데이터 활용 주체는 단순 조회 수준을 넘어 AI 기반 분석, 모델 개발, 운영 최적화, 서비스 혁신 등을 수행하는 중심 역할 담당
- » 다만 데이터 접근·활용은 제공자가 설정한 정책과 계약 조건을 준수하는 범위 내에서 이루어져야 하며, 활용 결과에 대한 설명 가능성, 책임성, 이력 관리 등도 함께 요구됨

③ 앱·서비스 제공 주체 (App / Service Provider)

- » 앱·서비스 제공 주체는 데이터와 AI 모델을 활용하여 최종 사용자에게 제공되는 애플리케이션, 분석 서비스, AI 기반 서비스, SaaS 형태의 응용 서비스를 개발·운영하는 주체
- » 이들은 데이터 스페이스 내부에서 생성된 데이터와 분석 결과를 산업 현장, 공공 서비스, 민간 시장 등 실제 수요 영역으로 연결하는 역할을 수행하며, 데이터 제공자와 활용 주체가 만들어낸 데이터 및 AI 기반 가치를 최종 서비스 형태로 구현하고 확산시키는 주체
- » 이 과정에서 앱·서비스 제공 주체는 특정 분야의 업무 요구를 반영한 사용자 인터페이스, API 기반 기능 제공, 분석 결과 시각화, 자동화 서비스, 의사결정 지원 서비스 등을 개발할 수 있음

■ 데이터 스페이스 공통 기반 제공 주체

① 데이터 스페이스 개발 주체 (Data Space Developer)

- » 데이터 스페이스 개발 주체는 데이터 스페이스 운영에 필요한 기술적 기반을 설계·구축·고도화하는 주체로서 데이터 스페이스 포털, 데이터 카탈로그 연계 기능, 커넥터 기반 분산형 데이터 교환 구조, 정책 집행 기능, 데이터 처리 및 AI 파이프라인, 사용자 인터페이스 등 다양한 기술적 요소 개발
- » 특히 데이터 스페이스는 중앙 집중형 플랫폼과 달리 다수의 기관과 시스템을 안전하게 연결해야 하므로, 개발 주체는 상호운용성 확보, 표준 적용, 보안 기능 연계, 확장 가능한 모듈형 구조 설계 등을 함께 고려해야 함
- » 한국형 데이터 스페이스가 단순 데이터 교환을 넘어 AI 활용과 서비스 확산까지 지원하고자 하는 만큼, 개발 주체는 데이터 전처리, AI Lab, 연합학습, 디지털 트윈 연계 등 확장 기능까지 단계적으로 구현할 수 있어야 함
- » 데이터 스페이스 개발 주체는 기술 아키텍처의 실질적 구현자이자, 데이터 스페이스의 성능과 확장성을 좌우하는 핵심 기반 제공자

② 검증 주체 (Conformity Assessment Body)

- » 검증 주체는 데이터, 서비스, 시스템, 참여기관 등이 데이터 스페이스의 정책, 표준, 기술적 요구사항을 충족하는지 확인하고 검증하는 역할을 수행하며 단순한 품질 점검을 넘어, 데이터 스페이스 내 신뢰성과 상호운용성을 확보하기 위한 적합성 평가 기능에 해당
- » 예를 들어 특정 기관이 데이터 스페이스에 참여하고자 할 때, 해당 기관의 시스템이나 커넥터가 요구되는 기술 표준을 만족하는지, 데이터 또는 서비스가 정해진 정책 및 컴플라이언스 기준을 준수하는지, 보안 및 접근 통제 체계가 적절히 구현되어 있는지 등을 검토할 수 있음
- » 또한 초기 온보딩 단계뿐 아니라 운영 중에도 정기 점검이나 재인증 절차를 통해 지속적인 적합성을 확인할 수 있으며, 이러한 검증 기능은 데이터 스페이스 내에서 교환·활용되는 자원에 대한 최소한의 품질과 신뢰 수준을 보장하는 장치로서 의미를 가질 수 있음

③ 신뢰·인증 인프라 제공 주체 (Trust Infrastructure Provider)

- » 신뢰·인증 인프라 제공 주체는 데이터 스페이스의 신뢰 체계를 기술적으로 지원하는 기반 인프라를 제공하는 주체로써 디지털 지갑, 전자증명서, 신원 인증 체계, 접근 권한 관리 체계, 거래 및 이용 이력 관리, 신뢰 검증 메커니즘 등이 포함될 수 있음
- » 데이터 스페이스는 다수의 참여기관이 분산 환경에서 협력하는 구조이므로, 누가 어떤 자격으로 참여하는지, 누가 어떤 권한을 가지고 어떤 데이터를 어떤 조건으로 활용했는지 확인할 수 있는 신뢰 인프라가 필수적임
- » 이 주체는 참여자의 신원과 권한을 검증하고, 이용 이력을 기록하며, 데이터 거래 및 활용의 신뢰성을 뒷받침하는 기술적 기반을 제공함으로써 데이터 스페이스 전반의 신뢰 환경을 조성함. 특히 분산형 구조에서는 중앙 운영기관의 통제만으로 신뢰를 보장하기 어렵기 때문에, 신뢰·인증 인프라는 데이터 스페이스가 안정적으로 작동하기 위한 핵심 기반 요소로 작동함

2.3.2 공통 정책·규칙 체계

- ◆ 데이터 스페이스는 참여 주체 간 사전 합의된 규칙을 바탕으로 데이터와 AI 활용을 빠르고 효율적으로 확산시키기 위한 협력 구조이기 때문에, 공통 정책과 규칙 체계가 필수적임
- ◆ 데이터 주권, 사용 조건, 역할과 책임, 기술 기준에 대한 합의가 선행될 때, 개별 협의나 반복적인 조율 없이도 데이터 결합과 AI 학습, 서비스 개발이 가능할 수 있음

① 참여 승인·관리 및 온보딩 절차

➤ 데이터 스페이스 참여를 위해 기관은 사전 정의된 참여 요건(기술, 보안, 정책 준수 등)을 충족해야 하며, 다음과 같은 온보딩 절차를 수행해야 함

- (1단계) 참여 신청: 기관 정보, 역할(Provider/Consumer 등), 시스템 구성 제출
- (2단계) 적합성 평가 및 인증: CAB(검증 주체)를 통해 기술·보안·정책 준수 여부 검증 및 인증
- (3단계) 디지털 신원 및 신뢰 자격 등록: DID/VC 발급 및 디지털 지갑 등록
- (4단계) 기술 연계 및 상호운용 검증: 커넥터, API, 카탈로그, 정책/계약, 엔드포인트 연동 시험
- (5단계) 운영 등록 참여 활성화: 정책 적용, 참여 승인, 운영 시작, 사후 모니터링·재평가 체계 적용

➤ 운영 주체는 참여기관의 역할, 권한, 책임을 정의하고 참여 유지 및 탈퇴 절차를 관리해야 하며, 정책 준수 수준, 데이터 기여도 및 활용 성과에 따라 인센티브 및 제재 기준을 적용해야 함

② 데이터 접근·이용 기본 정책

➤ 데이터는 분산 보유를 원칙으로 하며, 데이터 제공자가 설정한 조건에 따라 조건부 접근 방식으로 활용되어야 하며, 데이터 이용자는 카탈로그를 통해 데이터를 탐색하고, 이용 목적·범위·기간 등을 명시하여 접근 요청을 수행해야 함

③ 데이터 주권 및 사용 조건 관리

➤ 데이터 제공자는 이용 조건(목적, 범위, 기간, 재사용 여부, 정산 조건 등)을 직접 정의해야 하며 필요 시 이용 조건의 변경·갱신·철회 권한을 보유해야 함. 이용조건은 W3C ODRL(Open Digital Rights Language) 등 표준 기반 정책 표현 언어로 기술하고, 정책 집행 엔진(PDP/PEP)을 통해 자동 집행이 가능하도록 해야 함

» 이용자는 정의된 조건을 준수해야 하며, 위반 시에는 분쟁 해결 및 책임 규정에 따라 신뢰 수준 재평가, 접근 범위 제한, 조건부 참여 전환 등 단계적 제재가 적용됨

④ 데이터 표준·연계 정책 적용 절차

» 데이터 스페이스 내 데이터는 DCAT 기반의 메타데이터 구조를 적용하여 식별자, 설명, 품질 정보, 접근 조건 등을 포함한 형태로 관리되어야 하며, 「국가 데이터 통합 연계를 위한 데이터 카탈로그 표준 가이드(과기정통부)」를 준수하여 메타데이터를 작성해야 함

» API 및 커넥터 기반 연계를 통해 이기종 시스템 간 상호운용성이 확보되도록 설계되어야 하며 도메인별 특성을 반영한 표준은 확장 가능하도록 하되, 데이터 결합과 재사용을 고려하여 공통 구조와의 정합성을 유지해야 함

⑤ 신뢰 기반 데이터 교환 및 보안 체계

» 데이터 스페이스는 디지털 지갑 기반 신뢰 구조를 적용하며, 참여자는 DID/VC 기반 신원 인증 체계를 활용해야 하며, Trust Anchor는 인증서·자격증명 발급 및 검증 기준을 제공하는 주체로서 참여자의 신원 및 자격 검증의 기준점으로 작동함

» 디지털 지갑을 통해 참여자 신원 인증, 자격 증명(VC) 관리, 접근 권한 증명 등 수행

» 데이터 접근 및 이용 이력은 로그·감사 체계를 통해 기록되어야 하며, 추적 가능성을 확보해야 하며, 데이터 특성에 따라 Compute-to-Data, PETs(차등프라이버시, MPC 등) 등 보호 기술을 적용해야 함

⑥ 데이터 거래·계약 및 정산 체계

» 데이터 이용은 기본적으로 참여자 간 계약을 원칙으로 하되, 다음과 같은 거래 방식을 선택적으로 적용할 수 있음. 거래 유형에 따라 선택적으로 적용 가능하며, 도메인 특성과 데이터 민감도에 따라 유연하게 구성해야 함

- (1) 양자 계약 방식: 제공자-이용자 간 직접 계약
- (2) 표준 계약 기반 방식: 사전 정의된 이용 조건 템플릿 적용
- (3) 정책 기반 자동 승인 방식: 정책 엔진을 통해 사전 정의된 조건 충족 시 자동 이용 승인되며, 향후 스마트 컨트랙트 등 자동 집행 기술을 단계적으로 적용할 수 있음

- » 정산 체계는 데이터 이용량(호출 횟수, 처리량 등), 데이터 기여도(품질, 희소성 등), AI 활용 기여도(모델 성능 기여 등) 등을 종합적으로 고려하여 설계되어야 하며, 정산 방식은 정액형, 사용량 기반, 성과 기반 등 다양한 모델을 상황에 따라 혼합 적용할 수 있으며, 반복적이고 정형화된 정산을 효율적으로 처리하기 위해 자동화된 정산 체계의 적용 권장

[(예시) 데이터 스페이스 정산 모델]

정산 모델 유형	개념	적용 예시
사용량 기반 정산	데이터 또는 서비스 이용량에 따라 정산	데이터 조회 횟수, API 호출 수, 데이터 다운로드량 기반 정산
데이터 기여도 기반 정산	데이터 활용 결과에 대한 데이터 기여도를 고려하여 정산	복수 데이터셋 결합 분석 시 데이터 활용 비율 기반 정산
AI 성과 기여도 기반 정산	AI 모델 성능 또는 활용 성과에 기여한 데이터 또는 서비스에 대한 정산	AI 모델 학습에 기여한 데이터셋에 대한 성능 기여도 기반 정산
참여 주체 합의 기반 정산	참여주체 간 계약 또는 협의를 통해 정산 방식 결정	공동 프로젝트 또는 데이터 공동 활용 시 계약 기반 정산

- » 데이터 거래에 따른 대가는 초기에는 전자지급수단(디지털결제서비스) 및 참여자 간 직접 협의에 의한 결제 방식으로 운영하되, 데이터 스페이스의 성숙도에 따라 정산 포인트, 크레딧 등 정산 방식을 단계적으로 도입할 수 있음
- » 궁극적으로는 토큰 이코노미 기반의 자동화된 가치 교환 체계로 발전시켜, 데이터 기여·활용·정산이 연동되는 구조를 지향함. 다만 특정 화폐 또는 기술에 종속되지 않도록 설계하고, 관련 법·제도 및 회계 기준을 준수하는 범위 내에서 운영되어야 함

⑦ 데이터 결합·AI 활용·서비스 확산 절차

- » 데이터 전처리, 결합, AI 학습 및 분석은 데이터 제공자가 설정한 정책 조건을 준수하여 수행되어야 하며, AI 모델 및 분석 결과물에 대해서는 소유권, 재사용 조건, 서비스 적용 범위를 사전에 정의해야 함
- » AI 모델의 학습 데이터 출처와 이용 이력은 추적 가능하도록 관리되어야 하며, 이를 통해 설명 가능성과 책임성이 확보되어야 하며 아울러 데이터 스페이스 내에서 개발된 서비스는 등록 및 검증 절차를 거쳐 SaaS 또는 API 형태로 배포·활용될 수 있도록 설계되어야 함

⑧ 분쟁 해결 및 책임 규정

- ▶▶ 데이터 이용 및 거래 과정에서 발생할 수 있는 분쟁에 대비하여 사전에 정의된 분쟁 해결 절차를 마련해야 하며, 책임 판단은 이용 이력, 계약 조건, 정책 적용 기록 등을 기반으로 이루어져야 함
- ▶▶ 분쟁 발생 시에는 운영 주체 또는 별도의 분쟁 조정 기구를 통해 중재 및 조정이 수행될 수 있도록 체계 구축

2.3.3. 지속 가능한 운영 체계

- ◆ 데이터 스페이스의 운영 및 의사결정 체계는 참여 주체 간 합의를 통해 데이터와 AI 활용을 빠르고 지속 가능한 확산을 위한 핵심 기반
- ◆ 거버넌스, 참여 관리, Use Case·파이프라인, 검증된 SaaS·앱 활용, 운영 재원에 대한 기준이 마련될 때, 데이터 스페이스는 단발성 협력이 아닌 지속적인 AI 서비스와 혁신이 이루어지는 실행 환경으로 작동할 수 있음

① 거버넌스 협의체 및 운영 조직 구성

- ▶▶ 데이터 스페이스의 운영은 단일 기관의 판단에 의존하기보다, 주요 참여 주체가 함께 참여하는 독립적인 거버넌스 협의체 또는 운영위원회를 중심으로 이루어질 수 있도록 하며, 해당 조직은 데이터 스페이스의 운영 방향을 논의·조정하고, 공통 정책과 규칙의 적용 현황을 점검하는 역할을 수행함
- ▶▶ 이를 통해 데이터 제공자, 데이터 이용자, 운영 주체 간의 이해관계를 조율하고, 데이터와 AI 활용이 특정 주체에 편중되지 않도록 균형을 유지함

[(예시) 분야별 데이터 스페이스 운영 조직·협의체 구성]

조직	주요 역할
운영위원회	▶ 데이터 스페이스 운영 방향 및 주요 정책 결정
기술 운영위원회	▶ 데이터 스페이스 기술 구조, 데이터 교환 방식, 시스템 운영 관련 사항 검토
데이터 정책 협의체	▶ 데이터 등록 기준, 데이터 활용 규칙, 데이터 표준 및 정책 협의
데이터 표준 협의체	▶ 데이터 메타데이터, 데이터 포맷 등 데이터 표준 협의
데이터 품질 협의체	▶ 데이터 품질 관리 기준 및 데이터 검증 기준 논의
서비스 협의체	▶ 데이터 기반 서비스 및 애플리케이션 운영 기준 협의

② Use Case 공동 발굴 및 파이프라인 기반 적용·관리 기준

- ▶ 데이터 스페이스는 단순한 데이터 공유를 넘어 실질적인 활용과 가치 창출을 지향하므로, 참여 주체 간 협력을 통해 산업·도메인별 활용 수요 기반 Use Case를 공동으로 발굴·정의해야 함
- ▶ 발굴된 Use Case는 데이터 수집, 전처리, 연계, 결합, AI 학습 및 분석, 서비스화로 이어지는 전주기 데이터·AI 파이프라인으로 구체화되어야 하며, 해당 파이프라인은 표준화된 형태로 설계·관리되어 재사용성과 확장성을 확보해야 함
- ▶ 또한 Use Case는 필수 적용 영역과 선택 적용 영역으로 구분하여 단계적으로 확대될 수 있도록 설계함으로써, 데이터 스페이스의 활용 범위와 성과를 지속적으로 확장할 수 있도록 해야 함

[도메인별 데이터 스페이스 Use Case 예시]

분야	Use Case 예시	주요 내용	기대 효과
제조	품질 예측 및 불량 최소화	다수 제조기업의 생산·설비 데이터를 결합하여 AI 기반 품질 예측 모델 개발 및 공정 최적화	불량률 감소, 생산성 향상
제조	공급망 최적화	협력사·물류·재고 데이터를 연계하여 수요 예측 및 공급망 최적화	재고 비용 절감, 납기 안정성 확보
모빌리티	교통 흐름 최적화	차량·교통 인프라·위치 데이터를 결합하여 실시간 교통 흐름 분석 및 최적 경로 제공	교통 혼잡 완화, 이동 효율 개선
모빌리티	자율주행 데이터 공유	차량 센서 및 주행 데이터를 공동 활용하여 자율주행 AI 모델 학습 및 검증	자율주행 성능 향상
에너지	수요·공급 최적화	발전·소비 데이터를 결합하여 에너지 수요 예측 및 공급 최적화	에너지 효율 향상, 비용 절감
에너지	탄소 배출 관리	기업별 에너지 사용 및 배출 데이터를 연계하여 탄소 배출량 분석 및 감축 전략 수립	ESG 대응, 탄소 저감

③ 애플리케이션·서비스 생태계 활성화 정책

- ▶ 데이터 스페이스는 데이터 및 AI 활용 결과가 실제 서비스로 확산될 수 있도록, 공통 컴플라이언스를 충족한 애플리케이션 및 서비스 중심의 활용·연계 체계를 구축해야 함
- ▶ 이를 위해 보안, 접근 통제, 데이터 이용 조건 준수, 이력 관리 등 공통 기준을 충족한 앱·서비스를 등록·검증하고, 데이터 스페이스 내에서 연계·활용 가능하도록 관리해야 함
- ▶ 참여 주체는 이러한 검증된 서비스 및 애플리케이션을 활용함으로써 개별적인 검증 부담을 줄이고, 신뢰 기반 환경에서 일관된 방식으로 데이터 활용과 서비스 운영을 수행할 수 있어야 함
- ▶ 나아가 SaaS 및 API 형태의 서비스 제공 구조를 통해 데이터 스페이스 내에서 개발된 기능과 서비스가 외부로 확산될 수 있는 기반을 마련해야 함

④ 지속 가능한 운영을 위한 멤버십·수수료 정책

- ▶ 공통 정책·규칙 체계는 데이터 스페이스의 장기적 지속 가능성을 확보하기 위해 멤버십, 이용료, 거래 수수료 등 운영 자원과 관련된 정책 기준을 포함해야 함
- ▶ 파이프라인 개발, 분석환경 이용, 데이터·서비스 마켓플레이스 거래, 서비스·앱 이용 등에서 발생하는 비용과 수익의 원칙을 사전에 정의함으로써, 공공 재정 의존도를 점진적으로 낮추고 자생적인 운영 구조로 전환할 수 있도록 해야 함
- ▶ 이러한 정책은 참여 유인을 저해하지 않으면서도 데이터 스페이스의 안정적 운영을 뒷받침하는 기반이 될 수 있음

[(예시) 데이터 스페이스 멤버십 및 수수료 구조]

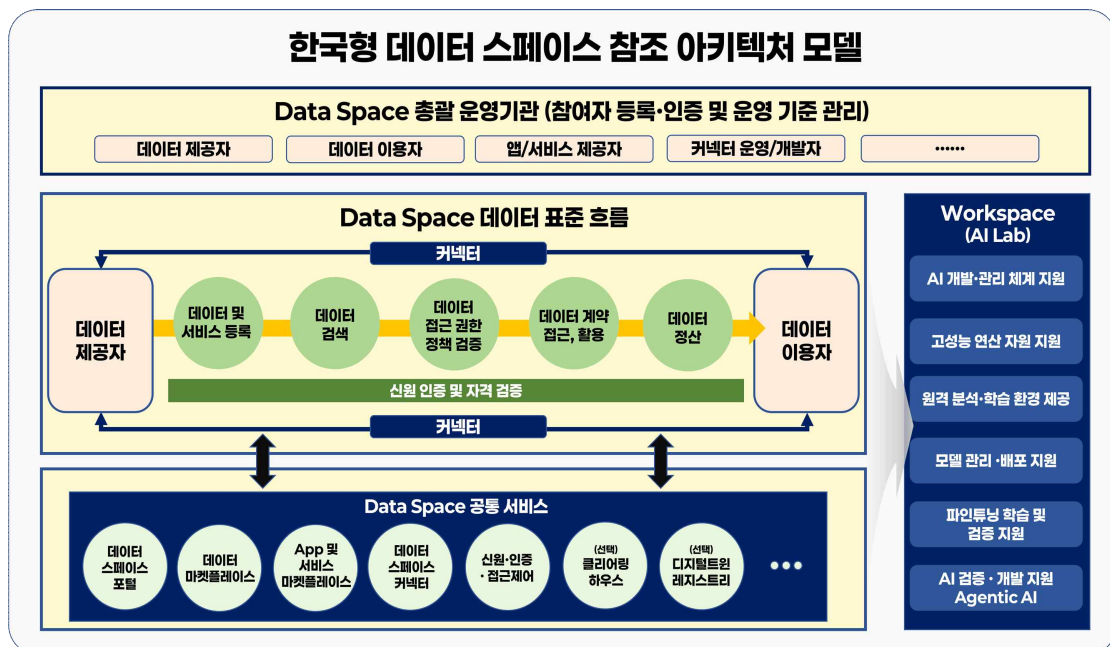
구분	서비스 내용	요금 구조
Basic Membership	▶ 데이터 스페이스 가입, 포털 이용, 데이터 검색	기본 가입비
Data Provider Service	▶ 데이터 등록, 메타데이터 관리, 데이터 제공	데이터 등록/ 관리 비용
Data Consumer Service	▶ 데이터 조회, 데이터 다운로드, API 이용	데이터 이용료
Data Exchange Service	▶ 데이터 거래, 계약 관리, 정산 서비스	거래 수수료
AI/Analytics Service	▶ 분석 환경, AI 학습 환경 이용	사용량 기반 요금
Premium Service	▶ 고급 분석, 공동 연구, 데이터 결합	별도 서비스 요금

2.4. 한국형 데이터 스페이스 아키텍처 모델

2.4.1 데이터 스페이스 아키텍처

- ▶▶ 한국형 데이터 스페이스 참조 아키텍처 모델은 데이터 제공자와 데이터 사용자가 각자의 시스템과 데이터 주권을 유지한 채, 표준화된 데이터 커넥터를 통해 안전하게 연결·교환·활용할 수 있도록 설계된 신뢰 기반 분산형 데이터 활용 구조를 제시함
- ▶▶ 본 아키텍처는 참여자 신뢰와 운영 원칙을 다루는 거버넌스 계층, 데이터 탐색·계약·이용·정산 등 가치 흐름을 다루는 비즈니스 계층, 그리고 이를 구현하는 커넥터, 연합 카탈로그, 정책·계약 관리, 신원·자격 검증, 클리어링하우스 등 공통 기술 요소로 구성됨
- ▶▶ 아울러 AI Lab 형태의 워크스페이스를 통해 Compute-to-Data 기반 분석·파인튜닝, MLOps, GPU 연산 자원, 모델 레지스트리 및 데이터 결합·가공 기능을 제공함으로써, 데이터의 안전한 공유를 넘어 데이터 기반 AI 개발과 서비스 창출까지 확장 가능한 실행 인프라로서의 데이터 스페이스 구현을 지향함

[한국형 데이터 스페이스 참조 아키텍처 모델 개념도]



2.4.2 데이터 스페이스 표준 흐름

단계	프로세스	주요 내용	핵심 기능/요소
1	온보딩	참여 주체의 식별·인증 및 데이터 스페이스 참여를 위한 신뢰 관계 형성	DID(분산 식별자), 자격증명(VC), 디지털 지갑, 신뢰수준(LoA) 부여
2	데이터 스페이스 커넥터 연결	조직 내부 시스템과 데이터 스페이스 간 분산형 데이터 연계 환경 구성	CaaS(Connector as a Service), Self-Host 커넥터, 데이터스페이스 프로토콜(DSP)
3	데이터·서비스 등록 및 검증	카탈로그 기반 메타데이터 등록 및 데이터 제공 기준·정합성 검증 수행	Self Description, ODRL 기반 이용 정책 등록, 스키마 및 메타데이터 검증
4	데이터 및 서비스 검색·발견	카탈로그 기반 데이터 및 서비스 탐색	카탈로그 조회, 시맨틱 허브, 온톨로지 기반 매핑
5	데이터 접근 권한 및 이용 정책 검증	데이터 이용 조건 및 정책 기반 접근 가능 여부 확인	PDP·PEP·PIP 아키텍처, ODRL 기반 정책 해석, ABAC(속성 기반 접근제어)
6	주체 간 데이터 계약	데이터 제공자와 이용자 간 사용 조건 협의 및 계약 체결	P2P 데이터 계약 방식, 결제 방식별 메커니즘 연계
7	데이터 전처리	계약 및 정책 조건에 따라 데이터 제공 형태로 가공 및 준비	데이터 프로파일링, 비식별화, 구조 표준화, 시맨틱 매핑
8	데이터 접근 및 활용	정책 및 계약 조건 하에서 데이터 교환 및 활용 수행	데이터 교환, C2D, 연합학습, API 교환
9	데이터 전주기 이력 관리	데이터 이용 전 과정의 추적 및 감사 가능성 확보	W3C PROV 기반 로그·이력 및 감사 관리
10	데이터 거래 정산·분쟁·사후관리	데이터 거래에 따른 정산 및 분쟁 대응 등 운영 관리	수익 분배, 이력 기반 정산, 분쟁 조정

Step 1. 온보딩

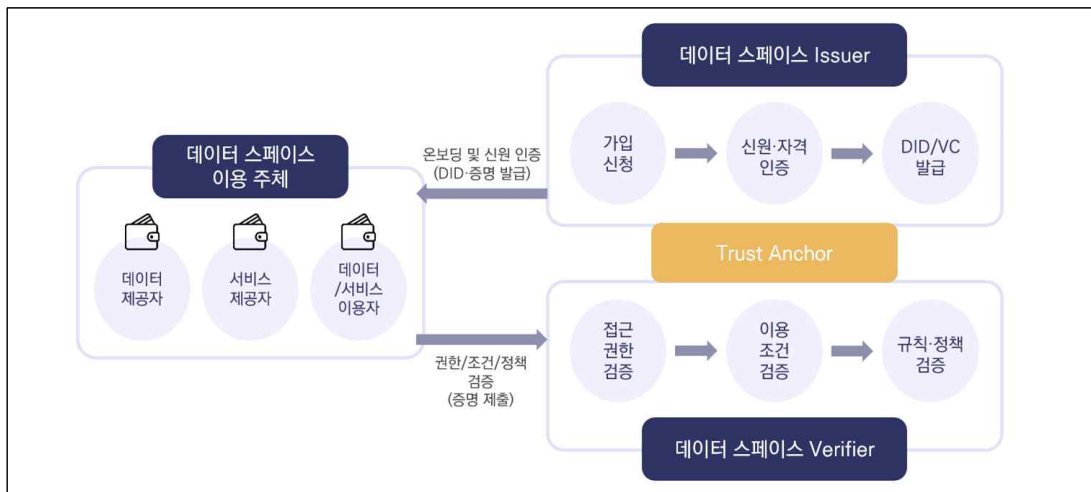
» 온보딩은 데이터 스페이스에 참여하려는 기관 또는 시스템이 디지털 신원과 자격을 기반으로 신뢰 관계를 형성하고, 역할·권한·정책 기준을 사전에 확정하는 단계임. 이 단계는 단순 계정 생성이 아닌, **디지털 지갑, DID, 검증 가능한 자격 증명(VC, Verifiable Credential)을 활용하여** 참여 주체의 신원을 검증하고, 데이터 접근 및 활용에 필요한 신뢰 기준을 설정하는 과정으로 구성됨. 온보딩 과정에서 설정된 신뢰 정보는 이후 정책 검증, 계약, 데이터 활용 전 단계에서 일관된 기준(Trust Anchor 기반)으로 활용됨

- ◆ 데이터 스페이스에서 온보딩 및 신원확인을 마치고 DID와 증명을 발급받은 참여 주체가 다른 데이터 스페이스를 이용하고자 할 경우 기 발급된 DID와 증명을 이용할 수 있어야 함
- ◆ W3C에서 제시하는 DID와 증명 포맷을 준수함으로써 규격에 대한 상호호환성을 확보하고, 증명에 대한 항목은 데이터 스페이스 상호호환 지침의 권고를 따름으로써 구성요소에 대한 상호 호환성을 준수해야 함

» 온보딩 과정에서는 누가 어떤 권한으로 운영, 요청, 승인 등의 행위를 수행할 수 있는지에 대한 역할과 책임의 기본 구조가 함께 정의되며, 이를 바탕으로 이후의 데이터 접근과 이용이 정해진 정책 및 책임 체계 안에서 이루어지도록 함

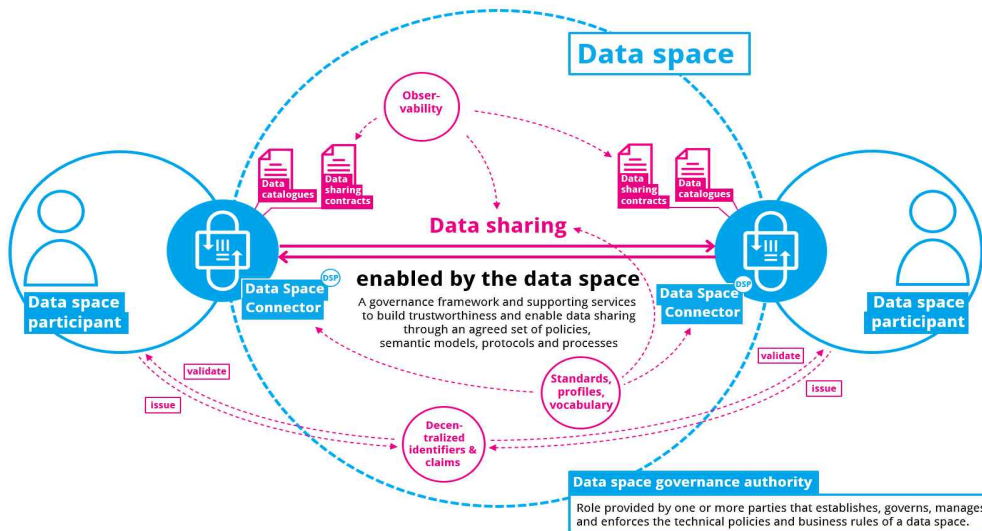
» 이러한 사전 등록·인증·권한 설정 절차를 통해 데이터 스페이스는 문제가 발생한 이후에만 통제하는 방식이 아니라, 참여 초기부터 신뢰성과 준수 요건을 확인하고 이후 운영 과정에서도 이를 지속적으로 검증·관리할 수 있는 기반을 마련함

[데이터 스페이스 온보딩 프로세스]



Step 2. 데이터 스페이스 커넥터 연결

- ▶▶ 데이터 스페이스 커넥터는 데이터 스페이스에 참여하는 기관이나 시스템을 연결하는 핵심 구성요소로, 참여자 간의 데이터 공유와 교환을 안전하고 신뢰할 수 있는 방식으로 가능하게 하는 기술
- ▶▶ 각 커넥터는 데이터 제공자와 이용자 사이에서 하나의 노드로 작동하며, 데이터가 실제로 저장된 위치를 변경하지 않으면서도 메타데이터 기반 탐색, 접근 제어, 데이터 전달을 수행함. 특히 커넥터는 한국형 데이터 스페이스 공통 규칙(아키텍처 및 표준 흐름)을 기반으로 표준화된 인터페이스(API)를 제공함으로써, 서로 다른 조직·시스템·도메인 간에도 일관된 방식으로 데이터 교환이 이루어지도록 지원하며, 국제적 상호운용성 확보를 위해 DSP TCK 등 검증 기준을 충족할 수 있도록 설계·운영되어야 함
- ▶▶ 커넥터의 운영 방식은 참여기관의 기술 역량과 인프라 환경에 따라 선택적으로 Self-host 방식 혹은 CaaS(Connector as a Service) 방식을 선택할 수 있도록 설계함. 자체 인프라를 보유한 기관은 커넥터를 직접 구축·운영하는 Self-Host 방식을 선택할 수 있으며, 기술 역량이나 인프라가 부족한 중소·중견기업 등은 데이터 스페이스 운영 주체 또는 공동 기반 제공 주체가 제공하는 CaaS 방식을 통해 별도의 시스템 구축 없이도 데이터 스페이스에 참여할 수 있음. 이를 통해 참여 주체의 진입 장벽을 완화하고, 데이터 스페이스 생태계 전반의 참여 확대를 유도함



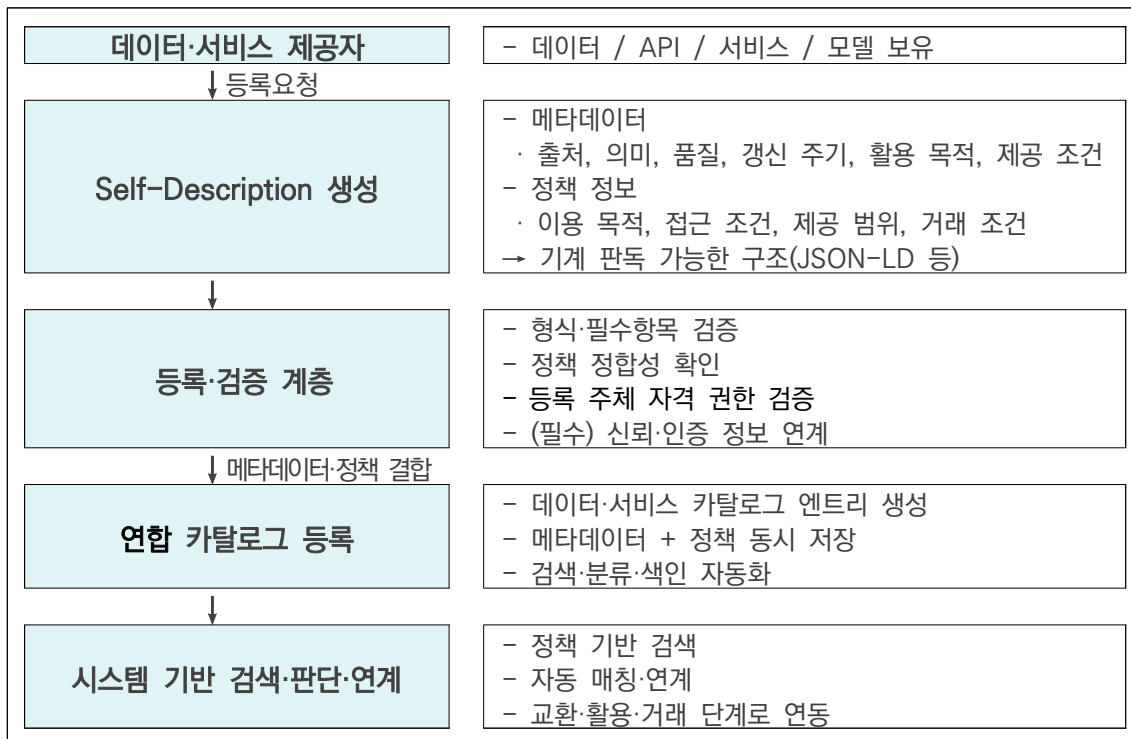
[데이터 스페이스에서 데이터 커넥터의 역할⁵⁾]

5) * (참고) Fundamental concepts of a data space, their roles and interactions. At its core: the Dataspace Protocol supporting the Data Space Connectors to ensure interoperability and bridge diverse systems (Data Space Connector Report, IDSA)

Step 3. 데이터·서비스 등록 및 검증

- » 데이터 및 서비스는 연합 카탈로그 또는 메타데이터 브로커에 등록되며, 이 과정에서 각 데이터와 서비스는 표준화된 메타데이터와 정책 정보를 포함한 자기서술 (Self-Description) 형태로 기술됨
- » Self-Description은 데이터와 서비스의 제공자, 출처, 의미, 품질, 갱신 주기, 접근 방식, 활용 조건 및 제공 정책 등을 구조화된 형태로 표현하는 기계판독 가능한 설명체계로서, 시스템이 해당 자산의 특성을 자동으로 발견하고 해석하며 비교·선택할 수 있도록 지원함
- » Self-Description은 JSON-LD 등 기계판독 가능한 구조로 작성되며, 데이터 스페이스의 성숙도와 신뢰 요구 수준에 따라 서명·검증 가능한 형태(W3C Verifiable Credentials 연계 등)로 확장하여 적용할 수 있음. Self-Description의 정책 정보는 ODRL 등 표준 기반 언어로 기술되며, 이후 Step 5(접근 권한 및 이용 정책 검증)의 PDP/PEP 정책 엔진에 의해 해석·집행됨
- » Self-Description은 최초 등록 이후에도 데이터의 품질·갱신 주기·접근 조건 등이 변경될 경우 갱신되어야 하며, 데이터가 폐기된 경우 카탈로그에서 비활성화 또는 삭제 처리되어야 함. 이러한 라이프사이클 관리를 통해 카탈로그의 최신성과 신뢰성을 지속적으로 확보함

[데이터·서비스 등록 프로세스]



- » 데이터 등록 시에는 표준 메타데이터 구조를 통해 데이터의 내용, 품질, 갱신 특성 및 활용 가능 범위가 기술되며, 데이터 제공 조건과 이용 제한 사항도 정책 정보 형태로 함께 등록되거나 연계될 수 있음. 이러한 정책 정보는 이후 계약, 접근 판단 및 이용 통제 과정에서 기계 판독 가능한 규칙으로 참조될 수 있음

 - ◆ 정책 정보에는 접근 조건, 이용 목적 제한, 결합·재이용 가능 여부, 의무 및 책임 조건 등이 포함될 수 있으며, 이를 통해 데이터 활용 조건을 사람의 해석에만 의존하지 않고 시스템이 처리 가능한 형태로 구조화할 수 있음. 다만 구체적인 계약 협상이나 도메인별 특수 조건은 별도의 합의 절차를 병행할 수 있음
 - ◆ 정책 표현에는 W3C의 ODRL과 같은 표준 기반 언어가 활용될 수 있으며, 데이터 스페이스 또는 도메인 특성에 따라 추가 어휘, 경량 정책 템플릿, 속성 기반 접근통제 (ABAC) 규칙 등과 조합하여 구현할 수 있음. 이때 ODRL은 주로 이용조건 표현에, ABAC는 접근권한 판단 및 집행에 활용되는 등 상호 보완적으로 적용될 수 있음
 - ◆ 이 단계에서는 데이터나 서비스를 등록할 수 있는 주체의 자격과 권한에 대한 신뢰 검증도 함께 수행되며, 해당 참여 주체가 데이터 제공자 또는 서비스 운영자로서의 역할을 부여받았는지와 등록하려는 데이터에 대한 소유권 또는 제공 권한 보유 여부를 확인함
 - ◆ 등록된 Self-Description 정보가 표준 형식과 정책 템플릿을 충족하는지, 필수 메타데이터와 책임 정보가 누락 없이 포함되어 있는지도 검증할 수 있으며, 이를 통해 카탈로그에는 신뢰 가능한 주체가 제공한 데이터와 서비스만 등록되도록 관리됨
- » 신뢰 기반 데이터 접근이 허용된 이후 데이터 품질 관리 서비스는 데이터 처리·분석 파이프라인에 자동 점검 단계를 연계하여 운영되며, 스키마 일관성, 결측치·중복·이상치 여부, 최신성 등 사전에 정의된 품질 지표에 따라 데이터를 점검함. 기준을 충족하지 못하는 경우 경고 또는 차단 상태로 분기되어 관리자가 확인·조치할 수 있도록 지원할 수 있음

 - ◆ 또한 데이터 스페이스 거버넌스 차원에서는 메타데이터 정합성 검증 기준과 데이터 검증 정책을 공통 규칙으로 관리하고, 검증 결과를 메타데이터 형태로 기록·공유함으로써 참여기관 간 데이터 해석 및 활용 기준의 일관성 확보, 데이터 신뢰성에 대한 객관적 판단 근거 제공, 데이터 활용 과정에서의 책임성 및 추적성 강화, 데이터 스페이스 전반의 안정적 운영을 지원할 수 있음

Step 4. 데이터 및 서비스 검색·발견

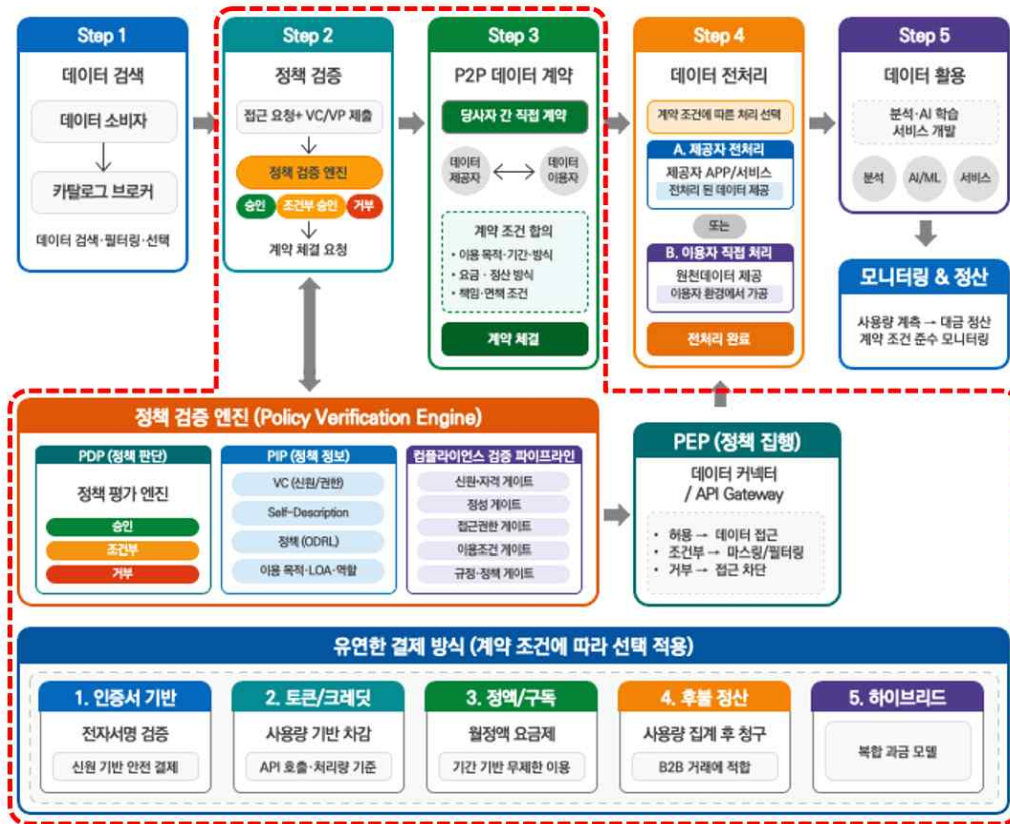
- » 이용자는 **연합 카탈로그**를 통해 데이터와 서비스를 탐색하며, 등록된 메타데이터와 정책 조건을 기반으로 자동 필터링을 수행하며 이 과정에서 **이용 목적, 이용자의 신뢰 수준, 보안 요구사항** 등이 **사전에 매칭되어, 실제로 계약 활용이 가능한 데이터와 서비스만 후보로 도출됨**. 이를 통해 탐색 단계에서부터 정책 위반 가능성을 제거하는 정책 인지형 검색·발견 구조가 구현될 수 있음
- » 데이터 스페이스의 검색·발견 기능은 **브로커(Broker)를 중심으로 구성되며**, 브로커는 분산된 참여기관이 등록한 데이터 및 서비스의 메타데이터를 집계·색인하고 정책 조건과 시맨틱 정합성을 고려한 검색 결과를 제공하는 중개 기능을 수행하며, 그 운영 방식은 데이터 스페이스의 규모와 운영 목적에 따라 브로커 설치 방식과 브로커 직접 참여 방식을 선택적으로 적용할 수 있음
- » 메타데이터 브로커 시스템 설치 방식은 데이터 스페이스 운영 주체가 직접 메타데이터 브로커 시스템을 구축·운영하여 각 **참여기관의 커넥터로부터 Self-Description 및 메타데이터를 수집·집계하고 통합 색인을 생성한 뒤, 이용자에게 정책 인지형 검색 기능을 제공하는 방식**이며, 이때 브로커 시스템 내에는 Semantic Hub(온톨로지 및 공통 어휘 체계 관리 모듈)를 함께 내장하여 시맨틱 상호운용성을 확보할 수 있음
- » 전문 데이터 브로커의 직접 참여 방식은 전문 데이터 브로커가 데이터 스페이스의 참여 주체로 직접 참여하여 복수의 데이터 제공자 간 정책 조건과 정산 조건을 비교·조정하고, 이용자의 활용 목적을 해석하여 적합한 데이터 조합을 탐색·추천하며, 데이터의 희소성·품질·결합 가능성 등에 대한 부가 정보를 제공하고, 특정 산업·분야에 대한 도메인 전문성을 바탕으로 데이터의 품질·적합성을 평가하여 계약 가능 후보를 정제·제시함으로써, 고부가가치 데이터 거래, 복수 데이터셋 결합이 필요한 AI 학습 프로젝트, 수요-공급 매칭이 복잡한 환경에 적용할 수 있음
- » 아울러 데이터·서비스 등록 및 관리 과정에서는 **온톨로지, 공통 어휘 체계, Vocabulary Hub와 같은 시맨틱 상호운용성 수단을 활용**하여, 서로 다른 제공자가 사용하는 용어와 분류 기준을 공통 의미 체계에 따라 정렬할 수 있음. 이러한 체계는 등록 단계의 메타데이터 작성·검증뿐 아니라 이후의 탐색·연계 과정에서도 활용될 수 있음

- » 이를 통해 동일한 개념이 서로 다른 명칭이나 데이터모델로 표현된 경우에도 의미 기반 검색과 매핑이 가능해지며, 이용자는 제공자별 표현 차이나 기술적 구조를 일일이 해석하는 부담을 줄이면서 목적에 부합하는 데이터와 서비스를 보다 쉽게 발견할 수 있음
- » 이러한 검색 과정은 기본적으로 시스템 규칙에 의해 자동 수행되지만, 필요에 따라 AI 에이전트나 데이터 브로커가 탐색을 보조할 수 있음. 예를 들어 AI 에이전트는 이용 목적을 해석해 적합한 데이터 조합을 추천하거나, 온톨로지 기반으로 유사·대체 데이터를 제안할 수 있으며, 데이터 브로커는 복수의 데이터 제공자 간 정책 조건을 비교·조정해 계약 가능 후보를 정제하는 역할을 수행할 수 있음
- » 이 구성요소들은 필수 기능이 아니라, 데이터 스페이스의 성숙도와 활용 시나리오에 따라 선택적으로 적용 가능한 확장 요소임

Step 5. 데이터 접근 권한 및 이용 정책 검증

- » 선택된 데이터에 대한 접근 여부는 정책 엔진을 통해 자동화된 방식으로 판단될 수 있음. 전형적인 구현에서는 접근 요청이 발생할 때 정책 판단 지점(PDP)이 호출되어, 이용자의 속성, 이용 목적, 적용 가능한 데이터 정책을 종합적으로 평가하고 승인, 제한, 거부 또는 조건부 승인 여부를 결정함
- » 이 과정에서 PDP는 정책 정보 지점(PIP)을 통해 수집된 신뢰 및 맥락 정보를 종합적으로 참조하며, 여기에는 이용자의 신뢰수준(LoA), 역할 및 권한 정보, 이용 목적의 적합성, 데이터 등급과 제공 조건, 보안·컴플라이언스 요구사항 등 온보딩 및 이전 단계에서 확정된 요소들이 포함됨
- » 정책 판단 결과는 정책 집행 지점(PEP)을 통해 데이터 커넥터에 즉시 반영되어 실제 데이터 접근, 활용 범위, 실행 조건을 기술적으로 강제할 수 있음. 이를 통해 데이터 접근은 담당자의 개별 판단이 아닌, 정책 엔진에 의해 일관되고 자동화된 방식으로 통제되며, 데이터 스페이스 전반에 걸친 신뢰 기반 운영이 가능해짐

[접근 권한 및 이용 정책 검증 프로토콜]



Step 6. 주체 간 데이터 계약

- ▶ 권한·정책 결정 단계에서 접근이 허용된 데이터에 대해서는, 해당 이용 조건을 법적·상업적으로 구체화하기 위한 데이터 계약 절차가 수행될 수 있음. 이 단계에서는 정책 판단 결과와 적용 가능한 이용조건을 바탕으로 데이터 제공자와 이용자 간의 권리·의무, 책임 범위, 이용 조건을 명확히 합의함으로써, 기술적으로 허용된 접근을 실제 데이터 이용의 법적·상업적 근거로 구체화할 수 있음
- ▶ 데이터 계약은 이용 목적, 이용 기간, 이용 방식, 재이용 및 결합 가능 여부, 책임 및 면책 조건, 요금·정산 방식 등을 포함하며, 데이터 스페이스 환경에서는 계약 조건을 코드화하여 자동으로 실행하는 스마트 계약 방식으로 체결될 수 있음
- ▶ 이를 통해 사전에 합의된 조건에 따라 데이터 접근 권한 부여, 이용 제한, 이용량 기반 정산 등 계약 이행이 자동으로 수행될 수 있으며, 이후 결제 및 정산 단계에서는 다양한 방식의 결제 메커니즘과 연계되어 처리될 수 있음

① 인증서 기반 결제 방식

- 전자서명 인증서를 기반으로 계약 당사자의 신원 검증, 거래 승인, 결제 수행까지 일관되게 처리하는 방식으로, 데이터 스페이스 내 참여자 간 신뢰 확보 및 법적 효력을 보장함
- 계약 체결, 데이터 이용 승인, 결제 요청·승인 과정 전반에 전자서명이 적용되어 거래의 무결성 및 부인 방지를 확보하며, 기존 전자계약 및 금융 인프라와의 연계를 통해 안정적인 정산 가능함

② 토큰 기반 결제 방식

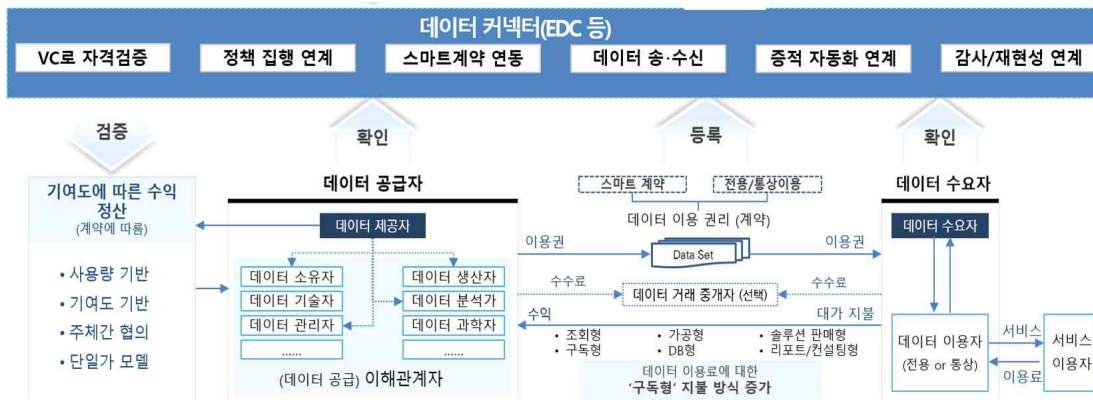
- 데이터 이용량, API 호출, 처리 건수 등 실제 사용 이벤트를 기반으로 토큰 또는 크레딧을 차감하는 방식으로, 데이터 이용과 정산이 직접적으로 연계되는 사용량 기반 과금 구조를 제공함
- 시스템이 이용 행위를 자동 계측하여 실시간 또는 주기적으로 정산을 수행하며, 다자간 참여 환경에서도 확장성과 자동화를 확보할 수 있어 데이터 스페이스에서 효과적으로 활용 가능함

Step 7. 데이터 전처리 자동화

- » 데이터 가공·전처리 자동화 기능은 데이터 스페이스의 부가 서비스로 제공되며, 커넥터를 통해 접근이 허용된 원천데이터에 대해 표준화·전처리 규칙을 정의하고 실행할 수 있도록 지원함
- » 해당 기능은 데이터를 외부로 이전하여 처리하는 방식이 아니라, 데이터가 위치한 제공자 환경에서 가공 작업이 실행되도록 지시하는 구조를 따르며 이를 통해 데이터 제공자는 원천데이터를 외부로 공개하지 않으면서도 활용을 허용할 수 있고, 데이터 사용자는 별도의 환경 구축 없이도 데이터 스페이스 내에서 일관된 전처리·가공 과정을 적용하여 분석과 AI 학습을 수행할 수 있음

Step 8. 데이터 접근 및 활용

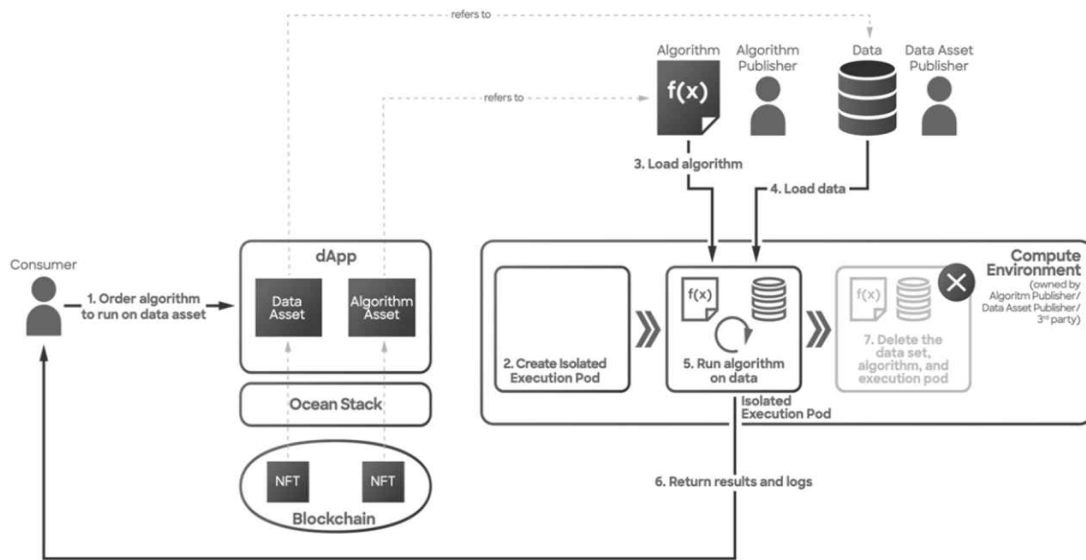
- » 데이터 계약을 통해 이용 조건이 확정되면, 데이터는 원칙적으로 물리적 이전 없이 접근·활용될 수 있음. 데이터 스페이스는 커넥터, API, 보안 게이트웨이 등과 같은 접근 중계·통제 수단을 통해 데이터 접근을 관리하며, 이용자는 계약과 정책에서 허용된 범위 내에서만 데이터를 조회하거나 처리할 수 있다. 커넥터는 이용 주체의 신원과 권한, 적용 정책을 확인한 뒤 접근을 중계하는 역할을 수행하며, 이를 통해 데이터는 제공자 환경에 유지된 상태로 안전하게 활용될 수 있음
- » 데이터 활용 방식은 접근 수단과는 별도로, 데이터의 위험도와 보안 요구 수준에 따라 차등적으로 적용됨. 저위험 데이터의 경우에는 이용 목적과 조건에 따라 제한적인 다운로드나 직접 접근이 허용될 수 있고, 중위험 데이터는 원본 데이터의 무단 확산을 방지하기 위해 커넥터 또는 API 기반 접근 방식으로 이용 범위가 통제될 수 있음. 고위험 민감 데이터는 원본 데이터를 외부로 이동시키지 않고, 분석 코드만을 데이터가 위치한 환경에서 실행하는 Compute-to-Data 방식이나 보호된 실행 환경(Secure Zone)에서만 활용되도록 제한됨



[(예시) 커넥터 기반 데이터 접근 및 활용 개념도]

- » 또한 데이터 스페이스와 같이 다수의 데이터 보유자가 분산된 환경에서는, 각 참여자의 데이터는 외부로 이동시키지 않고 모델 학습 또는 분석을 수행한 후 그 결과만을 공유·집계하는 연합학습(Federated Learning) 방식도 적용 가능하며, 이를 통해 개별 데이터의 직접 공유 없이도 다기관 간 협업 분석 및 공동 인사이트 도출이 가능함
- » 이러한 접근 방식과 활용 방식은 정책 엔진과 연계되어 자동으로 적용될 수 있음. 즉, 권한·정책 결정 및 데이터 계약 단계에서 확정된 조건에 따라 어떤 접근 수단을 사용할지와

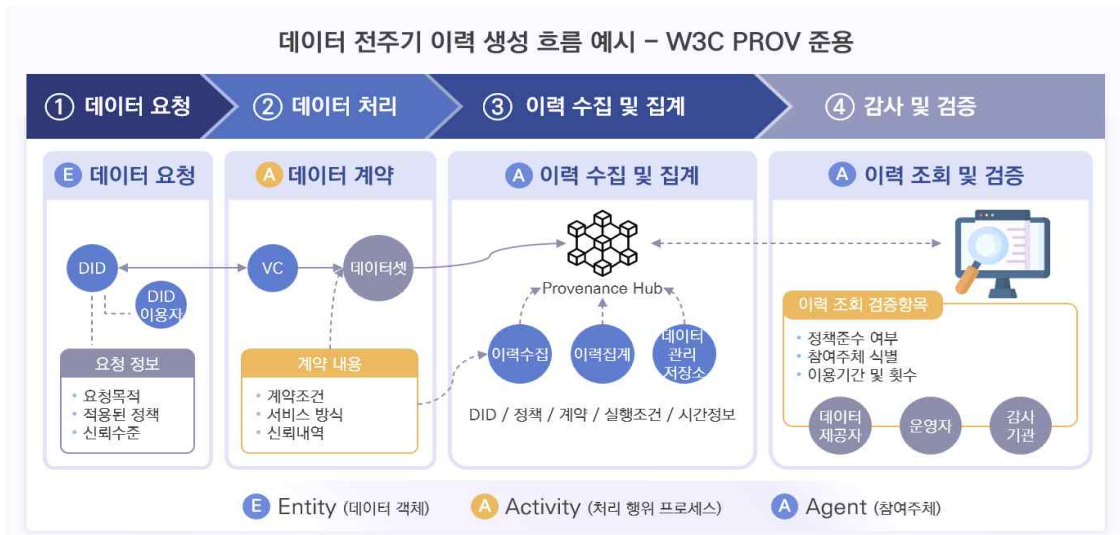
어떤 활용 방식을 적용할지가 함께 결정되며, 데이터 스페이스는 동일한 데이터라도 이용 목적과 리스크 수준에 맞는 활용 경로를 제공할 수 있음. 이를 통해 데이터 스페이스는 데이터 이동을 최소화하면서도, 신뢰와 보안을 전제로 한 유연한 데이터 활용을 가능하게 함



[Compute-to-data 기반 데이터 접근·학습 아키텍처]

Step 9. 데이터 전주기 이력 관리

- » 데이터 접근·처리·활용 전 과정은 시스템에 의해 자동으로 기록되며, 이용 이력과 정책 준수 여부는 사후 변경이 불가능한 형태로 관리되며 이 기록은 단순한 로그를 넘어, 누가 어떤 권한으로 어떤 데이터를 어떤 조건에서 이용했는지를 입증하는 컴플라이언스 증빙 자료로 활용될 수 있음
- » 이를 통해 데이터 스페이스 전반의 신뢰성과 투명성이 유지됨. 이력 정보는 데이터 접근 요청, 정책 판단 결과, 실제 집행 내역, 처리 방식, 이용 종료 시점 등 데이터 생명주기 전반에서 자동 생성되며 각 이력에는 주체 식별 정보, 이용 목적, 적용된 정책 조건, 신뢰수준, 시간 정보 등이 함께 결합되어 기록되며, 이를 통해 해당 이용이 데이터 스페이스별 컴플라이언스 기준과 이용 정책을 준수했는지를 사후에 검증할 수 있음



수집된 이력은 중앙 로그 저장소 또는 이력 관리 허브(Provenance Hub)에 집계되며, 권한을 가진 주체는 이를 전용 대시보드나 감사 인터페이스를 통해 조회할 수 있음. 데이터 제공자는 자신의 데이터가 어떻게 활용되었는지 확인할 수 있고, 운영자는 정책 위반 여부를 점검할 수 있으며, 감독·규제 기관은 필요한 범위 내에서 컴플라이언스 증빙 자료로 활용할 수 있음

기록·이력 관리는 데이터 스페이스의 성숙도와 리스크 수준에 따라 선택적으로 강화 될 수 있음. 예를 들어 고위험 데이터나 규제 민감 데이터의 경우, 주요 이력을 분산 원장이나 서명 기반 저장소에 기록하여 변경 불가능성을 강화할 수 있으며, 반대로 저위험 데이터의 경우 일반 로그 저장소 기반으로 운영할 수 있다. 또한 정책 준수 여부를 자동으로 점검하는 컴플라이언스 검사 규칙이나, 위반 가능성을 탐지하는 모니터링 기능을 선택적으로 연계할 수 있음

Step 10. 데이터 거래 정산·분쟁·사후관리

데이터 거래 정산은 데이터 제공자와 이용자 간 체결된 P2P 데이터 계약을 기반으로 수행되며, 계약에서 정의된 이용 조건과 실제 데이터 이용 이력을 연계하여 정산이 이루어짐

정산은 데이터 접근 및 활용 과정에서 발생한 이용 결과를 기반으로 수행되며 데이터 스페이스 초기 단계에서는 데이터 이용 행위 자체(접근, 처리, 호출 등)에 대한 정산 중심으로 운영

- » 이 과정에서 데이터 접근량, 처리 방식, 이용 기간, 호출 횟수 등의 이용 이력 정보가 정산 기준으로 활용되며, 정산 결과는 이력관리 체계(PROV 기반)에 기록되어 투명하게 관리함
- » 데이터 스페이스는 궁극적으로 데이터 특성 및 활용 목적에 따라 다음과 같은 3가지 정산 모델을 선택적으로 적용할 수 있도록 설계 지향

① **이용 기반 정산 모델 [기본 적용]**

- 데이터 접근량, API 호출 횟수, 처리 건수 등과 같은 실제 이용 행위 기반 정산으로, 데이터 스페이스 초기 단계에서 가장 기본적으로 적용되는 모델임. 이 방식은 명확하고 측정 가능한 정산 기준을 가지고 자동 계측 및 정산이 가능하며, 다양한 데이터 유형에 범용 적용이 가능함

② **구독·계약 기반 정산 모델**

- 일정 기간 또는 계약 조건을 기준으로 정산을 수행하며 정액 또는 구독형 데이터 서비스에 적합함. 이 방식은 안정적인 수익 구조 확보가 가능하고 반복적 데이터 이용 환경에 적합하며 계약 조건 중심 정산 구조를 가지고 있음

③ **성과·부가가치 기반 정산 모델 [확장 적용]**

- 데이터 활용 결과로 발생한 부가가치 또는 성과를 기준으로 정산하는 방식으로 AI 학습, 분석 결과, 서비스 매출 등과 연계할 수 있음. 이 방식은 데이터의 실질적 가치를 반영할 수 있으며 데이터 제공자와 이용자 간 수익 공유 구조로 고도화된 데이터 생태계에서 적용 가능함

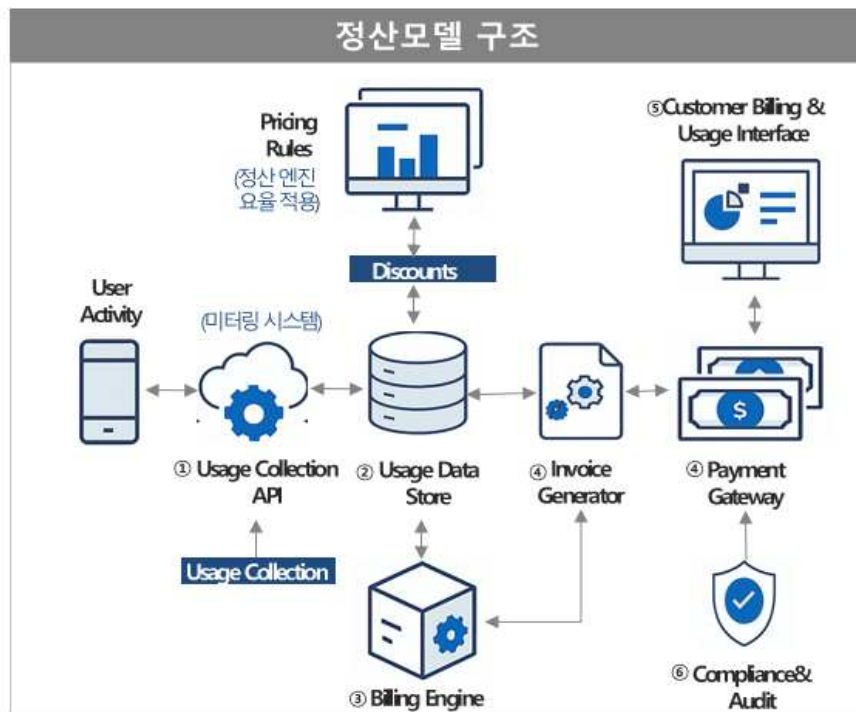
- » 분쟁이 발생할 경우에는 접근·처리·활용 전 과정에서 생성된 데이터 이용 이력(PROV 기반), 정책 검증 로그(PDP/PEP), 데이터 계약 조건이 객관적 판단 근거로 활용될 수 있으며, 이를 통해 이용행위의 사실관계를 명확히 확인할 수 있으며, 분쟁 해결 과정에서 주관적 해석이나 개별 협의에 대한 의존을 최소화함
- » 필요 시 독립적인 클리어링하우스(Clearinghouse) 또는 분쟁 조정 주체가 개입하여, 정산과 검증, 분쟁 조정, 책임 판단을 수행할 수 있음
- » 사후관리 단계에서는 반복적 정책 위반이나 계약 불이행 여부를 기반으로 참여 주체의 신뢰 수준을 재평가하고, 위반 정도에 따라 신뢰 수준 하향 조정, 접근 범위 제한,

조건부 참여 전환 등의 조치가 단계적으로 적용될 수 있으며, 이를 통해 데이터 스페이스의 신뢰를 지속적으로 유지함. 이러한 사후관리 결과는 이후 온보딩, 정책 검증, 계약 단계에서 다시 참조되어 데이터 스페이스 전반의 신뢰 계가 순환적으로 강화됨

참고 ▶ 데이터 스페이스 기반 데이터 정산 모델 예시

① 사용량 기반 정산

▶ 사용량 기반 정산 모델(Usage-based Settlement)의 아키텍처는 사용량 데이터를 기반으로 한 데이터 수집 및 저장, 요금 정책에 따른 사용량 및 금액 산정, 청구 및 결제, 모니터링 및 알림, 보안 및 규정 준수 계층으로 구성됨



[사용량 기반 정산 모델 적용 흐름]

구성 단계	적용 예시
사용량 데이터 수집	· 데이터 접근 API 호출 횟수, 조회 건수, 다운로드 용량, 연산 시간(GPU/CPU) 자동 기록
사용량·요금 산정	· 조회 1,000건당, GB당, 연산 분당 등 요금 정책 적용해 금액 자동 계산
청구·결제	· 월별 사용 내역 집계 후 기관·기업별 자동 청구 및 정산
모니터링·알림	· 사용량 임계치 초과 시 실시간 알림, 예상 비용 대시보드 제공
보안·준수	· 이용 목적·권한 기반 사용량 검증, 로그 암호화 및 감사 추적

② 기여도 기반 인센티브 정산

▶ 데이터품질, 신뢰성, 활용도, AI 학습 성능 개선 기여도 등 지표기반으로 보상을 분배하는 기여도 기반 인센티브 분배 모델(Contribution-based Incentive)은 기여도에 따라 보상을 차등 지급하는 모델로 적용 데이터 유형은 공동 AI 학습용 데이터, 연합 데이터셋, 오픈 이노베이션 프로젝트 등에서 적용

구성 단계	적용 예시
지표 정의	· 데이터 품질, 표준 준수율, 신뢰성, 활용도, AI 학습 성능 개선 기여
기여도 측정	· AI 학습 전·후 모델 성능 변화(정확도·손실·재현율 등)와 파라미터 업데이트 영향도를 결합하여 데이터·참여자별 기여도 점수 산정
보상 산정	· 기여도 점수에 가중치를 적용해 인센티브·수익 차등 산정
정산·분배	· 데이터 제공자·이용자·모델 개발자 간 기여 비율에 따른 정산
투명성·신뢰	· 기여도 산정 기준·가중치 공개, 학습 로그·결과 기반 검증 및 감사 추적

예시 시나리오 : 「공동 AI 모델 개발 프로젝트」
 ① 데이터 제공자 A, B, C가 각기 다른 데이터셋 제공
 ② 동일 모델에 순차적으로 학습 참여

① AI 학습 성능 변화 측정

구분	기준 모델	A 데이터 반영	B 데이터 반영	C 데이터 반영
정확도(Accuracy)	78.0%	80.5%	82.0%	82.3%
손실(Loss)	0.62	0.55	0.49	0.48
성능 개선 기여(▲)	-	+2.5%	+1.5%	+0.3%

② 파라미터 영향도 반영 (정성 → 정량)

참여자	파라미터 업데이트 영향	반영 방식
A	큼	성능 개선 폭과 일관성 반영
B	중간	성능 개선 지속성 반영
C	낮음	미세 개선으로 최소 가중치 적용

③ 기여도 점수 산정 (정규화)

참여자	성능 개선 기여	영향도 가중	기여도 점수
A	2.5	1.0	50점
B	1.5	0.9	35점
C	0.3	0.8	15점
합계	-	-	100점

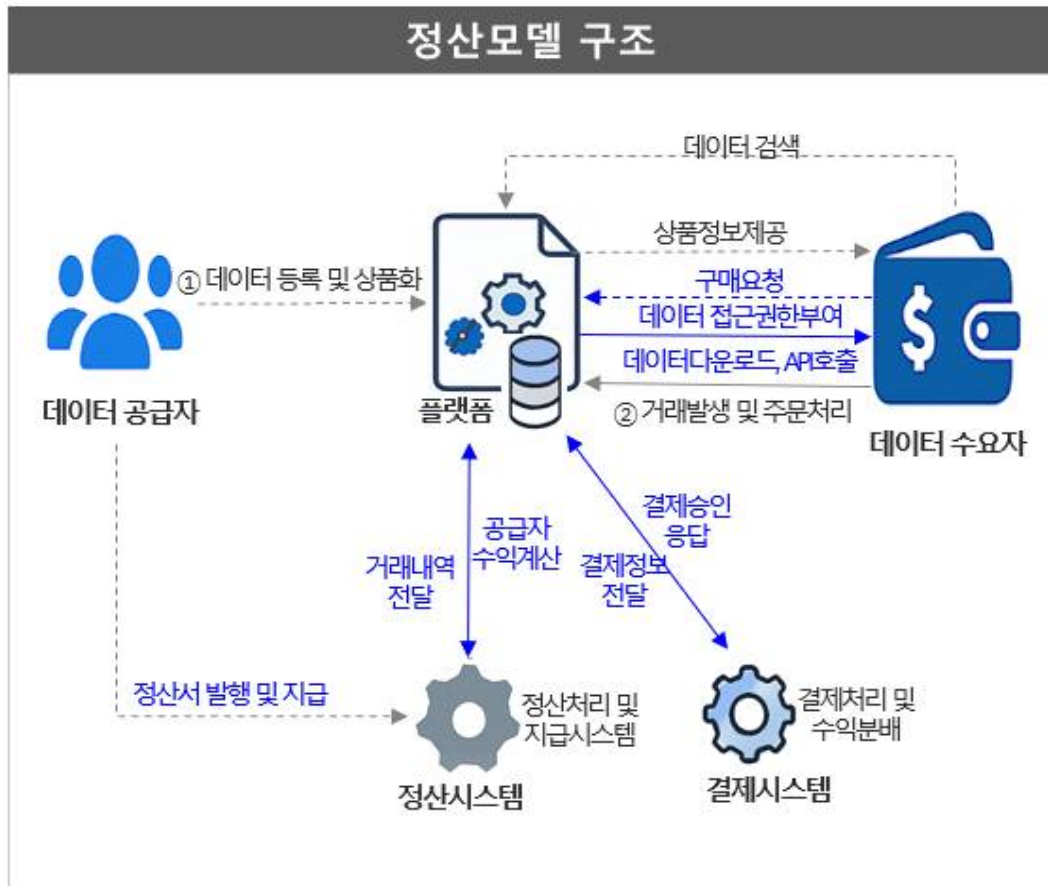
④ 인센티브 정산 예시

참여자	기여도 비율	정산 금액
A	50%	5,000만 원
B	35%	3,500만 원
C	15%	1,500만 원

③ 데이터 주체 간 협의 모델

▶ 데이터 제공자가 계약조건을 등록하면 중개플랫폼이 계약을 체결하고 API를 통해 데이터를 전달하고 정산 모듈이 사용량을 측정하고 스마트계약 등 방식에 따라 수익을 배분하며 블록체인에 거래 내역 저장으로 투명성 확보

구성 단계	적용 예시
계약조건 등록	· 데이터 제공자가 이용 목적, 사용 범위, 요율, 수익배분 조건을 플랫폼에 등록
계약 체결	· 중개플랫폼이 표준 계약 템플릿 기반으로 이용자와 계약 자동 체결
데이터 전달	· 인증된 이용자에게 API를 통해 데이터 접근·전달
사용량 측정	· API 호출 횟수, 처리량, 이용 시간 등 사용량 자동 집계
정산·배분	· 스마트계약 등 방식으로 사용량 기반 수익 자동 배분
투명성 확보	· 거래·정산 내역을 블록체인에 기록하여 추적·검증 가능

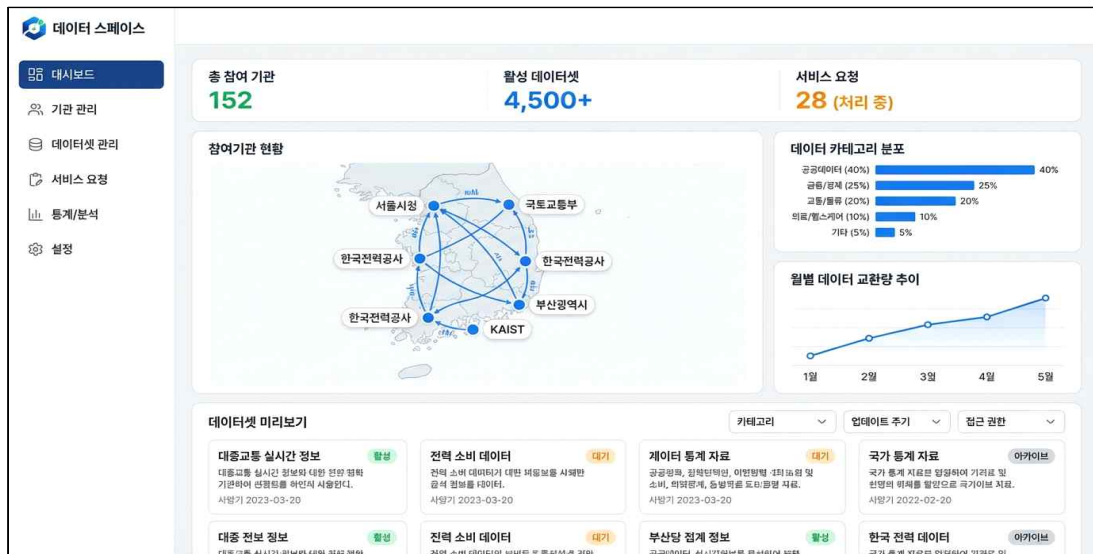


[데이터 주체 간 협의 모델]

2.4.3 데이터 스페이스 공통 서비스

○ 데이터 스페이스 포털

- ▶ 데이터 스페이스 포털은 다양한 이해관계자 데이터 스페이스에 처음 접속하여 전체 구조와 제공 기능을 이해하고, 실제 이용으로 이어지도록 지원하는 통합 진입 창구임. 포털은 데이터 스페이스의 목적, 운영 원칙, 참여 주체 간 역할과 책임을 명확히 제시함으로써, 이용자가 데이터 스페이스의 개념과 활용 방식을 직관적으로 이해할 수 있도록 함. 이를 통해 단순한 시스템 접속을 넘어, 데이터 스페이스에 대한 공통된 인식을 형성하는 역할을 수행함



[데이터 스페이스 포털 예시]

- ▶ 또한 포털은 참여 절차를 안내하고, 기관 및 사용자 등록, 서비스 신청, 데이터·서비스 탐색 등 주요 정보를 제공함. 분산된 데이터와 서비스가 각기 다른 위치에 존재하더라도, 포털을 통해 하나의 환경처럼 접근할 수 있도록 구성되며, 이용자는 포털을 중심으로 데이터 검색, 이용 신청, 접근 현황 확인 등을 수행할 수 있음
- ▶ 아울러 참여자 간 신뢰 기반 운영을 지원하는 관리 기능도 함께 수행함. 이용자는 자신의 참여 상태, 권한, 이용 이력 등을 확인할 수 있으며, 운영자는 참여 현황과 서비스 운영 상태를 종합적으로 관리할 수 있음. 이를 통해 포털은 단순한 화면 제공을 넘어, 데이터 스페이스 운영과 이용을 연결하는 핵심 허브로서 기능함

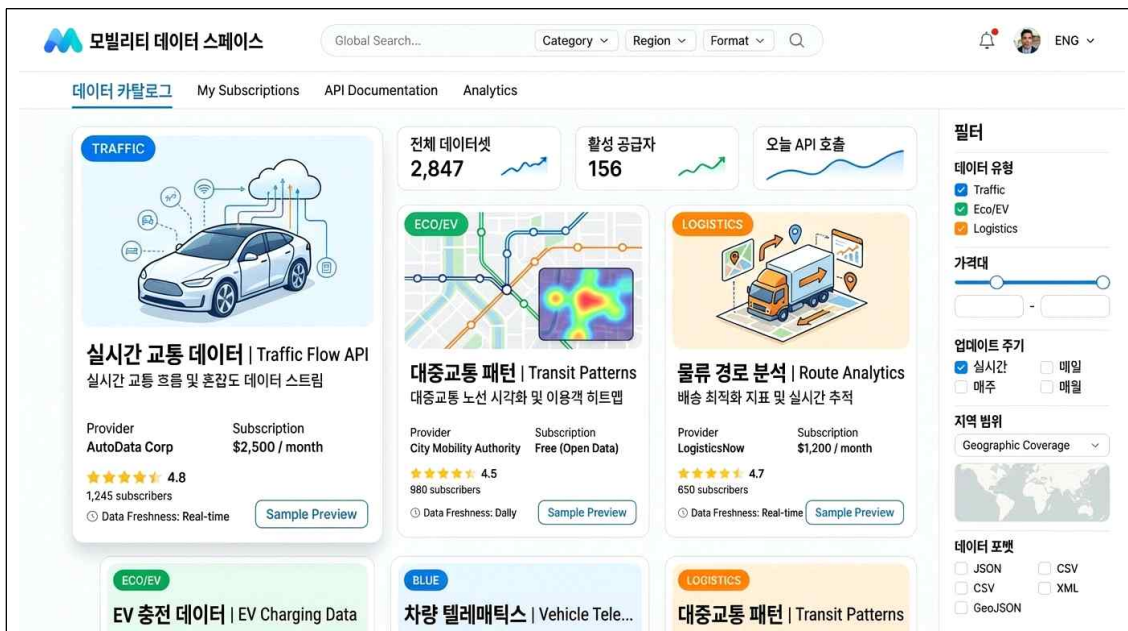
○ 데이터 마켓플레이스

» 데이터 마켓플레이스는 데이터 스페이스 내에 분산된 원천데이터와 이를 기반으로 생성된 고가치 파생데이터를 대상으로, 데이터의 물리적 이전이나 다운로드 없이 안전한 접근과 활용을 지원하는 유통 수단임

* 데이터와 데이터 상품은 「국가데이터인프라 데이터 카탈로그 표준 가이드」를 준수한 카탈로그 형태로 제공되어, 데이터의 내용, 활용 조건, 접근 방식 등을 일관된 기준으로 확인할 수 있음

» 데이터 제공자는 Compute-to-Data 및 원격 연산 기반 활용을 통해 분석·가공·모델 학습 결과를 데이터 상품 또는 파생 데이터 형태로 제공할 수 있으며, 모든 접근과 연산은 사전 인증과 정책 기반 통제를 거쳐 수행됨

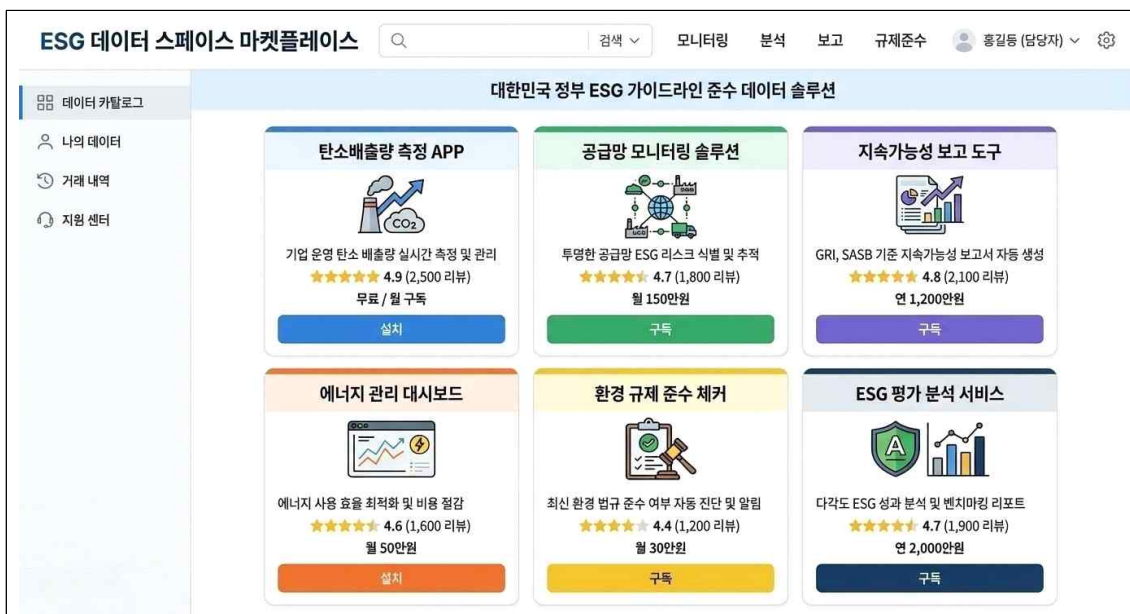
» 또한 마켓플레이스에서 유통되는 데이터 자산은 소재 정보 중심으로 국가데이터통합플랫폼 (One-윈도우)과 연계되어, 데이터 주권을 유지하면서도 신뢰 기반 데이터 거래 지원



[데이터 스페이스 기반 데이터 마켓플레이스 예시 (모빌리티 분야)]

○ 앱 서비스 마켓플레이스

- ▶ 앱 서비스 마켓플레이스는 데이터 스페이스 참여자가 데이터를 실제 분석, 운영, 의사결정, 서비스 개발에 활용할 수 있도록 관련 애플리케이션과 서비스를 게시·탐색·선택·연계하는 환경임. 이 과정에서 서비스의 등록 및 유통은 데이터 스페이스별 컴플라이언스와 신뢰 프레임워크에 따라 관리되며, 기준을 충족한 서비스가 우선적으로 활용될 수 있도록 운영할 수 있음
- ▶ 데이터 스페이스 내 서비스 제공자는 데이터 처리, 계산, 분석, 검증, 리포팅 등 특정 기능을 애플리케이션 또는 SaaS 형태로 제공하고, 다른 참여 주체들은 별도의 시스템 구축이나 복잡한 개발 없이 필요한 기능을 선택해 이용할 수 있음
- ▶ 이러한 서비스는 데이터 스페이스의 공통 데이터와 참여 주체의 자체 데이터를 연계하여 작동하도록 설계되며, 이용자는 자신의 데이터만 연결하면 필요한 업무 처리를 즉시 수행할 수 있음. 이를 통해 데이터 스페이스는 데이터 교환 중심 구조에서 나아가, 실무 기능을 제공하는 활용 중심 환경으로 확장될 수 있음
- ▶ 또한 서비스 이용 과정에서 발생하는 수익 일부를 수수료 형태로 환수해 데이터 스페이스 운영과 기능 고도화에 재투자함으로써, 공공 재정에 대한 의존도를 낮추고 데이터 스페이스가 지속적으로 확장·진화할 수 있는 자생적 운영 구조를 마련할 수 있음



[데이터 스페이스 기반 APP·서비스 마켓플레이스 예시(ESG 분야)]

○ 워크스페이스(AI Lab)

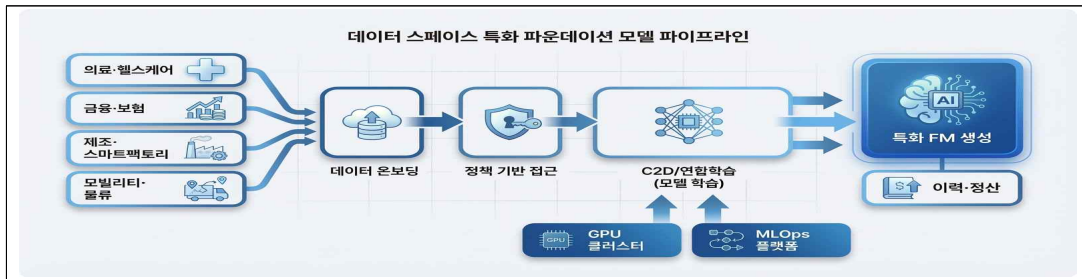
[워크스페이스(AI Lab) 구성 요소(안)]

구분	구성 요소	주요 기능 및 역할
표준 관리	공통 데이터 모델 연계	AAS·도메인 표준 기반 데이터 구조 적용
연산 환경	컨테이너 기반 실행 환경	분석·학습 작업을 격리된 환경에서 실행
연산 자원	GPU as a Service	AI 학습·추론을 위한 GPU·고성능 연산 자원 활용
분석·학습	AI 모델 학습 환경	파운데이션 모델 튜닝, 도메인 특화 모델 학습
MLOps	모델 학습·실험 관리	모델 버전 관리, 하이퍼파라미터 추적, 실험 이력 관리
	모델 레지스트리 및 레포지토리 관리	검증된 모델 저장·공유·재사용
	모델 검증·평가	성능 비교, 검증 시나리오 실행, 기준 충족 여부 판단

- » 워크스페이스(AI Lab)는 별도의 물리적 공간이나 개별 기관의 서버를 의미하는 것이 아니라, 데이터 스페이스 안에서 온톨로지 기반의 의미 체계 위에 데이터 활용과 AI 개발을 실행할 수 있도록 제공되는 기능형 실행 환경임. 워크스페이스를 통해 데이터 스페이스 참여자는 단순히 데이터를 조회하거나 분석하는 수준을 넘어, AI 모델·디지털 트윈·데이터 기반 서비스로 이어지는 실질적인 부가가치를 창출할 수 있음
- » 워크스페이스의 온톨로지는 참여 주체 간에 데이터의 개념·관계·맥락이 일관되게 해석되도록 하는 공통 의미 기반임. 서로 다른 도메인·기관에서 생산된 데이터가 단순한 포맷 변환이 아닌 의미적 정렬을 통해 통합되고, AI 모델 학습 시 피처의 의미가 명확히 정의되며, 데이터 카탈로그의 자산 분류와 이종 데이터 스페이스 간 교환이 공통 어휘 위에서 수행됨
- » 표준화되고 신뢰가 확보된 데이터를 온톨로지가 제공하는 의미 규칙에 따라 반복적인 모델 학습과 검증에 활용할 수 있으며, 이를 통해 고도화된 의사결정 지원, 설계·공정 최적화, 규제 대응 자동화, 신규 데이터 서비스 개발 등으로 활용 범위를 확장할 수 있음
- » 특히 개인화된 또는 주체 간 합의된 파이프라인을 통해 데이터 활용이 지속적으로 축적·고도화되면서, 데이터는 일회성 자원이 아닌 장기적인 경쟁력과 수익 창출의 기반으로 전환될 수 있음. 이러한 부가가치를 가능하게 하는 워크스페이스의 기본 기능과 실행 환경은 데이터 스페이스 운영 주체가 공통 인프라로 제공하는 것이 바람직함

- » 온톨로지 관리, 데이터 표준화, 가명처리, AI 학습·검증 핵심 기능은 공통 서비스로 구축하여 참여자의 진입 장벽을 낮추고, 실제 서비스 개발이나 도메인 특화 모델·응용 서비스(디지털트윈, 공정 시뮬레이션 등)는 민간 기업·연구기관 등 참여 주체가 공통 온톨로지 위에 도메인 어휘를 확장하여 주도적으로 수행하는 역할 분담이 필요함

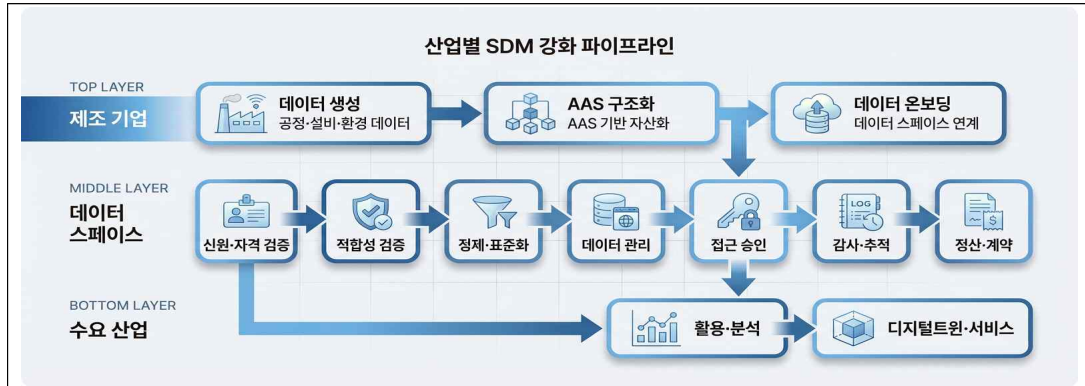
[예시 ① 분야별 데이터 스페이스 특화 파운데이션 모델 강화 파이프라인]



▶ 데이터 스페이스는 특정 도메인 데이터를 안전하게 활용할 수 있는 기반을 제공함으로써, 도메인 특화 파운데이션 모델이나 AI 모델의 학습·고도화·활용을 지원하는 환경으로 기능할 수 있음

단계	구분	주요 주체	주요 내용
1	데이터 준비	데이터 제공자	도메인 원천데이터를 기존 환경에 유지한 채 데이터 스페이스 참여
2	메타데이터 등록	데이터 제공자	도메인 특성·활용 조건·허용 연산 방식 메타데이터 등록
3	데이터 탐색	데이터 이용자	카탈로그를 통해 데이터 및 모델 학습 가능 여부 확인
4	활용·학습 요청	데이터 이용자	도메인 특화 모델 학습·강화 목적의 연산 요청 제출
5	인증·자격 검증	데이터 스페이스	이용 주체 신원·자격 검증
6	정책 검토·승인	데이터 스페이스	목적·민감도·규제 요건 기반 접근·학습 승인
7	연산·학습 전달	데이터 스페이스 → 제공자	승인된 분석·학습 작업 제공자 환경으로 전달
8	모델 학습 실행	데이터 제공자	C2D·연합학습 등으로 제공자 환경에서 학습 수행
9	결과 통제	데이터 제공자	허용된 모델 파라미터·지표만 반출
10	모델 검증·평가	데이터 이용자	반환 결과 기반 모델 성능 검증
11	기록·감사	데이터 스페이스	접근·학습 결과 전 과정 이력 기록
12	정산·보상	데이터 스페이스	사용량·계약 조건 기반 정산 수행

[예시 ② 산업별 Software-Defined 제조 강화 파이프라인]



- ▶▶ 제조·산업 전반에서 생성되는 데이터를 AAS 등 표준 구조로 자산화하고, 데이터 스페이스를 통해 정제·표준화·연계·활용하는 전 과정을 자동화된 흐름으로 통합
- ▶▶ 본 파이프라인은 참여 주체에 대한 인증과 데이터·시스템의 적합성 검증을 전제로, 데이터가 이동되지 않은 상태에서 안전한 분석과 서비스 활용을 가능하게 하며, 제조 데이터를 설계·시뮬레이션·규제 대응 등 다양한 산업 서비스로 확장할 수 있도록 지원

단계	구분	주체	주요 내용
1	데이터 생성	제조 기업	공정·설비·환경 데이터 생성
2	구조화	제조 기업	AAS 기반 데이터 자산화
3	신원·자격 검증	데이터 스페이스	참여 기업·시스템에 대한 신원 인증 및 참여 자격 확인
4	데이터 온보딩	제조 기업	인증된 주체만 데이터 스페이스 연계
5	적합성 검증	데이터 스페이스	데이터·시스템이 공통 보안·정책 기준 충족 여부 확인
6	정제·표준화	데이터 스페이스	표준 규칙에 따른 자동 처리
7	데이터 관리	데이터 스페이스	표준 데이터 자산 카탈로그 관리
8	접근 승인	데이터 스페이스	이용 목적·역할 기반 접근 권한 승인
9	활용·분석	수요 산업	승인된 방식으로 분석·연산 수행
10	디지털트윈·서비스	수요 산업	설계·시뮬레이션·규제 대응
11	감사·추적	데이터 스페이스	데이터 이용·연산 이력 기록
12	정산·계약	데이터 스페이스	사용량·계약 조건 기반 정산

신뢰 기반 데이터 유통 거래 활성화를 위한 한국형 데이터 스페이스 참조모델 v1.0

■ 발 행 : 2026. 3.

■ 발행인 : 황 종 성

■ 발행처 : 한국지능정보사회진흥원(NIA)

인공지능데이터본부 국가데이터인프라팀

■ 문의 : 오대훈 선임연구원(ohdang93@nia.or.kr), 053-230-4243
이윤수 선임연구원(ys@nia.or.kr), 053-230-4204