

AI·디지털 전환에 따른 공공부문 데이터 거버넌스 전략

A Study on Strategies for Public Sector Data
Governance in the Era of AI and Digital
Transformation

2025. 12

주관연구기관 : 정보통신정책연구원

연구기관 : (사)한국정책학회



과학기술정보통신부



정보통신기획평가원

ICT 진흥 및 혁신기반 조성사업 RS-2025-02173110

정책연구 25-28-09

AI·디지털 전환에 따른 공공부문 데이터 거버넌스 전략

(A Study on Strategies for Public Sector
Data Governance in the Era of AI and Digital
Transformation)

남태우 외

2025. 12

주관연구기관 : 정보통신정책연구원
연구기관 : (사)한국정책학회



과학기술정보통신부



정보통신기획평가원

이 보고서는 2025년도 과학기술정보통신부 방송통신발전기금 ICT
진흥 및 혁신기반 조성사업의 연구결과로서 보고서 내용은 연구자의
견해이며, 과학기술정보통신부의 공식입장과 다를 수 있습니다.

제 출 문

과학기술정보통신부 장관 귀하

본 보고서를 『디지털 대전환 메가트렌드 연구: AI·디지털 전환에 따른 공공부문 데이터 거버넌스 전략』의 연구결과보고서로 제출합니다.

2025년 12월

연구기관: (사)한국정책학회

총괄책임자: 남태우 교수

참여연구원: 기남훈 부교수

김준형 조교수

황한찬 조교수

신승윤 선임연구원

김기동 조교수

연구보조원: 하선우 연구원

조수향 연구원

김정원 연구원

목 차

요약문	xi
제 1 장 서 론	1
제 1 절 연구의 배경과 필요성	1
1. 연구의 배경	1
2. 연구의 필요성	5
제 2 절 연구 목적과 내용	10
1. 연구의 목적	10
2. 연구의 주요 내용	10
제 3 절 연구 방법과 추진체계	12
1. 연구방법과 분석의 틀	12
2. 연구 추진체계와 전략	14
제 2 장 이론적 고찰	17
제 1 절 데이터 거버넌스의 개념과 이론적 틀	17
1. 데이터 거버넌스 개념의 이론적 정립과 체계화	17
2. 데이터 거버넌스 모형과 프레임워크	30
제 2 절 주요국의 데이터 거버넌스 사례	53
1. 주요국의 AI 거버넌스 정책과 데이터 거버넌스 모형	53
2. 국제기구 차원의 데이터 거버넌스 프레임워크	63
3. 빅데이터 거버넌스의 유형별 실행 모델	68
제 3 절 데이터 거버넌스의 성공요인과 실행 전략	71
1. 데이터 거버넌스 성공요인	71
2. 데이터 거버넌스의 실행 전략	77

제 4 절	데이터 거버넌스의 사회적·윤리적 논의	94
1.	데이터 활용과 프라이버시 보호의 윤리적 균형	94
2.	프라이버시 중심 데이터 설계 원칙	95
3.	디지털 격차와 데이터 불평등 해소	97
4.	신뢰 형성과 수용성 확보를 위한 투명한 거버넌스	99
5.	글로벌 규범과 윤리 거버넌스의 국제화	100
제 5 절	소 결	103
1.	이론적 고찰 주요 내용	103
2.	이론적 고찰 시사점	106
제 3 장	공공기관 데이터 거버넌스 실제에 대한 탐색	107
제 1 절	데이터	108
1.	데이터 관리와 활용 현황	108
2.	주요 문제와 과제	117
3.	시사점	123
제 2 절	기술과 시스템	132
1.	데이터 활용에 있어서 사용되는 기술(AI, 통계패키지, 시각화툴)	132
2.	데이터 활용에 있어서 사용되는 시스템(플랫폼, 대시보드, 클라우드 등)	134
3.	기술과 시스템 차원에 있어서 장애요인과 성공요인	137
제 3 절	조직과 관리	142
1.	조직 구성원의 인식과 참여 수준	142
2.	리더십과 조직문화	146
3.	전문 인력과 역량 강화	154
제 4 절	법률과 제도	160
1.	데이터기반행정법 등 관련 법제도의 영향력	160
2.	법제도와 실무 간의 격차	162
3.	당위론적 방향성과 실제적 현실 간의 괴리	163
4.	법제도 차원의 장애요인과 성공요인	165

제5절 소 결	170
1. 데이터	170
2. 기술과 시스템	172
3. 조직과 관리	173
4. 법률과 제도	175
제4장 공공기관 데이터 거버넌스의 전략	178
제1절 프레임워크 및 전략 도출	178
1. OECD 기본 프레임의 수정과 고도화	178
2. 구체적인 거버넌스 전략	185
3. 데이터 거버넌스-AI 거버넌스 연계 프레임워크	188
4. 거버넌스의 층위화 multi-layered 프레임워크	194
5. 중앙-지방 역량 격차 해소 전략	196
제2절 프레임워크 전략에 대한 전문가 검증 및 보완	199
1. 대상자 및 실시 내용	199
2. 전문가 자문 내용	200
제5장 결 론	214
제1절 연구 결과	214
1. AI 시대 데이터 거버넌스의 진화와 핵심 성공 요인	214
2. 공공데이터 거버넌스 실태 진단과 주요 문제점	215
3. 통합 데이터 거버넌스 프레임워크 제시	215
4. 데이터-AI 거버넌스 연계 체계	216
제2절 정책적 시사점	219
1. 제도적 차원의 시사점	219
2. 조직적 차원의 시사점	220
3. 데이터 관리 시사점	221
4. 데이터 거버넌스 체계 구축 전략	222

참고문헌	224
[부록 1] 공공부문 데이터거버넌스 실제에 대한 서면 인터뷰 프로토콜	237
[부록 2] 공공부문 데이터거버넌스 프레임워크에 대한 서면 인터뷰 프로토콜	244

표 목 차

〈표 1-1〉 행정수단과 데이터 생애주기	8
〈표 2-1〉 AI 데이터 거버넌스의 범위	24
〈표 2-2〉 주요국 데이터정책 논의 비교	57
〈표 2-3〉 OECD AI 원칙	66
〈표 2-4〉 OECD 민첩한 AI 거버넌스를 위한 접근 방식	67
〈표 4-1〉 데이터 거버넌스 수행전략	187
〈표 4-2〉 데이터 거버넌스와 AI 거버넌스 간 공통점/차이점 정리	190
〈표 4-3〉 데이터 거버넌스-AI 거버넌스 연계 프레임워크	192

그 립 목 차

[그림 1-1] OECD 데이터 거버넌스 모델	2
[그림 1-2] 제2차 데이터기반행정 활성화 추진전략 및 10대 추진과제	4
[그림 1-3] 데이터 플랫폼 활용의 영향요인	6
[그림 1-4] 연구분석 틀	12
[그림 1-5] 제약요인 탐색 방법 예시	13
[그림 2-1] 전자정부, e-거버넌스, 빅데이터 거버넌스	20
[그림 2-2] AI 데이터 거버넌스의 핵심 요소(1)	23
[그림 2-3] AI 데이터 거버넌스의 핵심 요소(2)	25
[그림 2-4] 데이터 거버넌스의 문제 삼각형(Problem Triangle of Data Governance) ..	26
[그림 2-5] 대규모 언어 모델(LLMs)을 위한 데이터 거버넌스의 주요 과제	28
[그림 2-6] 전자정부, e-거버넌스, 빅데이터 거버넌스	30
[그림 2-7] ‘제도-조직-기술’ 3분법 분석틀	31
[그림 2-8] 데이터 생애주기 단계별 책임주체, 통제 메커니즘, 품질 기준	32
[그림 2-9] AI 기반 BDAS(Big Data Algorithmic Systems)	32
[그림 2-10] DMBOK(Data Management Body of Knowledge) Framework	33
[그림 2-11] Big Data Governance Framework	34
[그림 2-12] 빅데이터 거버넌스 유형화	35
[그림 2-13] A model of social data governance	35
[그림 2-14] Data governance modeling Platform society	36
[그림 2-15] Governance design by organizational/institutional perspective	37
[그림 2-16] 통합적 거버넌스 프레임워크	38
[그림 2-17] The First Step For Growth: Data Governance	39
[그림 2-18] Precision e-Policy Cycle	40

[그림 2-19] 캐나다의 국가 데이터 프레임워크 모델	41
[그림 2-20] EU의 GAIA-X 생태계 구성요소	42
[그림 2-21] 한국의 국가 데이터 카탈로그 목표 모델	43
[그림 2-22] 범정부 국가지식플랫폼	43
[그림 2-23] 데이터 미들 플랫폼 기반 프레임워크	44
[그림 2-24] AI Governance Framework	45
[그림 2-25] 전통 데이터 거버넌스로부터 데이터 & AI 거버넌스로의 변화	46
[그림 2-26] AI 거버넌스 계층구조	47
[그림 2-27] 지속가능한 AI를 위한 데이터 거버넌스 프레임워크	48
[그림 2-28] Open Data Governance Design	49
[그림 2-29] The 6-Step model in Open Government Data(OGD) Model design	49
[그림 2-30] Non-Invasive Data Governance(NIDG) 모델에서의 데이터 관리	50
[그림 2-31] 프론티어 데이터 거버넌스(Frontier data governance)에서의 데이터 관리	50
[그림 2-32] 데이터 분석 거버넌스	52
[그림 2-33] 2024년 IMD 세계 디지털 경쟁력 순위(1~15위)	53
[그림 2-34] 2023년 AI 민간투자 규모	54
[그림 2-35] 캐나다의 AI 거버넌스 조직도	59
[그림 2-36] 영국의 AI 데이터 거버넌스의 핵심 요소	60
[그림 2-37] 영국의 AI 거버넌스 조직도	61
[그림 2-38] 뉴질랜드 데이터 거버넌스 체계	63
[그림 2-39] OECD 데이터 거버넌스 모델	65
[그림 2-40] 국민건강주의 알람 서비스	68
[그림 2-41] 프레드폴(Predpol) 시스템	69
[그림 4-1] 통합 데이터 거버넌스 모형	180
[그림 4-2] 데이터 거버넌스 활성화를 위한 조건-이슈-제언 연계도	181
[그림 4-3] 데이터 거버넌스 활성화를 위한 로드맵	184
[그림 4-4] 데이터 거버넌스의 확장층으로서의 AI 거버넌스	194

요 약 문

1. 제 목

AI·디지털 전환에 따른 공공부문 데이터 거버넌스 전략

2. 연구 배경 및 목적

인공지능과 디지털 기술의 발전은 행정과 공공서비스의 디지털 전환을 가속화하며 데이터 기반 행정의 중요성을 높이고 있다. 정부는 공공데이터를 활용한 맞춤형 서비스 제공과 행정 효율성 제고를 위해 다양한 디지털 혁신 정책을 추진하고 있으나, 이러한 정책이 실질적 성과로 이어지기 위해서는 이를 뒷받침할 체계적 데이터 거버넌스의 확립이 전제되어야 한다. OECD와 주요국은 데이터 전 주기 관리, 개방성과 통제의 균형, 이해관계자 조정 등을 강조하며 통합 거버넌스 모델을 제시하고 있고, 국내에서도 데이터 상업적 활용 근거와 제도 강화가 필요하다는 지적이 나온다. 이를 극복하기 위해서는 한국적 상황에 맞춘 데이터 거버넌스 체계를 마련하고 궁극적으로 데이터 품질·표준화·개인정보 보호를 포함한 통합 운영 체계와 협력 인센티브 구조를 확립하는 것이 AI 기반 정책 실효성과 맞춤형 서비스 구현의 핵심 조건이 된다.

3. 연구의 구성 및 범위

본 연구는 공공부문 데이터 관리체계 강화를 위해 실행 전략 프레임워크를 제시하는 데 목적을 두고, 공공기관 간 데이터 연계 수준과 활용 현황을 진단하며 데이터 기반 정책 사례를 분석해 문제점을 도출한다. 아울러 미국·영국·독일·캐나다 등 선진국의 거버넌스 모델과 OECD의 3단계·6대 요소 프레임워크를 검토해 국내 적용 가능성과 시사점을 도출

하고 이를 기반으로 리더십, 법·제도, 조직 구조, 표준 및 관리 체계, 기술 인프라를 핵심으로 한 한국형 데이터 거버넌스 체계 구축의 전략을 제시한다.

4. 연구 내용 및 결과

가. 이론적 고찰

1) 데이터 거버넌스의 개념과 이론적 틀

데이터 거버넌스는 전통적 관리 개념에서 출발하여 e-거버넌스를 거쳐, AI 시대에는 통합적·참여적 관리 체계로 확장되고 있다. 초기에는 데이터 품질·표준화·보안·책임 분담 등 관리 요소 중심이었으나, 최근에는 네트워크형·플랫폼형·참여형 모델로 발전하며 사회적·윤리적 관점까지 포괄한다. 특히 AI 확산은 편향 완화, 공정성, 프라이버시, 추적성, 안전한 모델 배포 등 7대 핵심 범주를 요구하며, 알고리즘 투명성과 책임성이 새로운 과제로 부각되고 있다. OECD·EU는 이를 제도화하며 국가 차원의 데이터 거버넌스 프레임워크를 권고하고 있다.

즉, 데이터 거버넌스는 기본 관리 구조에서 출발해 참여·기술 융합·국제 제도화·AI 기반 통합으로 진화하는 다층적 체계라 할 수 있다.

2) 주요국의 데이터 거버넌스 사례

미국은 연방 차원에서 CDO(Chief Data Officer) 제도를 도입하여 기관별 책임성을 강화하고 데이터 개방과 활용을 촉진하였다. 영국은 국가데이터전략(National Data Strategy)을 수립해 민간과 공공의 협력 구조를 제도화하고 투명성과 신뢰성을 확보하였다. 독일은 엄격한 데이터 보호 규제를 바탕으로 산업 데이터 공유 플랫폼(Gaia-X)을 추진하여 보안과 활용의 균형을 추구하였다. 캐나다는 디지털 정부 전략을 통해 국민 참여형 데이터 서비스 확대와 기관 간 데이터 연계 강화를 추진하였다.

3) 데이터 거버넌스의 성공요인과 실행전략

도출된 성공 요인은 다음과 같다. 첫째, 데이터 관리 전담 조직과 리더십을 통한 책임성 확보, 둘째, 데이터 활용을 뒷받침하는 법·제도의 정비, 셋째, 상호운용성을 확보하기 위

한 기술 인프라와 표준화 체계 구축, 넷째, 공공성과 투명성 확보를 통한 국민 신뢰 형성이 핵심 요인으로 꼽힌다.

이러한 성공 요인을 바탕으로 실행 전략을 정리하면, 공공부문 내에서 데이터 거버넌스의 역할을 명확히 구분하고 책임성을 부여하는 체계를 마련해야 한다. 또한 데이터의 수집·관리·활용 전 주기를 포괄하는 법·제도를 제도화하고, 공통 표준과 기술 인프라를 확충하여 기관 간 효율적 연계를 보장해야 한다. 아울러 국민 참여와 민관 협력을 강화함으로써 데이터 거버넌스의 투명성과 지속 가능성을 제고하는 것이 필요하다.

4) 사회적·윤리적 논의

데이터 거버넌스는 활용과 프라이버시 보호의 균형을 핵심 과제로 한다. GDPR·OECD 지침은 사전 위험평가, 책임 주체 명시, 이해관계자 참여를 강조한다. 개인정보·가명정보·익명정보 구분, AI 학습 데이터의 민감정보 정제, 설명가능성과 공정성 확보가 요구된다. 또한 디지털 격차 해소와 데이터 리터러시 강화, 리빙랩 기반 시민 참여가 필요하며, 국제적으로는 초국경 데이터 흐름 속 윤리 규범 정착이 강조된다. 팬데믹과 의료 AI 사례는 사전 동의와 책임 메커니즘의 중요성을 보여준다.

나. 공공기관 데이터 거버넌스의 실제에 대한 탐색

1) 데이터

기관 간 데이터 공유 부족, 표준화·호환성 미비, 중복 수집, 비정형 데이터 처리 한계가 공통 문제다. 수도권·지방 간 역량 격차, 전문 인력 부족, 수작업 오류도 심각하다. 일부 기관은 메타데이터 관리, 통합 관리로 개선을 시도하나 여전히 보안·활용 균형과 제도 기반이 미흡하다. 해외의 통합 플랫폼, 민간데이터 활용, CDO 체계 등 사례에서 개선 방향을 찾을 수 있다.

2) 기술·시스템

기관별 편차가 크며, 연계·통합 부족으로 효율성이 떨어진다. 보안 규제에 AI 분석 툴 활용이 제한되고 데이터 포맷 불일치로 품질 저하가 발생한다. 해결을 위해 통합 거버넌스 플랫폼, 사용자 친화적 UI, 자동 검증 시스템, 중앙 데이터 카탈로그 구축이 필요하다.

3) 조직·관리

구성원의 낮은 인식과 소극적 참여가 한계이며, 가시적 성과와 보상체계가 있어야 참여가 활성화된다. 리더십 부재 시 제도화가 성과와 연결되지 못한다. 또한 전문 인력 부족과 교육 한계가 문제로, CDO 제도화·전담조직 강화·실무 역량 개발이 필수적이다.

4) 법률·제도

개인정보보호법과 활용 법령 간 충돌로 현장에서는 소극적 대응이 이루어지고 있다. 법령 간 정합성 부족, 데이터 소유권 불명확, 지역 간 제도 확산 한계가 문제다. 개선을 위해 개인정보 보호와 활용의 균형, 맞춤형 규제, 가이드라인 제공, 데이터 소유권 명확화가 필요하다. 기술 발전에 맞춘 유연한 법제 체계가 핵심 조건이다.

다. 공공기관 데이터 거버넌스의 전략

1) OECD 기본 프레임의 수정과 고도화

OECD의 통합 데이터 거버넌스 모형을 수정하여 환경 층위와 제도·조직·기술 층위라는 두 가지 핵심 층위로 구조화하였다. 또한, 데이터 거버넌스 활성화를 위한 조건-이슈-제언 연계도를 제도·조직·기술 차원에서 도출하고, 이를 단기-중기-장기 로드맵과 계획·조직·위험 기반 접근으로 실행 전략화했다. 그리고 앞서 제시된 데이터 거버넌스 활성화를 위한 제언들은 실행 전략(제도기반 강화, 기술기반 강화, 협력 활성화, 역량 강화) 및 실행 시기(단기, 중기, 장기)에 따라 단기, 중기, 장기 방안들로 재구성하였다.

2) 구체적인 전략

Janssen et al.(2020)이 제시한 데이터 거버넌스 전략을 본 연구의 프레임워크와 연결하여 재구성 하였다. 구체적으로 계획 기반 접근, 조직 기반 접근, 위험 기반 접근으로 나누어서 이론적 검토와 질적 분석을 통해 연구진이 검토한 주요 결과를 정리하였다

3) 데이터 거버넌스-AI 거버넌스 연계 프레임워크

데이터 거버넌스와 AI 거버넌스 간 공통점과 차이점을 정리하고 데이터 거버넌스와 AI 거버넌스 연계 프레임워크를 제시하였다. 이는 OECD 분류체계에 따라 데이터 거버넌스와 AI 거버넌스를 4개 영역(리더십, 법·제도, 표준화, 기술 인프라)으로 구분하여 연계 방안을

제시하였다.

4) 거버넌스의 층위화 multi-layered 프레임워크

거버넌스는 전통적 거버넌스에서 출발하여 e-거버넌스, 데이터 거버넌스, 그리고 인공지능 거버넌스로 이어지는 층위적 진화를 보여준다. AI 거버넌스는 데이터 거버넌스의 원칙을 윤리·사회적 책임으로 확장하며, 모델카드·알고리즘 평가·편향 지표 등 도구를 통해 구현된다.

5) 중앙-지방 역량 격차 해소 전략

중앙정부와 지방자치단체 간 역량 차이를 거버넌스의 주요 제약으로 보고, 국가 공통 기반 구축·역할 재정립·조직 역량 강화·활용 사례 확산 등을 통해 구조적 불균형을 완화할 전략적 방향을 제시하였다. 이러한 접근은 중앙의 표준화 능력과 지방의 실행력을 조화시켜 통합적 데이터·AI 거버넌스 체계의 실질적 작동을 뒷받침하고자 하는 데 목적이 있다.

5. 기대효과

본 연구는 공공부문의 데이터 거버넌스와 AI 거버넌스 체계를 통합적으로 정립함으로써 디지털 정부 혁신의 제도적 기반을 강화하는 데 기여할 것으로 기대된다. 우선 정책적 측면에서는 데이터기반행정법과 AI 기본법 간의 정합성을 확보하고, 중앙-지방 간 거버넌스 불균형을 해소하는 다층적 협력 체계를 구축함으로써 공공데이터와 AI 정책의 일관성과 실효성을 높일 수 있을 것이다. 조직적 측면에서는 CDO와 CAIO의 역할과 책임을 명확히 하고 협력 메커니즘을 제도화함으로써 데이터-AI 통합 관리체계의 효율성을 제고하며, 특히 중소 규모 기관에서도 적용 가능한 통합 운영 모델을 제시하여 조직 규모와 역량에 따른 차별화된 거버넌스 전략을 수립할 수 있게 될 것이다. 기술적 측면에서는 데이터 품질-AI 모델 품질-정책 품질로 이어지는 3단계 품질 연계체계를 통해 AI 시스템의 신뢰성과 공정성을 향상시키고, 데이터 표준화와 AI 품질검증 프로세스의 통합을 통해 공공서비스의 품질을 전반적으로 개선할 수 있을 것으로 기대된다. 나아가 시민 참여형 거버넌스 메커니즘을 통해 공공 AI의 투명성과 책임성을 강화하고, 데이터-AI 리터러시 교육 제도화로

공무원과 시민의 디지털 역량을 향상시켜 궁극적으로는 기술 중심 행정에서 민주적 AI 거버넌스로의 패러다임 전환을 실현하며, 국제 표준과 부합하는 한국형 데이터-AI 거버넌스 모델을 확립하여 글로벌 디지털 정부 혁신을 선도하는 데 기여할 것으로 기대된다.

SUMMARY

1. Title

A Study on Strategies for Public Sector Data Governance in the Era of AI and Digital Transformation

2. Research Objectives

To design a Korea specific, integrated data governance framework that supports AI driven, data based administration and personalized public services by strengthening leadership, laws and institutions, organizational arrangements, standards and management systems, and technical infrastructure.

3. Contents and Scope of the Research

The study diagnoses inter-agency data linkage and utilization; analyzes data-driven policy cases; benchmarks the Organisation for Economic Cooperation and Development three stage, six element framework and models from the United States, the United Kingdom, Germany, and Canada; and proposes phased recommendations for a Korean model centered on leadership, legal and institutional design, organizational structure, standards and management systems, and technological infrastructure.

4. Research Results

In this study, data governance is shown to have evolved toward participatory integration in the era of artificial intelligence, requiring fairness, privacy, transparency, traceability, and accountability. Case studies from advanced countries highlight Chief Data Officer systems, national data strategies, secure data sharing platforms, and citizen centric digital government. Success depends on accountable leadership, enabling laws, interoperable standards and technological infrastructure, and public trust. Ethical governance demands prior risk assessment, stakeholder participation, and protection by design. In the public institutions of the Republic of Korea, persistent gaps limited sharing, weak standardization, constrained analytical tools, shortages of skilled professionals, misaligned incentive structures, and incoherent legal frameworks necessitate an integrated governance platform, a central data catalog, automated quality control mechanisms, empowered Chief Data Officer and Chief Artificial Intelligence Officer roles, and harmonized regulations.

The research developed an integrated data governance framework through literature analysis and expert interviews, subsequently validated by professional feedback. The proposed model connects environmental demands with institutional, organizational, and technical capacities, thereby systematizing data and AI administration processes. It identifies activation conditions, key issues, and actionable suggestions across these dimensions, translating them into short-, medium-, and long-term roadmaps as well as planning-, organization-, and risk-based strategies. Furthermore, the study distinguishes data governance—focused on quality, standardization, and security—from AI governance—centered on fairness, explainability, and accountability—while interlinking them through leadership, policy, capability, process, and education mechanisms. The evolution of governance is conceptualized as a continuum from traditional to electronic, data-driven, and ultimately AI-based systems, with AI governance positioned as a complementary extension. Finally, the research proposes practical enhancement measures, including

prioritization of tasks, clarification of responsible entities, establishment of circular feedback loops, alignment with legal frameworks, CDO-CAIO collaboration, integration of standardization and ethics, and institutionalization of citizen participation.

5. Expectations

The results of this study will strengthen the institutional base for digital government innovation, align the Data Based Administration Act with the Artificial Intelligence Basic Act, reduce imbalances between central and local governments through multi level cooperation, improve efficiency by clarifying the roles of the Chief Data Officer and the Chief Artificial Intelligence Officer, raise service quality through a three tier quality chain that links data to artificial intelligence models and then to policy, enhance transparency and accountability through citizen participation, and establish an internationally aligned data and artificial intelligence governance model tailored to the Republic of Korea.

CONTENTS

Chapter 1. Introduction

1.1 Background and Need

- 1. Background of the Study
- 2. Necessity of the Study

1.2 Purpose and Content of the Study

- 1. Purpose of the Study
- 2. Main Contents of the Study

1.3 Research Methods and Implementation System

- 1. Research Methods and Analysis Framework
- 2. Research Implementation System and Strategy

Chapter 2. Theoretical Review

2.1 The Concept and Theoretical Framework of Data Governance

- 1. Theoretical Establishment and Systematization of the Data Governance Concept
- 2. Data Governance Models and Frameworks

2.2 Data Governance Cases in Major Countries

- 1. AI Governance Policies and Data Governance Models in Major Countries
- 2. Data Governance Frameworks at the International Organization Level

2.3 Success Factors and Implementation Strategies for Data Governance

- 1. Success Factors for Data Governance
- 2. Implementation Strategies for Data Governance

2.4 Social and Ethical Discussions on Data Governance

- 1. Ethical Balance between Data Utilization and Privacy Protection
- 2. Privacy-Centered Data Design Principles
- 3. Resolving the Digital Divide and Data Inequality
- 4. Transparent Governance for Building Trust and Ensuring Acceptance
- 5. Internationalization of Global Norms and Ethical Governance

2.5 Conclusion

- 1. Key Contents of the Theoretical Review
- 2. Implications of the Theoretical Review

Chapter 3. Exploring the Practice of Data Governance in Public Institutions

3.1 Data

- 1. Current Status of Data Management and Utilization
- 2. Key Issues and Challenges
- 3. Implications

3.2 Technology and Systems

- 1. Technologies Used in Data Utilization(AI, Statistical Packages, Visualization Tools)
- 2. Systems Used in Data Utilization(Platforms, Dashboards, Cloud, etc.)
- 3. Barriers and Success Factors at the Technology and System Level

3.3 Organization and Management

- 1. Awareness and Participation Levels of Organizational Members

- 2. Leadership and Organizational Culture
- 3. Strengthening Professionals and Capacity Building

3.4 Laws and Systems

- 1. The Influence of Related Legal Systems, Including Data-Based Administrative Law
- 2. The Gap Between Legal Systems and Practice
- 3. The Disparity Between Idealistic Direction and Actual Reality
- 4. Barriers and Success Factors at the Legal System Level

Chapter 4. Strategies for Data Governance in Public Institutions

4.1 Framework and Strategy Development

- 1. Revision and Enhancement of the OECD Basic Framework
- 2. Specific Governance Strategies
- 3. Data Governance–AI Governance Linkage Framework
- 4. Multi-Layered Governance Framework
- 5. Central-Local Capacity Gap Mitigation Strategy

4.2 Expert Verification and Supplementation of the Framework Strategy

- 1. Target Audience and Implementation Details
- 2. Feedback Reflection and Supplementation Details

Chapter 5. Conclusion

5.1 Research Results

- 1. Evolution and Key Success Factors of Data Governance in the AI Era
- 2. Diagnosis of the Current Status and Major Issues in Public Data Governance
- 3. Proposal of an Integrated Data Governance Framework

- 4. Linkage System between Data Governance and AI Governance

5.2 Policy Implications

- 1. Implications at the Institutional Level
- 2. Implications at the Organizational Level
- 3. Implications for Data Management
- 4. Implementation of Data Governance

References

제1 장 서론

제1 절 연구의 배경과 필요성

1. 연구의 배경

인공지능(AI)과 디지털 기술의 급속한 발전은 행정과 공공서비스 영역에서 디지털 전환을 가속화시키고 있으며, 이에 따라 데이터 기반 행정의 중요성도 지속적으로 강조되고 있다. 최근 정부는 공공부문의 데이터 및 AI의 활용을 극대화하여 국민에게 맞춤형 서비스를 제공하고 행정의 효율성을 높이기 위한 다양한 디지털 혁신 정책을 추진하고 있다. 이를 위해 행정안전부, 과학기술정보통신부 등 주요 부처들은 범정부 차원의 데이터 연계 및 활용 기반을 강화하고, 공공서비스 혁신을 위한 다양한 정책들을 다각도로 추진하고 있다.

이러한 데이터 기반 행정을 확립하기 위해서는 공공부문이 단순히 데이터를 활용하는 수준을 넘어, 그 활용을 체계적으로 뒷받침할 수 있는 거버넌스 체계의 구축이 반드시 필요하다. OECD는 데이터 거버넌스를 ‘데이터의 생성부터 삭제까지 전 과정에 걸쳐 기술적, 정책적, 규제적 프레임워크를 통해 관리하는 것’으로 정의하며, 이는 공공행정뿐 아니라 보건, 연구, 금융 등 다양한 정책 영역에 적용되는 개념이라고 설명한다. 특히 2022년에 발표된 OECD의 『Going Digital Guide to Data Governance Policy Making』 보고서에서는 데이터 개방성과 통제 간의 균형, 데이터 관련 이해관계자 간의 충돌 조정, 그리고 데이터에 대한 투자 및 효과적인 재사용을 유도하기 위한 인센티브 제공 등, 데이터 거버넌스와 관련된 정책 접근 방식들이 다각도로 제시되고 있다.

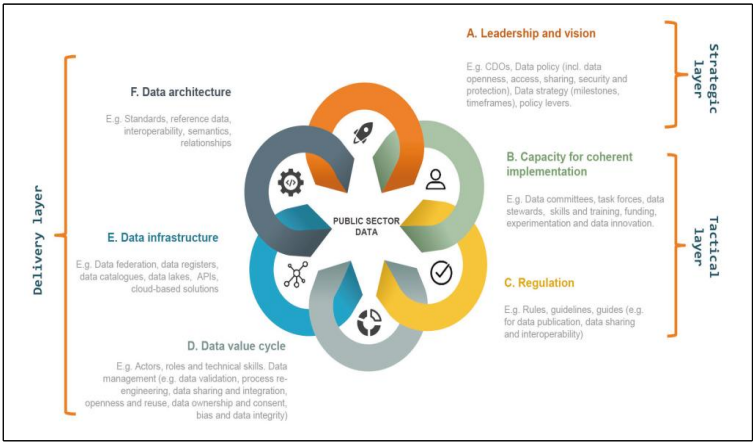
이와 함께 OECD는 『Digital Government Policy Framework(2020)』을 통해 공공데이터의 활용을 정부 역량 강화의 핵심 요소로 명시하고 있으며, 데이터를 기반으로 하는 공공부문이 정책의 적시성, 정확성, 그리고 대응력을 제고할 수 있다는 점을 강조한다. 국내에서도 정보통신정책연구원(KISDI)은 『EU 데이터 거버넌스 법안 주요 내용 및 시사점(2021)』

보고서를 통해 공공기관이 보유한 데이터를 기존의 목적을 넘어 상업적 목적으로도 활용할 수 있는 법적 기반을 마련할 필요성을 제기하였다. 나아가 데이터 거버넌스 체계를 정립함으로써 공공데이터 활용이 보다 활성화될 수 있음을 강조하였다.

이러한 국제적 논의를 종합할 때, 한국의 행정·정책 환경에 적합한 데이터 거버넌스 체계를 구축하기 위한 노력이 시급하다. 공공데이터를 둘러싼 국내의 논의는 단순히 데이터의 개방 여부를 넘어, 실질적인 활용성과 부처 간 협력 체계의 고도화를 요구하고 있다. 이를 실현하기 위해서는 기존의 직관이나 경험 중심의 정책 수립 방식에서 벗어나, 데이터에 기반한 과학적 정책 설계로의 전환이 요구된다. 특히 우리나라는 현재 공공부문 내 데이터 활용을 위한 제도적, 운영적 기반이 미흡하다는 점이 꾸준히 지적되고 있다.

OECD를 비롯한 주요국들은 통합 데이터 거버넌스 모델을 통해 공공데이터 혁신을 추진하고 있으며, 전략(Strategy), 전술(Tactics), 전달(Delivery)이라는 3단계 구조와 여섯 가지 구성 요소를 갖춘 모델을 제시하고 있다. 이 모델은 정책 혁신을 촉진하고 정책 성과를 평가하는 지표로도 활용될 수 있다는 점에서 유용하며, 이러한 국제적 기준을 바탕으로 한국의 행정환경에 적합한 데이터 거버넌스 모델을 수립하는 것이 시급한 과제로 부각되고 있다.

[그림 1-1] OECD 데이터 거버넌스 모델



출처: OECD(2019), The Path to Becoming a Data-Driven Public Sector, OECD Digital Government Studies

인공지능 정부를 구현하기 위한 데이터 기반 행정은 크게 세 가지 측면에서 추진되고 있다. 이는 기관 간 데이터 칸막이 해소, 과학적 행정의 가속화, 그리고 데이터 기반 행정 문화의 정착을 통해 실현되고 있다.

우선, 기관 간 데이터 칸막이 해소는 행정 효율성을 높이기 위한 핵심 전략이다. 이를 위해 예외 적용 대상 외의 모든 데이터를 공유하고, 비정형 정보나 가명정보와 같은 다양한 데이터 유형을 확대함으로써 데이터의 다양성을 확보하고 있다. 또한, 모든 데이터를 연계·통합하여 활용도를 높이며, 데이터 개방과 공유를 일원화하여 부처 간 협업을 강화하고자 한다. 이러한 기관 간 칸막이 해소의 의의는, 공공부문이 보유한 데이터뿐만 아니라 민간부문에서 구매한 데이터까지 범정부 차원에서 공동 활용할 수 있는 통합적인 관리 체계를 마련하는 데 있다. 더 나아가, 공동 활용 데이터 등록과 메타데이터 관리 등의 기능을 갖춘 데이터 통합 관리 플랫폼을 구축함으로써, 정부 내 데이터 공동 활용을 실질적으로 지원하고자 하는 점에서 그 중요성이 크다.

다음으로, 과학적 행정의 추진을 가속화하기 위한 노력도 병행되고 있다. 이를 위해 데이터 분석의 표준 모델을 확장함으로써 다양한 영역에서 효율적인 데이터 분석이 가능하도록 하고 있으며, 정책 수립과 행정 운영에 즉각 활용할 수 있는 고품질 데이터를 체계적으로 관리·제공하고 있다. 또한, 누구나 손쉽게 데이터를 분석하고 활용할 수 있도록 기술 인프라를 지원하고 있다. 이러한 과학적 행정 추진의 의의는 국민 체감형 과제, 국제 공동 협력 과제, 범정부적 중장기·긴급 정책 과제, 그리고 민원인 대상 행정서비스 등을 효과적으로 수행할 수 있도록 행정서비스를 지능화하는 데 있다.

마지막으로, 데이터 기반 행정 문화의 정착은 정부 전반에 걸쳐 데이터를 중심으로 한 행정 운영을 일상화하는 데 목적을 두고 있다. 이를 위해 데이터 기반 행정을 각 부처의 업무에 자연스럽게 정착시키고, 공직자들의 데이터 이해력과 분석·활용 역량을 강화하고자 한다. 또한, 데이터 전문 인력을 양성하고, 실무 수준과 대상별로 맞춤형 교육과정을 운영하는 동시에, 기관별 데이터 분석 및 활용 우수 사례를 적극적으로 공유하고 있다. 이러한 문화 정착의 전략적 의미는, 데이터 기반 행정을 정부 전반에 내재화하고 이를 일상적이며 보편적인 업무 처리 방식으로 정착시키는 데 있다고 할 수 있다.



출처: 관계부처 합동(2024), 제2차 데이터기반행정 활성화 기본계획(안)

기관 간 데이터의 연계, 공유, 활용에 대한 필요성이 점차 높아지고 있으며, 이는 공공서비스의 맞춤화, 인공지능(AI) 기반 정책 수립 등 주요 국정과제를 효과적으로 추진하기 위해 반드시 해결해야 할 과제로 떠오르고 있다. 그러나 현재 공공기관 간에는 데이터 연계의 부족, 데이터 품질 저하, 표준화 미흡 등의 문제가 지속적으로 지적되고 있는 실정이다. 공공서비스를 보다 정교하게 맞춤화하기 위해서는 데이터에 기반한 정책 수립이 요구되

며, 이를 위해 문제 발생 시 신속하고 유연하게 대응할 수 있는 체계 역시 필요하다. 이러한 역량은 궁극적으로 데이터 거버넌스 체계를 확립함으로써 실현될 수 있다.

아울러 공공서비스의 혁신을 안정적으로 추진하기 위해서는 민간과 공공의 데이터 및 서비스를 상호 연결하고 융합·활용할 수 있는 기반의 마련이 요구된다. 더불어 AI 기반 정책을 효과적으로 구현하기 위해서는 고품질 데이터의 수집과 저장, 분석, 공유, 관리, 그리고 보안까지 포함하는 전 주기의 데이터 거버넌스 체계 정립이 필수적이다.

2. 연구의 필요성

데이터 기반 행정은 추진 과정 전반에서 다양한 차원의 제약에 직면하고 있다. 먼저, 자료 및 정보 차원에서는 데이터의 비일관성이나 부정확성, 부적절한 데이터 내용, 불명확한 데이터 정의와 구조, 데이터 수요조사의 미실시, 그리고 내용 및 유형의 불충분성과 같은 문제가 주요 제약요인으로 작용한다. 이러한 요소들은 데이터 자체의 품질과 활용 가능성에 직접적인 영향을 미쳐 정책 수립과 행정 운영의 기반을 약화시킨다.

기술 및 시스템 차원에서는 시스템 간의 비호환성, 기술적 복잡성과 신기술 도입에 대한 부담, 시스템 설비의 부족, 파편화되고 분산적인 시스템 운영, 그리고 보안 기술의 미도입이나 취약성 등이 제약요인으로 나타난다. 이는 데이터의 안전한 관리와 효과적인 연계를 어렵게 만들며, 궁극적으로 행정의 디지털 전환을 저해하는 요소로 작용한다.

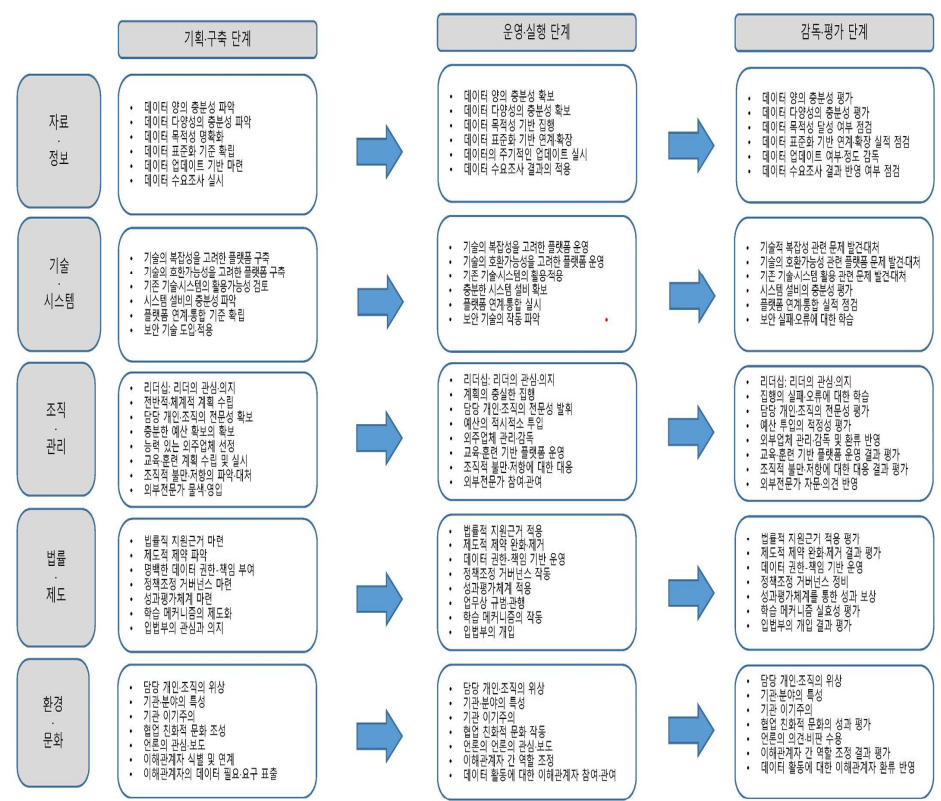
조직 및 관리 차원에서도 다양한 장애요인이 존재한다. 리더십의 부재, 인력·조직·예산 등 행정 자원의 부족, 조직의 본래 목표와 데이터 기반 행정이 지향하는 목표 간의 괴리, 교육과 훈련을 포함한 역량 강화 기회의 부족 등이 대표적이다. 이러한 제약들은 데이터 기반 행정이 실질적으로 뿌리내리기 어려운 구조적 문제로 이어질 수 있다.

법률 및 제도 차원에서도 제약은 분명히 존재한다. 데이터 기반 행정을 지원하는 법적·제도적 근거가 충분하지 않고, 성과를 평가할 수 있는 체계가 미비하며, 다양한 이해관계자를 포괄할 수 있는 거버넌스 운영 체계가 확보되지 않은 상황이다. 이는 제도적 신뢰를 약화시키고, 정책 실행의 정당성과 효율성을 떨어뜨릴 수 있다.

마지막으로 환경 및 문화 차원에서는 담당 개인 또는 조직의 위상이 낮거나, 데이터 기반 행정 도입을 어렵게 만드는 업무 특성, 협업에 친화적인 조직 문화의 부재, 기관 간 이

기주의, 이해관계자의 낮은 참여 수준 등이 장애로 작용하고 있다. 특히 이러한 환경·문화적 제약은 단기간에 인위적으로 개선하기 어려운 특성을 지니고 있기 때문에, 데이터 기반 행정의 초기 단계부터 장기적인 관점에서 이 문제들을 인식하고 적극적으로 대응할 필요가 있다.

[그림 1-3] 데이터 플랫폼 활용의 영향요인



출처: 오윤경·차남준(2023)

공공부문에서 데이터 기반 행정을 실현하기 위해서는 먼저 현장의 실태를 정확히 진단하고, 이에 기초한 실행 전략을 마련하는 것이 필요하다. 인공지능과 디지털 기술 환경이 빠르게 고도화됨에 따라 공공 영역 역시 정교하고 신속한 정책 대응 역량을 요구받고 있으며, 이를 가능하게 하기 위해서는 데이터 기반의 의사결정 구조 확립이 필수적이다. 그

러나 현재까지는 기관별로 데이터의 관리 및 연계 수준이 상이하며, 정책 수립과 집행 과정에서도 실효성 있는 데이터 활용 구조가 충분히 마련되지 못한 상황이다. 이에 따라 체계적인 실태 분석과 문제 구조화 작업이 선행되어야 한다.

정책 현장에서 실질적으로 작동할 수 있는 데이터 거버넌스 체계를 수립하기 위해서는, 기관 간의 데이터 연계 실태, 실제 활용 수준, 그리고 제도적·기술적 기반 등 다양한 차원에 대한 정밀한 진단이 요구된다. 이는 단순한 기술적 점검을 넘어, 데이터 기반 행정이 실질적으로 작동하는 구조를 설계하기 위한 필수 단계이다.

또한 데이터 기반 행정의 전면적 실현을 위해서는 데이터 활용의 사각지대를 해소하고, 정책 전 주기에 걸쳐 데이터를 연속적으로 활용할 수 있는 구조로의 전환이 필요하다. 현재 대부분의 정책이나 공공서비스 분야에서는 데이터가 기획, 집행, 평가 단계에 걸쳐 유기적으로 연계되지 못하고 있으며, 특히 특정 분야에서는 데이터 요구 유형과 흐름이 명확히 규정되지 않아 실무적 활용에 큰 제약이 존재한다. 이에 따라 정책 및 서비스별 데이터 흐름을 유형화하고, 활용의 사각지대를 구조적으로 발굴함으로써 향후 데이터 기반 행정 고도화를 위한 실행 기반을 확보해야 한다.

나아가 국내에 실제 적용 가능한 데이터 거버넌스 체계를 구체적으로 설계하는 일도 시급하다. OECD를 비롯한 주요국은 전략(Strategy)-전술(Tactics)-전달(Delivery)이라는 일관된 구조의 데이터 거버넌스 모델을 제시하고 있으나, 이를 한국의 행정 환경에 그대로 이식하기에는 제도적·운영적 한계가 존재한다. 한국의 행정 특성과 조직 운영 구조, 정보 보안 수준, 기술 인프라 등을 종합적으로 고려한 ‘맞춤형’ 거버넌스 체계 설계가 필요하며, 단순한 모형의 수용이 아니라 ‘적응(adaptation)’ 또는 ‘재설계(redesign)’의 관점에서 접근해야 한다.

특히 중앙과 지방, 부처와 산하기관 간의 협업 체계가 복잡한 한국의 행정 현실에서는 각 단계별 실행 방안과 협력 메커니즘을 포함하는 실천 가능한 프레임워크 도출이 요구된다. 예를 들어 국내 데이터 자치권 거버넌스 구축 방안은 데이터의 생애주기에 걸쳐, 이를 뒷받침할 행정 수단인 제도, 정책, 조직과 연계하여 구체적이고 체계적으로 제시할 수 있을 것이다.

〈표 1-1〉 행정수단과 데이터 생애주기

행정 수단 데이터 생애주기	제도 (제도적 강화 방안)	정책 (정책적 강화 방안)	조직 (조직적 강화 방안)
데이터 생성·수집	- 지방자치단체 데이터 배분 권한 확인 규정 마련 - 지방자치단체의 사무 범위 확장을 통해 법령상 데이터 자치의 범위 확대	- 기존 데이터 파악·접근을 위한 공통 플랫폼 활용 - 메타데이터플랫폼 구축을 통한 지방자치단체 데이터 생성·수집 활성화	- 데이터 전문 인력 채용 및 처우 개선 - 민간 전문가 채용 분야 확대 및 처우 개선을 통한 데이터 전문 인력 확대
데이터 제공·가공	- 공공데이터 제공 활성화를 위한 담당자·절차 규정 마련 - 공공데이터 개방 담당자 지정 및 심의위원회 운영 관련 규정 마련	- 민간 데이터 구매를 통한 다양한 데이터 제공 활성화 - 데이터 거래사 제도 활용을 통한 민간 데이터 구매 활성화	- 내부 전문가 양성을 위한 교육훈련 과정 마련 - 데이터 생애주기 관점에서의 전문가 양성을 위한 교육·훈련 실시
데이터 분석·활용	- 법률상 데이터 분석·활용 범위 확대 - 데이터 분석·활용 제한 관련 법률의 예외 규정 확대	- 데이터 호환성 및 분석·활용 강화를 위한 표준 정립 - 원활한 데이터 분석·활용을 위한 데이터 표준 시스템 마련	- 공공부문 데이터 분석·활용 도구 제공 및 개선 - 결과·성과 및 실무 위주의 데이터 분석·활용 관련 교육·훈련 전개
데이터 전 생애주기	- 데이터 자치 업무 관련 공공부문 인센티브 제도 강화 - 데이터 생성·개방, 데이터 기반 서비스 개발·제공을 성과관리와 연계	- 데이터 자치 컨트롤타워 수행 기관 지정 및 역할 수행 - 객관적·체계적 기준 아래 데이터 자치 컨트롤타워 지정 및 역할 수행	- 데이터 자치권 민관 거버넌스 체계 구축 및 운영 - 데이터 자치를 위한 다중 이해관계자 중심 협력적 거버넌스 모델 구축

출처: 경제·인문사회연구회(2024)

데이터 기반 정책을 효과적으로 구현하기 위해서는 무엇보다 신뢰할 수 있는 데이터 품질, 일관된 표준 체계, 그리고 민감한 데이터에 대한 적절한 보안이 전제되어야 한다. 그러

나 현재까지 공공기관별로 상이한 품질관리 기준과 표준체계가 적용되고 있어, 데이터의 상호운용성과 연계 가능성을 저해하고 있는 실정이다. 이러한 문제를 해결하기 위해서는 품질관리 프로세스, 메타데이터 관리, 개인정보보호 기준 등을 통합적으로 포함한 협력형 데이터 운영 체계를 정립해야 한다. 더불어 이러한 체계를 실현하기 위해서는 거버넌스 체계 구축 전략과 함께 기관 간 협력을 유도할 수 있는 인센티브 구조의 마련도 병행되어야 한다.

또한 인공지능(AI) 기반 정책의 실행을 위해서도 고품질 데이터의 확보와 안정적인 활용을 보장하는 데이터 거버넌스 체계가 필수적이다. AI 정책은 대용량의 고품질 데이터를 안정적으로 확보하고 분석할 수 있을 때 비로소 실효성을 갖출 수 있으며, 이는 데이터 거버넌스의 성숙도와 직접적으로 연결된다. 한편, 통합적이고 맞춤형 공공서비스의 실현 역시 다양한 기관의 데이터를 기술적·제도적으로 연계 및 활용할 수 있는 기반 없이는 불가능하다.

따라서 본 연구는 이러한 AI 기반 정책과 데이터 기반 행정 전략의 실행력을 제고하기 위한 핵심 기반을 마련하고자 한다. 이는 정책 추진의 실효성을 높이는 동시에, 국민이 체감할 수 있는 행정 혁신을 실현하는 데 기여할 수 있을 것이다.

제2절 연구 목적과 내용

1. 연구의 목적

인공지능과 디지털 기술 환경의 고도화에 대응하여, 공공부문에서의 데이터 활용을 활성화하고 체계적인 데이터 거버넌스 체계를 구축할 필요성이 커지고 있다. 이를 통해 정책 수립부터 행정 집행 전반에 걸쳐 데이터 기반의 의사결정 환경을 조성하고, 공공데이터를 체계적으로 관리하며 기관 간의 연계, 공유, 활용을 적극적으로 촉진하고자 한다.

이를 위한 첫 단계로 공공기관 간 데이터 관리 현황을 확인하고 각 기관의 연계 수준과 활용 실태를 분석하여 구조적인 문제점을 파악할 필요가 있다. 구체적으로는 데이터 관리 체계와 연계 방식, 활용 사례 등을 면밀히 조사하고 데이터 기반 행정의 실현을 저해하는 제도적·기술적 장애요인을 도출하고자 한다.

또한 OECD와 주요국의 데이터 거버넌스 모델을 벤치마킹하여 국내 행정환경에 적합한 프레임워크를 도출하는 작업도 병행된다. 이를 위해 전략(Strategy)-전술(Tactics)-전달(Delivery) 단계로 구성된 통합 모델을 분석하고, 국내 적용 가능성과 성공 요소를 반영하여 한국형 데이터 거버넌스 체계를 설계할 예정이다. 이 과정에서 잠재적 활용 가능 영역을 도출하고, 데이터 기반 행정 고도화를 위한 전략적 방향도 함께 제시할 계획이다.

마지막으로, 데이터 품질의 향상과 표준화, 그리고 협력 기반의 실효성 있는 데이터 관리체계를 구축함으로써 데이터 중심의 행정 혁신 기반을 확보하고자 한다. 기관 간 협업이 가능하도록 거버넌스 체계 구축의 전략을 제시하고, 행정의 효율성과 국민 체감도를 높이기 위한 방안도 함께 마련할 것이다.

2. 연구의 주요 내용

본 연구는 공공부문의 데이터 관리체계에 대한 실행 전략 프레임워크 제시에 중점을 두고 있다. 공공기관 대상으로 기관 간 데이터 연계 수준과 방식, 활용 현황을 진단한다. 아울러 데이터 기반 정책 수립 및 집행 사례를 분석하여 관련 문제점을 도출함으로써 현장의 구체적 상황을 파악하고자 한다.

해외 주요국과 OECD의 데이터 거버넌스 프레임워크도 함께 분석한다. 미국, 영국, 독

일, 캐나다 등 선진국의 거버넌스 모델을 벤치마킹하고, OECD가 제시한 전략-전술-전달의 3단계 모델과 6대 핵심 요소를 검토한다. 이를 바탕으로 국내 행정환경에 적합한 적용가능성과 시사점을 도출하여 한국형 데이터 거버넌스 체계 설계에 활용할 계획이다.

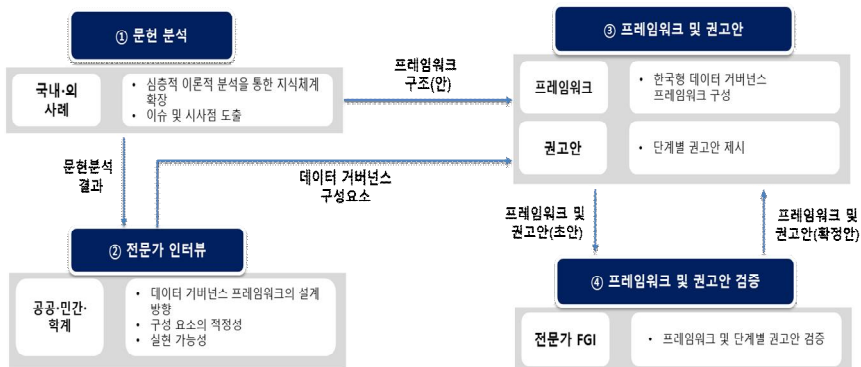
국내형 데이터 거버넌스 프레임워크는 리더십, 법·제도, 조직 구조, 표준 및 관리 체계, 기술 인프라 등 핵심 구성 요소를 중심으로 체계화할 예정이다. 이를 통해 중앙정부와 지방자치단체, 부처와 산하기관 간의 실질적이고 효과적인 거버넌스 체계 구축을 도모한다.

제3절 연구 방법과 추진체계

1. 연구방법과 분석의 틀

본 연구는 인공지능 및 디지털 기술 환경에 대응하여 공공부문에서의 데이터 활용을 활성화하고, 체계적인 데이터 거버넌스 체계를 구축하는 것을 목표로 한다. 이를 위해 공공기관 간 데이터 관리 실태를 진단하고, 제도적·기술적 장애요인을 분석하였으며, OECD 및 주요국의 사례를 바탕으로 한국형 거버넌스 프레임워크를 설계하였다. 아울러 공공서비스 전 과정에서의 데이터 활용 실태를 분석하고 사각지대를 해소하기 위한 전략을 제시하였다. 나아가 데이터 품질 향상, 표준화, 협력 기반의 관리체계 구축을 통해 데이터 중심 행정 혁신의 실행 기반을 마련하고자 하였다. 이를 위해 아래와 같은 연구 분석틀을 제시하였다.

〔그림 1-4〕 연구분석 틀



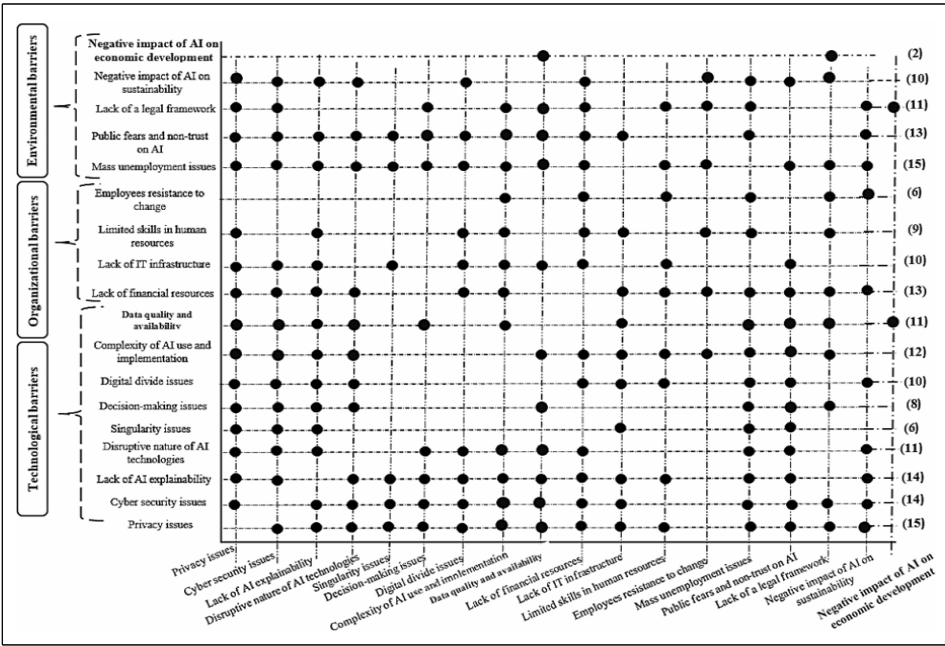
가. 국내외 문헌 및 사례 분석

OECD 및 주요 국가들, 특히 뉴질랜드, 노르웨이, 네덜란드 등의 데이터 거버넌스 프레임워크와 정책 동향, 그리고 실행 전략을 분석하고, 이들 국가의 구성 요소와 추진 체계를 유형화하는 데 중점을 둔다. 이를 통해 국내 데이터 거버넌스 전략 수립에 필요한 시사점을 도출하고자 한다.

또한 국내의 데이터 거버넌스 관련 현행 정책인 데이터기반행정 기본계획 등과 법률적 제도인 데이터기반행정법, 공공데이터법 등을 면밀히 분석하여 전략 수립에 필요한 핵심 이슈를 파악한다. 아울러 정보통신정책연구원(KISDI), 한국디지털정부학회, 한국정책학회 등 국내 주요 연구기관과 학술지에 발표된 관련 연구 동향도 검토하여 국내 학계 및 정책 연구의 최신 흐름을 반영한다.

이와 더불어 국내외 사례 간의 유사점과 차이점을 비교 분석하여 한국 행정환경에 적합한 정책 설계 기준을 도출하고, 벤치마킹 가능한 요소를 확보한다. 마지막으로 기관 간 데이터 연계, 공유, 활용을 촉진하는 요인과 장애요인을 체계적으로 도출하여, 이를 바탕으로 실현 가능한 정책 모형 설계의 기반을 마련하는 데 기여하고자 한다.

[그림 1-5] 제약요인 탐색 방법 예시



출처: Rjab et al.(2023), Barriers to artificial intelligence adoption in smart cities, Government Information Quarterly

나. 전문가 인터뷰 및 FGI(Focus Group Interview)

문헌분석으로 얻은 데이터 거버넌스 이슈에 대해 전문적인 의견 수렴을 위해 학계, 연구기관(한국행정연구원, KISDI 등), 현업 실무자 등으로 구성된 전문가 집단을 대상으로 자문회의 및 인터뷰를 개최하였다. 1차 인터뷰는 2025년 7월 22일부터 8월 12일까지 중앙부처 2명, 지방자치단체 3명, 공공기관 4명, 국책연구기관 전문가 2명, 교수 3명을 대상으로 서면 및 대면 인터뷰를 진행하였다. 1차 인터뷰에서는 문헌분석 결과를 통해 얻은 데이터, 기술과 시스템, 조직과 관리, 법률과 제도 등 4개 범주에 대한 질문을 실시하였다.

전문가 인터뷰에서는 정책 실행 관점에서 데이터 거버넌스 프레임워크의 설계 방향과 구성 요소의 적정성, 그리고 실현 가능성에 대해 심도 있는 검토하였다.

아울러 중앙부처, 지방자치단체, 공공기관 실무자를 대상으로 표적집단면접(FGI, Focus Group Interview)를 실시하여 기관 간 데이터 연계 및 공유를 저해하는 실제적이고 구조적인 원인을 도출하였다. 이를 통해 데이터 활용 현황뿐만 아니라 제도, 기술, 인력 등 현장 상황에 기반한 실무적 개선 요구도 함께 수렴하였다.

대면과 비대면 방식을 병행한 전문가 FGI와 인터뷰를 통해 다양한 주체들의 의견을 균형 있게 반영하였으며, 그룹별 응답 간 비교 분석을 수행한다.

다. 정책 프레임워크 및 권고안 제시

국내외 문헌, 사례 그리고 전문가 인터뷰 및 전문가 FGI를 바탕으로 확인한 데이터 거버넌스의 핵심 구성 요소를 중심으로 한국형 데이터 거버넌스 프레임워크를 구성하였다. 한국형 데이터 거버넌스 프레임워크 설계를 위해 데이터 거버넌스의 정의, 범위 설정하였으며 문헌분석과 인터뷰를 통해 핵심요소를 도출하였다. 이를 바탕으로 계층적 구조를 설계하였으며 전문가 FGI와 인터뷰를 통해 설계안에 대한 피드백을 받아 논리적 일관성과 완결성을 검증받았다. 아울러 구성된 프레임워크를 바탕으로 권고안을 설계하였으며 이를 전문가 FGI와 인터뷰를 통해 검증하여 보완하였다.

2. 연구 추진체계와 전략

가. 연구 추진체계

데이터 거버넌스와 공공서비스 혁신 분야의 전문 연구진을 중심으로 수행 체계를 구성

하였다. 연구팀은 공공행정, 디지털정부, 정보정책, 데이터 과학 등 관련 분야의 전문가들로 편성되었으며, 정책 분석, 제도 설계, 사례 조사, 실증 연구 등 각자의 역할을 분담하여 협업 체계를 구축하였다.

또한 연구의 단계별 수행을 위해 기획, 진단, 분석, 설계, 전략 제안 등의 각 단계에 책임자를 지정하고, 자문회의 및 전문가 워크숍을 통해 외부 의견을 수렴하는 체계적인 조직적 대응 방안을 마련하였다. 이와 함께 중앙정부, 지방자치단체, 공공기관 실무자들과의 인터뷰 및 실무 협의를 적극 추진하여 현장 중심의 연구 수행에 만전을 기하였다.

나. 연구 추진 전략

본 연구는 공공부문 데이터 거버넌스 체계 구축을 위해 단계적이고 체계적인 접근 방식을 채택하여 수행되었다. 먼저 중앙행정기관, 지방자치단체, 공공기관 등 국내 공공부문을 대상으로 데이터의 생성·수집, 제공·가공, 분석·활용, 폐기에 이르는 생애주기별 관리 실태와 활용 현황을 진단하였다. 이러한 진단은 현재 공공부문 내 데이터 관리 구조의 문제점을 파악하고 개선 방향을 모색하는 데 기초자료로 활용되었다.

이어 해외 국제기구 및 주요국의 데이터 거버넌스 프레임워크와 정책 동향, 실행 전략 등을 비교 분석하였다. 이를 통해 도출된 시사점을 바탕으로, 국내 행정환경과 법·제도적 여건을 고려한 데이터 거버넌스 프레임워크 설계 시 반영해야 할 핵심 고려사항들을 도출하였다.

이러한 분석 결과를 바탕으로 리더십, 법·제도, 표준화, 기술 인프라 등 데이터 거버넌스의 핵심 구성요소를 중심으로 국내에 적용 가능한 프레임워크를 설계하였다. 아울러 신규 조직 신설, 기존 조직 개편, 법률 제·개정, 인적 전문성 강화 등 실행 가능성을 높이기 위한 구체적인 정책 및 거버넌스 체계 구축의 전략을 제시한다.

특히 본 연구는 실효성을 확보하기 위해 문제 중심의 접근 방식을 채택하였다. 이를 통해 기관 간 연계와 공유를 저해하는 구조적 원인을 분석하고, 데이터 품질 저하 및 표준화 미비 문제를 진단하였다. 또한 데이터 수요와 공급 간의 미스매칭 문제를 해결하기 위한 실천적 전략을 도출하였다.

궁극적으로 본 연구는 정책 활용성을 높이는 데 초점을 맞추어, 정책 수립과 행정서비스 개선에 실질적으로 적용 가능한 전략을 마련하였다. 법·제도의 정비, 기술 인프라 구

축, 조직 운영 체계 개선 등 다층적인 추진 방안을 포함하며, 인센티브 설계 및 성과관리 체계를 포함한 실행 기반 마련을 통해 지속 가능한 데이터 거버넌스 체계 정착을 도모하고자 하였다.

제 2 장 이론적 고찰

제 1 절 데이터 거버넌스의 개념과 이론적 틀

1. 데이터 거버넌스 개념의 이론적 정립과 체계화

가. 데이터 거버넌스의 개념적 토대

1) 데이터 거버넌스 전통적 개념화

거버넌스는 정부뿐 아니라 비영리, 민간 등 다양한 행위자들이 상호 협력하여 공공의 문제를 하는 새로운 통치 메커니즘을 의미한다. 현대 사회 문제가 복잡해지고 다원화됨에 따라 정부를 포함한 하나의 단독 조직이나 행위자로는 그러한 문제들을 해결하는데 한계가 있어, 거버넌스라는 개념이 등장하게 되었다(McGuire & Agranoff, 2011). 이러한 거버넌스는 수직적 상명하달의 체제를 넘어서, 공동 책임성, 책무성, 조정, 협력, 등의 정책 과정 전반에 걸친 협업을 핵심 원리로 한다. 특히 정책의 설계와 집행에 이르는 전 과정에서 참여의 확대와 투명성이 최근 효과적인 거버넌스의 필수 요소가 되었다(Harrison, T. M., & Sayogo, 2014).

이러한 거버넌스 패러다임은 정보통신기술(ICT)의 발달과 결합하면서 'e-거버넌스'라는 형태로 진화하였다(Heeks, 2001). e-거버넌스는 전자적 수단을 통해 정부 운영을 개선하고, 특히 시민과의 쌍방향 커뮤니케이션을 증진시켜 투명성과 참여를 통한 다양한 행위자의 협력(Cooperation), 조정(Coordination)의 메커니즘을 통한 협업(Collaboration)을 제고하는 것을 핵심으로 한다(Mcnamara, 2012; Misuraca, 2007). 단순한 행정 자동화를 넘어서, 정책결정의 개방성과 데이터 기반 행정을 가능케 하는 디지털 전환의 핵심 축으로 작동하고 있다. 이와 동시에, 전자정부 시스템을 통해 축적되는 대규모 공공데이터는 새로운 형태의 정책자산으로 떠오르고 있으며, 이 데이터의 활용과 보호를 둘러싼 새로운 규범과 조정이 요구되고 있다.

인공지능(AI)의 도입은 이러한 e-거버넌스의 기반 위에서 더욱 복잡한 정책 환경을 만들

어내고 있다. AI는 단순한 정보처리를 넘어 예측, 판단, 의사결정을 자동화하는 기술로서 공공데이터를 학습 재료로 삼아 정책 집행과 평가를 근본적으로 변화시키고 있다. 그러나 이 과정에서 데이터의 편향성, 알고리즘의 불투명성, 프라이버시 침해와 같은 윤리적·사회적 문제도 함께 대두되고 있다. 따라서 AI를 기반으로 한 거버넌스, 즉 AI 데이터 거버넌스는 기술적 효율성만을 추구하는 것이 아니라, 데이터의 생성부터 활용, 폐기까지의 전 주기에 걸쳐 책임성과 공정성을 확보하려는 제도적 노력이다(Floridi et al., 2018). 이는 전통적 거버넌스와 e-거버넌스가 강조해온 가치들(참여, 투명성, 책임성)을 기술 환경에서도 구현하려는 확장된 거버넌스 체계라 할 수 있다. 이러한 AI 기반 거버넌스의 확장된 요구는 곧 데이터 자체를 어떻게 관리하고 조정할 것인가라는 보다 근본적인 문제로 이어진다. 이때 핵심 개념으로 등장하는 것이 바로 데이터 거버넌스(data governance)이다.

데이터 거버넌스는 전통적으로 조직의 효율적인 운영과 관리를 위하여 조직 내외부의 데이터 관리와 관련하여 조직 및 조직구성원의 책임과 권한, 역할을 정의하고, 적용하고, 감독하는 것과 관련된다(Harrison et al., 2019; Janssen et al., 2020; Nielsen, 2017). 데이터 거버넌스는 조직이 데이터를 관리하는 것이 단순히 기술적인 측면에서 논의할 문제가 아니라 조직 구조적, 제도적 차원과 같은 비기술적인 측면을 같이 고려해야 한다는 문제의식을 공유하고 있다(Janssen et al., 2020; Scholl et al., 2012). 구체적으로, 조직 내 어떤 구성원이 데이터 관리의 구체적인 업무를 담당할 것인지, 어떤 조직구성원이 데이터 제공자인지, 관리자인지, 이용자인지, 감독자인지, 조직구성원이 데이터 관리와 관련된 규칙과 지침을 따르도록 보상과 처벌을 어떻게 설계할 것인지 등이 모두 비기술적인 측면에서 데이터 거버넌스의 핵심적인 논의에 포함된다(Janssen et al., 2020). 구체적으로 다음과 같은 논의들이 있다.

Nielsen(2017)은 데이터 거버넌스를 단순한 IT 문제를 넘어서 조직 전체의 비즈니스 문제로 인식되어야 한다고 강조한다. 이를 위해 기술 부서 영역에 국한 되지 않고 조직 차원의 전략적 접근을 요구한다. 저자는 데이터 거버넌스를 데이터 사용에서 바람직한 행동을 장려하기 위한 의사결정 권한과 관련된 책임의 배분(allocation of decision-making rights and related responsibilities)으로 정의하고 있다. 이는 단순한 기술적 통제를 넘어 조직 내 권한 구조와 책임의 명확화, 데이터 활용까지 포함하는 거버넌스 체계로 기능하다.

Harrison et al.(2019) 이들은 정부가 AI를 성공적으로 활용하기 위해 데이터 관리의 중요

성을 강조하며 데이터 거버넌스를 데이터 자산의 관리에 대한 권한과 통제를 행사하기 위해 필요한 구조와 실천으로 정의한다. 이들은 데이터를 수집하고 저장하는 절차뿐 아니라 어떤 방식으로 사용될 수 있는지를 규정하는 권한과 책임의 배분 구조를 포함한다는 점에서 보다 보다 구조화된 개념을 제시한다. 이는 조직이 데이터에 관한 결정을 내리고 데이터 관련 행동 정책을 수립하는 체계로 이해되며, 특히 AI 환경에서 데이터 관리 역량이 모든 종류의 데이터 분석에 기본이 되는 핵심 요소로 작용한다고 지적한다.

2) 확장된 데이터 거버넌스 개념화

데이터 거버넌스의 최근 논의들은 데이터 거버넌스의 논의의 범위를 다양한 측면으로 확장하고 있다. 첫째, 조직 차원의 전략이자 행동 체계로서 데이터 거버넌스가 순수하게 어느 한 조직에만 국한해서 논의하기 어렵다는 문제의식이 있다(Kroll, 2018). 어느 한 조직은 데이터 거버넌스에 관한 구체적인 지침을 조직을 넘어서 사회적·제도적 차원의 윤리, 규범 등의 영향을 강하게 받는다.

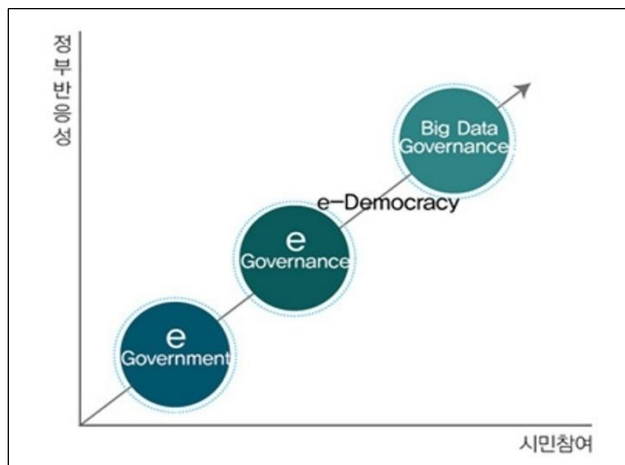
구체적으로 살펴보면, Kroll(2018)은 포괄적 책임 체계로 데이터 거버넌스를 개념화하였다. 데이터 기반 자동화된 의사결정 과정에서 발생할 수 있는 보안, 프라이버시, 공정성, 투명성, 책임성 문제를 포괄적으로 관리하는 조직적 틀로 정의하였다. 이는 단순한 데이터 관리를 넘어 사회적 책임, 윤리적 규범, 법적 사항을 아우르는 것이다. 전통적으로 데이터 거버넌스가 데이터 수집, 저장, 보안, 품질 관리 등을 의미했지만, 저자는 이를 넘어 책임 있는 데이터 활용 체계(responsible use of data)로 개념을 확장하였다. 데이터 거버넌스가 단일 시점이나 특정 기능에만 작동하는 것이 아닌 데이터 수집, 정규화, 전처리, 분석, 결정, 보관, 삭제까지 모든 단계에서 데이터 거버넌스가 작동해야 한다고 말한다. 데이터 거버넌스는 단계별 정책과 기준의 집합이며, 조직 내 의사결정 전반에 내재화되어야 하는 통합적 구조로 볼 수 있다.

둘째, 조직 차원에서 데이터 거버넌스 도입은 다양한 조직 외부의 행위자들과의 협력을 고려하는 것이 중요하다. 이러한 조직 외부의 행위자로는 데이터를 생산하거나 이용하는 데 직접 참여하거나 영향을 줄 수 있는 다양한 기업이나 공공기관들을 포함하여 직접 서비스의 혜택을 받을 시민들을 포함할 수도 있다.

예를 들어, 송정재 외(2018)는 수평적 네트워크 거버넌스로서의 거버넌스에 대해 설명하

였다. 빅데이터 시대의 데이터 거버넌스는 다양한 행위자들과의 수평적 네트워크에서 대등하게 참여하는 방식으로 정책결정에 빅데이터를 활용하는 구조이다. 거버넌스에서 e-거버넌스, 그리고 빅데이터 거버넌스로의 개념적 발전이 이루어졌다. 통찰력 있는 시민(insightful citizen)의 등장을 기반으로 하는 ICT 공간에서의 실시간 협의와 조정이 가능한 탈중앙화된 통치구조가 특징이다.

〔그림 2-1〕 전자정부, e-거버넌스, 빅데이터 거버넌스



자료: 송경재·장우영·조인호(2018)

마지막으로, 데이터 거버넌스는 단순히 데이터를 관리하는 것을 넘어서서 데이터와 알고리즘을 포함하는 것으로 확장되고 있다(Janssen et al., 2020). 이는 데이터와 알고리즘이 상호 간에 상당한 영향을 미치면서 조직의 데이터 활용에 영향을 미치고 있기 때문이다. 이러한 맥락에서 최근에 ‘신뢰할 수 있는’ 대규모 데이터를 확립하는 것이 중요한 과제가 되고 있다(Janssen et al., 2020). 무엇보다도 신뢰할 수 없는 데이터는 블랙박스 알고리즘과 결합되어 잘못된 불투명한 판단과 정보의 견잡을 수 없는 확산을 야기할 수도 있다(Janssen et al., 2020).

3) AI 및 디지털 전환이 초래한 데이터 거버넌스의 도전과제

디지털 전환과 인공지능(AI)의 도입은 조직의 데이터 활용 방식뿐만 아니라, 데이터 거

버넌스 전략 전반에 구조적인 재설정을 요구하고 있다. 특히 AI 시스템은 더 이상 수동적인 데이터 소비에 머무르지 않고, 데이터를 생성하고 변환하며 해석하는 능동적 주체로 기능함에 따라, 기존의 정태적이고 내부지향적인 데이터 거버넌스 체계로는 이러한 변화에 대응하는 데 한계가 존재한다(Pahune et al., 2025; Schoeman, 2023). 더불어 데이터는 이제 더 이상 단순한 참고자료나 기록물이 아닌, 공공의사결정을 자동화하고 미래를 예측하는 핵심 자산으로 자리매김하고 있다(Park, 2018). 특히 AI 기반의 정책 집행 및 서비스 운영이 확산됨에 따라, 데이터는 정책 과정의 중심적 요소로 기능하며, 그 관리와 통제 방식이 거버넌스의 핵심 의제로 부상하고 있다. 이러한 맥락에서 AI와 디지털 전환이 초래하는 새로운 도전은 다음과 같은 영역에서 나타나고 있다.

첫째, 학습데이터의 품질과 편향성 문제이다(Boppiniti, 2023). 생성형 AI를 포함한 고도화된 AI 시스템은 대규모의 데이터셋에 기반해 학습되며, 이때 데이터의 정확성뿐 아니라 대표성, 다양성, 공정성이 중요한 품질 기준으로 작용한다(Pitoura, 2020). 데이터셋에 편향이 존재할 경우, AI 모델의 예측이나 판단 결과가 왜곡되어 사회적 차별이나 오류 가능성을 증폭시킬 수 있다. 이러한 윤리적 문제는 단순히 기술적 보정으로 해결되기 어려우며, 데이터 수집 및 선별 단계에서부터 책임 있는 기준과 관리 원칙을 반영한 데이터 거버넌스 구조가 필요하다(김윤명, 2024).

둘째, 데이터와 알고리즘 간 상호작용으로 인한 책임소재 불분명성이 대두되고 있다(김민주·김현경, 2025). AI는 특정 데이터를 기반으로 판단과 의사결정을 자동화하는 기능을 수행하지만, 그 결과에 대한 책임이 데이터 제공자, 알고리즘 개발자, 운영 주체 중 누구에게 있는지를 명확히 하기 어렵다. 이로 인해 데이터 거버넌스는 단순한 운영 규칙을 넘어, 권한과 책임의 분배에 대한 제도적 설계, 투명한 감사 체계, 윤리적 감독 메커니즘을 포함하는 포괄적 체계로 확장되어야 한다. 실제로 최근 AI의 안전성과 책임성 이슈가 부각되면서, 관련 논의는 대부분 AI 모델의 블랙박스적 특성과 함께 입력 데이터의 품질 및 통제 문제로 귀결되고 있다(신재형, 2024). 이는 곧 데이터 거버넌스가 기술의 성능 확보뿐 아니라 사회적 수용성과 신뢰성 보장의 핵심 기제로 기능해야 함을 시사한다.

셋째, 데이터 처리 속도와 규제 대응 간의 비대칭성 문제이다. AI 기반 의사결정은 실시간 또는 준실시간 데이터 흐름 속에서 작동하지만, 이에 대응하는 법·제도적 프레임워크와 내부 규정은 상대적으로 느리게 작동한다. 이러한 시간적 비동기성은 사전 예방적 규

제나 책임 추적을 어렵게 만들며, 특히 개인정보 보호, 데이터 오용, 알고리즘 리스크 등에 대한 실효적 통제를 저해한다. 따라서 선제적이고 유연한 데이터 거버넌스 체계가 AI 환경에서는 필수적으로 요구된다.

넷째, 설명가능성과 신뢰성 확보의 어려움 역시 핵심적인 과제로 지적된다(Chamola et al., 2023). 특히 딥러닝 기반의 AI 모델은 그 내부 작동 원리가 블랙박스로 간주되는 경우가 많아, 외부 이해관계자에게 모델의 판단 근거를 설명하거나 정당성을 부여하기 어렵다. 이는 AI 시스템의 책임성과 투명성을 확보하는 데 심각한 제약으로 작용하며, 설명가능성(explainability)을 중심으로 한 데이터 라인리지, 메타데이터 기록, 외부 검증 등의 기능을 데이터 거버넌스 구조에 내재화할 필요성이 강조된다(Verma et al., 2024).

이와 동시에, 데이터 주권(Sovereign AI)의 관점에서 공공데이터는 점차 국가의 전략적 자산으로 인식되고 있으며, 디지털 행정과 공공서비스의 핵심 인프라로 활용되고 있다. 특히 LLM 기반의 행정 자동화, 맞춤형 복지, 위험 예측 시스템 등에 공공데이터가 직접 활용되면서, 데이터의 접근권, 관리권, 활용권에 대한 국가적 수준의 통제 메커니즘이 중요해지고 있다. 이는 단지 기술적 문제를 넘어, 데이터 거버넌스가 국가의 정책 주권과 직결되는 정치·제도적 사안으로 확장되고 있음을 보여준다.

요컨대, 디지털 전환과 AI의 도입은 데이터 거버넌스를 단순한 기술적 관리 도구가 아닌, 사회적 책임성과 제도적 정당성을 포괄하는 전략적 프레임워크로 재정립할 것을 요구한다. 데이터의 생성부터 활용, 폐기에 이르는 전 주기에 걸쳐 책임성과 투명성을 확보하고, 다양한 이해관계자 간 권한과 책임의 조화를 구현하는 방향으로 거버넌스 전략이 재편되어야 한다. 이는 AI 기반 정책 환경의 신뢰성 확보뿐만 아니라, 민주적 거버넌스 원칙의 지속 가능성을 담보하는 핵심 조건이 된다.

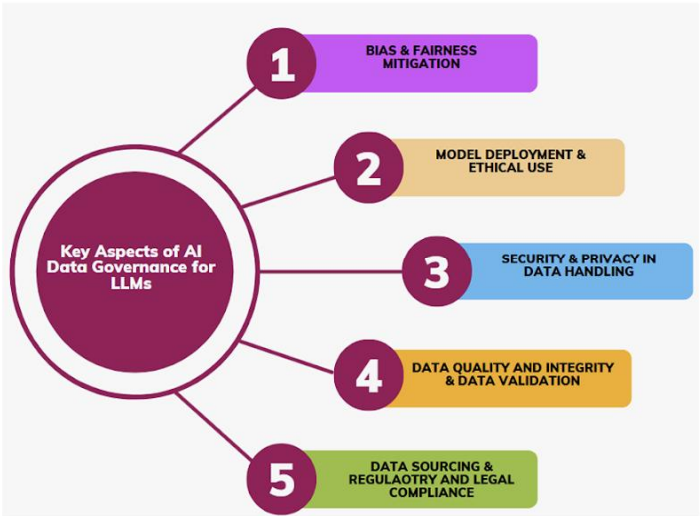
나. AI 시대의 데이터 거버넌스 개념 진화

1) AI 데이터 거버넌스의 전주기적 개념화

Pahune et al.(2025)은 AI 데이터 거버넌스를 전체 생애주기(lifecycle)를 아우르는 통합적 관리체제로 개념화하였다. 특히 대규모 언어 모델(LLMs)과 같은 고도화된 AI 시스템에서는 단순한 데이터 관리 차원을 넘어 보다 정교하고 책임 있는 데이터 거버넌스 구조가 요구된다. 저자는 LLM 기반 AI 거버넌스를 효과적으로 설계 및 운영하기 위해 다음 다섯 가

지 요소를 제시한다. 이는 △편향과 공정성 완화 △모델 배포 및 윤리적 활용 △데이터 보안 및 프라이버시 확보 △데이터 품질 및 검증 △데이터 출처 관리와 규제 준수 등으로 구성되며 AI 거버넌스의 전반적인 책임성과 공공 신뢰 확보를 위한 필수 조건으로 설명된다.

[그림 2-2] AI 데이터 거버넌스의 핵심 요소(1)



자료: Pahune, S., Akhtar, Z., Mandapati, V., & Siddique, K(2025)

이러한 핵심 요소들을 바탕으로 AI 모델의 개발부터 학습, 검증, 배포 및 사후 운영에 이르기까지 모든 단계에서 데이터를 지능적이고 체계적으로 관리할 수 있는 프레임워크를 제안한다. 해당 프레임워크는 다음의 일곱 가지 핵심 범주로 구체화 된다.

첫째, 데이터 생애주기 관리(Data Lifecycle Management)이다. 둘째, 규제 준수 및 법적 체계로서 GDPR, HIPAA, AI Act 등 글로벌 규제를 포함한다. 셋째, 윤리적이고 공정한 AI 운영을 통해 인종, 성별, 종교, 나이 등과 무관하게 투명성, 공정성, 차별 없는 방식을 보장한다. 넷째, 데이터 프라이버시 및 보안을 위해 암호화 기술과 위협 대응 메커니즘을 활용한다. 다섯째, 데이터 품질, 무결성, 검증을 통해 AI 모델 신뢰성 확보에 핵심으로 작용한다. 여섯째, 데이터 계보 및 추적성(Data Lineage and Traceability)은 데이터 사용 흐름을 기록하고, 문제가 발생했을 때 원인을 분석하거나 디버깅할 수 있는 기반을 제공한다. 일

곱재, 보다 안전한 종단 간 모델 배포(More Secure End-to-End Model Deployment)는 DevOps, MLOps, LLMops 등의 파이프라인을 활용하여 초기 설계부터 학습, 검증, 운영, 사후관리까지 전 과정에 걸쳐 AI 시스템의 보안성과 신뢰성을 확보한다.

〈표 2-1〉 AI 데이터 거버넌스의 범위

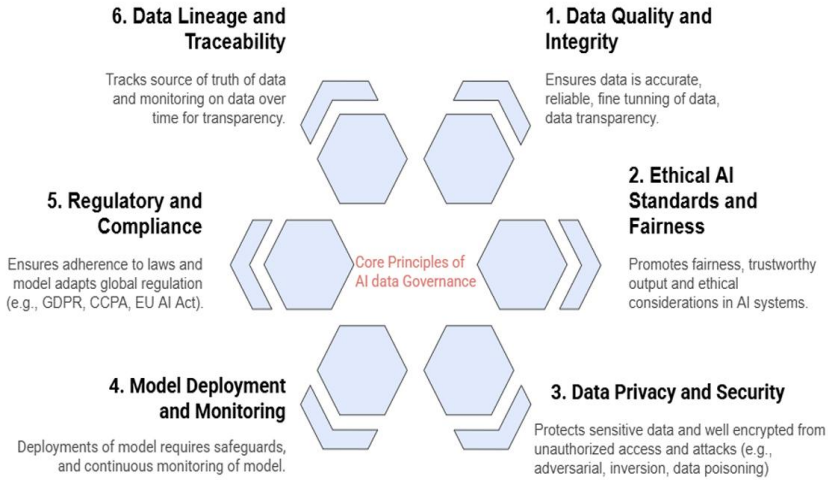
구분	설명
데이터 생애주기 관리	지능형 데이터 거버넌스를 활용하여 AI 모델의 개발 단계부터 배포 종료 단계까지 전 생애주기에 걸쳐 데이터를 관리
규제 준수 및 법적 프레임워크	GDPR, CCPA, HIPAA, AI Act, AIRMF 등 글로벌 규제를 반영한 확장 가능한 유연한 거버넌스 모델
윤리적이고 공정한 AI 운영	투명성과 공정성 기반 인종, 성별, 종교, 나이 차별 없는 AI 모델
데이터 프라이버시 및 보안	- 데이터 프라이버시를 보호하고 암호화된 메커니즘을 활용한 데이터 유출 방지 - 적대적 공격, 모델 역추론, 데이터 변조 등 다양한 사이버 위협 예방
데이터 품질, 무결성, 검증	데이터 품질, 무결성, 검증은 데이터 거버넌스의 핵심 요소로 AI 모델의 신뢰성에 직접적인 영향을 미침
데이터 계보 및 추적성	AI 거버넌스 감사자가 데이터 사용 경로를 추적하고 문제 원인 분석을 도움
보다 안전한 종단 간 모델 배포	DevOps, MLOps, LLMops 등 다양한 파이프라인을 통해 AI 모델을 초기 단계부터 학습, 검증, 배포, 사후 단계까지 안전하고 신뢰성 있게 배포하는 데 기여

자료: Pahune, S., Akhtar, Z., Mandapati, V., & Siddique, K(2025)

2) AI 거버넌스와 데이터 거버넌스의 관계 정립

Schoeman(2023)는 통합적 거버넌스 개념을 제시하였다. 그는 AI 거버넌스를 조직 내·간·외부에서 AI를 관리하여 신뢰성, 설명가능성, 책임성을 보장하고 윤리적 기준을 준수하도록 하는 활동으로 정의하였다. 이 과정에서 데이터 거버넌스가 AI 거버넌스의 토대나 AI 특화 영역의 독립적 관리가 필요하다고 강조한다. AI 모델 품질, AI 아키텍처, AI 메타데이터, AI 라이프사이클 관리가 핵심 요소이다. 또한 사회적 책임 영역을 추가하여 조직 경계를 넘어 AI 시스템의 사회적 영향까지 포괄하는 개념으로 확장하였다.

[그림 2-3] AI 데이터 거버넌스의 핵심 요소(2)



자료: Pahune, S., Akhtar, Z., Mandapati, V., & Siddique, K(2025)

다. 사회적·관계적 관점의 데이터 거버넌스 개념화

1) 사회적 데이터 거버넌스 개념

현대 사회에서 데이터는 단순한 기술적 자산이나 조직 내부의 운영 도구를 넘어, 사회 구조와 문화적 맥락 속에서 끊임없이 생성 및 유통되고 있다. 데이터 거버넌스 역시 기존의 기술·조직 중심 접근을 넘어서는 보다 확장된 개념화가 요구되고 있으며 최근 학계에서는 사회적·관계적 관점에 기반한 새로운 이론적 틀들이 제시되고 있다. Liu(2022)는 사회적 데이터 거버넌스 개념을 제시하였다. 그는 기존 조직 중심적 관점의 한계를 비판하면서, 데이터 관리, 인프라, 알고리즘, 정책 등 조직 수준 요소에 편중된 접근을 문제로 지적하였다. 대신 거시적 수준의 사회적, 정치적, 문화적 요소들을 포괄하는 개념으로 확장할 것을 제안하였다. 이 관점에서 데이터 거버넌스는 단순히 사회를 반영하는 것이 아니라 해당 사회의 정책과 실행을 생산하고 재생산한다.

2) 관계적 데이터 거버넌스 이론

데이터 거버넌스를 거시적인 사회적 구조의 일부로 바라보면서 Viljoen(2021)은 관계적 접근법을 제시하였다. 그는 개인의 데이터 소유권 부여나 동의 중심 모델의 한계를 지적

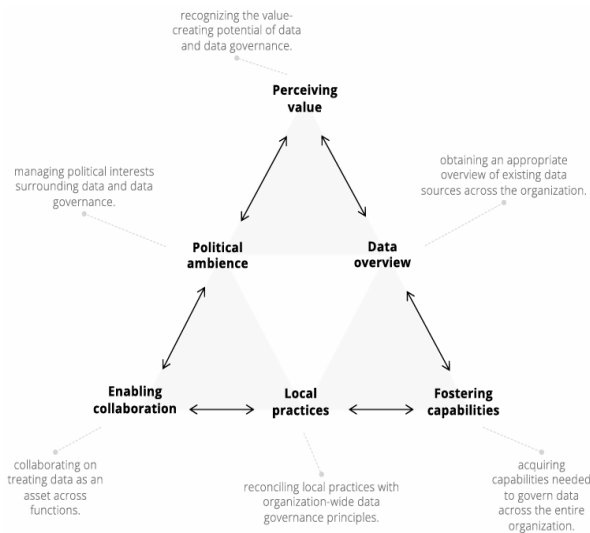
하면서, 데이터가 개인 간, 개인과 조직 간 관계 속에서 생성되고 활용됨을 강조하였다. 이 접근법의 핵심 원칙은 다음과 같다. 첫째, 데이터는 개인의 소유물이 아니라 사회적 상호작용의 부산물이라는 것이다. 둘째, 개인 대 기업의 이분법을 넘는 다층적 책임 관계를 고려해야 한다는 점이다. 셋째, 집단적 권리, 공공의 책임, 제도적 설계 중심의 거버넌스가 필요하다는 것이다. 마지막으로 민주주의, 사회정의, 정보정의와 긴밀히 연결된 법적·윤리적 차원을 고려해야 한다고 주장하였다.

3) 집합행동 문제로서의 데이터 거버넌스

집합행동 이론을 적용한 접근법도 제시되었다. 이 관점에서 데이터 거버넌스는 조직 구성원 간 협력 실패라는 집합행동의 실패로 개념화된다. 이러한 실패는 6가지 상호연결된 도전과제로 나타난다.

첫째, 가치 인식 부재(Perceiving Value)는 데이터 또는 거버넌스의 장기적 가치 인식이 부족한 상황이다. 둘째, 협업 유도 실패(Enabling Collaboration)는 부서 간 협업이 부족하고 사일로 구조가 해소되지 않는 문제이다. 셋째, 역량 편차(Fostering Capabilities)는 데이

[그림 2-4] 데이터 거버넌스의 문제 삼각형(Problem Triangle of Data Governance)



자료: Pahune, S., Akhtar, Z., Mandapati, V., & Siddique, K(2025)

터 관리 기술·조직적 능력의 부서별 편차를 의미한다. 넷째, 데이터 파악의 어려움(Data Overview)은 조직 내 데이터 존재와 위치 파악의 기술적 한계를 말한다. 다섯째, 지역적 실무 관행의 다양성(Local Practices)은 부서별 데이터 처리 방식이 상이하여 통일 원칙 도입에 저항이 발생하는 현상이다. 여섯째, 정치적 환경(Political Ambience)은 거버넌스를 조직 의제로 끌어올리는 정치적 지지나 인식이 부족한 상황을 의미한다.

라. 사회현상의 데이터화와 민주적 거버넌스

1) 사회현상의 데이터화의 개념

Redden(2018)은 데이터화된 거버넌스 개념을 제시하였다. 이는 사회현상과 시민을 데이터화된 정보로 전환하는 과정으로, 정부가 시민과 사회 문제를 이해하고 관리하는 방식의 근본적 변화를 초래하는 사회기술적 현상이다. 시민을 추적 가능하고 분석 가능한 데이터 주체(data subjects)로 변환시키며, 국가-시민 관계에 새로운 권력 역학을 초래한다.

2) 데이터화된 거버넌스의 특징과 함의

정책 패러다임의 전환이 핵심 특징이다. 전통적 인과관계 중심 정책에서 상관관계와 예측 중심 정책으로의 전환이 일어난다. 지식/권력(power/knowledge) 개념으로 해석되는 데이터와 알고리즘의 사회 통제 및 규율 도구화가 진행된다. 데이터 기반 시스템의 블랙박스화로 인한 투명성과 책임성 확보가 어려워진다. 공공-민간 협력 확대와 맞물린 정부 권한의 재분배 및 공공성 약화 우려가 제기된다.

마. 글로벌·국제적 데이터 거버넌스 개념화

1) AI·데이터 시대의 초국경 거버넌스 패러다임

Arner(2022)는 초국가적 접근법을 제시하였다. 그는 디지털 시대 데이터가 국경을 초월해 활용되는 현실에서 기존 국가 중심 규제 체계의 구조적 한계를 지적하였다. 이에 따라 데이터 거버넌스를 정치적 정당성, 권력 분산, 시장 구조, 윤리적 책임이 결합된 복합적 문제로 재정의할 것을 제안하였다.

MacFeely et al.(2022)은 글로벌 공공재 접근법을 제시하였다. 이들은 글로벌 공공재로서의 데이터 보호 및 활용을 통한 지속가능한 개발과 인권 보호 실현을 강조하였다. 데이터의 비경합적·비배제적 특성과 현실적 통제·제한 상황의 모순을 해결해야 한다고 주장하였다. 이를 위해 국제적 사회적 계약(Social Contract)의 필요성을 강조하였다. 또한 디지털

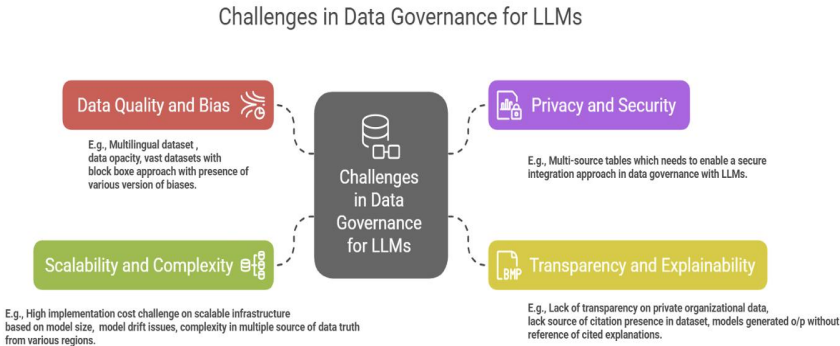
시대 데이터와 AI 등 신기술 환경 변화에 대응하는 유연하고 포괄적인 원칙 및 규제 체계가 필요하다고 제시하였다.

바. 정부·공공부문 데이터 거버넌스 개념화

1) LLM 시대의 정부 데이터 거버넌스

LLM 연계 거버넌스는 디지털 시대 데이터가 국경을 초월해 활용되는 현실에서 기존 국가 중심 규제 체계의 구조적 한계를 인식한다. 데이터 거버넌스를 정치적 정당성, 권력 분산, 시장 구조, 윤리적 책임이 결합된 복합적 문제로 재정의한다. 개인정보, 기업 데이터, 공공데이터 등 다양한 데이터 유형을 전 생애주기 동안 법률, 관리체계, 표준, 기술 도구 등을 통해 효과적으로 관리하는 것이 핵심이다. 정부의 LLM 적용 방향 명확화를 위한 관련 정책과 문서 발간이 필요하고, 데이터 수집·처리·공유 전 과정에서 투명성과 책임성의 강화가 강조된다.

[그림 2-5] 대규모 언어 모델(LLMs)을 위한 데이터 거버넌스의 주요 과제



자료: Pahune, S., Akhtar, Z., Mandapati, V., & Siddique, K(2025)

2) 엔터프라이즈 데이터 거버넌스

Luna-Reyes와 Harrison(2024)은 전사적 접근법을 제시하였다. 이들은 조직 내 이질적인 데이터 소스를 통합하고, 데이터 품질을 검증하며, 공통 데이터 모델과 메타데이터를 관리하는 것의 중요성을 강조하였다. 조직 전체에 걸쳐 단일 진실 버전(single version of truth)을 확보하는 것이 목적이다. 정부 조직의 특수성상 단일 조직 내 관리를 넘어 여러 기관과

민간 파트너 간 협업을 포함하는 다자간 협력 모델이 필요하다. 이는 협력 거버넌스, 공공 가치 거버넌스, joined-up government, whole-of-government 접근법과 연계된다.

3) 데이터기반행정과 연계된 거버넌스 개념

윤건(2023b)은 협력적 의사결정 구조로서 데이터 거버넌스를 제시하였다. 그는 데이터 정책에 관여하는 수많은 이해관계자들이 협력적으로 상호작용하는 정책적 의사결정구조이자 관리·감독의 체계로 정의하였다. 이는 데이터기반행정이나 데이터 융합과 같은 소기의 목적을 달성하기 위한 수단으로 이해된다. 정부의 데이터 정책의 의사결정구조로서 다양한 정책목표와 정책수단적 요소들을 포함한다.

사. 확장된 데이터 주권과 사회적 수용성 기반 거버넌스 구성

1) 데이터 주권의 확장된 개념화

데이터 주권은 다층적 개념으로 이해된다. 전통적 정의는 특정 국가·권역에서 수집된 데이터는 해당 국가 또는 권역의 데이터 법률이 적용된다는 것이었다. 확장된 정의는 데이터의 생성, 저장, 유통, 활용에 대한 주권 국가의 배타적인 권리와 자신의 데이터에 대한 개인의 권리를 모두 포함한다. Lähteenoja(2021)는 이를 특정한 디지털 영역과 관련된 개인에 가장 우선시되는 권위로 정의하였다. 전통적 주권 개념의 디지털 공간 확장으로 개인의 데이터 권리를 폭넓게 인정하는 기본권으로 해석된다. 유승현(2022)은 데이터 통제력 관점의 독점 개념을 제시하였다. 데이터를 통제하는 힘 또는 데이터를 바탕으로 경쟁을 좌우하는 힘을 나타내는 표현으로, 경쟁법적 개념보다는 통제력의 관점에서 이해된다. 일정한 거래 분야에서 가격이나 거래 조건 등을 좌우할 수 있는 시장 지위의 데이터 영역 확장으로 개념화된다.

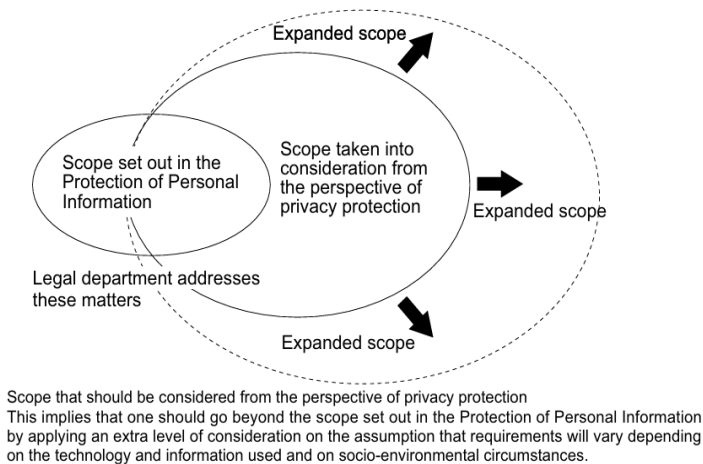
2) 실무적 데이터 거버넌스의 구성요소

데이터 거버넌스의 핵심 구성요소로 5가지 구성요소 체계가 제시되었다. 실사 수행은 데이터의 흐름, 접근 주체, 외부 벤더의 정책 및 위험성을 식별하는 것이다. 데이터 인벤토리 구축은 민감 데이터의 위치, 보호 조치, 접근 권한을 명시하는 것이다. 통제 체계 구축은 최소 권한 원칙 및 데이터 이동·보관·처리 정책을 수립하는 것이다. 데이터 거버넌스 패널은 부서 간 협의 기구로 정책·접근 방식을 결정한다. 정책 소통은 정책의 명확한 전달, 이해관계자 교육 및 인식 제고를 담당한다.

3) 프라이버시 거버넌스와 애자일 거버넌스

Watanabe(2024)는 사회적 수용성 중심 거버넌스를 제시하였다. 프라이버시 거버넌스는 단순한 법적 준수를 넘어 사회적 맥락과 기술 변화에 유연하게 대응할 수 있도록 설계되어야 한다. 애자일 거버넌스는 서비스 도입 후 사회적 반응이나 기술 변화에 따라 빠르게 규칙을 조정하는 구조이다. 사회적 수용성 확보가 기술적 문제보다 핵심 실행 요소이며, 이해관계자에게 충분한 설명과 투명성 제공이 필수적 조건이다.

[그림 2-6] 전자정부, e-거버넌스, 빅데이터 거버넌스



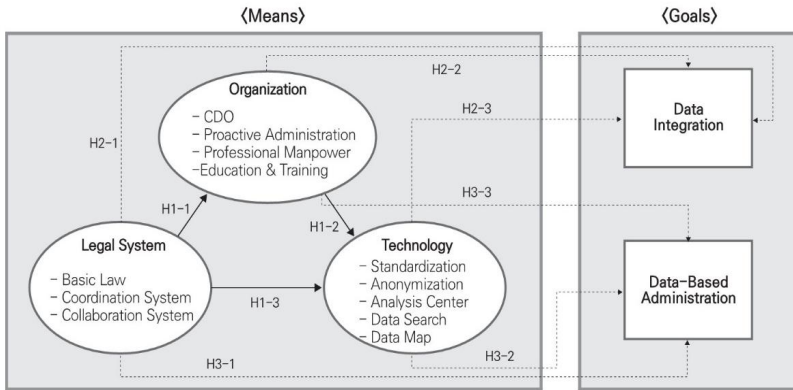
자료: Watanabe, T.(2024)

2. 데이터 거버넌스 모형과 프레임워크

가. 데이터 거버넌스 핵심요소: 조직·정책·기술·프로세스

데이터 거버넌스의 핵심 요소로는 조직 구조, 정책 및 제도, 기술 인프라, 실행 프로세스의 통합적 관리가 중요하게 부각되고 있다. 윤건(2023), 윤건·김윤희(2020), Mahanti(2018), Harrison et al.(2019)은 제도, 조직, 기술의 3분법과 정책, 프로세스, 책임이라는 3대축 프레임워크가 효과적인 데이터 거버넌스의 기반이라고 강조하였다.

[그림 2-7] '제도-조직-기술' 3분법 분석틀



자료: 윤건(2023)

Janssen et al.(2020)은 신뢰할 수 있는 AI 기반 BDAS(Big Data Algorithmic Systems)를 구축하기 위해 다음의 세 가지 접근을 개별적으로 또는 종합적으로 사용할 수 있다고 주장한다. 먼저, 계획 및 통제 기반 접근은 조직의 계획(목표, 방향 설정), 실행, 평가의 일련의 과정을 통해 시스템을 구축하는 접근이다(Janssen et al., 2020). 이것은 정기적이고 장기적인 변화에 적합하지만 유연한 대응에 적합하지 않을 수 있다(Janssen et al., 2020). 둘째, 조직 기반 접근은 조직 차원에서 역할, 규칙, 책임, 권한 등을 공식적으로 부여하는 접근방법이다(Janssen et al., 2020). 예를 들어, 최고데이터관리자(CDO)나 최고알고리즘관리자(CAO)와 같이 데이터와 알고리즘의 역할과 책임을 명확히 하는 것을 들 수 있다. 셋째, 위험 기반 접근은 데이터 생애주기에서 언제 발생할지 모르는 위험(risk)을 예방하고 대응하는 것을 중심으로 하는 접근이다(Janssen et al., 2020). 예를 들어, 데이터 폐기 단계에서는 민감정보가 폐기되지 않은 채 남아있거나 하는 위험을 관리하는 것이 이 접근에서 중요한 과제가 된다.

〔그림 2-8〕 데이터 생애주기 단계별 책임주체, 통제 메커니즘, 품질 기준

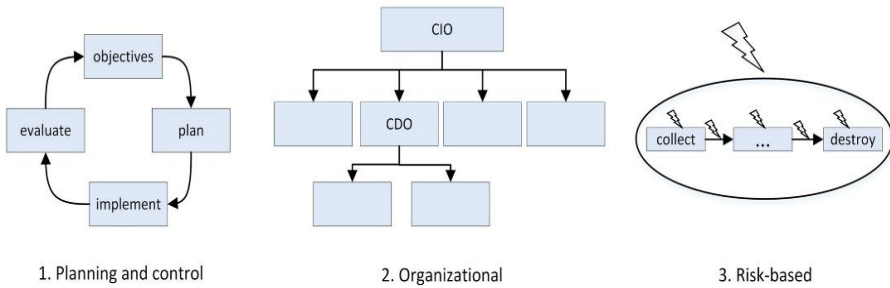
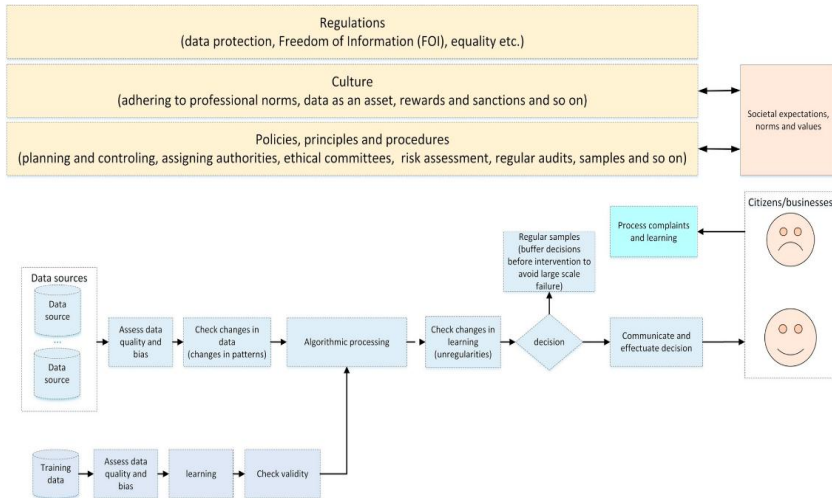


Fig. 1. Tripartite approaches to data governance.

자료: Janssen et al.(2020)

〔그림 2-9〕 AI 기반 BDAS(Big Data Algorithmic Systems)



자료: Janssen et al.(2020)

한편, Luna-Reyes et al(2024)는 데이터 관리, 조직 구조 및 역할, 정책과 프로세스, 협력 및 참여의 4대축 프레임워크를 제시하였다. 이 프레임워크는 구성 요소들이 반복적이고 상호 보완적으로 작동하는 시스템으로서, 데이터와 AI 프로젝트 관리에서 계획-실행-모니터링-평가의 순환 프로세스를 견고히 하며, 기술적·관리적 통합뿐만 아니라 사회·정책적 맥락을 고려한 다층적 거버넌스 메커니즘을 반영하여, AI 시스템의 공정성, 투명성, 프라

이버시 보호를 실현하는 데 중점을 둔다. 또한, 단일한 거버넌스 모델이 아니라 상황에 맞는 계획 및 통제 접근(planning and control), 조직적 접근(organizational), 위험기반 접근(risk-based approach)을 조합하는 ‘복합적 거버넌스 체계’임을 강조한다. 이 체계는 AI 특성상 데이터 준비와 알고리즘 검증 단계마다 발생하는 위험요소를 지속적으로 분석·관리하는 리스크 기반 접근이 핵심적임을 시사한다. 종합하면, Luna-Reyes와 Harrison의 데이터 거버넌스 프레임워크는 정부 AI 응용에서 데이터의 기술적 품질 확보와 함께 복잡다단한 조직 간 협력과 윤리·법적 요구사항을 균형 있게 반영하는 시스템적이고 협력적인 거버넌스 모델로서 기능한다. 이 프레임워크는 AI의 신뢰성 확보와 지속 가능한 공공서비스 혁신을 위한 실천적 설계 방향을 제시한다.

Harrison et al.(2019)은 DAMA International의 DMBOK(Data Management Body of Knowledge)¹⁾ 프레임워크를 정부 데이터 관리의 기준으로 제시하였다. 이 프레임워크를 통

〔그림 2-10〕 DMBOK(Data Management Body of Knowledge) Framework



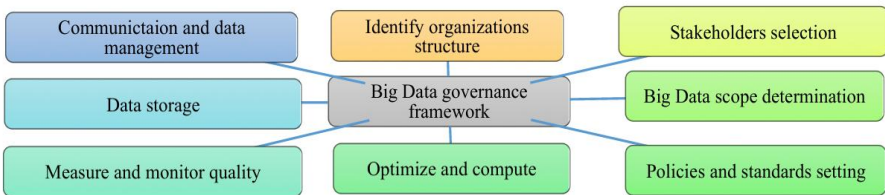
자료: Harrison et al.(2019)

1) DMBOK는 데이터 거버넌스(Data Governance), 데이터 아키텍처 및 설계(Data Architecture & Design), 데이터베이스 관리(Database Management), 데이터 접근 관리(Access Management), 데이터 품질 관리(Quality Management), 마스터 데이터 관리(Master Data Management), 데이터 웨어하우스 및 비즈니스 인텔리전스 관리(Data Warehouse & Business Intelligence Management), 기록 관리(Records Management), 메타데이터 관리(Metadata Management) 등 9가지 핵심 차원으로 구성된다(Harrison et al., 2019).

해 데이터 거버넌스를 조직이 데이터의 생성, 접근, 사용, 보안, 품질, 적합성에 대해 정책적으로 통제할 수 있게 하는 기반이자 AI 도입의 전제 조건임을 강조하였다. 또한, 데이터 아키텍처, 데이터 품질, 메타데이터, 데이터 접근 및 보안 등 각 차원이 AI 시스템의 신뢰성, 투명성, 품질 확보에 결정적 역할을 수행한다고 평가하였다.

Al-Badi et al.(2018)는 기존의 다양한 빅데이터 거버넌스(Big Data Governance) 프레임워크를 비교 분석하여 새로운 개념적 프레임워크를 제안하였다. 이들이 제안한 새로운 프레임워크는 데이터 수명 주기 전반을 포괄하며, 정책, 기술, 조직문화가 통합된 다차원적 거버넌스 모델로 향후 실제 조직에의 적용을 통해 타당성을 추가적으로 검증해나갈 필요성이 제기된다. 이와 더불어 민첩성(agility), 자동화(automation), 협업(collaboration), 실시간 품질관리 등 적응형 데이터 거버넌스가 빅데이터·디지털 전환 시대에 중요하게 부각되고 있으며(Nookala, 2024), 개방적 시민참여와 정부 반응성의 결합, 수평적 협력·임파워먼트 모델(2×2)이 제시되기도 한다. 이는 기준에 따라 정부주도압출(Push) 모델, 시민주도압출(Push) 모델, 인터랙티브(Interactive) 모델로 나뉜다(송경재 외, 2018).²⁾

[그림 2-11] Big Data Governance Framework



자료: Al-Badi et al.(2018)

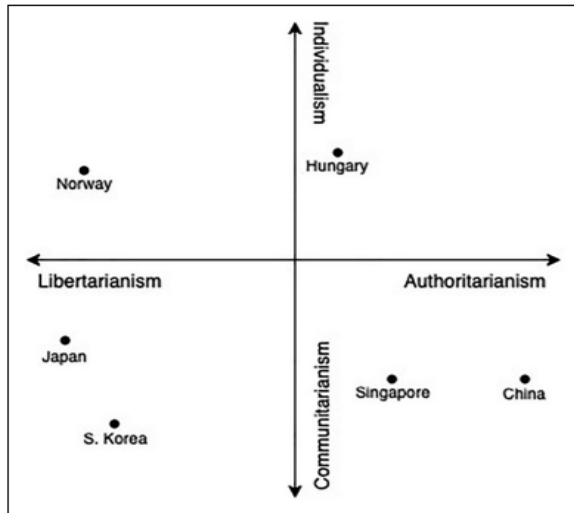
2) 시민참여와 정부반응성 모두 낮은 ‘신민형(subject)’ 모델은 엄밀한 의미에서 빅데이터 거버넌스라고 할 수 없으며, 전통적 거버넌스 이전 단계인 ‘거버먼트(government)’로 보아야 한다고 본다(송경재 외, 2018).

[그림 2-12] 빅데이터 거버넌스 유형화



자료: 송경재 외(2018)

[그림 2-13] A model of social data governance



자료: Liu(2022)

이와 더불어 Liu(2022)는 사회적 데이터 거버넌스 모델(A model of social data governance)을 제시하였는데, 이는 국가 개입과 사회적 자율성의 차원과 국가문화의 두 가지 차원으로

로 구성된다. 국가개입과 사회적 자율성의 양극단에는 권위주의와 자유민주주의(또는 자유주의)가 위치한다. Liu(2022)가 제시하는 이차원 모델은 전 세계의 다양한 사회적 데이터 거버넌스의 변이를 설명하고 유럽적 관점을 넘어서는 개념화와 분류의 보완된 관점을 제공한다.

나. 네트워크·플랫폼·참여형 데이터 거버넌스

1) 협력적 데이터 네트워크의 중요성과 플랫폼·참여 기반 거버넌스 프레임워크 등장
 Micheli et al.(2020)은 데이터 인프라 개념을 활용하여 네 가지 주요 데이터 거버넌스 모델(데이터 공유 풀, 데이터 협동조합, 공공데이터 신탁, 개인 데이터 주권)을 분석하였다. 이들은 각 모델별 이해관계자 역할, 가치 창출 및 재분배 방식, 거버넌스 원칙과 메커니즘을 사회과학적 관점에서 개념화하였다.

[그림 2－14] Data governance modeling Platform society

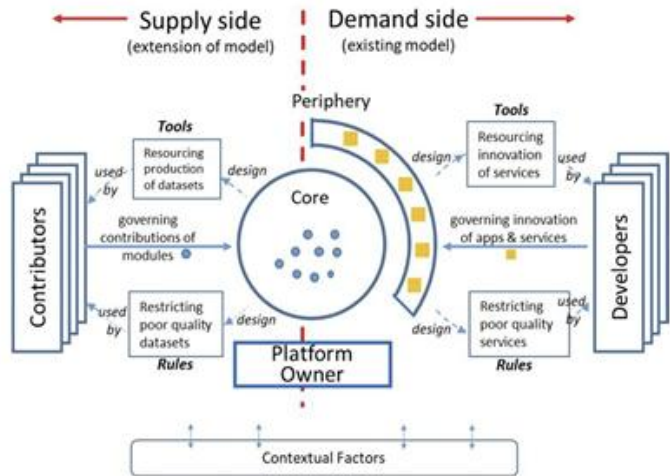
Model	Key actors	Goals	Value	Mechanisms
Data sharing pools (DSPs)	• Business entities • Public bodies	• Fill knowledge gaps through data sharing • Innovate and develop new services	• Private profit • Economic growth	• Principle of 'data as a commodity' • Partnerships • Contracts (e.g. repeatable frameworks)
Data cooperatives (DCs)	• Civic organisations • Data subjects	• Rebalance power unbalances of the current data economy • Address societal challenges • Foster social justice and fairer conditions for value production	• Public interest • Scientific research • Empowered data subjects	• Principles from the cooperative movement • Data commons • 'Bottom-up' data trusts • GDPR Right to data portability
Public data trusts (PDTs)	• Public bodies	• Inform policy-making • Address societal challenges • Innovate • Adopt a responsible approach to data	• Public interest • More efficient public service delivery	• Principle of 'data as a public infrastructure' • Trust building initiatives • Trusted intermediaries • Enabling legal framework
Personal data sovereignty (PDS)	• Business entities • Data subjects	• Data subjects self-determination • Rebalance power unbalances of the current data economy • Develop new digital services centred on users need	• Empowered data subjects • Economic growth • Private profit • Knowledge	• Principle of 'technological sovereignty' (e.g. MyData) • Communities and movements • Intermediary digital services (personal data spaces) • GDPR Right to data portability

자료: Micheli et al.(2020)

Bonina & Eaton(2020)은 플랫폼 기반 데이터 생태계에서의 거버넌스를 공급자(Supply side)와 수요자(Demand side), 즉 조직 내 다양한 이해관계자의 역할과 규칙, 도구 설계 등을 중심으로 설명하는 조직·제도적 관점의 거버넌스 모델을 시각화하였다. 이 프레임워크 내에는 플랫폼 운영자(Platform Owner)가 있으며, 이들은 데이터 생산자(Contributors)와 서비스 개발자(Developers) 양측을 연결·조정하는 핵심 역할을 수행한다. 이러한 설계는

도구(tools)와 규칙(rules)을 통해 구체화 되며, 이 거버넌스 설계는 궁극적으로 맥락적 요인(contextual factors)에 따라 조정되어야 함을 시사한다(Tan, 2021).

[그림 2-15] Governance design by organizational/institutional perspective



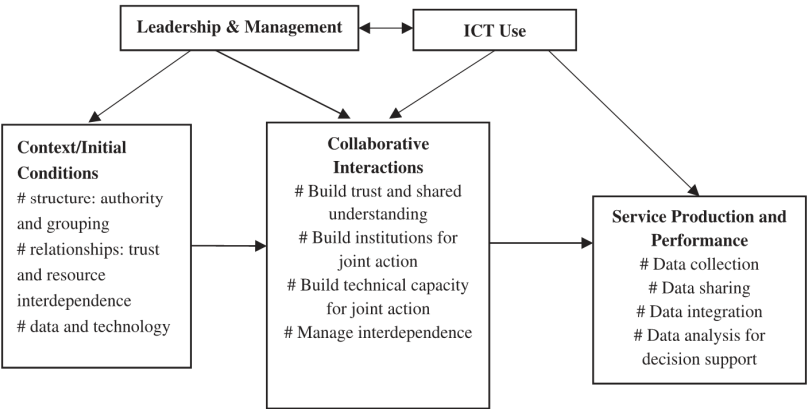
자료: Bonina & Eaton(2020); Tan(2021)

Jiong'en & Yingliang(2024)은 빅데이터 시대를 맞아 정부의 데이터 공유 및 교환 문제를 해결하기 위해 클라우드 플랫폼 기반의 '정부 종합 데이터 자원 관리 프레임워크(Government Comprehensive Data Resource Management Framework)'를 제안하였다. 이 프레임워크는 클라우드 플랫폼 인프라, 데이터 공유 및 교환 플랫폼 데이터, 데이터 개방 서비스의 3계층 구조로 설계되어 있다. 추가적으로 이 프레임워크는 5개의 주제별 논리적 데이터베이스(5 thematic logical databases)와 데이터 자원 인터페이스 풀(data resource interface pool)이 구성요소로 제시하고 있는데, 이는 데이터가 기존 시스템에 유지된 채 논리적 접근만 허용하는 방식이다. 이를 통해 데이터 이동 없이 실시간 데이터 공유가 가능하며, 부서 간 데이터 협력에 유리한 구조를 제공한다(Jiong'en & Yingliang, 2024).

이와 더불어, Chen & Lee(2021)는 통합적 거버넌스 프레임워크를 제시하였는데 이는 기존 전자정부, 협력 거버넌스, 네트워크 관리 연구를 통합한 것으로 데이터 공유 및 의사결정 지원 서비스가 조직 간 경계를 넘어 효과적으로 운영될 수 있는 기반을 마련하였다. 이

프레임워크는 네트워크의 서비스 생산 및 제공 성과 측정을 포함하여, 정보 품질, 데이터 표준화, 시의성, 이용자 만족도를 종합적으로 평가할 수 있도록 설계되었다.

[그림 2-16] 통합적 거버넌스 프레임워크

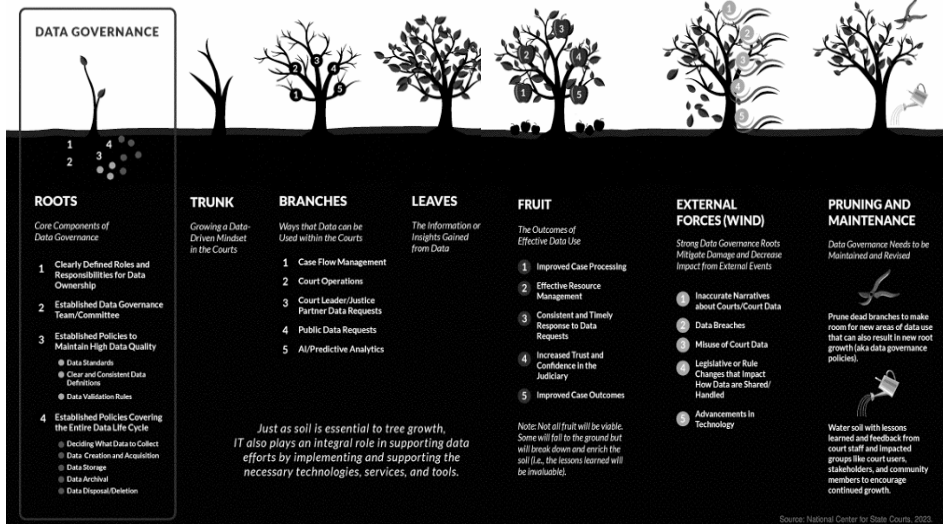


자료: Chen & Lee(2021)

2) 전통적 거버넌스와 AI 데이터 기술의 결합

최근 데이터 거버넌스가 단순한 관리 절차가 아니라 데이터 기반 조직문화를 뿌리내리고 성장시키기 위한 토대임을 강조한다(Data Governance and AI in State Courts, 2024). 데이터 거버넌스가 단순한 기술적 관리가 아니라 사람, 정책, 절차, 기술이 결합된 프레임워크로서 데이터의 품질과 보안, 일관성을 확보하는 데 중점을 둔다. AI 기술이 확대됨에 따라, 데이터 거버넌스의 중요성은 과거보다 부각되고 있으며 이는 데이터 활용의 ‘뿌리’임이 강조된다.

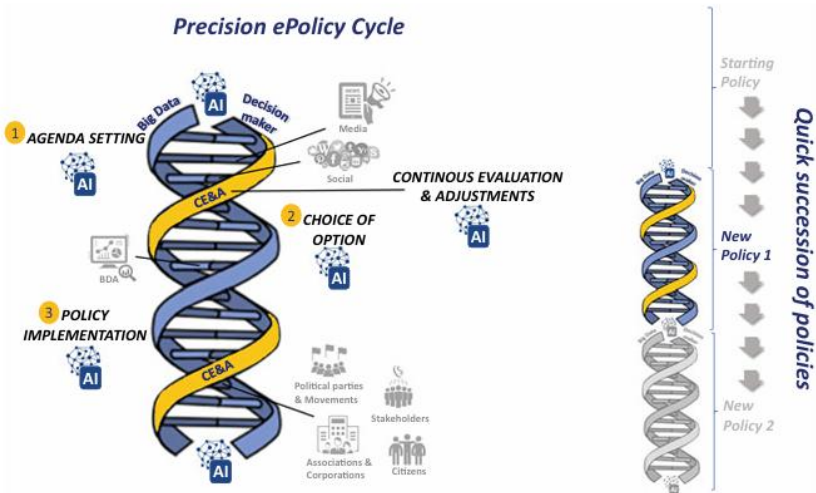
[그림 2-17] The First Step For Growth: Data Governance



자료: Data Governance and AI in State Courts(2024)

이러한 흐름에서 Carlizzi & Quattrone(2022)는 전통적인 정책 결정 과정(policy cycle)을 인공지능(AI)과 데이터 기반 거버넌스 기술을 통해 재구성한 새로운 이론 모형으로 정밀 전자정책 순환모형(Precision e-Policy Cycle)을 제안하였다. 이 모형에서 이중나선의 한 축은 데이터 기반의 분석과 예측을 담당하고, 다른 축은 시민과 정책결정자 간의 피드백 루프와 참여를 의미한다. 이 구조를 통해 정책 과정은 선형(linear)이 아니라 동시적이고 반복적이며, 실시간으로 조정 가능한 체계로써 재정의된다.

[그림 2-18] Precision e-Policy Cycle



자료: Carlizzi & Quattrone(2022)

다. 글로벌 및 국가 데이터 거버넌스 프레임워크

1) 글로벌 데이터 거버넌스의 메타-프레임워크 접근과 핵심 원칙

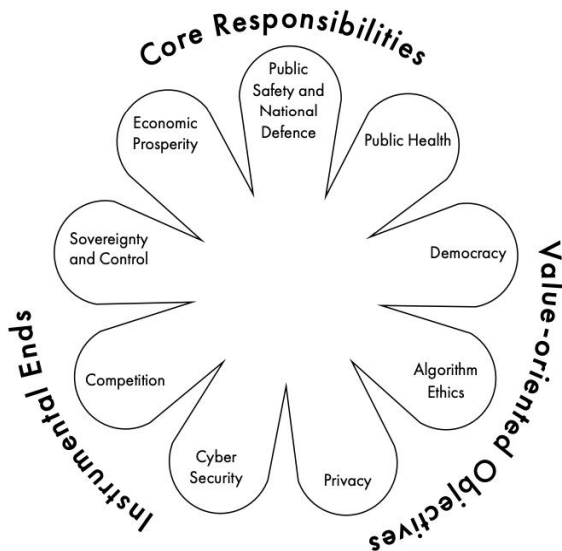
글로벌 데이터 거버넌스는 신뢰, 투명성, 공공성, 프라이버시, 접근성 등 핵심 원칙을 중심으로 다양한 국가 및 이해관계자 간 협력을 강조하는 메타-프레임워크로 발전하고 있다. Marcucci et al.(2023)는 데이터 거버넌스를 구성하는 6가지 핵심 축 -목적, 원칙, 법적 근거, 데이터 정의, 생애주기, 거버넌스 프로세스 및 실천, 데이터 스튜어드십-을 중심으로 글로벌 데이터 생태계에 적용가능한 종합적인 모델을 제시하였다. 이는 국가 간 조화를 위한 일종의 메타-프레임워크로 간주될 수 있다.

MacFeely et al.(2022)은 '글로벌 데이터 컴팩트(Global Data Compact, GDC)'의 필요성을 강조하며, 데이터의 공공성(public good)과 민간 소유권(data proprietary) 간 균형, 프라이버시 보호와 혁신 촉진 사이의 균형을 달성하기 위한 정책적·기술적 프레임워크를 제안하였다. Aaronson et al.(2021)은 데이터 거버넌스를 국가 전략, 민주적 책임성, 제도적 구조, 시민 참여, 국제 협력까지 포괄하는 종합적 체계(comprehensive system)로 재정립할 것을 주장하며, 이를 위해 'DataGovHub 패러다임'이라는 정량적 모형을 제시하였다.

2) 국가별 데이터 거버넌스 모델과 전략의 특성

국가 차원에서는 각국의 정책목표와 가치 체계를 반영한 다양한 데이터 거버넌스 모델이 제시되고 있다. Medhora et al.(2018)은 캐나다의 국가 데이터 프레임워크를 핵심 책임(Core Responsibilities), 가치 지향적 목표(Value-oriented Objectives), 도구적 목적(Instrumental Ends)이라는 세 층위로 구성하여, 단순한 기술적 관리가 아닌 국가의 핵심 가치와 정책목표를 통합적으로 고려해야 함을 강조하였다.

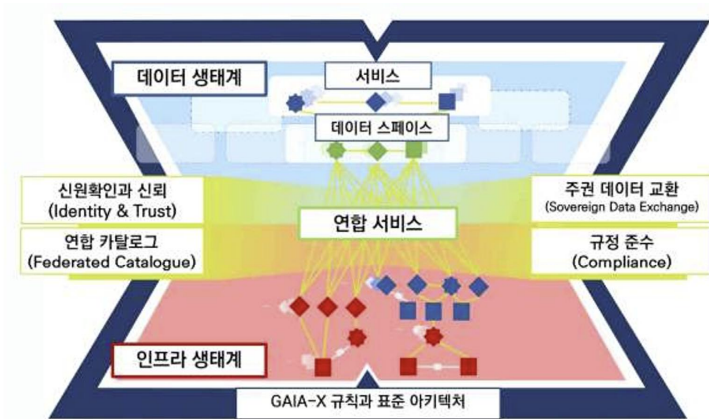
〔그림 2-19〕 캐나다의 국가 데이터 프레임워크 모델



자료: Medhora et al.(2018)

유럽연합(EU)은 데이터 거버넌스 법(Data Governance Act, DGA), AI 법(AI Act), 데이터 법(Data Act) 등을 통해 수평적·부문별 규제 체계를 마련하고, FAIR 원칙 및 신뢰중개기관 등 구조적 관리 체계를 제도화하고자 한다(Bhouri, 2025). 이러한 접근은 데이터의 재사용성과 책임성을 높이기 위한 노력의 일환이라고 할 수 있다. 또한, 유럽연합(EU)은 민간 주도의 차세대 유럽 데이터 인프라와 이를 기반으로 한 독자적인 디지털 생태계 구축을 위한 'GAIA-X 프로젝트'를 추진하고 있다.

[그림 2-20] EU의 GAIA-X 생태계 구성요소

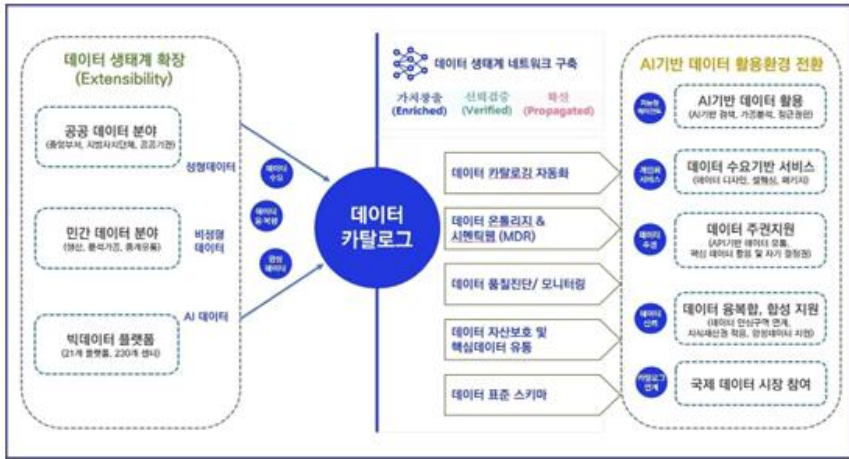


출처 : GAIA-X Architecture Document ('22.4)

- ▶ (데이터 생태계) 데이터 경제의 활성화를 위한 데이터 교환방식, 상호운용성 등에 대한 기준, GAIA-X 적용시나리오 등에 해당
 - ▶ (인프라 생태계) EU 내의 다양한 IT 인프라 제공업체의 상호운용성 확보를 위해 표준화·결합·연계 서비스를 위한 기술적인 방안
 - ▶ (연함 서비스) 참여자 관리·신뢰 메커니즘·데이터 접근·사용 권한 보장을 통한 데이터 주권 확보, EU GDPR·데이터보호 규정 준수, 인프라·응용서비스·데이터 상호운용성 등
- 자료: 한국지능정보사회진흥원(2024)

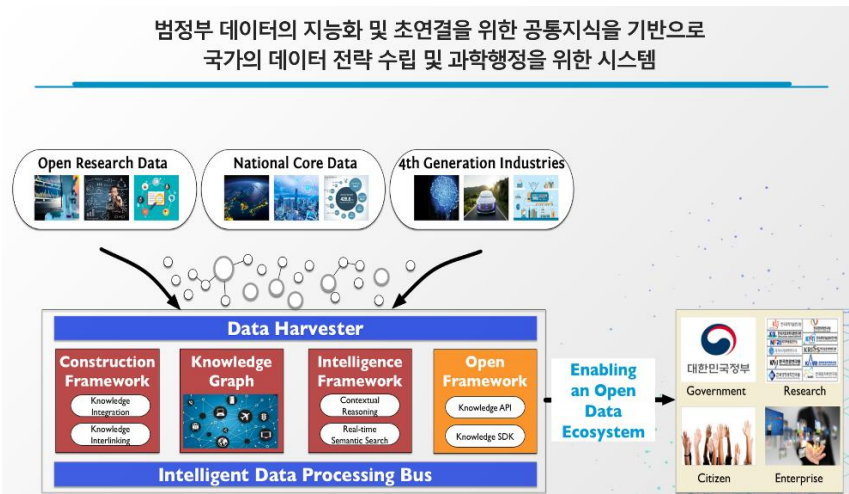
한국은 경제적 가치와 사회적 가치 간 균형을 강조하면서, 데이터 주권, 데이터 표준화, 플랫폼 기반 연계, 공공-민간 협력 강조하고 있다(한국지능정보사회진흥원, 2024; 김학래, 2020). 특히 한국지능정보사회진흥원(2024)는 국가 데이터 카탈로그 목표 모델을 통해 유럽의 GAIA-X와 같은 국제 협력모델과 같은 한국형 모델 구축의 방향성을 제시하였다. 이 모델은 표준화 체계 정립, 분산형 데이터 저장 및 공유체계 설계, 자율적 데이터 관리환경 조성, 신뢰 기반의 데이터 유통 체계 확립 등을 통해 국가 데이터 인프라가 필수적인 전략 자산이며, 공공정책 수립, 산업혁신, 시민의 디지털 권리 보장을 위한 핵심 인프라임을 강조한다.

〔그림 2-21〕 한국의 국가 데이터 카탈로그 목표 모델



자료: 한국지능정보사회진흥원(2024)

〔그림 2-22〕 범정부 국가지식플랫폼

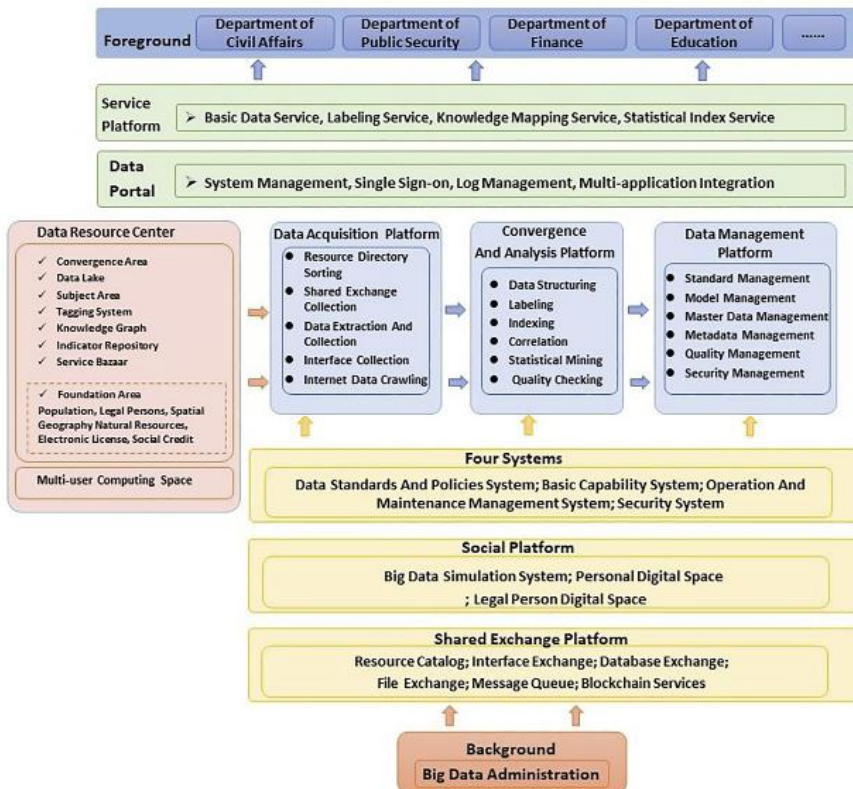


자료: 김학래(2020)

Mao et al.(2022)는 데이터 미들 플랫폼(Data Middle Platform) 기반 프레임워크를 통해 부처 간 데이터 연계와 공동활용을 가능하게 하는 기술적·조직적 기반을 마련하였다. 이

들은 데이터 파편화 문제(Khatri & Brown, 2010), 데이터 연계의 복잡성(Abraham et al., 2019), 그리고 정보시스템 기반의 공공데이터 통합 문제(Winter & Davidson, 2019) 등을 비판적으로 검토하여 프레임워크 설계에 반영하였다. 이들은 중국 하이난 지방의 사례를 통해 미들 플랫폼 기반 거버넌스가 행정서비스 혁신과 정책 개발에 기여하는 바를 시나리오 기반으로 설명하고 있으며, 데이터 연계 프로세스, 표준화 절차, 품질관리 체계, 그리고 부처 간 협력 메커니즘의 유기적 메커니즘을 시각적으로 도식화하였다. 다양한 정부 부처가 플랫폼을 통해 데이터를 주고받는 과정에서 발생할 수 있는 중복 방지, 책임성 확보, 실시간 피드백 구조 구축 등 실무적 고려사항을 실증적으로 검토하였다.

[그림 2-23] 데이터 미들 플랫폼 기반 프레임워크



자료: Mao et al.(2022)

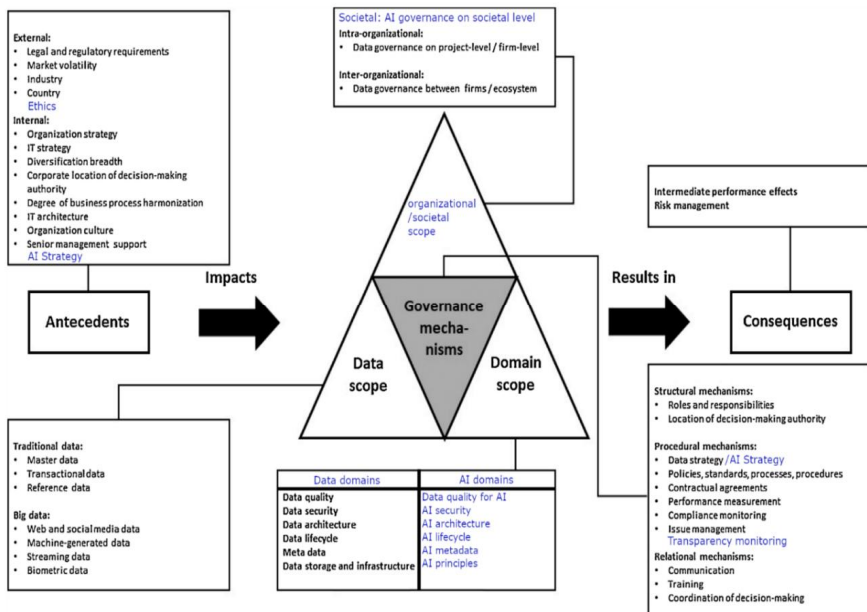
라. 인공지능(AI) 및 신기술 통합 데이터 거버넌스

1) AI 거버넌스 프레임워크의 구성요소 및 모듈형 접근

인공지능(AI)의 확산으로 인해 데이터 거버넌스는 신뢰성, 설명가능성, 공정성, 책임성, 윤리 등 다양한 개념을 통합적으로 다루게 되었다. AI 거버넌스 프레임워크(AI Governance Framework)는 데이터 품질 관리, 메타데이터, 리스크관리, 모델 평가, 편향 및 차별 제거, MLOps 관리 등 다양한 요소들을 포함하고 있다.

Schoeman, P.(2023)은 이러한 AI 거버넌스를 데이터 범위, 조직 및 사회 범위, 도메인 범위, 거버넌스 메커니즘, 선행조건과 결과 등 6가지 구성요소를 중심으로 설계하였다. AI 거버넌스는 데이터 거버넌스의 일부로 기능하지만, 모델 품질, AI 보안, AI 아키텍처, AI 라이프사이클, AI 메타데이터 등 독자적 관리 영역을 포함하며, 사회적 책임 영역(societal scope)까지 다루고 있다.

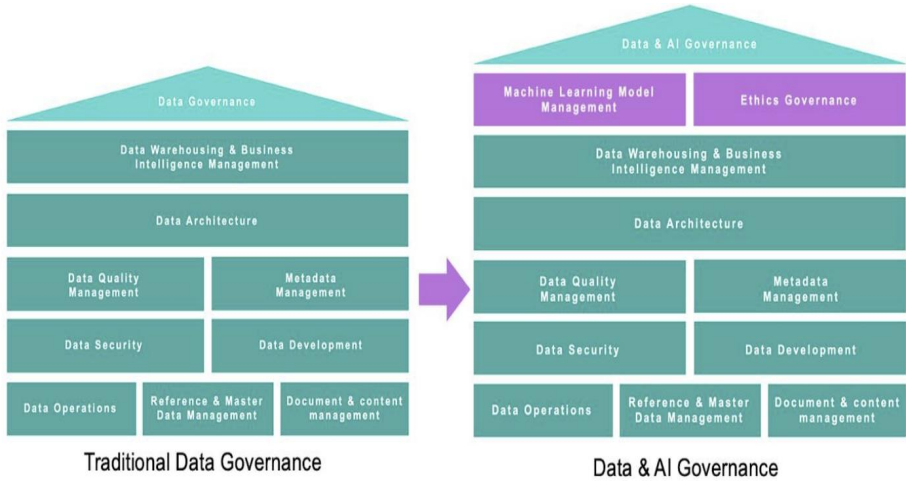
[그림 2-24] AI Governance Framework



자료: Schoeman, P.(2023)

Alqudah(2021)는 AI 거버넌스를 조직 내부 거버넌스 구조 및 통제 메커니즘, AI 의사결정 모델 정의 및 리스크 평가 시스템, 운영 프로세스 및 알고리즘 성능 관리 체계, 고객 및 이해관계자와의 투명한 커뮤니케이션 및 신뢰 구축 전략의 네 가지 핵심 구성 요소로 정의하였다. 이러한 프레임워크는 알고리즘 편향성, 개인정보보호, 설명 가능성(Explainability) 같은 윤리적 요소를 체계적으로 반영하며, 기술과 제도의 조화를 추구한다.

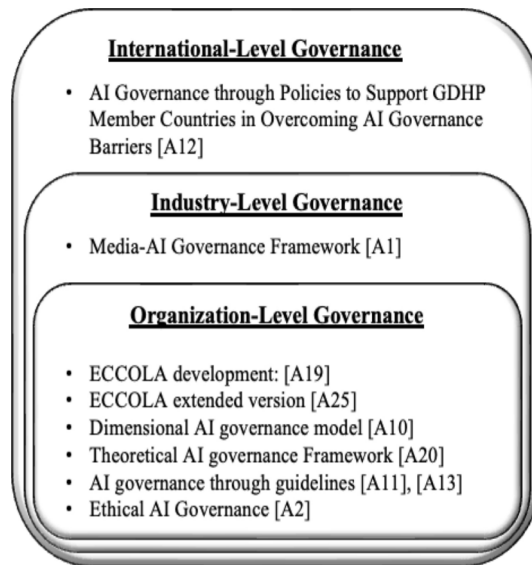
[그림 2-25] 전통 데이터 거버넌스로부터 데이터 & AI 거버넌스로의 변화



자료: Alqudah(2021)

Batool et al.(2025)은 AI 거버넌스 프레임워크를 팀-조직-산업-국가-국제의 다층적 구조로 접근하고, 윤리위원회·영향평가 등 다양한 통제 메커니즘을 포함한 모듈형 접근을 제안하였다. 이는 AI 거버넌스 프레임워크가 단일화된 모델로 수렴되는 것이 아닌 법적 환경, 산업 특성, 조직의 성숙도, 기술 수준에 따라 적합한 거버넌스 프레임워크를 선택하고 조합하는 ‘모듈형 접근(modular approach)’이 필요하다는 점을 강조한다. 또한, 인공지능의 발전이 가져오는 사회적 영향과 그 기반이 되는 데이터 관리의 중요성을 조망하며, 지속 가능한 AI 실현을 위한 데이터 거버넌스 프레임워크가 제시되기도 한다. 이러한 프레임워크에서는 AI 및 데이터 정책 통합 운영, ESG 기반 성과관리, 사회적 책임(포용성 등) 등이 강조된다(Tukia, 2022).

[그림 2-26] AI 거버넌스 계층구조



자료: Batool et al.(2025)

[그림 2-27] 지속가능한 AI를 위한 데이터 거버넌스 프레임워크

	AI Governance Elements		
	AI System (technical)	Organisation	Ecosystem
Objectives & Key Results	To include the prevention of technical risk and vulnerabilities of AI system data in strategies	To include the alignment of data activities, business goals and organisational values related to AI in strategies	To include AI related legal, regulatory, and societal value and norm compliance to strategies
Data Governance Activities	To ensure adequate decision rights and accountabilities to govern data and AI systems	To ensure adequate decision rights and accountabilities to reach business goals and value alignment related to AI systems	To have the necessary decision right and accountabilities to ensure positive impact on AI ecosystems, and alignment with laws, regulations, values, and norms
Decision Rights and Accountabilities			
Data Policies, Rules, Definitions, and Standards	To have the necessary data policies, rules, definitions, standards to ensure human control over AI system data, prevent technical vulnerabilities, and foster expertise	To have the necessary data policies, rules, definitions, standards to reach business goals and to ensure alignment with organisational values	To have the necessary data policies, rules, definitions, standards to ensure positive impact on AI ecosystems, and alignment with laws, regulations, values, and norms
Roles and Responsibilities	To ensure adequate roles and responsibilities to govern data and AI systems	To ensure adequate roles and responsibilities to reach business goals and value alignment related to AI systems	To ensure adequate roles and responsibilities to positively impact AI ecosystem, and alignment with laws, regulations, values, and norms
Data Processes	To have standardised, documented, and repeatable processes for AI system data	To have the necessary data processes to ensure alignment with organisational values	To have the necessary data processes to ensure positive impact on AI ecosystems, and alignment with laws, regulations, values, and norms
Data Governance Metrics	To measure the impact of data governance activities on AI systems	To measure the impact of data governance activities on organisations and businesses developing or using AI	To measure the impact of data governance activities on AI ecosystem and the level of alignment with laws, regulations, values, and norms
Controls	To prevent technical risks and vulnerabilities of AI system data	To prevent business and value compliance related risk regarding AI systems	To prevent legal, regulatory, value, and norm compliance risks related to AI system data
Middle section: what is the purpose of each activity from AI governance perspective			

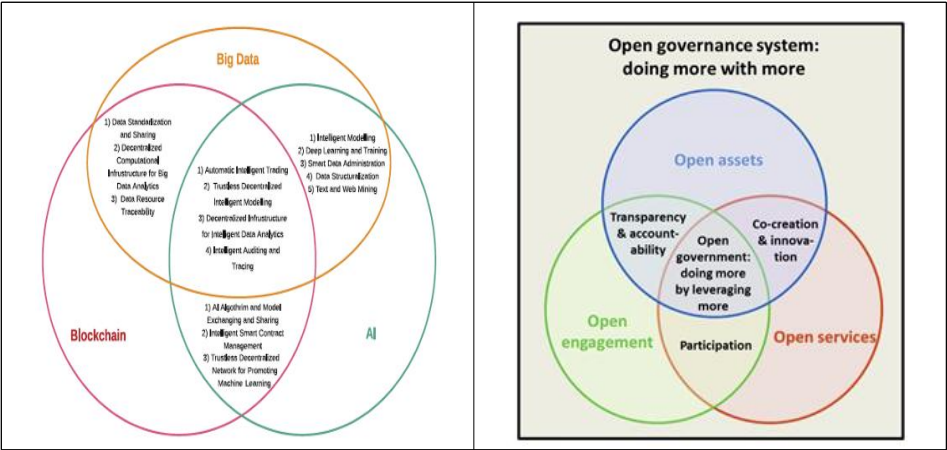
자료: Tukia(2022)

2) 신기술 융합 기반 데이터 거버넌스 전략

AI의 확산은 블록체인(Blockchain), 클라우드(Cloud) 등 신기술과 결합하여 데이터 거버넌스의 새로운 형태를 요구한다. Tan(2021)과 GENERATIVE AI FOR DATA STEWARDS (2024)는 신기술과 결합된 개방형 데이터 거버넌스(Open Data Government, ODG) 거버넌

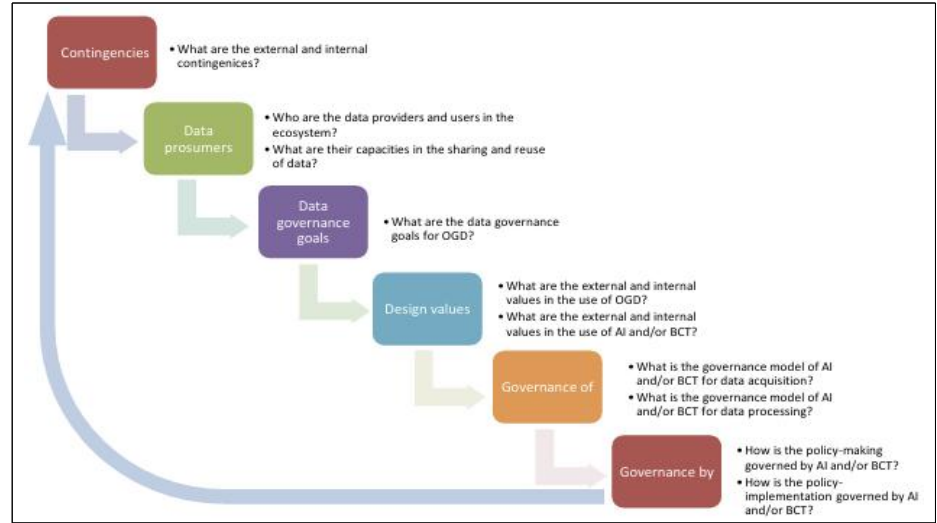
스 구조와 Non-Invasive Data Governance(NIDG) 모델 기반의 전사적 데이터 스튜어드십을 통해 데이터 관리 책임의 분산을 강조했다.

[그림 2-28] Open Data Governance Design



자료: Hassani et al.(2019); Millard(2018); Tan(2021)

[그림 2-29] The 6-Step model in Open Government Data(OGD) Model design



자료: Tan(2021)

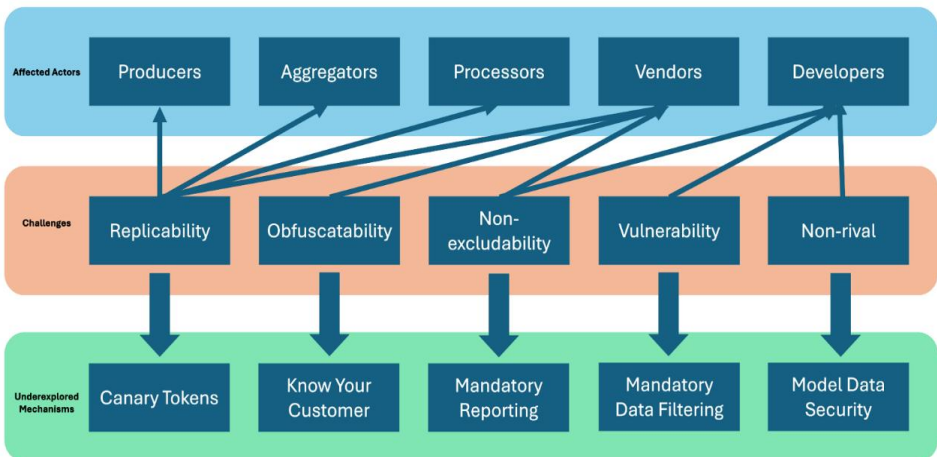
[그림 2-30] Non-Invasive Data Governance(NIDG) 모델에서의 데이터 관리



자료: Sugureddy, A. R.(2024)

Hausenloy et al.(2024)는 ‘프론티어 데이터 거버넌스(Frontier data governance)’ 개념을 통해 AI 개발 파이프라인 전반에서 데이터를 체계적으로 모니터링하고 규제하며 통제할 수 있는 정책과 관행, 메커니즘을 제안하였다. 특히 데이터의 비경합성(non-rival), 비배제성(non-excludable), 복제 용이성(replicability) 등의 특성으로 인한 데이터 통제의 어려움을

[그림 2-31] 프론티어 데이터 거버넌스(Frontier data governance)에서의 데이터 관리

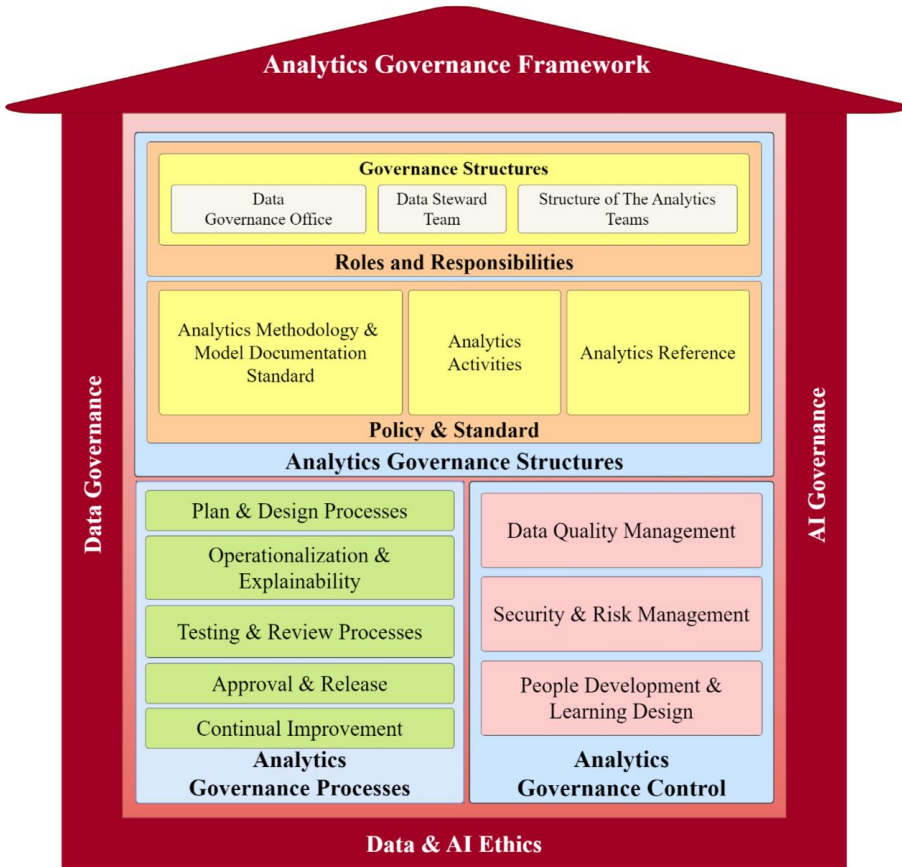


자료: Hausenloy et al.(2024)

해결하기 위한 정책 메커니즘(전략)으로 카나리아 토큰(canary tokens), 데이터 필터링, 보고 의무화, 데이터 보안 강화, 데이터 제공자 대상의 신원확인(Know Your Customer, KYC) 규제 등을 제시하였다. 이러한 전략들은 AI 및 신기술 융합 시대에 효과적인 데이터 거버넌스를 구축하기 위한 핵심 요소로 평가된다.

Kanying et al.(2023)은 데이터 분석 거버넌스(Analytics Governance)의 필요성을 강조하며, 기존의 데이터 거버넌스와 AI 거버넌스를 통합한 포괄적이고 실용적인 분석 거버넌스 프레임워크를 구축하였다. 이 프레임워크는 크게 구조(Structure), 프로세스(Process), 통제(Control)의 세 가지 주요 요소로 구성된다. 먼저 구조 측면에서는 명확한 역할 및 책임 규정과 정책 표준이 중심이다. 데이터 거버넌스 오피스(Data Governance Office)가 데이터 표준과 정책 수립, 데이터 전략을 총괄하며, 데이터 스튜어드 팀(Data Steward Team)이 실제 데이터 품질과 규정 준수를 관리한다. 데이터 분석 팀(Analytics Teams)은 프로젝트 매니저, 데이터 엔지니어, 데이터 분석가, 데이터 과학자 등으로 구성되어 조직 내 데이터 분석 작업의 일관성과 전문성을 확보한다(Kanying et al., 2023). 다음으로 프로세스는 계획·설계, 운영화, 테스트 및 평가, 승인 및 배포, 지속적 개선의 5단계로 구성된다. 계획 및 설계 단계는 목표 설정과 분석방법 및 기술 선정이 이루어지며, 운영화 단계는 분석 결과를 실제 업무에 적용하고 설명 가능성을 확보하는 것이 핵심이다. 테스트 및 평가는 규정 준수 여부를 점검하고 품질을 유지하며, 승인 및 배포 단계는 명확한 프로세스에 따라 이뤄진다. 마지막으로 지속적인 개선 과정은 피드백을 통해 분석 체계의 질적 향상을 지속적으로 도모한다(Kanying et al., 2023). 마지막으로 통제 영역은 데이터 품질 관리, 보안 및 위험 관리, 인력 개발 및 학습 설계 등 세 가지로 구분된다. 데이터 품질 관리에서는 데이터 정확성과 일관성을 평가하고 유지한다. 보안 및 위험 관리 부문에서는 데이터 분석 과정에서 보안 취약성과 위험을 평가 및 관리하여 분석의 신뢰성을 보장한다. 마지막으로 인력 개발 및 학습 설계는 조직 내 인력의 분석 역량을 지속적으로 향상시키는 교육과 역량 강화 프로그램을 설계하여 데이터 분석 작업의 전문성을 제고한다(Kanying et al., 2023).

[그림 2-32] 데이터 분석 거버넌스



자료: Kanying, T., Thammaboosadee, S., & Chuckpaiwong, R.(2023)

제 2 절 주요국의 데이터 거버넌스 사례

1. 주요국의 AI 거버넌스 정책과 데이터 거버넌스 모형

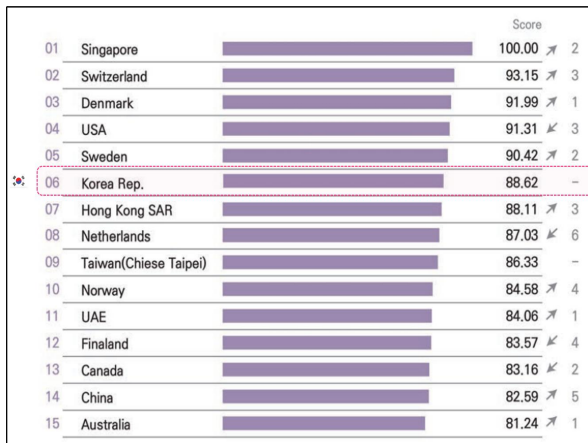
가. 국가별 데이터 거버넌스 접근법의 차별화

1) AI·데이터 거버넌스의 글로벌 확산과 의미

디지털 전환과 인공지능(AI) 기술의 급속한 발전은 세계 각국에 거버넌스 구조의 근본적 변화를 요구하고 있다. 전통적 산업사회에서 정보사회로의 변화로 데이터와 AI는 경제적 경쟁력의 원천이자 핵심 자원이 되었다. 이에 따라 각국은 데이터 보호, 활용, AI 윤리와 같은 여러 과제를 동시에 해결해야 하는 ‘거버넌스 혁신’을 추진 중이다.

주요 정책지표를 통해 각국의 데이터 거버넌스 역량을 비교할 수 있었다. 대표적으로 스위스(International Institute for Management Development, IMD)의 ‘세계 디지털 경쟁력 지수’(2024년 기준)는 한국이 64개국 중 종합 6위에 오르는 등 디지털 경쟁력 강국임이 확인된다. 이외에도 OECD, UN, Stanford AI Index 등은 데이터·AI 기반 경제·사회의 기술적 역량과 제도환경, 윤리 프레임워크의 성숙도를 다층적으로 평가하여 국가별 특수성을 드러내고 있다.

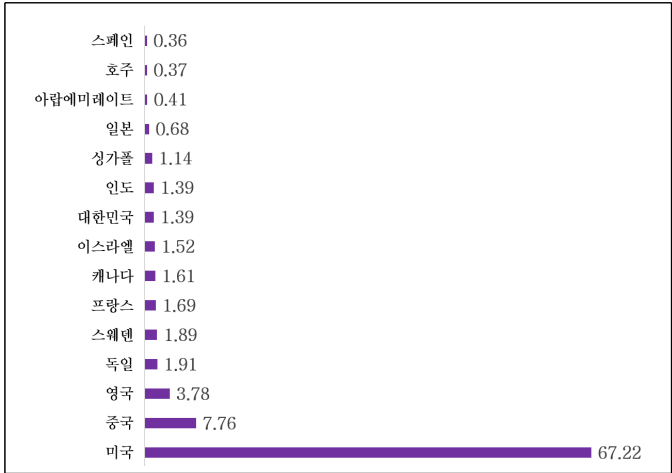
[그림 2-33] 2024년 IMD 세계 디지털 경쟁력 순위(1~15위)



자료: IMD(2024), 저자 재구성

2023년 기준 민간 AI 투자 규모를 살펴보면 미국이 672억 달러로 전 세계에서 가장 높은 투자를 기록하고 있으며 그 뒤를 중국, 영국, 프랑스, 독일, 캐나다 등이 잇고 있다.

[그림 2-34] 2023년 AI 민간투자 규모
(단위: 십억 USS)



자료: Artificial Intelligence Index Report(2023). 저자 재구성

글로벌 데이터·AI 거버넌스 지표를 통해 국가별 정책 및 제도적 성숙도의 현황을 객관적으로 파악할 수 있지만 수치와 순위만으로는 각국이 처한 고유한 정치·경제·사회적 맥락의 차이와 실제 정책 실행 메커니즘까지 충분히 설명하기 어렵다. 이에 본 절에서는 지표상 대표성을 갖춘 주요국(예: 영국, 미국, EU, 캐나다 등)을 선정하여 글로벌 순위 이면에 존재하는 각국의 거버넌스 구조, 정책 추진 방식, 실행 사례 그리고 각국의 전략을 심층 분석하고자 한다. 이러한 비교 분석은 지표 기반의 추상적인 논의를 넘어 실제 정책 설계와 제도 고도화에 필요한 실질적인 시사점을 도출하는 데 의의가 있다.

2) 국가별 데이터 거버넌스 모델의 구조적 차이

각국의 데이터 거버넌스 모델은 정치적, 경제적, 사회적 배경에 따라 현저한 구조적 차이를 보이고 있다. 유럽연합(EU)은 GDPR 중심의 강력한 규제 접근법을 채택하고 있다. 2018년 GDPR 시행을 통해 시민의 개인정보 보호를 기본권으로 규정하였으며 개인정보 오남용으로부터 보호를 핵심 목표로 설정하였다. 특히 역외적용 효력을 통해 EU와 거래하는

모든 기업에 전 세계적으로 유럽 데이터 보호법을 적용하도록 강제하고 있다.

미국은 CLOUD Act와 디지털 패권 추구를 통한 접근법을 보여준다. 2018년 CLOUD Act를 통해 서버의 지리적 위치에 관계없이 미국 기업이 소유한 모든 데이터센터에 대한 접근권을 확보하였다. 이 법률은 미국의 공공안전 보호와 가장 심각한 위반, 범죄, 테러리즘과의 싸움을 명분으로 제정되었다. 전 세계 디지털 스토리지 시장의 85%를 미국 기업이 운영하는 현실에서 이 법률은 광범위한 영향력을 행사하고 있다.

중국은 사이버 주권과 국가 주도 전략을 추진하고 있다. 2013년 시진핑 집권 이후 빅데이터 국가전략을 구축하고 사이버 강국을 위한 포괄적 접근법을 추진하고 있다. 2017년 국가정보법 제정을 통해 국가의 중요한 데이터 자원을 보호하는 능력 강화, 관련 입법 가속화, 데이터 재산권 보호 개선을 명시하였다. 중국의 데이터 거버넌스는 4대 구성축으로 이루어져 있다. 이는 디지털 기업과의 전략적 협력, 정부 통계에 빅데이터 통합, 빅데이터의 국가전략 수준 격상, 대규모 국가 데이터센터 구축이다.

3) 글로벌 데이터 정책 동향

주요국의 데이터 정책 전략을 비교 분석한 결과 각국은 디지털 전환과 정부혁신을 중심으로 데이터 활용의 체계를 구축하고 있다는 점에서 공통된 방향성을 공유하는 것을 확인할 수 있었다. 윤상필·권현영(2021)의 연구에 따르면 대한민국은 디지털 정부혁신 전략 아래 데이터 공유, 개방, 보안, 품질 제고, 행정 활용, 민관 협력, 산업 기반 조성 등 다층적 과제를 통합적으로 추진하고 있다. 다만 데이터 윤리 측면에 있어 상대적으로 언급이 제한적이며 정책문서에서 '디지털 포용'의 하위 요소로 간헐적으로 언급되는 수준에 그친다. 반면 미국과 영국은 데이터 윤리의 독립적 프레임워크를 마련하고, 평가체계와 정책기관(예: 데이터윤리혁신센터 등)을 통해 제도적으로 정착시키고 있다. EU는 디지털 단일시장의 구축을 위한 전략을 중심으로 고품질 데이터의 개방, 역내 공동 데이터 공간 마련, 허위정보 대응 등을 포함한 데이터 거버넌스 체계를 전방위적으로 구축하고 있으며, 디지털 경제와 정보주권을 아우르는 규범 기반 접근이 특징적이다.

거버넌스 구조에 있어서도 각국은 상이한 체계를 보인다. 대통령직속 국가인공지능위원회와 행정안전부 주도의 데이터기반행정 활성화위원회, 중앙행정기관으로 격상된 개인정보보호위원회 등을 통해 데이터 정책을 분산적이고 다원적으로 운영하고 있다. 미국은 대

통령 직속 관리예산처(OMB)를 중심으로 연방데이터정책위원회, 데이터최고책임자위원회(CDO Council)를 두고 민간부문은 자율을 존중하되 정부가 공정경쟁과 지원을 보완하는 방식을 취하고 있다. 영국은 기존의 부처 단위 정책을 내각사무처로 이관하며 중앙집중형 거버넌스로 재편하고 있으며, EU는 집행위원회를 통해 전략을 수립하고, 규정과 지침을 병행하여 회원국 법체계에 영향을 미치는 다층적 통합 방식으로 거버넌스를 구성한다.

법제화 측면에서 한국은 공공데이터법, 데이터기반행정법, 데이터산업기본법 등을 중심으로 공공과 산업을 아우르는 입법체계를 형성하고 있다. 미국은 증거기반정책수립법(2018), OPEN 데이터법, 주 정부 차원의 개인정보보호법(예: CCPA)을 통해 정책과 법제를 연결하고 있으며 EU는 데이터 거버넌스법(Data Governance Act)을 통해 역내 데이터 유통과 공동 활용을 제도화하고, 향후 데이터서비스법 등 후속 입법을 예고하고 있다. 특히, EU의 입법은 ‘데이터 이타주의’ 개념을 통해 데이터 공유와 재사용을 위한 조건과 절차를 구체화하고 있다는 점에서 제도적 진화를 보여준다.

전반적으로 한국의 데이터 정책은 기술적 활용과 산업진흥 중심의 기조가 강한 반면 주요 선진국은 윤리, 신뢰, 제도적 책임 등 비기술적 요소를 보다 적극적으로 제도화하고 있다. 이러한 차이는 향후 데이터 기반 거버넌스의 지속 가능성과 사회적 수용성 확보에 있어 정책 설계의 핵심적 변수가 될 것이다. 특히 한국이 국가 차원의 데이터 윤리 거버넌스를 확립하지 못할 경우 공공신뢰 기반의 데이터 정책 실현에는 구조적 제약이 따를 수 있다(윤상필·권현영, 2021). 이를 고려할 때 국가 차원의 CDO 체계 확립, 대통령 직속 데이터혁신기구 설립, 데이터윤리 프레임워크 구축은 정책적 시급성을 가지는 과제로 판단된다.

〈표 2-2〉 주요국 데이터정책 논의 비교

구분	대한민국	미국	영국	유럽연합
국가 전략의 핵심 요소	- 데이터 기반 정책 - 데이터 공유 - 데이터 활용 - 데이터 보안 - 데이터 품질 - 데이터 개방성 - 데이터 역량	- 데이터 기반 정책 - 데이터 윤리 - 데이터 보안 - 데이터 품질 - 데이터 역량 - 신뢰성 - 투명성	- 데이터 기반 정책 - 데이터 윤리 - 데이터 공유 - 데이터 활용 - 데이터 개방성	- 디지털 경제 - 데이터 품질 - 데이터 공유 - 데이터 보안 - 데이터 개방성 - 데이터 역량 - 신뢰성
거버넌스 체계	- 대통령직속 국가 인공지능 위원회 - 데이터특별위원회 - 데이터기반행정 활성화위원회 - 개인정보보호위원회	- 관리예산처 - 대통령실 산하 과학기술정책국 - 연방데이터정책위원회 - 연방데이터최고책임자위원회	- 디지털·문화·미디어·체육부 - 국무조정실	- 유럽연합 집행위원회 - 데이터보호위원회
관련 법제도	- 데이터기본법 - 데이터 3법(개인정보보호법, 정보통신망법, 신용정보법)공공데이터제공 및 이용활성화에 관한 법률 등	- 연방 프라이버시법 - 증거기반 정책 결정법(2018) - 전자정보자유법 - 전자정부법 - 공공데이터 개방법 - 캘리포니아 소비자 개인정보 보호법(CCPA)	- 디지털경제법 (2017)	- 일반개인정보 보호법(GDPR) - 데이터 거버넌스법 (DGA) - 디지털서비스법 (DSA)

자료: 윤상필·권현영(2021), 유승현(2022), 저자 재구성

4) 공공부문 AI 거버넌스 실행 사례

캐나다는 전략 수립부터 실행, 윤리, 규범까지 아우르는 중앙집중형 AI 거버넌스 체계를 구축하고 있으며, 이는 공공부문에서의 책임 있는 인공지능 운영 사례로도 구체화되고 있다. 특히 재정위원회 사무처(TBS)가 2019년 제정한 「자동화 의사결정 지침」은 그러한 실행

기반의 대표적 제도로, 공공영역에서의 AI 시스템 도입에 따른 위험을 체계적으로 관리하고 있다. Kuziemski & Misuraca(2020)의 연구에 따르면 해당 지침은 알고리즘의 사회적 영향력과 위험성에 따라 자동화된 의사결정 시스템을 4단계로 분류하는 평가 프레임워크를 제시한다. 위험 등급은 ‘저위험’, ‘중위험’, ‘고위험’, ‘극심위험’으로 구분되며, 등급에 따라 설명가능성, 감사 가능성, 공개 수준이 달라진다. 예컨대, 이민심사 시스템의 경우 극심위험에 해당하여 SHAP(Shapley Additive Explanations) 기반의 설명가능성이 요구되었고 이에 따라 시스템의 소스코드 공개와 외부감사가 의무화되었다. 이처럼 위험 수준에 따라 투명성과 책임성을 실질적으로 강화하는 메커니즘을 마련함으로써, 캐나다는 공공부문 AI 거버넌스에 있어 규범적 원칙을 실행 가능한 제도로 전환하는 데 선도적 사례를 보여주고 있다.

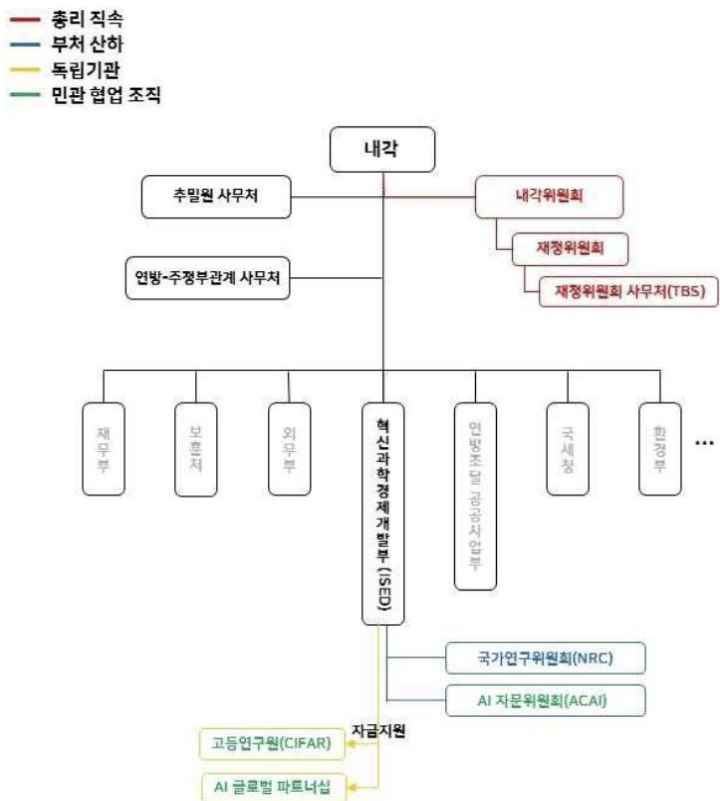
이는 앞서 설명한 캐나다의 거버넌스 구조와도 일관된 방향성을 보인다. 혁신과학경제개발부(ISED)를 중심으로 인공지능(AI) 정책의 총괄, 실행, 조정 기능을 수행하는 중앙집중형 거버넌스를 구축하고 있다. ISED는 2015년 설립 이래 범국가 AI 전략의 수립과 이행을 주도하고 있으며 연구개발(R&D) 투자와 관련 조직에 대한 자금 지원을 통해 전략 실행을 촉진하고 있다. 특히 ISED는 캐나다 고등연구원(CIFAR)과의 협력을 통해 국가 AI 전략(Pan-Canadian AI Strategy)을 세계 최초로 수립한 바 있으며, 토론토(Vector), 에드먼턴(AMII), 몬트리올(MILA)에 위치한 3대 AI 연구소를 중심으로 실험적이고 도전적인 연구를 지원하고 있다.

캐나다 AI 거버넌스의 특징은 민관 협의체 중심의 자문 구조와 국제 협력 기반의 윤리적 거버넌스에 있다. 대표적으로 AI 자문위원회(ACAI)는 산업계, 학계, 시민사회 전문가가 개인 자격으로 참여하는 민관 합동 자문기구로 ISED 장관에게 캐나다의 AI 정책, 글로벌 파트너십, 지식재산, 응용연구 및 상용화 등에 관해 권고를 제공한다. 해당 위원회는 외무부, 재정위원회 사무처(TBS), 추밀원 사무처 등과의 연계를 통해 정부 내 다양한 부처 간 전략적 일관성을 확보하는 역할도 수행한다. 동시에, AI 글로벌 파트너십(GPAI)를 통해 국제 사회와의 다자 협력을 추진하며, 책임성과 인권, 포용성과 다양성에 기초한 AI 활용 원칙을 국제적 차원에서 확산시키고 있다.

이와 같은 캐나다의 AI 거버넌스 체계는 중앙부처의 전략적 조정 기능을 중심으로, 민관

협의체와 국제 파트너십의 윤리적 틀, 공공연구기관의 기술 개발 및 상용화 지원, 규범 기관의 제도 정비가 유기적으로 결합된 구조라 할 수 있다. 각 구성 기관은 상호보완적 역할 분담을 통해 정책-연구-윤리-표준 간의 유기적 연계를 형성하고 있으며, 이는 캐나다 AI 정책의 책임성 기반 다층적 거버넌스 모델을 특징짓는다.

[그림 2-35] 캐나다의 AI 거버넌스 조직도

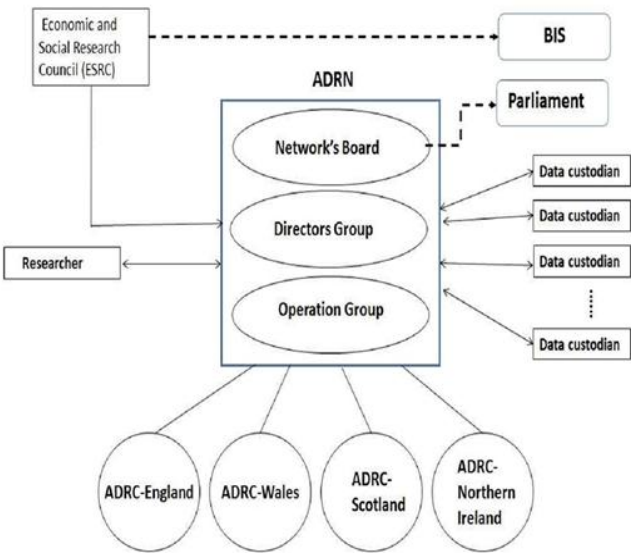


자료: 한국지능정보사회진흥원(2022)

5) 데이터 활용과 보호 모델: 영국 ADRN(Administrative Data Research Network) 사례
영국의 ADRN(Administrative Data Research Network) 사례는 데이터 활용과 보호의 균형 추구를 모범적 모델이다. ESRC 주도로 설립된 이 네트워크는 국가 차원의 분산된

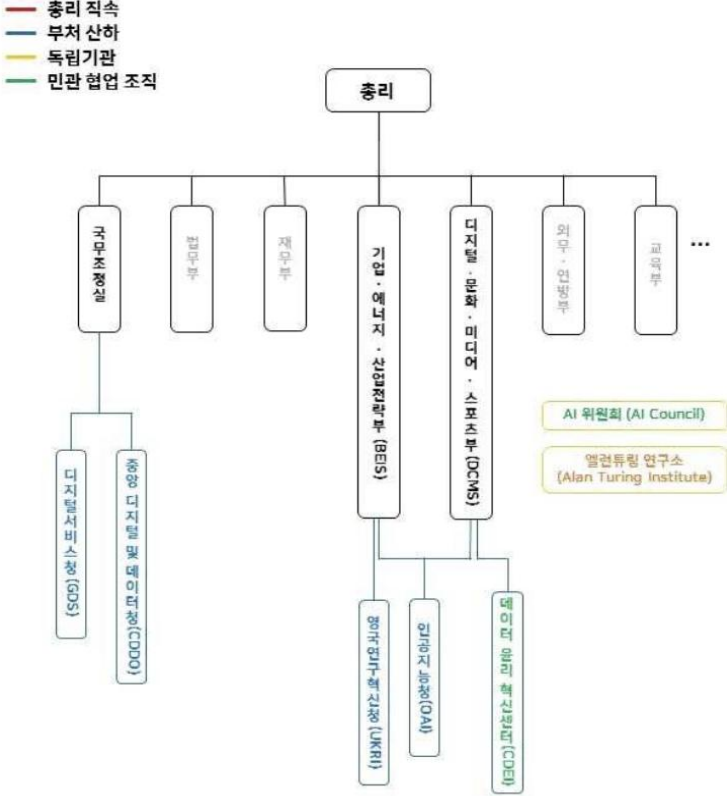
행정데이터를 연구자가 안전하게 활용할 수 있도록 지원하는 인프라이다. 4개 주요 구성 요소로 이루어져 있다. 이는 ADS, Secure Labs, 각국 데이터 소유기관, 독립 윤리 자문기구이다. ADRN은 데이터 비식별화 처리, 보안성 강화, 다기관 간 협업, 시민사회 수용성 제고를 위한 커뮤니케이션 전략이 복합적으로 작동한다. 핵심은 안전한 환경에서의 통제된 접근이라는 신뢰 기반 데이터 활용 메커니즘을 구축한 것이다. ADS는 데이터 제공자와 사용자 간의 신뢰를 조율하는 중간 조정기구 기능을 수행하며, 데이터 관리 정책, 보안 가이드라인, 접근 승인 절차를 총괄한다.

[그림 2-36] 영국의 AI 데이터 거버넌스의 핵심 요소



자료: 윤광석, 이건(2017)

이러한 모델을 바탕으로 영국의 AI 거버넌스 체계는 통합적 정책 조정과 분화된 기능적 역할을 하고 있다. 중앙정부 차원의 통합 전략 조정기구와 부처 간 협업 체계, 그리고 전문 기관별 분화된 기능 수행이라는 이중 구조를 통해 운영되고 있으며, 총괄 부처로는 디지털·문화·미디어·스포츠부(DCMS)가 있으며, 기업·에너지·산업전략부(BEIS)와 함께 영국 AI 국가 전략의 수립과 이행을 공동으로 주관한다. DCMS는 데이터 공유, 보호, 윤리 등



도, 자금조달 경로 마련 등도 그 범주에 포함된다. UKRI는 BEIS 산하의 공공기관으로, 9개의 전문 연구기관(예: MRC, EPSRC, AHRC 등)과 Innovate UK 및 Research England를 아우르는 대형 연구지원 체계이다. 이 기관은 기초연구부터 산업 협력까지 전 주기의 AI 연구 생태계를 지원하고 있다.

이와 같이 영국의 인공지능 거버넌스는 정책 총괄, 전략 실행, 기술 연구, 윤리 감시, 공공 혁신 등 각 기능별 전문 기관들이 상호 연계되면서도 분명한 역할 분담을 통해 작동한다. 이는 중앙집중적 거버넌스와 분권적 실행 체계 간의 균형 속에서 고도화된 AI 정책을 추진하는 영국의 전략적 특징을 반영하는 구조라 할 수 있다.

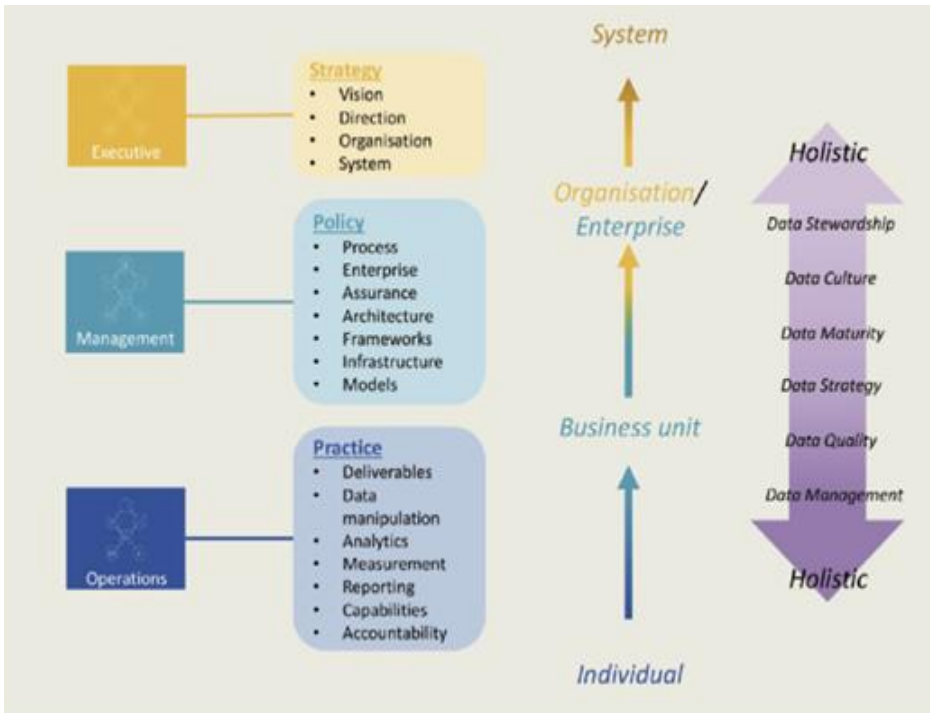
6) 인간 중심 AI 서비스 설계 모델

핀란드의 AuroraAI는 생애주기 기반 접근법을 통한 인간 중심 AI 서비스의 모범 사례이다. 핀란드는 공공서비스 AI 활용 시 반드시 ‘인간 중심성, 설명 가능성, 공정성’ 기준을 적용하도록 규정하였다. 모든 공공기관은 특정 AI 시스템을 도입하기 전에 AI 윤리 사전 평가서를 작성해야 하며, 모델이 실제 운영된 이후에도 정기적으로 관련 리스크를 모니터링해야 한다. 이 모델은 공공 AI의 신뢰성을 확보하려는 대표적 사례라고 할 수 있다.

2019년 시범 운영된 이 시스템은 12개 생애주기 이벤트를 중심으로 맞춤형 공공서비스를 제공한다. 출산, 취업, 은퇴 등의 주요 생애 사건에 따라 서비스를 구성하였다. 강화학습을 활용한 서비스 추천 시스템은 사용자 15,000명을 대상으로 한 설문에서 89%의 접근성 개선 효과를 보였다. 그러나 알고리즘 투명성 지수(ATI)는 54점으로 부분적 개선 필요성이 제기되었다. AuroraAI의 혁신적 측면은 MyData 플랫폼과의 연계에 있다. 개인데이터 주권 강화를 위해 시민에게 직접 데이터 접근, 수정, 삭제 권한을 부여하였다. 이는 GDPR의 데이터 이동권을 넘어선 적극적 데이터 통제 권한을 제공하는 혁신적 모델로 평가된다.

뉴질랜드 정부는 원주민 마오리 공동체와의 공동설계를 통해 지역 특수성을 반영한 데이터 거버넌스 모델을 개발하였다. ASEAN은 데이터 생애주기 전 단계에 걸친 원칙과 행동지침을 포함하는 데이터 관리 프레임워크를 구축하였다. 이러한 사례들은 보편적 원칙과 지역별 사회·문화적·법적 특수성을 고려한 유연한 적용의 필요성을 확인해준다.

[그림 2-38] 뉴질랜드 데이터 거버넌스 체계



자료: OECD(2019)

2. 국제기구 차원의 데이터 거버넌스 프레임워크

가. OECD 및 유엔기구 중심 글로벌 데이터 거버넌스 모델의 특성과 시사점

Marcucci et al.(2023)의 글로벌 벤치마킹 연구에서는 다양한 국제기구와 국가 단위에서 개발된 데이터 거버넌스 프레임워크를 분석하였다. 유엔 산하 기관, 비정부기구, 정부 부처, 지역 기구 등 다양한 국제·국가 단위에서 데이터 거버넌스 프레임워크가 분석되었다. UNHCR, UNICEF, WHO, IOM 등 유엔계 기관들은 인권과 프라이버시 보호에 중점을 두면서도 기관별 특성을 반영한 다채로운 접근 방식을 보여주었다.

또한 OECD가 2018-2019년에 33개 회원국을 대상으로 실시한 디지털정부 평가에서 한국이 종합지수 1위를 기록하였다. OECD는 데이터 기반 정부 구축을 위한 5대 핵심 과제를 제시하였다. 이는 데이터를 전략적 자산으로 인식하고 가치를 정의하는 것, 데이터 관

리·공유·재사용의 장벽을 제거하는 것, 데이터를 활용한 혁신적 정책 설계와 모니터링을 실시하는 것, 데이터 공개와 조직 간 활용을 중시하는 것, 윤리적 행위와 개인정보 보호 측면에서 데이터 권리를 이해하는 것이다.

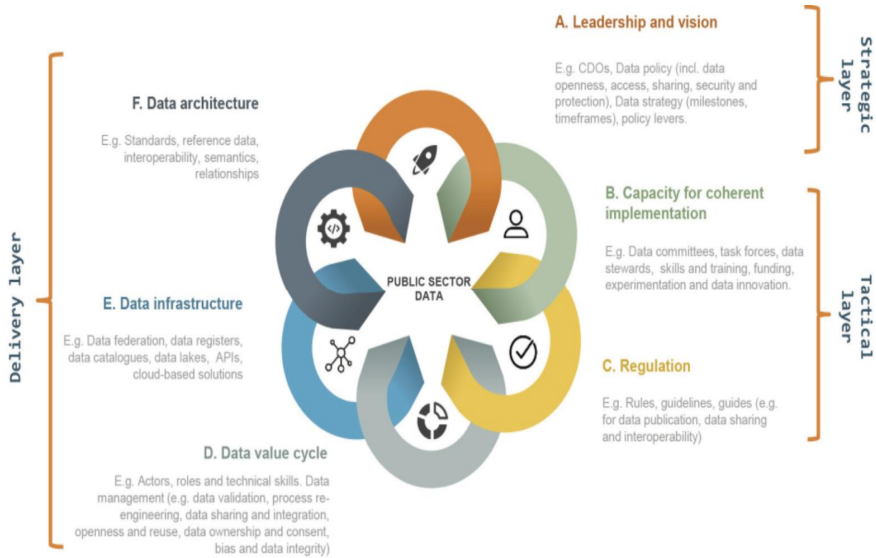
OECD 데이터 거버넌스 공통모델은 전략-전술-전달의 3개 핵심단계로 구성되어 있다. 전략 단계에서는 리더십과 비전을 통해 데이터 기반 정책의 방향성을 확립한다. 전술 단계에서는 일관된 구현과 법률·규제 체계를 통해 정책을 조정한다. 전달 단계에서는 데이터 가치 순환, 인프라, 아키텍처를 통해 실질적인 데이터 활용을 구현한다.

최상위 층위인 전략적 기획 단계에서는 조직 전반의 데이터 활용 철학과 목표 체계를 정립하는 데 중점을 둔다. 이 과정에서 가장 핵심적인 요소는 데이터 관리에 대한 명확한 리더십 구조의 확립이다. 정부 조직 내에서 데이터 정책을 총괄할 수 있는 거버넌스 체계를 마련하되, 이는 기존 행정 체계 내 새로운 역할 부여 방식이나 별도의 전담 조직 신설 방식 모두 가능하다. 더불어 데이터 기반 행정을 통해 달성하고자 하는 정책적 목표와 사회적 가치를 명문화하여, 조직 구성원들이 공유할 수 있는 비전을 제시해야 한다.

중간 층위인 전술적 실행 단계에서는 상위 전략을 구체적인 실행 체계로 전환하는 역할을 담당한다. 이 단계의 핵심은 부처 간 또는 기관 간 데이터 정책의 일관성을 확보하고, 상호 협력할 수 있는 메커니즘을 제도화하는 것이다. 특히 데이터 관련 업무를 조정하는 중앙 기구의 설립과 함께, 각 부문 간 지식 공유와 협업을 촉진할 수 있는 네트워크 체계를 구축해야 한다. 아울러 데이터의 수집, 처리, 공유, 보호에 관한 법적 규범과 가이드라인을 정비하여, 데이터 활용의 투명성과 책임성을 담보할 수 있는 규제 환경을 조성한다.

마지막 운영적 실현 단계에서는 데이터의 실제적 활용을 위한 기술적 기반과 운영 체계를 구축한다. 데이터 생명주기 관리 측면에서는 생산된 정보가 개방, 재이용, 통합되는 전 과정에서 상호 호환성과 표준화를 보장하는 체계를 마련해야 한다. 기술적 인프라 구축에서는 다양한 형태의 데이터를 통합적으로 관리할 수 있는 플랫폼을 개발하고, 정책 목표 달성을 위한 분석 도구와 처리 시스템을 구축하는 것이 중요하다. 이러한 기술적 구성요소들은 전체 거버넌스 체계에서 가장 가시적이고 직접적인 성과를 창출하는 영역으로 작용한다.

[그림 2-39] OECD 데이터 거버넌스 모델



자료: OECD(2019)

최근 OECD는 2025년 2월 보고서 「Steering AI's Future」에서 인공지능(AI) 정책환경에 구체적으로 적용하는 분석을 시도하였다. 해당 보고서는 AI의 급속한 발전이 야기하는 다양한 사회적 도전과 기술적 기회를 보다 효과적으로 대응하기 위해, 예측적이고 미래지향적인 접근 방식이 AI 거버넌스 전반에 필요함을 강조하고 있다.

첫째, 가치(values) 측면에서 효과적인 AI 거버넌스는 인권과 민주주의를 수호하면서도 신뢰할 수 있는 기술 개발을 보장하는 명확한 원칙을 필요로 한다. 이에 따라 OECD는 2019년 최초로 수립하고, 2024년에 최신화한 「AI 원칙(OECD AI Principles)」을 통해 포용적 성장, 인권 존중, 투명성, 안전성, 책임성 등 다섯 가지 가치 기반 원칙과 함께 정책 입안자에게 주는 다섯 가지 실천 권고를 제시하였다. 이러한 가치들은 OECD.AI 정책 관측소(Policy Observatory) 및 AI 도구·지표 목록(Catalogue of Tools & Metrics)을 통해 정책 설계, 실무 적용, 평가 과정 전반에 통합되고 있다.

〈표 2-3〉 OECD AI 원칙

가치기반 원칙	정책 입안자에 대한 제언
포용적인 성장, 지속가능한 개발과 웰빙	• AI 연구개발에 대한 투자
인권과 민주적 가치 (공정성, 프라이버시 등)	• 포용적이며 AI를 활성화하는 생태계 조성
투명성과 설명가능성	• 상호운용이 가능한 AI 거버넌스 및 정책 환경
강건성(robustness), 보안과 안전	• 인적 역량 구축과 노동 시장 전환 대비
책임성	• 신뢰할 수 있는 AI를 위한 국제적 협력

자료: 한국과학기술기획평가원(2025)

둘째, 전략적 인텔리전스(strategic intelligence)는 신흥 기술의 장기적 파급효과를 분석하기 위한 예측 도구로 작동하며, AI의 잠재적 이익과 위협을 식별하는 데 핵심적이다. 특히 OECD는 센티넬 모니터링과 실시간 추적을 활용해 AI 관련 사건과 위협을 식별하는 AI 사고 모니터(AI Incidents Monitor)를 운영하고 있다.

셋째, 이해관계자 참여(stakeholder engagement)는 AI 거버넌스의 수용성과 정당성을 제고하는 데 필수적인 요소로, OECD는 참여의 수준에 따라 정보 제공(informative), 자문적(consultative), 협력적(collaborative) 참여로 구분한다. 이는 AI 정책 초기 단계부터 다양한 사회적 집단의 의견을 수렴하고, 파트너십을 통해 공통 과제를 해결하는 데 기여한다. OECD.AI의 블로그와 공개 자문 절차, GPAI(글로벌 AI 파트너십), 전문가 네트워크(ONE AI) 등이 이러한 참여 구조를 실현하고 있는 구체적 사례들이다.

넷째, 민첩한 규제(agile regulation)는 AI 기술의 발전 속도에 대응하면서도 윤리성과 혁신을 동시에 촉진하기 위한 정책적 유연성을 지향한다. 이를 위해 OECD는 규제 샌드박스, 표준 및 by-design 접근, 그리고 기업의 책임경영(RBC)이라는 세 가지 축을 중심으로 규제 실험, 시스템 설계 초기단계에서의 윤리 내재화, 기업 내부의 위험관리 체계를 강조한다. 특히 미국 NIST의 AI 위험관리 프레임워크, G7 히로시마 프로세스 등은 이러한 원칙의 실천적 사례로 언급된다.

〈표 2-4〉 OECD 민첩한 AI 거버넌스를 위한 접근 방식

접근방식	주요 내용
통제된 환경과 규제 샌드박스	• 혁신 테스트베드, 리빙랩, 정책 프로토타입, 규제 샌드박스과 같은 규제 실험적 접근은 새로운 경제적, 제도적, 기술적 접근을 실험해보기 위해 활용됨 • 노르웨이나 싱가포르 등의 국가에서는 신뢰할 수 있는 AI의 개발을 위해 샌드박스 이니셔티브를 시행
표준과 설계에 의한 (by-design) 접근	• 표준과 by-design 접근은 민첩하며 상호운용이 가능한 AI 거버넌스 생태계 구성에 필수적인 요소로, 표준은 일관된 기준과 가이드 라인을 제공하고, by-design 접근은 프라이버시, 윤리, 보안, 투명성 등의 원칙이 처음부터 AI 시스템에 내재화되도록 함 • OECD는 AI 수명 주기 각 단계의 위험을 관리하기 위한 프레임워크를 제시한 바 있고, 미국의 국립표준기술연구소(NIST)의 AI 위험 관리 프레임워크(RMF), G7 히로시마 AI 프로세스의 사례
기업의 책임 경영 (Responsible Business Conduct, RBC)	• OECD는 2018년 기업 책임경영 실사 지침을 발간하고, 혁신 초기 단계에서 신흥기술의 위험을 예측·대응할 수 있는 주요 주체인 기업이 적용할 수 있는 방법을 제시 • 제시된 방법에는 ① 내부 정책 수립, ② 심각한 영향의 정의와 우선순위 설정, ③ 적절한 안전장치의 설계, ④ 이해관계자들의 참여 등이 있음

자료: 한국과학기술기획평가원(2025)

마지막으로, 국제 협력(international cooperation)은 AI 기술의 범지구적 영향력을 고려할 때 필연적인 요소이다. UN, 세계은행(World Bank), UNESCO 등 글로벌 기관은 안전하고 신뢰할 수 있는 AI 정책 가이드라인을 제시하고 있으며 유럽연합이나 아프리카연합 등은 지역적 맥락에 맞는 전략을 개발하고 있다. GPAI, EU-US 무역기술위원회(TTC), NATO AI 전략 등 다자 간 협의체 역시 기술의 경계를 넘어서는 거버넌스 조율에 중요한 역할을 수행하고 있다.

이와 같이, OECD의 예측적 거버넌스 프레임워크는 단지 원칙의 나열을 넘어서서, AI 정책의 실제 환경에서 가치 내재화, 위험 예측, 다자 협력, 실증 기반 규제, 포용적 참여의 구체적 실행을 위한 포괄적이고 상호보완적인 체계로 정립되고 있다. 이는 기술의 불확실성과 복잡성을 제도적 신뢰와 윤리적 정당성으로 전환하기 위한 국제사회의 공동 대응 전략이라 평가할 수 있다.

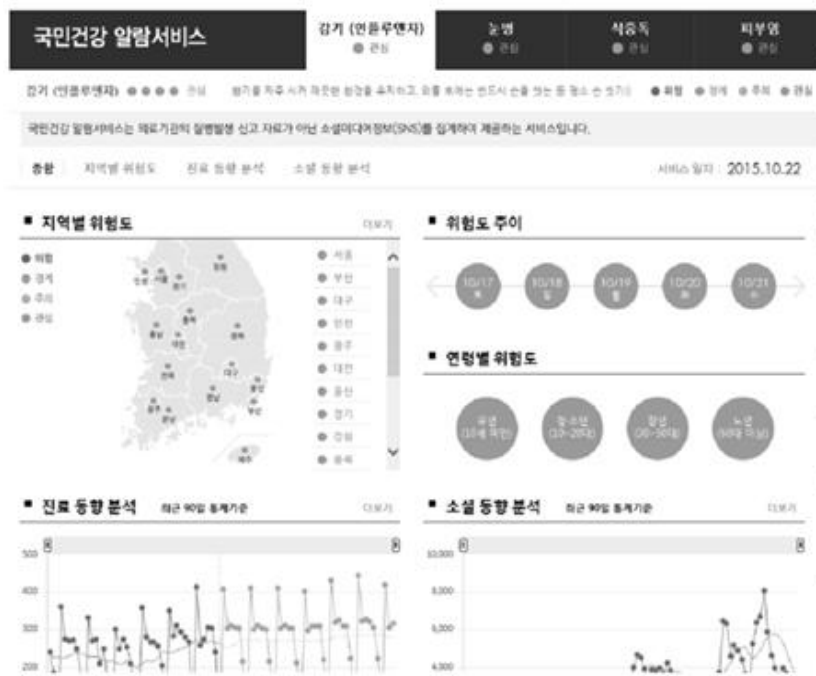
3. 빅데이터 거버넌스의 유형별 실행 모델

가. 데이터 기반 공공정책의 진화: 압출 모델의 유형

1) 정부주도 압출(Push) 모델

한국의 국민건강주의 알람 서비스는 앞서 설명한 빅데이터 거버넌스 유형 중 정부주도 압출 모델의 대표적 사례이다. 국민건강보험공단이 보유한 국민건강정보 공공 빅데이터와 ㈜다음소프트의 소셜미디어 정보인 트위터 데이터를 융합하였다. 2008년부터 2012년까지의 진료 데이터와 최근 3년간의 SNS 문서 37억 건 중 872만 건의 키워드를 분석하였다. 4단계 위험 정보 체계를 통해 위험, 경계, 주의, 관심 단계로 구분하여 감기, 눈병, 식중독 등 주요 질병에 대한 위험도 알람을 제공한다. 이 모델은 정부의 높은 반응성을 보여주지만, 시민들의 개방적 참여보다는 키워드 분석에 의존하는 한계가 존재한다.

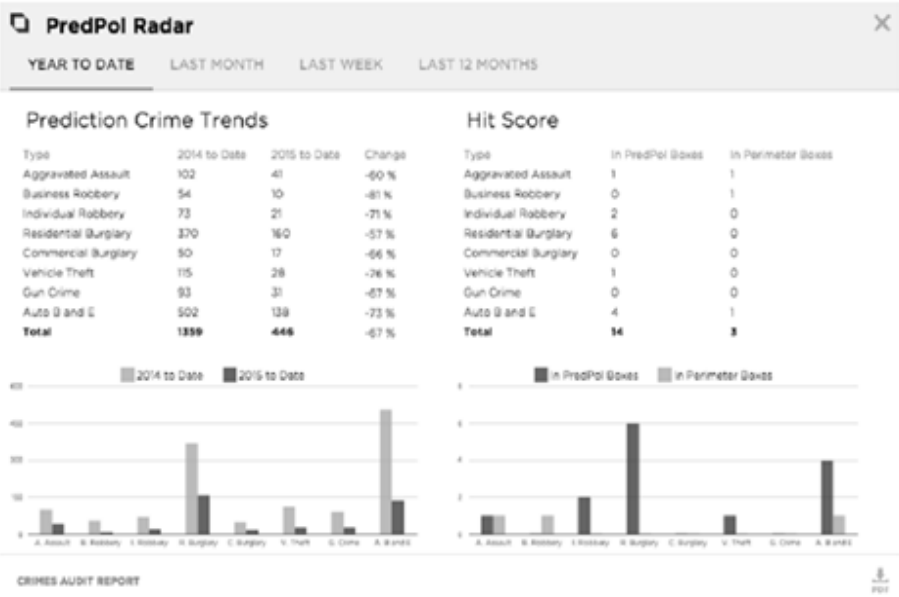
[그림 2-40] 국민건강주의 알람 서비스



자료: 송경재·장우영·조인호(2018)

미국 LA경찰의 프레드폴(Predpol) 시스템은 예측적 치안(Predictive Policing) 개념을 활용한 사례이다. 과거 범죄 데이터를 바탕으로 범죄 발생 가능성이 높은 시간대와 장소를 예측한다. 시스템 도입 후 LA지역 내 절도, 폭행 등이 30% 이상 감소하는 효과를 보였다. 플로리다 포트로더데일과 영국 켄트주 등 다른 지역에서도 범죄율 감소 효과가 보고되었다.

[그림 2-41] 프레드폴(Predpol) 시스템



자료: 송경재·장우영·조인호(2018)

2) 시민주도 압출 모델

한국의 MERS Map 사례는 시민주도 압출 모델을 보여준다. 2015년 MERS 사태 당시 관계당국의 병원 정보 비공개에 대응하여 시민들이 자발적으로 정보를 수집하고 지도화하였다. 소셜미디어와 스마트 디바이스를 활용하여 병원 정보를 수집하였고, 이는 정부의 병원 공개를 유도하며 일정한 영향력을 발휘하였다. 이는 정부 반응성이 낮은 상황에서 시민의 독자적 플랫폼이 형성된 사례이다. 빅데이터 시대에 정보 폐쇄의 어려움과 소셜미디어·스마트폰을 활용한 시민 참여 요구의 강력함을 실증적으로 보여준다.

3) 인터랙티브(Interactive) 모델

한국의 T맵 택시 포트홀 감지 시스템은 인터랙티브 모델의 대표적 사례이다. 2015년 10월부터 경기도에서 택시 2천여 대에 스마트폰 앱을 장착하여 포트홀 충격 데이터를 실시간으로 수집·전송하였다. 반복적으로 큰 폭의 진동 데이터를 분석을 통해 포트홀 복구 사업의 우선순위를 결정하였다. 이는 시민의 자발적 참여와 정부의 신속한 반응성이 결합된 단일 이슈형 빅데이터 거버넌스 플랫폼이다. ICT를 활용한 개방적 시민참여와 정부 반응성을 기반으로 한 인터랙티브 모델의 대표 사례로 평가된다.

미국의 Challenge.gov 플랫폼은 정책참여형 빅데이터 거버넌스의 사례이다. 정부의 정책 중심 문제 해결을 위해 상금과 과제를 통한 시민 참여를 유도하는 정책 아이디어 공모 플랫폼이다. 2015년 기준으로 80개 정부기관이 참가하여 450개 공모를 실시하였고, 2만 개의 해법을 도출하였다. 1억 5천만 달러의 상금을 제시하였고 350만 명이 참여하였다. 빅데이터의 집합지성을 활용한 정책참여형 빅데이터 거버넌스의 대표적 사례로서, 정부의 높은 반응성과 시민의 자발적 참여가 결합된 정책결정과정에서 시민 아이디어 공모 체계를 보여준다.

제3절 데이터 거버넌스의 성공요인과 실행 전략

1. 데이터 거버넌스 성공요인

가. 전략적 리더십 및 조직지원

전략적 리더십 및 조직지원은 최고이사결정자가 조직구성원에게 데이터 거버넌스에 대한 조직목표의 필요성과 기대효과를 설득력 있게 제시하여 조직구성원의 공감과 추종을 유발하고, 데이터 거버넌스에 대한 조직목표를 달성하기 위해 조직이 보유하고 있는 인적·물적 자원을 적재적소에 지원하는 과정을 의미한다(조세현·차세영, 2022). 데이터 거버넌스의 성공적인 추진을 위해 최고이사결정자의 전략적 리더십과 명확한 비전 설정, 단계적이고 지속적인 추진이 필수적이다. Mahanti(2018)는 데이터 거버넌스가 조직의 핵심 전략으로 자리 잡기 위해서는 최고이사결정자의 적극적 의지와 예산, 인력, 기술 등 자원의 충분한 지원이 필요하다고 강조하였다. 이는 단순한 데이터 거버넌스를 IT 프로젝트가 아닌 조직의 운영효율성과 경쟁우위를 위한 전략적 사업 기능으로 위치시켜야 함을 의미한다. 이를 위해서는 조직 전반에 걸쳐 데이터 거버넌스의 중요성을 전파해야 할 필요성이 있다.

조직 전반에 걸쳐 데이터 거버넌스의 중요성을 전파하기 위해서는 변화관리와 단계적 실행 전략을 통한 데이터 거버넌스의 점진적 확대 및 반복적 개선이 요구된다. 조직 전체에 데이터 거버넌스를 일시에 적용하기보다 단계적 확대 및 지속적 개선을 통해 성숙도를 높여나가는 것이 중요하다고 할 수 있다. 이와 관련하여 Alqudah(2021)는 상향식(실무진, 이해관계자의 참여)과 하향식(최고결정자의 리더십)의 균형적 추진을 제안하였다. 특히 전담 조직의 구성이 중요한데, 데이터 거버넌스 위원회, 데이터 책임관(Chief Data Officer), 데이터 거버넌스 관리자 등의 역할을 명확히 정의하고 이들에게 적절한 권한을 부여하여야 한다. 데이터 거버넌스가 조직문화에 정착되기 위해서는 실무진의 참여를 유도하고, 이들이 일상 업무에서 데이터 거버넌스의 가치를 체감할 수 있도록 지원해야 한다.

또한, 조직 전반의 협력 체계 구축이 필수적이다. 공공부문에서 데이터가 기관·영역·분야별로 각기 다른 형식으로 생성되고 데이터 공유를 위한 제도와 문화가 형성되지 않았거나 불확실한 경우, 데이터 사일로(칸막이 행정으로 인해 모든 데이터 생애주기에 걸쳐 개

인·조직·기관 간의 협력에 제약이 발생하는 현상)가 발생한다(김준형 외, 2020). 그러므로, 부처 및 부서 간 데이터 사일로를 해소하고, 위험 관리 및 컴플라이언스, 사이버 보안, 데이터 분석, 개인정보보호 등 다양한 영역 간의 협력을 촉진해야 한다. 이를 위해서는 조직 내 전사적 커뮤니케이션 채널 구축 및 부서 간 협업 체계 구축, 실무진 참여 유도, 정기적인 업데이트 시스템 구축 등이 중요하다.

나. 정책·제도 기반 및 법적 구조

데이터 거버넌스의 실행을 위해서는 정책, 조직, 제도가 적절하게 결합되어야 한다. 정책은 데이터 거버넌스가 지향하는 공공 가치와 방향 달성에 필요한 목표와 수단의 결합으로서의 계획을 의미하고, 조직은 정책목표 달성에 투입되는 인적·물적 자원과 이러한 자원의 운영·관리를 의미하고, 제도는 정책의 정당성, 지속가능성, 안정성을 부여하는 공식적·비공식적 규정 및 관행을 의미한다(경제·인문사회연구회, 2024). 제도는 정책의 안정성과 예측가능성을 증가시킨다는 점에서 중요하고, 조직은 정책의 실행가능성과 적용가능성을 증가시킨다는 점에서 중요하다. 특히, 제도 중에서 공식적 제도인 법은 정책목표 달성에 필요한 조건들을 조직이 의무적으로 실행하도록 강제함으로써 정책의 안정성과 예측가능성을 확실하게 증가시킨다.

데이터 거버넌스의 성공을 위해서는 명확하고 실행가능한 정책 프레임워크가 구축되어야 한다. Mahanti(2018)는 데이터 거버넌스 정책은 명확하고 집행 가능해야 하며, 정책의 정기적 갱신과 일관성 있는 실행 체계가 구축되어야 한다고 강조하였다. 이는 데이터 수집부터 활용, 보관, 폐기까지 전체 생명주기에 걸친 법적 근거와 내부 규정을 포함해야 한다. 즉, 데이터 분류 및 관리 기준의 표준화가 필요하다는 것이다. 조직 내 모든 데이터 자산에 대한 일관된 분류 체계를 구축하고, 각 데이터 유형에 따른 보안 등급과 접근 권한을 명확히 정의해야 한다. Marcucci et al.(2023)은 데이터 관리 전 과정(수집부터 활용, 폐기까지)에 대한 법적 근거 및 내부 규정 마련과 권한·책임·역할 명확화를 요구하였다. 이는 개인정보보호, 기밀정보 관리, 공개 가능 정보 등으로 세분화되어야 한다.

또한, 데이터의 권한, 책임, 역할의 명확화가 중요하다. 데이터 소유자(Data Owner), 데이터 관리자(Data Steward), 데이터 관리인(Data Custodian) 등의 역할을 명확히 구분하고, 각자의 책임과 의무를 구체적으로 정의하여야 한다. 이를 통해 데이터 거버넌스 활동에 대한 책임소재를 명확히 하고, 효과적인 의사결정 구조를 구축할 수 있다. 이와 더불어

GDPR, CCPA, HIPAA 등 국제적 규제뿐만 아니라 국내 관련 법규에 대한 준수 체계를 구축해야 한다. Hausenloy et al.(2024)은 데이터 특성에 맞는 법·제도적 통제장치를 제안하며, 이는 단순한 규정 준수를 넘어서 지속적인 모니터링과 감사를 통해 컴플라이언스 상태를 유지하고 개선하는 시스템을 포함해야 함을 강조한다.

마지막으로, 정책의 정기적 갱신과 일관성 확보가 중요하다. 기술 발전과 규제 변화에 대응하여 정책을 지속적으로 업데이트하고 조직 전반에 걸쳐 일관된 적용을 보장해야 한다. 차오신·오세환(2024)은 공공데이터 플랫폼에서 부처 간 메타데이터 표준화, 데이터 책임자 지정, 운영 규정 보완의 필요성을 강조하였다. 이를 위해서는 정책 생명주기 관리 프로세스와 버전 관리 시스템을 구축하는 것이 필요하다.

정책 생명주기 관리 프로세스와 버전 관리 시스템의 실효성과 지속가능성을 확보하기 위해서는 데이터 거버넌스 관련 정책의 정기적 갱신과 일관성 확보를 전담하는 개인·조직을 지정하고, 이러한 개인·조직의 업무 처리 결과를 기관 전체 차원에서 수용하도록 할 필요가 있다(Rollings et al., 2019). 데이터 거버넌스 관련 정책의 전담 주체는 기술 발전과 규제 변화에 대한 효과적인 대처가 가능하도록 데이터에 대한 기술전문성(기술적 차원의 데이터 생성·수집·관리에 대한 전문성)과 현장전문성(정책적·제도적 차원의 데이터 활용 및 공공가치 창출에 대한 전문성)을 겸비한 개인·조직으로 구성되어야 한다.

다. 윤리·투명성·책임성 확보

AI의 등장과 확산으로 인해 데이터 수요가 폭발적으로 증가하면서, 데이터를 안전하게 관리하고 활용하기 위한 데이터 윤리의 중요성이 대두되고 있다. 데이터 윤리는 데이터의 오남용이 발생하지 않도록 데이터를 안전하게 관리하고 활용하기 위한 가치, 원칙, 기준 등을 의미한다(김정남 외, 2023). 특히, AI로 인한 가능해진 자동화된 데이터 수집과 의사결정으로 인해 데이터 투명성과 책임성 확보가 중요한 과제로 등장하고 있다. 무엇보다도, 개인정보 주체에게 자신에 대한 데이터의 처리 과정을 투명하게 고지하고, 데이터 관련 문제점 및 오남용에 대한 책임을 강화할 필요가 있다.

데이터 거버넌스에서 윤리적 데이터 사용은 조직의 신뢰성과 지속가능성을 위한 핵심요소이다. Kroll(2018)은 데이터 사용 검토 위원회 설치와 내외부 감사·검증 체계의 도입을 제안하였고, 조직은 데이터 수집, 처리, 활용의 모든 과정에서 윤리적 기준을 내재화하고,

이를 실천하기 위한 구체적인 가이드라인을 마련해야 한다. 이는 개인정보보호를 넘어서 편향성 방지, 알고리즘 공정성, 사회적 책임 등을 포괄하는 종합적인 접근이 필요하다.

또한, 투명성 확보를 위한 체계적 접근이 필요하다. 데이터 수집 정책, 처리 방법, 활용 목적을 명확히 공개하고, 이해관계자들이 쉽게 이해할 수 있는 형태로 데이터를 제공하여야 한다. 또한 데이터 처리 과정의 추적 가능성(Data Lineage)을 확보하여, 데이터의 출처와 변환 과정을 투명하게 공개할 수 있어야 한다.

이와 더불어 다중 이해관계자 참여 메커니즘 구축이 강조된다. 시민, 고객, 직원, 규제기관 등 다양한 이해관계자가 데이터 거버넌스 과정에 참여할 수 있는 채널이 마련되어야 한다. Schoeman(2023)과 Redden(2018)은 빅데이터·AI 거버넌스 분야에서 다중 이해관계자 참여와 유연한 참여 설계가 강조되며 이는 정기적인 피드백 수집, 공청회, 시민 패널 등 다양한 형태로 구현될 수 있으며 이를 통해 사회적 신뢰를 구축할 수 있다고 보았다. 또한, 감시 및 감사 체계 강화가 필요하다. 내부 감사팀과 외부 전문기관을 통한 정기적 감사를 실시하고, 데이터 사용 검토 위원회 등을 설치하여 지속적인 모니터링을 수행하여야 한다. Redden(2018)과 Safarov(2019)는 사용자 피드백과 민주적 감시의 사회적 신뢰 구축 중요성을 언급하였다. 이러한 감사 결과는 투명하게 공개되어야 하며, 발견된 문제점에 대한 개선 조치를 신속하게 실행하여야 한다.

마지막으로 공공데이터의 사회적 가치 창출과 데이터 문해력 증진이 중요하다. 데이터 거버넌스가 단순한 규제 수준을 넘어 사회적 가치를 창출할 수 있도록 하고, 조직 구성원과 시민들의 데이터 문해력을 높이는 교육과 훈련을 지속적으로 실시해야 한다. 차오신·오세환(2024)과 Safarov(2019)는 공공데이터의 사회적 가치 확산과 데이터 문해력 증진, 민관 협력도 주요 과제임을 강조하였다. 이를 통해 데이터 거버넌스에 대한 이해를 높이고, 자발적인 참여를 유도할 수 있다.

데이터 문해력을 높이기 위한 교육과 훈련은 데이터 관리의 다양한 차원을 포괄하는 프로그램으로 구성되어야 한다(강정목 외, 2014). 데이터 문해력은 데이터와 디지털 기기를 학습하고 이해할 수 있는 역량(데이터 학습), 데이터와 데이터 분석 도구를 특정한 목표를 달성하기 위한 수단으로서 적절하게 활용할 수 있는 역량(데이터 활용), 데이터로 인해 발생할 수 있는 역기능 또는 부작용을 차단하거나 대처할 수 있는 역량으로 구성된다(데이터 책임). 데이터 문해력을 높이기 위한 교육과 훈련은 조직 구성원과 시민의 데이터 학습

· 활용·책임 역량이 균형 있게 함양될 수 있도록 실행되어야 한다.

라. 데이터 품질·보안 및 기술·운영 역량

데이터 거버넌스의 성공을 위해서는 포괄적인 데이터 품질관리 체계가 필수적이다. 이는 정확성(Accuracy), 완전성(Completeness), 일관성(Consistency), 적시성(Timesliness), 유일성(Uniqueness) 등 다차원적인 품질 지표를 포함해야 한다. 데이터 품질 관리는 단순한 오류 검출을 넘어 예방적 품질 보증 시스템을 구축하여 데이터 생성 단계부터 품질을 확보해야 한다.

이를 위해 고도화된 보안 및 개인정보보호 체계가 중요하다고 할 수 있다. 데이터 암호화, 접근 제어 체계, 백업 및 복구 시스템을 포함한 종합적인 보안 인프라가 구축되어야 한다. Hausenloy et al.(2024)과 Kroll(2018)은 AI, 머신러닝, 자동화 등 첨단기술 도입을 통한 단계별 품질보증 체계가 요구된다고 하였고, 특히 AI와 머신러닝 기술을 활용한 자동화된 보안 모니터링과 이상 탐지 시스템을 도입하여 실시간 위협 대응 능력을 강화해야 한다고 강조하였다. 이와 더불어 자동화 및 지능화 기술 도입의 중요성도 강조된다. 반복적이고 규칙 기반의 데이터 거버넌스 업무는 자동화하고, AI와 머신러닝을 활용한 지능형 데이터 관리 시스템을 구축해야 한다는 것이다. 이를 통해 데이터 프로파일링, 품질 검증, 메타데이터 관리 등의 업무효율성을 크게 향상시킬 수 있을 것으로 기대된다.

한편, 이를 활용할 수 있는 기술적 역량과 전문인력의 확보가 필수적이라고 할 수 있다. 데이터 거버넌스 관련 전문 인력을 충분히 확보하고, 이들에 대한 지속적인 교육과 역량 개발을 실시해야 한다. Schoeman(2023)은 데이터 관련 전문 인력의 확보와 이들에 대한 지속적 교육, 역량 개발 및 전담팀 운영은 거버넌스 운영 역량 강화에 중요하다고 강조하였다. 데이터 관리자, 데이터 스튜어드, 데이터 분석가 등 각 역할에 필요한 기술적 역량과 비즈니스 지식 등을 균형있게 갖춘 인력을 육성해야 한다.

마지막으로 사용자 중심의 데이터 서비스 체계를 구축해야 한다. 데이터 카탈로그, 셀프 서비스 분석 도구, 직관적인 사용자 인터페이스(UX/UI, 검색 기능 등) 등을 제공하여 조직 구성원들이 쉽게 데이터에 접근하고 활용할 수 있도록 해야 한다. 또한 데이터 공급망 내 각 참여자 역할과 책임 명확화, 위험관리 또한 필수적이라고 할 수 있다. 특히 지속 가능한 ODG(Open Data Government) 환경 조성을 위해서는 제도적 지원과 시민 인식, 기술 역

량, 정책 일관성이 상호 보완적으로 작동해야 할 필요가 있다(Safarov, 2019). 이는 데이터 민주화를 통한 조직의 데이터 활용 역량 제고로 이어진다.

마. 모니터링·평가 및 지속적 개선

데이터 거버넌스의 지속가능성을 위해서는 체계적인 모니터링과 평가 시스템이 필요하다. Mahanti(2018)는 데이터 및 알고리즘 성능을 정기적으로 검증하고, 성과 측정 및 품질 점검을 통해 운영 실효성을 평가할 것을 제안하였다. 이는 데이터 품질, 활용, 만족, 효과를 포함하는 다양한 측면에서 성과를 측정하고 추적할 수 있는 KPI(Key Performance Indicator)를 포함해야 한다(엄석진 외, 2020). 품질은 데이터의 정확성·완결성·시의적절성과 같은 데이터 특성의 우수성을 의미하고, 활용은 데이터가 사용자들에 의해 양적으로 많이·자주 활용되는 정도와 질적으로 적절하게 활용되는 정도를 의미한다. 만족은 데이터가 다양한 이해관계자들에게 제공하는 만족의 정도를 의미하고, 효과는 데이터가 다양한 이해관계자들에게 제공하는 금전적·비금전적 편익을 의미한다.

이와 더불어, 피드백 루프 시스템 구축이 중요하다. 차오신, 오세환(2024)은 시민과 사용자 피드백을 수집하고 공공데이터 플랫폼의 이용 현황을 분석하는 중요성을 강조하였다. 사용자, 데이터 스튜어드, 관리자 등 다양한 이해관계자로부터 지속적으로 피드백을 수집하고, 이를 정책 개선과 시스템 개선에 반영하는 체계를 구축해야 한다. 이러한 피드백은 자동화된 시스템을 통해 수집되고, 실시간으로 분석되어 신속한 의사결정을 지원할 수 있도록 해야 한다. 나아가, 시민이 직접 데이터 관리의 과정에 참여하는 기회를 확대하여 피드백 루프 시스템에 대한 사용자의 효능감을 개선할 필요가 있다(엄석진 외, 2020). 특히, 데이터 생성 과정과 평가 과정에 각각 시민을 대상으로 하는 수요조사와 시민이 참여하는 시민평가단 활동을 실시하여 그 결과를 실질적으로 반영할 필요가 있다.

이를 위해서는 조직의 적응적 학습체계 구축이 필요할 것이다. 기술 발전, 규제 변화, 조직 환경 변화에 대응하여 데이터 거버넌스 전략과 실행 방안을 지속적으로 조정하고 개선해야 한다. 이는 조직학습을 통한 데이터 거버넌스의 진화와 발전을 의미한다. 이와 관련하여 데이터 거버넌스 성숙도 진단을 정기적으로 실시할 필요가 있다. Marcucci et al.(2023)과 Nookala(2024)는 정책·조직·기술 성숙도 진단 및 개발도상국 역량 강화 지원 필요성을 제기하였다. 조직의 데이터 거버넌스 수준을 객관적으로 평가하고, 업계 표준이

나 모범사례와 비교하여 개선 방향을 설정해나가는 것이다. 이를 통해 조직의 데이터 거버넌스 역량을 지속적으로 고도화할 수 있다.

아울러 교육과 훈련의 지속적 실시가 중요하다. 데이터 거버넌스 관련 지식과 기술은 빠르게 변화하므로, 조직 구성원들에 대한 지속적인 교육과 훈련을 실시해야 한다. 교육, 홍보, 데이터 리터러시 향상과 현장 기반 개선 활동도 병행해야 할 필요가 있다(Nookala, 2024). 이는 데이터 문해력 향상, 새로운 기술 도입, 정책 변화 등을 포함하는 종합적인 교육 프로그램으로 구성되어야 한다. 이와 더불어 데이터 거버넌스 활동이 단순한 규제 수준을 넘어 조직의 가치 창출에 기여하는지를 측정하고 관리하여야 한다. 종합해보면, 데이터에 대한 교육과 훈련은 데이터의 역기능 또는 부작용을 방지하거나 대처하기 위한 교육 프로그램과 데이터의 효과 또는 편익을 구현하기 위한 교육 프로그램을 포괄하는 방향으로 실시될 필요가 있다.

2. 데이터 거버넌스의 실행 전략

가. 제도적 기반과 정책 실행 체계 정비

1) 국가 데이터 정책 및 법·제도 기반 강화

데이터 거버넌스의 실행 전략에서 가장 핵심적인 전제는 명확하고 포괄적인 법적 기반과 정책적 정합성을 구축하는 것이다(정용찬, 2018). 디지털 전환이 전방위적으로 확산되면서, 국가 차원에서 데이터를 공공 자산으로 인식하고 이를 체계적으로 관리·활용하기 위한 법적·제도적 기반은 정책 효과성과 행정 신뢰도를 높이는 핵심 인프라로 작용한다(행정안전부·한국지능정보사회진흥원, 2024). 특히 한국은 세계적으로도 빠르게 데이터 기반 정책을 제도화한 국가 중 하나이며, 법령 및 정책 계획의 정합성을 통해 데이터 활용과 보호의 균형을 도모하고 있다(OECD, 2022).

우선 2020년 제정된 「데이터기반행정 활성화에 관한 법률」은 공공기관이 데이터를 정책 결정, 행정업무 개선, 대국민 서비스 향상에 적극적으로 활용할 수 있도록 제도적 토대를 마련하였다. 동 법률 제4조는 중앙행정기관, 지방자치단체, 공공기관이 데이터를 공동 활용하도록 규정하고 있으며, 제8조에서는 행정안전부 장관이 데이터기반행정 기본계획을 수립·시행할 수 있도록 명시하고 있다(행정안전부, 2024). 또한, 「공공데이터의 제공 및 이

용 활성화에 관한 법률」은 2013년 제정되어 공공데이터의 개방 및 민간 활용 촉진을 핵심으로 하고 있으며, 2020년 개정된 「개인정보 보호법」은 데이터의 보호와 활용 간 균형을 위한 일반법으로 기능하고 있다. 「AI 기본법」은 현재 제정이 추진 중이다(행정안전부·한국지능정보사회진흥원, 2024). 정용찬(2018)은 이러한 법·제도 기반이 부처 간 데이터 통합과 공동 활용을 제약에서 가능성으로 전환시키는 조건이 되며, 단순한 정보공개 차원을 넘어 데이터의 가치 창출을 제도적으로 뒷받침해야 함을 강조하였다. 특히, 데이터 생태계는 상호 연계된 구조이므로, 하나의 법이나 부처 단위 접근이 아닌 다부처 연계형 법·제도가 요구된다. 정책적 측면에서는 「제4차 공공데이터 기본계획(2023~2025)」(행정안전부, 2022), 「제2차 데이터기반행정 기본계획(2024~2026)」(관계부처 합동, 2024), 그리고 각 부처의 「데이터기반행정 시행계획」들이 종합적 거버넌스 체계를 뒷받침하고 있다. 이들 계획은 데이터 생산·개방·활용 전주기 체계화, 표준·품질·보안 기준 정립, 민관협력 및 국제 연계 확대 등을 포괄하며, 데이터 정책의 목표와 실행 방안을 구조적으로 연결짓는 역할을 한다.

한편, 미국은 「Foundations for Evidence-Based Policymaking Act(2019)」와 함께 Federal Data Strategy를 통해 데이터 중심의 정책 결정을 촉진하고 있으며, 유럽연합은 「Data Governance Act(DGA, 2020)」를 통해 공공 및 민간 데이터의 안전하고 신뢰 가능한 공유를 촉진하고 있다(European Commission, 2020). 이와 같은 해외 사례는 우리나라의 정책·법제 설계에 있어 글로벌 정합성을 고려한 제도 보완의 필요성을 시사한다. 김지현(2023)은 최근 전 세계적으로 데이터 규제 정책이 보호 중심에서 활용 중심으로 전환되는 경향을 보이며, 국내 정책도 이와 보조를 맞춰야 함을 지적하였다. 특히 디지털 대전환기에 국가 간 데이터 이동과 신뢰 확보가 중요해지는 만큼, 국경 간 데이터 거버넌스를 고려한 유연하고 신뢰성 있는 법제도 설계가 핵심 과제가 된다. 종합하면, 국가 데이터 정책의 법제도 기반은 데이터 생태계 전반의 신뢰성, 투명성, 효율성을 뒷받침하는 구조적 요소로서 기능하며, 이를 위해 법률 간 정합성 확보, 다부처 협력체계 정립, 국제 표준과의 연계가 병행되어야 한다. 이는 공공데이터의 전략적 활용뿐만 아니라, 디지털 주권과 데이터 주도 혁신 역량 강화를 위한 핵심 토대라 할 수 있다(유승현, 2022; OECD, 2022).

2) 전담 조직 및 리더십 체계 구축

데이터 거버넌스의 실효적 실행을 위해서는 제도적 기반뿐만 아니라 전담 조직과 리더십 체계의 확보가 필수적이다(행정안전부, 2024). 특히, 부처 간 정책 이행의 조정, 책임성 확보, 실행 동력을 뒷받침하기 위해 각 기관 내·외부의 역할 분담과 리더십 구성이 중요하다(Marcucci et al., 2023). 이와 관련해 국제적으로는 최고데이터책임자(Chief Data Officer, CDO) 제도와 데이터 전담 조직의 구축이 대표적인 전략으로 자리 잡고 있다(Federal Data Strategy Development Team, 2021). 한국은 「데이터기반행정 활성화에 관한 법률」 제12조에 따라 중앙행정기관과 지방자치단체에 CDO 지정을 의무화하고 있으며, 실질적인 데이터 정책 총괄과 실행을 위한 리더십 체계를 제도화하였다(행정안전부, 2024). 행정안전부는 각 기관의 CDO를 연결하는 CDO 협의체를 운영하여 기관 간 정보공유, 공동 기획, 우수사례 확산을 촉진하고 있다(행정안전부, 2025). 이러한 협의체는 기관 간 정책 추진 격차를 줄이고, 중복투자 방지 및 공동 프로젝트 발굴에 실질적 기여를 하고 있다.

국제적으로는 미국의 Federal Chief Data Officers Council과 EU의 European Data Innovation Board가 대표적 사례로, 각각 「Foundations for Evidence-Based Policymaking Act(2019)」과 「Data Governance Act(2020)」를 근거로 설립되었다(Federal Data Strategy, 2021; European Commission, 2020). 이들 조직은 국가 데이터 전략 수립, 데이터 표준 제정, 윤리적 활용 가이드라인 제안 등에서 중추적인 역할을 수행한다. Hausenloy et al.(2024)은 AI 및 프런티어 기술의 확산으로 인해 데이터 운영의 복잡성과 리스크가 증가함에 따라, 전담조직의 기술 전문성과 법제 조정 역량이 갈수록 중요해지고 있음을 강조하였다. 이러한 조직은 단순히 데이터를 수집·관리하는 기능을 넘어, 거버넌스 원칙 내재화, 품질·보안 체계 연계, 외부 감사 기능까지 수행할 수 있어야 한다. 또한 Charles et al.(2022)은 CDO의 리더십이 단기적 정책 실행뿐만 아니라, 장기적 비전 제시와 조직문화 개선에 기여할 수 있다고 분석하였다. 실제로 일부 국가에서는 데이터 거버넌스 관련 인력의 역량을 제고하기 위한 전문교육 프로그램, 인증제도, 윤리 가이드라인 교육을 함께 운영하고 있다.

국내에서도 유사한 흐름이 나타나고 있으며, 「2024년 데이터기반행정 실태점검 결과」(행정안전부, 2025)에 따르면 CDO가 존재하는 기관이 그렇지 않은 기관에 비해 데이터 분석 활용도, 타 부서 협업 수준, 품질관리 이행 정도에서 상대적으로 우수한 평가를 받았다.

이는 CDO 제도가 단순한 형식이 아닌 실질적 거버넌스 실행의 핵심 구성요소임을 시사한다. 요컨대, 데이터 거버넌스를 실효적으로 운영하기 위해서는 전담조직 구성과 CDO 중심의 리더십 확보가 핵심이다. 이를 위해서는 CDO의 권한과 역할 명확화, 다기관 협업 네트워크 제도화, 기술·법률 복합 역량을 갖춘 전문인력 양성 등이 종합적으로 추진되어야 한다(Marcucci et al., 2023; Harrison et al., 2019).

3) 정책 일관성과 다기관 협업 강화

데이터 거버넌스는 단일 기관의 독립적인 노력만으로는 실효성 있는 정책 수행이 어렵다. 다양한 기관 간 유기적인 협업 체계와 정책의 일관성 확보가 필수적이며, 이는 데이터 정책의 지속가능성을 좌우하는 핵심 요소다(행정안전부, 2024; 관계부처 합동, 2024). 한국과 같이 다부처 및 다기관 구조를 가진 행정 체계에서는 특히 중앙정부 부처 간, 중앙과 지방정부 간, 그리고 공공과 민간 부문 간 협력 메커니즘의 정교한 설계가 요구된다(Marcucci et al., 2023; Federal Data Strategy Development Team, 2021). 행정안전부(2024)는 「2024년 데이터기반행정 활성화 시행계획」을 통해 부처 간 데이터 전략의 연계성과 일관성을 강화할 필요성을 제시하였으며, 「2차 데이터기반행정 기본계획(2024~2026)」에서는 이를 정책 실행의 핵심과제로 명시하였다(관계부처 합동, 2024). 이러한 협업 구조는 단지 조정의 차원을 넘어 공동 목표 설정, 과제의 분담 및 통합 실행을 가능하게 한다. 첫째, 정부는 데이터 정책의 조율을 위해 CDO(Chief Data Officer) 협의체, 데이터정책위원회, 데이터기반행정 전담조직 등 공식적 기제를 활용하고 있다. 이러한 구조는 각 부처 및 기관의 역할을 명확히 하고 상호 협업을 통해 정책 시너지를 창출하는 기반을 마련한다(행정안전부, 2025). 둘째, 데이터 표준화 및 메타데이터 체계 정립을 위한 다기관 공동작업도 중요한 협업 영역이다(과학기술정보통신부·NIA·TTA, 2024). 예를 들어, 한국정보화진흥원(NIA), 한국인터넷진흥원(KISA), 한국정보통신기술협회(TTA)는 「데이터 품질관리 가이드라인」과 「공공데이터 표준화 지침」을 공동으로 마련하여 법정부적으로 적용하고 있다. 이는 데이터 간 상호운용성을 높이고 품질관리를 표준화하는 데 기여하고 있다. 셋째, 지방정부 차원의 협업도 활발히 전개되고 있다. 일부 자치단체는 자체 데이터관리위원회를 구성하여 시민, 민간기업, 전문가가 참여하는 지역 거버넌스 네트워크를 운영하고 있으며, 이는 지역 맞춤형 정책 수립과 디지털 격차 해소에 기여하고 있다(윤건, 2023).

국제적으로도 유사한 구조가 나타난다. 유럽연합(EU)은 「European Data Innovation Board」를 통해 회원국 간 데이터 협력 메커니즘을 제도화하고 있으며(European Commission, 2020), 미국의 「Federal Data Strategy(2021)」는 기관 간 데이터 공유 프레임워크, 공통 표준 적용, 데이터 윤리 가이드라인 정립 등을 핵심 실행 전략으로 제시한다(Federal Data Strategy Development Team, 2021). Marcucci et al.(2023)은 이러한 다기관 협업의 구조가 데이터 품질, 접근성, 윤리성 등의 복합적 문제를 해결할 수 있는 유일한 전략이라고 분석하며, 특히 '신뢰 기반 상호운용성'이 현대 데이터 거버넌스 체계의 성패를 가르는 핵심 요소를 강조한다. 결론적으로, 데이터 거버넌스는 다양한 주체 간 전략적 협력과 정책 일관성을 기반으로 작동하며(Marcucci et al., 2023). 이를 위해 제도화된 협의 기구, 표준화된 기준, 참여적 의사결정 메커니즘이 통합적으로 작동해야 한다(행정안전부, 2024; 과기정통부·NIA·TTA, 2024). 이는 단순한 정보 교환을 넘어 공동 기획, 실행, 평가로 이어지는 통합 거버넌스의 실현을 의미한다(Federal Data Strategy Development Team, 2021; European Commission, 2020).

나. 기술 기반 실행 메커니즘과 품질관리 체계

1) 생애주기 기반 품질관리 체계 정비

데이터의 품질은 데이터 거버넌스의 성패를 좌우하는 핵심 요소로, 신뢰성과 활용도를 결정짓는 기준이 된다(정용찬, 2018; Kroll, 2018). 공공부문에서는 데이터가 생성·처리·저장·공유·활용되는 전 생애주기(lifecycle)를 기반으로 품질을 체계적으로 점검하고 보장하는 체계가 요구된다(행정안전부, 2024; 성욱준, 2017). 한국은 「공공데이터 품질관리 중장기계획(2019~2023)」과 「공공데이터 품질관리 수준진단·평가 세부 안내서(2024)」 등을 통해 각 단계별 품질지표와 검증 절차를 마련하고 있다(행정안전부·NIA, 2024). 이 계획은 데이터 품질의 4대 핵심 지표로 '정확성, 표현성, 최신성, 완전성'을 제시하고 있으며(행정안전부·한국지능정보사회진흥원, 2024), 이를 기반으로 전 기관에 대한 품질진단 및 개선 활동이 정례적으로 수행된다(행정안전부, 2025). 예를 들어, 2023년 기준 약 1,200여 개 기관에 대해 품질진단이 실시되었고, 70% 이상이 주요 오류를 개선하거나 포맷 일관성을 확보한 것으로 평가되었다(행정안전부, 2024).

국제적으로도 유사한 움직임이 있다. OECD는 데이터 품질을 '지속가능한 데이터 거버넌스'의 핵심 조건으로 간주하며, 특히 AI 활용 및 자동화된 의사결정이 확산됨에 따라 데

이터의 오류 가능성 및 편향성 문제가 더욱 부각되고 있음을 경고한다(OECD, 2022). 이에 따라 미국, 유럽, 호주 등은 데이터 생성 시 메타데이터 구조화, 품질 이력 관리, 자동화 점검 도구 활용 등을 제도화하고 있다(Federal Data Strategy Development Team, 2021; European Commission, 2020; NSW Government, 2021). 한국의 「AI 학습용 데이터 품질관리 가이드라인(2024)」도 이러한 흐름에 발맞춰 데이터 표현의 의미적 정확성, 범주 간 균형성, 대표성 확보, 노이즈 제거 등의 기준을 명확히 설정하고 있으며, 민간 데이터셋에도 점차 확대 적용되고 있다(과학기술정보통신부·NIA·TTA, 2024).

Kroll(2018)은 데이터 품질이 단순한 기술적 문제가 아니라, 데이터 생산 및 활용의 사회적 책임성과 직결된다고 지적하였다. 특히 공공정책에서 데이터에 기반한 결정이 시민의 삶에 직접적인 영향을 미치는 만큼, 데이터 품질관리 체계는 윤리적이고도 거버넌스적인 책무로 인식되어야 한다. 또한 Harrison et al.(2019)은 품질관리 체계가 실질적으로 작동하기 위해서는 부서 간 품질 기준의 일관성 확보, 자동화된 검증 도구 도입, 전담 인력의 전문성 강화, 품질 개선 결과에 대한 피드백 루프 구축이 중요하다고 제시하였다. 따라서 품질관리 체계는 단순한 '사후 점검' 수준에 머무르지 않고, 데이터 기획 단계부터 포함된 '사전 예방적 관리'로 전환되어야 한다. 이를 위해서는 첫째, 데이터가 생성되는 단계부터 폐기까지 생애주기별로 품질관리 기준을 설정하고 이를 표준화하여 적용해야 한다. 둘째, 자동화된 진단 도구와 인공지능 기반의 오류 탐지 시스템을 적극 도입함으로써 품질 점검의 효율성과 신뢰성을 확보할 필요가 있다. 셋째, 품질 정보는 외부에 투명하게 공개되어야 하며, 사용자들로부터의 의견과 피드백을 수렴할 수 있는 구조를 구축함으로써, 정책 수요자와 공급자 간의 지속적이고 쌍방향적인 품질 개선이 이루어져야 한다. 이러한 전 생애주기 기반 품질관리 체계는 데이터 거버넌스를 신뢰 기반으로 작동하게 하는 중요한 기술 메커니즘이자, 행정서비스 품질과 시민 수용성을 제고하는 토대가 된다(Hausenloy et al., 2024). 특히 향후 AI·빅데이터 행정의 고도화와 함께, 품질 기준의 정량화와 국제 기준 정합성 확보가 더욱 중요해질 것이다(OECD, 2022).

2) 기술 기반 메타데이터 및 추적관리 체계 구축

데이터의 투명성과 신뢰성을 확보하기 위해서는 각 데이터의 출처, 맥락, 변경 이력 등을 추적할 수 있는 메타데이터 관리 체계와 추적 관리(traceability) 메커니즘이 필수적이다(Hausenloy et al., 2024). 특히 대규모 데이터셋이 AI 학습, 정책 결정, 공공서비스에 활용

되는 오늘날, 데이터의 흐름을 정확하게 기록하고 모니터링하는 기술 기반 구조는 데이터 거버넌스의 핵심 기반이 된다(OECD, 2022; Yukhno, 2024). 메타데이터는 데이터의 생성 시점, 작성자, 출처, 처리 이력, 접근 권한, 품질 상태 등을 기술하는 정보로, 데이터의 맥락을 해석하고 적절하게 활용하기 위한 핵심 도구이다(Harrison et al., 2019). 이에 따라 한국 정부는 「공공데이터 메타데이터 관리 지침」을 통해 표준화된 메타데이터 구조를 마련하고, 각 공공기관에 적용을 의무화하고 있다(행정안전부·한국지능정보사회진흥원, 2024). 이 지침은 데이터 등록 시 필수 메타데이터 입력 항목을 지정하고 있으며, 추후 관리 및 교차연계를 위해 공통 포맷을 요구한다(행정안전부, 2024).

국제 사례로는 EU의 「Data Governance Act」가 메타데이터의 상호운용성과 추적 가능성 확보를 위한 기술 요건을 명시하고 있으며, 데이터셋 이동, 복제, 가공 등의 전체 이력을 관리하는 ‘데이터 로그 시스템’ 도입을 권고하고 있다(European Commission, 2020). AI 분야에서도 투명성과 안전성을 위해 메타데이터 기반 추적 체계가 강화되고 있다. Hausenloy et al.(2024)은 Frontier AI 모델의 책임성과 안전성을 확보하기 위해 ‘카나리아 토큰(Canary Tokens)’ 및 ‘데이터 체인 추적 시스템’과 같은 기술적 장치를 활용해야 함을 주장한다. 이는 모델 학습에 사용된 원천 데이터의 변조 여부를 탐지하고, 유해한 데이터가 포함된 경우 소급 추적을 가능하게 한다. 또한 Harrison et al.(2019)은 메타데이터 기반의 투명한 데이터 흐름 관리가 공공부문에서 신뢰 기반 의사결정의 전제가 된다고 보았다. 이들은 효과적인 데이터 흐름 관리를 위해 자동으로 수집되고 정기적으로 갱신될 수 있는 메타데이터 관리 시스템의 구축이 필요하며, 모든 변경 이력을 기록하고 추적할 수 있는 로그 시스템이 필수라고 지적하였다. 아울러 시스템 간 상호운용성을 보장하기 위해 공통 메타데이터 표준의 정합성 유지가 중요하며, 메타데이터를 바탕으로 데이터 품질을 평가하고 위험 요인을 사전에 감지할 수 있는 체계의 마련이 요구된다고 제안하였다.

국내 실태점검 결과에 따르면, 일부 기관은 메타데이터 등록을 단순 기술정보 수준에 그치고 있어 활용도와 추적 가능성이 낮은 것으로 나타났다(행정안전부, 2025). 이에 따라 메타데이터를 단순 문서화가 아닌, 실시간 분석 및 통합 거버넌스를 위한 구조적 자산으로 인식하는 패러다임 전환이 필요하다. 이를 종합하면, 기술 기반의 메타데이터 및 추적 관리 체계는 다음과 같은 방향에서 체계적으로 구축될 필요가 있다. 첫째, 메타데이터의 수집과 활용이 자동화될 수 있도록 시스템을 설계하고, 범정부적으로 통일된 메타데이터

표준을 정립함으로써 기관 간 상호운용성과 데이터 품질의 일관성을 확보해야 한다. 둘째, 데이터 생성부터 변경, 폐기에 이르기까지의 모든 과정을 추적할 수 있도록 변경 이력에 대한 통합 로그 시스템을 도입하여 투명한 관리가 가능하도록 해야 한다. 셋째, 민감정보를 포함하거나 인공지능(AI) 분석에 활용되는 데이터를 대상으로는 보다 정밀한 추적 기술을 적용하고, 이상 징후나 품질 저하 요소를 조기에 탐지할 수 있는 고도화된 기술 역량을 확보해야 한다. 마지막으로, 메타데이터를 기반으로 한 데이터 책임성 평가 프레임워크를 구축하여 각 기관의 데이터 관리 수준을 진단하고, 개선책을 제시할 수 있는 구조로 운영되어야 한다. 이러한 체계는 단순한 기술적 관리 도구를 넘어서서, 데이터 거버넌스의 핵심 원칙인 투명성, 설명 가능성, 그리고 신뢰성 확보를 실질적으로 뒷받침하는 기반이 된다.

3) 자동화 기반 데이터 표준화 및 연계체계 정비

데이터의 상호운용성과 통합 활용을 위해서는 표준화된 구조와 시스템 간 연계성을 전제로 한 기술적·제도적 체계가 필요하다(OECD, 2022). 특히 다양한 부처와 기관에서 생성·보유하는 데이터의 구조와 형식이 상이한 현실에서, 자동화를 기반으로 한 데이터 표준화와 연계체계 구축은 데이터 거버넌스의 실행력을 제고하는 핵심 전략으로 간주된다(Harrison et al., 2019). 한국 정부는 「공공데이터 표준화 지침」과 「데이터 연계 표준 프레임워크」를 통해 메타데이터, 데이터셋 구조, 코드 체계, 속성값 등에 대한 공통 표준을 제시하고 있으며(과학기술정보통신부·NIA·TTA, 2024), 이를 바탕으로 기관 간 연계 API, 연계 데이터베이스(DB) 통합 작업을 확대하고 있다(행정안전부, 2024). 또한 「2025년 데이터 기반행정 활성화 시행계획」에서는 데이터 연계 활성화를 위한 기술 인프라 고도화와 API 기반 플랫폼 확장 전략을 강조하고 있다(행정안전부, 2025).

기술적으로는 RPA(Robotic Process Automation), ETL(Extract-Transform-Load), 데이터 파이프라인 자동화 솔루션이 점차 도입되고 있으며, 데이터 변환·정제·통합 과정을 자동화함으로써 표준화 효율성과 정확성이 크게 향상되고 있다. 이와 함께 클라우드 기반의 데이터 허브 및 연계 중계 플랫폼 구축도 활발히 진행 중이다(과기정통부·NIA, 2024). 국제 사례로는 미국의 「Federal Data Strategy」가 대표적이다. 이 전략은 '데이터 표준화와 자동화 기반 연계'를 핵심 원칙으로 내세우며, 각 정부 기관에 메타데이터 표준 준수 및 시스템 간 API 연계를 의무화하고 있다(Federal Data Strategy Development Team, 2021).

또한 호주의 NSW 주정부는 'OneGov 데이터 플랫폼'을 통해 분산된 부처 데이터를 통합·표준화하여, 전자정부 서비스의 일관성과 효율성을 제고하고 있다(NSW Government, 2020).

한편, 윤건(2023)의 연구에 따르면 한국의 공공기관들은 데이터 표준화 수준이 편차를 보이며, 부처 간 연계율이 낮은 편에 속한다. 이에 따라 국가 차원의 통합 거버넌스 체계를 통한 연계 인프라 통합, 통일된 속성 정의서 제공, 자동변환 도구 보급 등이 요구된다. 향후 자동화 기반의 표준화 및 연계체계는 다음과 같은 방향으로 발전되어야 한다. 첫째, 범정부적으로 사용되는 표준용어와 데이터 속성 체계를 통합하고, 이를 일관되게 관리할 수 있는 기술적·제도적 시스템을 갖추어야 한다. 둘째, 부처 간 원활한 데이터 연계를 가능하게 하기 위해 통합 메타허브와 이를 효율적으로 운영할 수 있는 중앙화된 포털을 구축해야 한다. 셋째, 데이터 연계 과정에서 발생할 수 있는 오류를 실시간으로 탐지할 수 있도록 연계 로그를 분석하고, 자동 오류 감지 기능을 포함한 모니터링 체계를 마련해야 한다. 넷째, 클라우드 인프라를 기반으로 하는 API 자동 연동 시스템을 도입함으로써, 다양한 기관 간의 데이터 흐름을 실시간으로 연계하고, 업무 간소화 및 효율성을 극대화할 수 있어야 한다. 이러한 기술 기반의 연계 체계는 단지 부처 간 협업을 넘어, 데이터 중심의 정책 연계성 강화를 가능하게 하며, 나아가 공공서비스의 개인화, 맞춤형 행정 구현 등을 위한 핵심 역량으로 작용할 것이다.

다. 참여 기반 운영 구조와 민관협력 체계

1) 시민 참여 기반의 데이터 정책 운영 구조

데이터 거버넌스의 정당성과 수용성을 제고하기 위해서는 정책의 수립 및 실행 과정에서 시민의 참여가 제도적으로 보장되어야 한다(MacFeely et al., 2022). 이는 단순한 피드백 수렴 수준을 넘어서, 데이터의 수집, 분석, 활용 전 과정에 걸쳐 시민의 목소리를 반영하는 '참여 기반 데이터 거버넌스(participatory data governance)' 체계를 의미한다. Van Zoonen (2020)은 디지털 복지국가에서 시민참여형 데이터 거버넌스가 데이터 편향을 방지하고 정책 신뢰성을 제고하며, 사회적 연대를 강화하는 데 기여한다고 보았다. 특히 복지, 교육, 의료 등 공공서비스 분야에서 이러한 효과는 더욱 두드러진다. Tennison(2024) 또한 시민 참여가 데이터 규범 형성과 윤리적 책임 확보의 수단이 될 뿐만 아니라, 기술 관료주의적

결정구조에 대한 민주적 견제 장치로 기능한다고 분석한다. 한국에서도 일부 지자체와 공공기관은 시민위원회, 시민평가단, 시민감시단 등 다양한 참여 조직을 구성하여, 데이터 개방 범위, 민감정보 활용, AI 서비스 설계 등 민감한 사안에 대해 시민 의견을 수렴하고 있다. 예를 들어 서울시의 「데이터 거버넌스 민관협의회」는 시민, 전문가, 공무원이 공동으로 참여하여 데이터 정책의 우선순위와 가치기준을 합의하는 방식으로 운영되고 있다(윤건, 2023).

한편, 시민 참여를 효과적으로 지원하기 위한 기술 기반도 중요한 요소다. 예를 들어, 개방형 플랫폼(Open Policy Platform), 온라인 의견수렴 포털, 시각화 기반 의사결정 도구 등은 시민들이 데이터 기반 정책 과정에 보다 쉽게 접근할 수 있는 경로를 제공하며, OECD(2022)는 이를 ‘데이터 민주화 인프라’로 정의한다.

효과적인 참여형 데이터 거버넌스를 구현하기 위해서는 몇 가지 핵심 조건이 필요하다. 첫째, 시민 참여를 제도화하기 위한 법적 기반이 마련되어야 한다(MacFeely et al., 2022; OECD, 2022). 예컨대, 공공데이터법 내에 시민 참여 조항을 명시하거나, 데이터 정책 결정 구조 내에서 시민의 참여 권리를 명문화하는 것이 포함된다(Van Zoonen, 2020). 둘째, 정책 수립 초기 단계부터 시민 의견을 반영할 수 있는 절차적 통로가 확보되어야 하며, 이는 단순한 사후적 피드백을 넘어서 공동기획 과정까지 포함해야 한다(Van Zoonen, 2020). 셋째, 시민의 데이터 이해도를 높이기 위한 교육 프로그램과 리터러시 향상 전략이 병행되어야 하며, 이는 참여의 질과 지속가능성을 보장하는 기반이 된다(Tennison, 2024). 넷째, 다양한 이해관계자 간의 협의 구조를 통해 의견 충돌을 조정하고, 소수 의견도 포용할 수 있는 거버넌스 설계가 필요하다(이진우, 2020). 이러한 조건들이 충족될 때, 시민 참여 기반 데이터 정책 운영 구조는 단순한 민주적 형식을 넘어서, 데이터 활용의 투명성과 정당성을 실질적으로 제고하는 수단으로 작용하게 된다(윤건, 2023). 나아가 시민은 데이터 정책의 단순한 수요자가 아니라 공동설계자(co-designer), 나아가 ‘데이터 공동통치자(co-governor)’로 자리매김하게 되며, 이는 궁극적으로 신뢰 기반의 지속가능한 데이터 거버넌스 체계를 실현하는 토대가 된다(MacFeely et al., 2022; Tennison, 2024).

실제 한국의 일부 지자체에서는 시민참여형 데이터위원회, 주민 패널단, 공공데이터 활용 워크숍 등의 방식을 통해 시민 의견을 정책에 반영하고 있다(윤건, 2023). 예를 들어 서울시와 성남시는 시민 제안 데이터를 정책화하는 절차를 제도화하고, 데이터 기반 정책

설계 과정에서 시민·전문가·공무원이 함께 참여하는 협의 모델을 운영 중이다. 이진우 (2020)는 이러한 시민 참여 구조가 정부에 대한 신뢰를 높이고, 데이터 기반 정책의 사회적 수용성을 향상시키는 데 기여한다고 평가하였다. 특히 참여 기반 데이터 거버넌스는 취약계층이나 소수집단 등 데이터로부터 소외될 가능성이 있는 집단의 목소리를 제도 내에 통합함으로써, 데이터 공정성과 형평성 실현의 실질적 수단이 된다(Tennison, 2024). 따라서 데이터 거버넌스 실행 전략에서는 참여 채널을 제도화하고, 시민이 정책 결정에 실질적 영향을 미칠 수 있는 설계가 필요하다(MacFeely et al., 2022). 구체적으로는 시민 패널 운영(MacFeely et al., 2022), 거버넌스 위원회 내 시민 대표 포함(Tennison, 2024), 시민 데이터 실험실(Civic Data Lab) 운영(OECD, 2022) 등 다양한 참여 모델이 요구되며, 이를 위한 법적·재정적 지원 또한 병행되어야 한다(이진우, 2020; 윤건, 2023).

2) 민관 데이터 협력체계와 디지털 파트너십 구축

데이터 기반 정책환경의 복잡성과 기술의 급속한 발전은 정부 단독의 거버넌스로 대응하기 어려운 한계를 노정시키고 있다. 이러한 배경에서 공공과 민간 간의 협력적 데이터 생태계 구축은 필수적인 전략으로 부상하고 있으며, 이는 ‘디지털 파트너십’이라는 개념으로 제도적 정당성을 확보하고 있다(Fernandes, 2022). 민간이 보유한 기술력, 혁신 역량, 방대한 데이터 자산은 공공부문과의 연계를 통해 상호 보완적인 데이터 거버넌스 체계로 통합될 수 있으며, 이는 행정의 혁신성과 정책의 효과성을 동시에 제고하는 기반이 된다. 한국 정부는 「데이터기반행정 기본계획」과 「데이터경제 활성화 전략」 등을 통해 민관 데이터 연계 및 공동 활용 체계를 제도화하였으며, 한국지능정보사회진흥원(NIA), 한국데이터산업진흥원 등 중간지원기관을 중심으로 다양한 협업 프로젝트를 확장하고 있다(행정안전부, 2025; 관계부처 합동, 2024).

민관 협력은 크게 세 가지 유형으로 구분된다. 첫째, 민간 데이터의 공공 활용은 통신사, 카드사, 유통사 등 민간 부문에서 수집된 데이터를 기반으로 공공정책을 설계하거나 행정 서비스를 개선하는 방식으로 나타난다. 대표 사례로는 유동인구 데이터를 활용한 대중교통 최적화 정책이 있다. 둘째, 공공데이터의 민간 제공 확대는 API 기반의 공공데이터 개방을 통해 기업의 서비스 혁신과 산업 경쟁력 제고를 유도하는 방식이다. 예를 들어 기상, 환경, 건강 관련 정보는 기업의 상품 개발과 데이터 분석에 광범위하게 활용되고 있다. 셋

제, 민관 공동의 데이터 생산·분석 플랫폼 구축은 정부와 민간이 협력하여 데이터 수집, 표준화, 분석 시스템을 공동으로 설계하고 운영함으로써, 상호 활용 가능한 구조를 마련하는 데 목적이 있다. 지역통계 허브, 스마트시티 데이터 연합 등이 그 대표적 사례이다(행정안전부·한국지능정보사회진흥원, 2024).

국제적으로도 디지털 파트너십의 제도화가 활발히 진행되고 있다. 유럽연합은 ‘공공-민간 데이터 스페이스’ 전략을 통해 민간이 보유한 데이터셋에 대한 공공 접근권을 제도화하고 있으며(European Commission, 2020), 싱가포르와 핀란드는 국가 주도의 공동 데이터 거버넌스 프레임워크를 도입하여 민관 간 협력체계를 구조화하고 있다(Fumega, 2024). 이러한 민관 협력형 데이터 거버넌스의 성공적 구축을 위해서는 몇 가지 조건이 수반되어야 한다. 우선, 데이터 접근권 및 소유권에 관한 법적 규정을 명확히 정비함으로써 민관 간 책임성과 권한 분장을 명료히 할 필요가 있다. 둘째, 개인정보 보호와 데이터 활용 간 균형을 유지할 수 있는 협약 체계 구축이 필수적이다. 셋째, 메타데이터 구조와 표준 연동을 기반으로 한 기술적 연계성 확보가 중요하며, 마지막으로 민간의 자발적 참여를 유도할 수 있는 세제 혜택, 인센티브 도입 등 인센티브 설계가 필요하다(KISA, 2023). 김지현(2023)은 이러한 협력 구조의 제도화를 위해 정책의 일관성 확보, 중간지원기구의 전문성 강화, 기업·비영리·시민단체 등 다중 이해관계자의 협의체 구성이 선결되어야 함을 강조하고 있다. 결국 민관 협력체계는 디지털 전환 시대에 공공성과 혁신성을 동시에 확보하기 위한 제도적·기술적·조직적 연결고리로 기능하며, 지속가능한 데이터 생태계를 구현하는 핵심 기반으로 자리매김하게 된다.

3) 지역기반 데이터 협치와 분권형 거버넌스 구조 구축

중앙집중적 데이터 거버넌스 체계는 지역의 특수성과 다양성을 반영하는 데 한계를 가지며, 이에 따라 지역 기반의 데이터 협치 구조 구축이 새로운 전략 과제로 부상하고 있다. 지역 단위의 데이터 거버넌스는 주민의 삶에 밀접한 데이터를 기반으로 정책의 실효성을 제고하고, 참여 기반의 분권적 데이터 생태계를 조성하는 데 중점을 둔다(Fernandes, 2022). 한국에서는 「데이터기반행정 기본계획」을 통해 지방자치단체의 데이터 기반 정책 역량 강화를 위한 방안을 제시하고 있으며, 시범사업으로 ‘지역 데이터 거버넌스 네트워크(Local Data Governance Network)’ 구축을 지원하고 있다(행정안전부, 2024). 이러한 분권

형 데이터 거버넌스는 세 가지 핵심 요소를 중심으로 구성된다. 첫째, 지역 데이터 관리위원회의 구축이 요구된다. 이는 지방정부, 시민사회, 민간 전문가 등이 함께 참여하는 협의체로, 지역 차원의 데이터 활용 원칙, 개방 수준, 인공지능 적용 범위 등을 공동으로 논의하고 결정하는 역할을 수행한다(윤건, 2023). 둘째, 지역 데이터 허브 및 연계 시스템의 구축이 필수적이다. 이는 지역 내 공공기관과 민간이 보유한 데이터를 연계·통합할 수 있도록 기술적 기반을 마련하는 것으로, 대표적으로 경기도의 '데이터 허브'나 서울시의 '스마트도시 데이터 플랫폼'과 같은 사례가 있다(행정안전부·한국지능정보사회진흥원, 2024; 김지현, 2023). 셋째, 시민참여형 협치 구조를 통한 참여 기반 강화가 요구된다. 이는 지역 주민이 직접 데이터의 제공과 활용 과정에 참여하거나, 공공서비스 개선에 필요한 데이터를 제안할 수 있는 구조를 의미하며, 데이터 청구 플랫폼이나 지역 단위의 데이터 공론장 등이 이에 해당한다(MacFeely et al., 2022; OECD, 2022).

국제적으로도 지역 기반 데이터 거버넌스의 모범 사례가 등장하고 있다. 예컨대 영국의 Greater Manchester Combined Authority는 지역 주도의 데이터 전략을 수립하고 이를 복지, 보건, 교통 등 다양한 정책 영역에 적용함으로써 실질적인 데이터 기반화를 구현하고 있다(NSW Government, 2020). 또한 캐나다 몬트리올시는 도시 데이터 플랫폼을 구축하여 시민의 직접 참여를 기반으로 한 협치형 데이터 거버넌스 구조를 운영하고 있다(Charles et al., 2022). 지역 기반의 거버넌스가 성공적으로 작동하기 위해서는 몇 가지 선결 조건이 충족되어야 한다. 먼저, 지역정부 내에 데이터 전담 조직과 전문 인력을 확보하는 것이 중요하다. 이는 데이터 기획, 운영, 분석, 품질관리 등의 업무를 안정적으로 수행하기 위한 기본 인프라로 기능한다(윤건, 2023). 다음으로, 중앙정부와 지방정부 간에는 데이터 형식과 전송 방식에 대한 표준화된 연계 프로토콜이 마련되어야 하며, 이를 통해 상호운용성과 정책 일관성을 동시에 확보할 수 있다(행정안전부·NIA, 2024). 또한 지역 간 데이터 기반 역량 격차를 고려하여 기술적·재정적 지원이 차등적으로 제공되어야 하며, 이로써 지역 특성을 고려한 균형 잡힌 발전이 가능해진다(OECD, 2022). 마지막으로, 각 지역의 특수성과 수요를 반영한 맞춤형 데이터 활용 규범과 실행 가이드라인이 병행적으로 개발되어야 한다(MacFeely et al., 2022). 윤건(2023)은 이러한 분권형 데이터 거버넌스를 제도적으로 뒷받침하기 위해 지방정부가 데이터 자율성과 책임성을 균형 있게 확보할 수 있도록 법제도 차원의 구조 개편이 필요하다고 주장한다. 그는 특히 중앙정부는 통제자가 아니라

조정자 및 지원자로서의 역할에 집중해야 한다고 강조한다. 결론적으로, 지역기반 데이터 협치 및 분권형 거버넌스 구조는 중앙의 일방적 통제를 탈피하고, 지역 사회의 특수성을 반영한 데이터 생태계 구축을 가능하게 한다. 이는 데이터 민주주의 실현과 포용적 디지털 전환의 핵심 기반으로 작용할 수 있다(OECD, 2022; MacFeely et al., 2022).

라. 데이터 책임성과 윤리 거버넌스 확보

1) 데이터 책임성 강화를 위한 제도적 장치 마련

데이터 기반 사회에서는 오류, 차별, 불투명성 등 다양한 문제가 발생할 수 있으며, 이에 대한 예방과 사후 책임 규명을 위한 제도적 장치의 마련이 필수적이다(Fumega, 2024). 특히 인공지능과 자동화된 결정 시스템이 공공정책에 활용되는 경우, 데이터의 투명성과 설명 가능성을 확보하는 것은 정책의 정당성과 공공성 담보를 위한 전제조건이다(OECD, 2022). OECD(2022)는 데이터 거버넌스의 핵심 원칙 중 하나로 ‘책임성과 정당성(accountability and legitimacy)’을 강조하면서, 이를 실현하기 위한 구체적 제도적 방안을 제시한다. 첫째, 알고리즘이 공공정책 결정에 미치는 영향을 사전에 점검할 수 있도록 ‘알고리즘 영향 평가(Algorithmic Impact Assessment, AIA)’ 제도의 도입이 필요하다(OECD, 2022). 둘째, 데이터 활용과 관련된 책임 주체를 명확히 하기 위해, 데이터의 생성부터 처리, 활용에 이르기까지 전 과정을 추적할 수 있는 ‘데이터 책임 추적 체계’를 구축할 필요가 있다(OECD, 2022). 셋째, 데이터의 오류나 부정확성으로 인해 시민이 피해를 입었을 경우, 이를 시정하거나 이의를 제기할 수 있도록 ‘사후 정정 및 이의신청 절차’를 제도화해야 한다(OECD, 2022).

한국의 경우 「개인정보보호법」 및 「데이터기반행정법」은 개인정보 처리 책임자 지정, 오남용에 대한 제재 조항 등을 규정하고 있으나, 일반 비식별 데이터나 공공정책용 데이터에 대한 포괄적 책임 프레임워크는 아직 미비한 실정이다(개인정보보호위원회, 2025). 이에 따라 데이터 활용의 전 과정에서 책임 주체를 명확히 규정하고, 정책 실패나 사회적 부작용이 발생했을 때 이에 대응할 수 있는 사후 책임 메커니즘이 필요하다(개인정보보호위원회, 2025). 특히 AI 학습용 데이터의 경우, 데이터 신뢰성과 책임성을 확보하기 위한 방안으로 ‘데이터 사용 내역 기록 시스템’과 ‘데이터 신뢰지수 제도’ 도입이 제안되고 있다(Hausenloy et al., 2024). 아울러 윤건(2023)은 데이터의 품질뿐 아니라 출처, 사용 이력, 윤리적 검토 여부 등을 포괄적으로 평가하는 ‘데이터 거버넌스 인증제도’의 도입을 통해 공공성과 책임성을 강화할 수 있다고 주장한다.

이러한 맥락에서 데이터 책임성을 제도적으로 확보하기 위해 다음과 같은 정책 방향이 제시될 수 있다. 첫째, 정부 및 각 기관별로 최고 데이터 책임자(CDO)의 역할을 강화하고 법제화를 추진하여 조직 내 책임 구조를 명확히 해야 한다(Tennison, 2024). 둘째, 정책 수립이나 인공지능 모델 개발 시 활용되는 데이터셋에 대해 사전검증 및 사후검토 절차를 의무화함으로써 데이터의 신뢰성과 적정성을 제고해야 한다(Hausenloy et al., 2024). 셋째, 데이터 오남용이나 예기치 못한 피해가 발생했을 때 분쟁을 해결할 수 있는 중재기구나 이의제기 절차 등 사후적 대응 메커니즘을 마련해야 한다(OECD, 2022). 넷째, 공공기관이 수행하는 데이터 활용과 관련하여 ‘데이터 책임 이행보고서’를 정기적으로 발간함으로써, 데이터 운영의 투명성과 국민에 대한 설명 책임을 강화할 필요가 있다(Fumega, 2024). 결과적으로, 이러한 제도적 장치는 데이터 활용의 윤리성과 책임성을 확보하고, 사회 전반의 신뢰 기반 위에서 데이터 거버넌스를 실효성 있게 추진하기 위한 핵심 조건이라 할 수 있다(MacFeely et al., 2022; OECD, 2022).

2) 프라이버시 중심 설계와 개인정보 보호 강화

디지털 전환과 인공지능 기술의 급속한 확산은 대규모의 개인정보 수집 및 분석을 가능하게 하였으며, 이는 개인의 프라이버시 침해 가능성을 심화시키는 결과를 초래하고 있다(Marcucci et al., 2023). 이러한 환경 변화 속에서 데이터 거버넌스 전략은 단순한 기술 관리 차원을 넘어, 프라이버시 보호를 핵심 원칙으로 내재화하는 ‘설계에 의한 보호(privacy-by-design)’ 접근을 채택해야 한다(European Commission, 2020). ‘프라이버시 중심 설계’는 데이터 수집 이전 단계에서부터 보호조치를 구조적으로 설계하는 방식으로, 최소한의 데이터 수집, 정보의 목적 명시 및 이용 제한, 익명화·가명화 조치 등을 기술적·제도적으로 구현하는 것을 의미한다(Marcucci et al., 2023). 유럽연합의 GDPR 및 한국의 「개인정보보호법」은 모두 이 원칙을 명시하고 있으며, 특히 인공지능 알고리즘이 개인정보를 활용할 경우에는 데이터 주체에게 자동화된 결정에 대한 설명권(right to explanation)을 보장해야 한다고 규정하고 있다(European Commission, 2020).

국내에서도 이와 같은 흐름에 발맞추어 정책적 조치가 강화되고 있다. 개인정보보호위원회(2025)는 「2025년 주요업무계획」을 통해 공공기관이 시스템을 설계할 때 프라이버시 영향평가(Privacy Impact Assessment, PIA)를 의무적으로 시행하도록 하고 있으며, 민감정보를 처리할 경우에는 기술적 보호조치와 사전 검토 체계를 도입하도록 요구하고 있다(개

인정보보호위원회, 2025). 기술적으로는 최근 ‘차등 개인정보 보호(differential privacy)’나 ‘연합 학습(federated learning)’과 같은 방식이 프라이버시 중심 데이터 분석 기술로 주목받고 있으며, 이는 원천 데이터를 외부로 노출하지 않으면서도 분석 가능성을 확보할 수 있는 방법으로 평가받고 있다(Marcucci et al., 2023). 특히 AI 학습용 데이터 구축 과정에서 프라이버시 보호는 더욱 중요해지고 있다. 과학기술정보통신부와 한국지능정보사회진흥원(2024)은 「AI 학습용 데이터 품질관리 가이드라인」을 통해 비식별 조치, 민감정보의 분리 저장, 접근 통제 등의 구체적 기준을 제시하였으며, 이 기준은 공공 부문은 물론 민간 AI 기업에게도 확대 적용되고 있다(과학기술정보통신부·한국지능정보사회진흥원, 2024).

프라이버시 중심 데이터 거버넌스를 효과적으로 구현하기 위해서는 몇 가지 핵심 조건이 선행되어야 한다. 첫째, 시스템 설계 초기 단계에서부터 프라이버시 영향평가(PIA)를 의무화해야 하며(개인정보보호위원회, 2025), 둘째, 데이터 수집을 최소화하고 민감정보에 대한 별도 처리 기준을 명확히 해야 한다(Marcucci et al., 2023). 셋째, 익명화 및 가명화 조치 기준을 고도화함으로써 데이터 주체의 재식별 가능성을 원천적으로 차단해야 하며(European Commission, 2020), 넷째, 암호화, 접근제어와 같은 기술적 보호조치와 사전승인·외부검토 등 제도적 보호장치를 병행하는 종합적인 거버넌스 체계를 마련할 필요가 있다(과학기술정보통신부·한국지능정보사회진흥원, 2024). 결국 프라이버시 보호는 단순한 규제 문제가 아니라, 데이터 활용에 대한 사회적 신뢰를 구축하고, 데이터 주체의 권리를 실질적으로 보장하는 ‘디지털 권리 기반’ 데이터 거버넌스의 핵심 축이다(Marcucci et al., 2023; OECD, 2022). 따라서 데이터 거버넌스 실행 전략은 기술적 수단과 제도적 장치를 유기적으로 결합하여, 프라이버시 보호를 전략적 우선순위로 내재화하는 방향으로 설계되어야 한다(European Commission, 2020; 개인정보보호위원회, 2025).

3) 윤리 기준 내재화 및 글로벌 규범 연계

데이터의 수집, 분석, 활용 과정 전반에 윤리적 기준을 내재화하는 것은 신뢰 기반의 데이터 거버넌스 체계를 구축하기 위한 핵심적 요소로 평가된다. 특히 인공지능(AI) 기술과 결합된 데이터 활용이 사회적 불평등, 알고리즘 편향, 차별적 결과를 초래할 수 있다는 점에서, 윤리적 가이드라인과 실행 규범의 정립은 데이터 정책의 필수적인 구성요소로 부각되고 있다(Schoeman, 2023). 이와 관련하여 OECD(2022)는 데이터 윤리의 네 가지 핵심 원칙으로 인간 중심성(human-centricity), 공정성(fairness), 투명성(transparency), 책임성

(accountability)을 제시하고 있다. 또한 이들 원칙을 실현하기 위한 정책적 수단으로는 윤리영향평가(Ethical Impact Assessment, EIA), 윤리 자문위원회, 알고리즘 감사(auditing), 윤리 인프라 구축 등이 포함된다(OECD, 2022). 국제적 차원에서도 윤리적 데이터 거버넌스 구조는 구체적으로 제도화되고 있다. 유럽연합은 「AI Act」를 통해 고위험 AI 시스템에 대한 윤리 기준을 명시하였으며(European Commission, 2021), 미국은 NIST(National Institute of Standards and Technology)를 중심으로 AI 시스템의 윤리 프레임워크를 마련하고 있다(NIST, 2023). 한국의 경우에도 「2025년 개인정보보호위원회 업무계획」과 과학기술정보통신부의 데이터 가이드라인을 통해 AI 데이터에 대한 윤리 검토 절차와 품질 기준이 제시되고 있으나, 전반적인 윤리 거버넌스 체계는 아직 초기 단계에 머물고 있다(개인정보보호위원회, 2025; 과학기술정보통신부, 2024).

윤리 기준의 내재화를 위한 구체적인 실행 방안으로는 다음과 같은 제도적 장치가 요구된다. 첫째, 공공 데이터 및 AI 시스템 개발 단계에서 윤리영향평가(EIA)를 의무화하여, 잠재적 사회적 영향을 사전에 검토해야 한다(OECD, 2022). 둘째, 데이터 처리 및 알고리즘 개발 전반에 걸쳐 편향성과 차별 가능성을 점검할 수 있는 기술적·절차적 체계를 구축해야 한다(Tennison, 2024). 셋째, 독립적인 윤리 자문기구 및 외부 검토 위원회를 제도화함으로써 객관적 감시와 책임성 확보를 병행해야 한다. 넷째, 공공기관은 데이터 윤리보고서를 정기적으로 발간하고, 그 성과와 이행 내용을 대중에 공개함으로써 투명성과 정당성을 확보할 수 있어야 한다(개인정보보호위원회, 2025). Effoduh et al.(2024)은 아프리카의 데이터 윤리 사례를 분석하며, 포용성과 형평성을 기준으로 윤리 기준을 지역 맥락에 맞게 조정할 필요성을 강조하였고, Tennison(2024)은 윤리 원칙이 선언적 차원에 머무르지 않고, 실질적인 운영 규범과 제도적 구조로 통합되어야 한다고 주장한다. 따라서 데이터 거버넌스에 윤리 기준을 내재화하는 과정은 단순한 선언적 원칙을 넘어서, 법·제도적 보완과 함께 실효적인 절차 설계 및 책임 주체의 명확화를 포함하는 구조적 거버넌스의 일부로 통합되어야 한다. 더 나아가, AI 및 데이터 활용의 초국경적 특성을 고려할 때, 국내 윤리 기준은 국제적 규범과의 정합성을 확보하는 방향으로 확장되어야 할 것이다(OECD, 2022; European Commission, 2021).

제 4 절 데이터 거버넌스의 사회적 · 윤리적 논의

1. 데이터 활용과 프라이버시 보호의 윤리적 균형

가. 개인정보 보호와 데이터 활용 간의 갈등 구조

디지털 전환의 가속화는 데이터 기반 정책의 실효성과 효율성을 제고하는 한편, 개인정보 보호와 관련한 윤리적 우려를 동반한다. 한편으로, 여러 기관이 공공 데이터로서 보유하는 개인정보는 공유 및 연계·통합될 때 데이터 활용의 목적을 달성하기 용이하다(조세현·차세영, 2022). 동일한 영역 또는 분야의 데이터라고 하더라도, 공공부문과 민간부문에 걸친 여러 기관이 보유하는 데이터를 결합할 때, 수요자의 필요와 요구를 충족시키는 창의적이고 혁신적인 공공 서비스를 개발·제공할 수 있다.

다른 한편으로, 개인정보가 엄격한 기준 없이 광범위하게 활용되는 경우, 프라이버시 침해를 포함한 다양한 사회문제를 초래할 수 있다. MacFeely et al.(2022)은 데이터 거버넌스가 인권 보호, 프라이버시 존중, 형평성 증진이라는 윤리적·사회적 책임을 내포하고 있으며, 특히 민감정보가 포함된 경우 포괄적 인권의 일환으로 접근해야 함을 강조한다. 이러한 관점은 데이터 활용의 투명성과 책임성 확보를 위한 국제적 윤리 기준 정립이 요구된다는 점을 시사한다. Marcucci et al.(2023) 또한 데이터가 본질적으로 인권과 연결된다는 점을 강조하며, UN과 EU의 인권 규범이 윤리적 앵커로 작동해야 함을 주장한다.

한편, 공공 데이터를 통한 예측 기반 정책 수립은 사회복지나 보건정책 등에 기여할 수 있지만, 동의 피로(consent fatigue), 낙인 효과(stigmatization), 데이터 사일로(data silo) 등의 부작용으로 인해 시민 신뢰를 약화시킬 수 있다(Malomo & Sena, 2017). 공공부문의 빅데이터 활용은 효율성과 위험성이 공존하는 이중성을 지니며, 목적 외 사용, 정보 비대칭 등의 문제가 반복적으로 지적되고 있다(김기환, 2013). 따라서, 데이터의 활용성과 보안성을 균형적으로 추가하기 위한 제도와 규범이 설계되고 준수되어야 한다.

나. 법·제도적 기준과 글로벌 규범

이러한 맥락에서 데이터 보호에 관한 국제적 규범은 국내 거버넌스 설계의 중요한 기준이 된다. Kroll(2018)은 GDPR의 핵심 조항인 접근권, 삭제권, 자동화된 결정에 대한 이의권이 윤리적 데이터 거버넌스를 위한 법적 기초를 제공한다고 보았다. 그는 데이터 기반 시

시스템이 단순히 법을 준수하는 데 그치지 않고, 사회적 책임을 내포한 체계로 운영되어야 한다고 강조한다. OECD(2022) 또한 「Going Digital Guide」에서 정책 수립·집행·평가의 전 과정에서 프라이버시 보호와 데이터 활용의 균형을 제시하는 기준으로 ‘사전 위험 평가’, ‘책임 주체 명시’, ‘이해관계자 참여’ 등을 제시하고 있다. 이러한 국제 기준들은 국내 제도의 정비 방향과 연계되어야 하며, 특히 개인정보보호위원회(2025)가 제안한 AI 자동결정에 대한 설명요구권, 디지털 취약계층 보호, 자기결정권 강화 등은 시민권 기반의 데이터 거버넌스 정착을 위한 구체적 전략으로 주목된다.

2. 프라이버시 중심 데이터 설계 원칙

가. 민감정보의 기준 설정과 활용 제한

데이터의 활용 과정에서 발생할 수 있는 프라이버시 침해를 예방하기 위해서는 데이터의 민감도에 따라 접근권, 재사용 조건, 공유 범위 등을 차등적으로 설정하는 체계적인 설계가 필요하다. Kanying et al.(2023)은 AI 기반 데이터 거버넌스에서도 ‘보안과 프라이버시’, ‘공정성’, ‘책임성’ 등 윤리적 요소가 제도적으로 내재화되어야 한다고 강조하며, AI 시스템이 데이터 주체의 권리를 구조적으로 보장할 수 있어야 한다고 주장한다. 이와 같은 관점은 AI 알고리즘이나 자동화 의사결정 시스템이 가져올 수 있는 권리 침해 위험을 사전적으로 차단하기 위한 기반 설계를 요구한다. Hausenloy et al.(2024)은 프런티어 AI 모델이 학습에 사용하는 데이터셋이 유해 정보나 차별적 콘텐츠를 포함하지 않도록, 데이터 수집 단계에서부터 정제·검토 과정이 내재화되어야 한다고 주장한다. 이는 프라이버시 중심 설계가 단순히 민감정보를 ‘보호’하는 수준을 넘어, AI의 윤리적 정당성 자체를 확보하기 위한 핵심 전제로 작동해야 함을 시사한다.

국내에서는 데이터의 민감도에 따라 데이터 유형을 개인정보, 가명정보, 익명정보로 구분하고, 각 데이터 유형에 따라 데이터 보호와 활용의 범위를 차등적으로 적용한다(개인정보보호위원회, 2023). 한편으로, 개인정보는 다른 정보와의 결합 없이도 그 자체로서 개인을 식별할 수 있는 정보이기 때문에, 활용가능성은 높지만 그만큼 프라이버시 침해가능성도 높다. 따라서, 개인정보에 대해서는 데이터 활용보다는 보호에 초점이 맞추어진다. 다른 한편으로, 익명정보는 그 자체로서 개인을 식별할 수 없고 다른 정보와의 결합에도 불

구하고 개인을 식별할 수 없는 정보이기 때문에, 프라이버시 침해가능성은 낮지만 그만큼 활용가능성도 낮다. 따라서, 익명정보에 대해서는 데이터 보호보다는 활용에 초점이 맞추어진다. 이러한 맥락에서, 최근 그 자체로서 개인을 식별할 수 없지만 다른 정보와의 결합을 통해 개인을 식별할 수 있는 가명정보를 데이터 활용과 보호의 균형을 위한 데이터로 간주하고 가명정보의 안전한 활용을 위한 제도적 설계 방안이 논의된다.

나. 기술적·제도적 보호조치

기술적 수단을 통한 프라이버시 보호는 데이터 거버넌스의 실효성을 담보하는 중요한 축이다. Micheli et al.(2022)는 지리정보 기반 플랫폼에서 위성·드론 영상 및 자동 지도화 기술이 활용될 경우, 익명화·가명화 기술만으로는 사회적 수용성을 확보하기 어려우며, 지역사회와의 합의 및 맥락 기반의 보호 전략이 병행되어야 한다고 지적한다. Schoeman (2023) 또한 프라이버시 보호는 단순한 기술적 차원에 머물러서는 안 되며, 조직 내 윤리 위원회 구성, 사회적 책임의 명확화, 데이터 윤리 규범의 제도화가 동시에 이루어져야 한다고 강조한다. 특히 설명가능성, 공정성, 자율성 존중과 같은 원칙은 설계와 실행의 전 단계에 내재되어야 하며, 투명한 감사 체계와 피드백 시스템을 통한 점점이 지속적으로 이루어져야 한다.

특히, AI 시대에 적합한 제도적 개인정보보호 조치가 마련되고 시행되어야 한다(명승환, 2020). 국내에서는 「개인정보보호법」에서 정보 주체의 동의를 받아 개인정보를 수집하고 그 수집 목적의 범위에서 이용할 수 있도록 규정함으로써 개인정보 자기 결정권을 확보하도록 한다. 하지만, AI 기술을 기반으로 한 자동화된 정보 수집 및 분석에서는 정보 주체가 자신도 인지하지 못한 상태에서 개인정보가 수집되고, 수집된 개인정보에 대하여 인간이 직접 조작하지 않은 상태에서 AI 기반 기계나 장치가 스스로 작동하여 처리하기 때문에, 개인정보 자기 결정권이 확보되기 어렵다. 그러므로, AI 시대에는 개인정보가 자동화된 방식을 통해 가명정보를 포함한 다양한 형태로 수집 및 이용될 수 있기 때문에, AI 서비스에 적합한 개인정보보호가 이루어질 수 있도록 하는 규정이 필요하다.

3. 디지털 격차와 데이터 불평등 해소

가. 접근성의 지역·계층 간 격차

디지털 격차는 디지털 기반 사회의 다양한 개인 및 집단 사이에서 데이터와 디지털 기술에 대한 접근 기회 및 활용에서 발생하는 차이를 의미한다(Van Dijk, 2020; 김문조·김종길, 2002). 디지털 격차의 유형은 크게 접근 격차, 활용 격차, 수용 격차로 구분된다. 디지털 기기와 데이터 자체에 대한 대가를 지불해야 하는 디지털 기반 사회에서 특히 사회적 취약계층은 디지털 기기와 데이터에 대한 접근 기회가 제한된다. 디지털 기기와 데이터에 대한 접근 기회가 동등하다고 하더라도, 디지털 기기와 데이터를 활용하여 다양한 가치를 창출하는 능력에 차이가 있는 경우, 개인 및 집단에 따라 활용 수준 및 정도에 차이가 발생한다. 접근 격차 및 활용 격차와 별도로, 디지털 기기와 데이터를 부작용 또는 역기능 없이 문명의 이기로서 지혜롭게 수용하는 능력에 차이가 있는 경우, 개인 및 집단에 따라 수용 수준 및 정도에 차이가 발생한다.

실제로 디지털 격차는 다양한 유형에 걸쳐 다양한 유형의 개인 및 집단 사이에서 발생하고 있다. 과학기술정보통신부·한국지능정보사회진흥원(2022)에 따르면, 사회적 취약계층(예: 장애인, 고령층, 저소득층, 농어민)의 디지털 정보화 접근 수준, 역량 수준, 활용 수준은 일반인에 비해 전반적으로 상당히 낮은 것으로 파악된다. 구체적으로, 사회적 취약계층의 PC 보유율과 스마트폰 보유율이 일반인에 비해 10%p 이상 낮은 것으로 확인되고, 사회적 취약계층의 PC 이용 능력과 모바일 기기 이용 능력이 일반인에 비해 20%p 이상 낮은 것으로 확인되고, 사회적 취약계층의 검색·이메일·콘텐츠 서비스 이용률이 일반인에 비해 10%p 낮은 것으로 확인된다. 특히, 고령층은 디지털 정보화 접근 수준, 역량 수준, 활용 수준에 걸쳐 다른 사회적 취약계층에 비해서도 가장 낮은 것으로 파악된다.

데이터 기반 행정이 확산되면서 공공데이터에 대한 접근성과 활용 능력이 새로운 형태의 사회적 불평등을 야기할 수 있다는 우려가 제기된다. Koltay(2016)는 데이터 리터러시가 단순한 기술 습득이 아니라, 정보의 구조, 품질, 맥락을 비판적으로 이해하는 능력을 포함한다고 정의하며, 이러한 능력의 결여는 시민뿐 아니라 공공기관 내부에서도 제도적 소외를 초래할 수 있음을 경고하였다. 유승현(2022) 또한 지방정부의 경우, 데이터 자산을 보유하고 있음에도 불구하고 이를 분석하거나 정책에 활용할 수 있는 전문인력과 조직 역량이 부족한 현실을 지적하였다. 이러한 구조적 제약은 수도권과 비수도권 간의 정책 설계

역량 격차를 심화시키고, 장기적으로는 행정 서비스의 질적 불균형으로 이어질 수 있다.

나. 참여 기반의 역량 강화 전략

이러한 격차를 해소하기 위해서는 단순한 기술적 지원을 넘어, 참여 기반의 역량 강화 전략이 병행되어야 한다. Koltay(2016)는 데이터 활용 역량 강화를 위한 교육은 기술 교육을 넘어서 정책적 리터러시를 포함하는 포괄적 구조로 설계되어야 한다고 보았다. Malomo and Sena(2017)도 지방정부에서 데이터 활용이 저조한 이유는 단순한 기술 부족이 아니라, 분석 결과와 정책 설계 간의 연계 부재 및 조직 내부 학습 구조의 미흡 때문이라고 지적하였다.

무엇보다도 수요자의 수준과 필요를 고려하는 수요자 맞춤형 교육을 통해 데이터 역량이 강화되어야 한다. 조세현·차세영(2022)은 수요자 맞춤형 교육을 위해 교육에 대한 수요 조사 및 데이터 역량 수준 진단을 바탕으로 한 교육 프로그램 운영과 현장에서 일반인 수준의 지식과 경험만으로도 활용할 수 있는 데이터 분석·개발 도구의 제공 및 교육을 강조한다. 데이터에 대한 기술전문성과 사업전문성을 보유한 개인 또는 기관 간 협업을 통해 새로운 가치를 창출하는 우수 사례를 학습하도록 유도하고 이러한 사례를 현장에 적용할 수 있는 교육 프로그램도 분석 결과와 정책 설계 간의 연계성 증진 및 조직 내부 학습 구조 강화를 위한 효과적인 수단이다.

한편, 시민의 데이터 참여 역시 실질적으로 확대되어야 한다. 유승현(2022)은 시민이 데이터의 생산자이자 수요자로 기능할 수 있도록 데이터 공개 구조를 설계해야 하며, 이는 단순한 정보 개방을 넘어 시민 주도형 정책 설계의 기반이 되어야 한다고 주장하였다. 특히 리빙랩(living lab)과 같은 지역사회 기반 실험 모델은 시민 참여를 제도화할 수 있는 효과적인 수단으로 평가된다. Marcucci et al.(2023)은 시민 참여 기반의 거버넌스가 단순한 공청회 수준을 넘어서, 시민이 실제 데이터 분석과정과 정책 의사결정에 관여할 수 있는 구조로 설계되어야 한다고 강조한다.

다양한 시민의 데이터 참여 수단 중에서 리빙랩은 시민들이 데이터를 기반으로 직접 공공문제를 해결하기 위한 아이디어를 창출하고 대안을 계획하고 실행함으로써 참여 기반의 역량을 강화하기 위한 효과적인 수단으로 평가된다(성지은 외, 2014). 특히 지역사회에서 시민들은 자신들이 거주하고 있는 근린지역의 다양한 현안과 문제에 대처하기 위해 지역

사회의 다양한 영역(예: 교통, 환경, 복지, 의료)에 대한 데이터 분석을 바탕으로 현황 진단, 문제점 파악, 대안 발굴을 스스로 시도할 수 있다. 이러한 시도의 반복과 확산은 리빙랩을 통해 시민들이 데이터의 생산자이자 수요자로 기능할 수 있는 다양한 잠재력을 현실에서 구현하는 데 기여한다.

4. 신뢰 형성과 수용성 확보를 위한 투명한 거버넌스

가. 투명성 중심의 데이터 운영

AI 기술을 바탕으로 한 자동화된 데이터 수집과 의사결정이 증가할수록, 데이터 신뢰성 및 수용성을 확보하기 위한 필요가 증가한다(정원섭, 2020). 구체적으로, 자동화된 데이터 수집과 의사결정을 가능하게 하는 알고리즘(algorithm)이 작동하는 컴퓨터 연산 과정은 근본적으로 비가시성(invisibility)을 특징으로 한다. 그러므로, 알고리즘을 설계한 전문가가 일반인에게 알고리즘을 공개해도, 일반인은 알고리즘 양의 방대함으로 인해 알고리즘을 가시적으로 확인하기 어려울 뿐만 아니라, 알고리즘이 어떤 과정을 거쳐 특정한 결론에 도달하게 되었는지를 기술적으로 이해하기 어렵다. 따라서, 특히 AI 기술을 바탕으로 한 자동화된 데이터 수집과 의사결정 영역에서는 알고리즘을 포함한 데이터 과정을 단순히 공개하는 것만으로는 일반인으로부터 데이터 신뢰성 및 수용성을 확보하기 어렵다.

AI를 포함한 데이터 기반 정책이 시민의 신뢰를 얻기 위해서는, 데이터 수집·처리·활용의 전 과정을 투명하게 공개하고 설명할 수 있는 체계가 필요하다. MacFeely et al.(2022)은 데이터 생애주기의 투명성 확보가 공공의 신뢰 형성에 핵심 요소라고 강조하였다. 수집 주체, 처리 기준, 활용 목적 및 결과를 명확하게 공개함으로써 시민이 이해 가능한 방식으로 데이터를 접할 수 있도록 해야 한다는 것이다. Marcucci et al.(2023) 또한 단순한 정보공개 수준을 넘어서, 데이터에 대한 의미 있는 해석과 문맥 설명이 제공되어야 하며, 특히 자동화된 의사결정 과정에 사용되는 알고리즘의 작동 원리와 기준은 필수적으로 공개되어야 한다고 지적하였다. 이러한 투명성 확보는 알고리즘의 신뢰성과 결과의 정당성을 높이기 위한 전제조건이 된다.

나. 시민참여형 의사결정과 사회적 합의

투명한 운영뿐 아니라, 시민의 실질적 참여가 수용성 확보의 또 다른 핵심 축이다.

Micheli et al.(2022)는 지리정보 기반 행정의 사례를 통해, 데이터 활용 전 과정에서 시민 의견이 반영될 수 있는 구조적 피드백 메커니즘이 필요함을 강조하였다. 특히 민감정보가 사용될 경우, 사전 설명과 의 절차는 신뢰 확보의 출발점으로 작동한다. 유승현(2022)은 시민이 단순히 데이터 제공자나 정책 수용자에 그치는 것이 아니라, 데이터 분석과 정책결정에 직접적으로 참여할 수 있는 구조를 마련해야 한다고 주장하였다. 이러한 참여는 정책의 수용성과 실행 가능성을 높이는 데 기여하며, 데이터 기반 행정이 민주적 정당성을 확보하는 데 핵심적 역할을 한다.

세계 각국의 정부는 AI 및 데이터에 대한 시민참여형 의사결정을 AI 또는 데이터 관련 법률, 지침, 현장 등에 명시하고 있다(고학수 외, 2021; 관계부처 합동, 2020). 예를 들어, 한국 정부는 「사람이 중심이 되는 인공지능(AI) 윤리기준」을 통해 사회적 신뢰 형성을 위해 AI 활용 상황에 적합한 수준의 투명성과 설명가능성을 높이려는 노력을 기울여야 한다는 원칙(투명성)뿐만 아니라, 인공지능 전 주기에 걸쳐 다양한 주체들의 공정한 참여 기회를 보장해야 한다는 원칙(연대성)을 10대 핵심요건의 구성요소로 제시한다. 또한, 유럽연합은 「인공지능법」을 통해 사람과 교류하고 있는 인공지능 시스템은 사람이 인공지능 시스템과 교류하고 있다는 사실을 적시에 고지해야 하고, 감정인식 및 생체인식 분류 시스템 사용자는 그 시스템의 작동 사실을 시스템에 노출되어 있는 사람에게 적시에 고지하고 사전에 동의를 획득해야 한다는 원칙(투명성 의무)를 제시한다.

5. 글로벌 규범과 윤리 거버넌스의 국제화

가. 초국가적 데이터 규범과 민주적 정당성

국경을 초월한 데이터 흐름이 일상화되면서, 국가 단위의 규제만으로는 글로벌 데이터 생태계의 윤리적 위험을 통제하기 어려운 상황이다. Arner et al.(2022)은 이를 ‘초국가적 데이터 거버넌스 문제’로 개념화하고, 글로벌 플랫폼 기업들이 규범의 실질적 창출자로 작동하면서 전통적인 국가 규제 체계의 민주적 정당성이 약화되고 있다고 지적하였다. 특히 규범 간 충돌과 조정 실패는 국제 데이터 정책의 일관성과 책임성을 저해하고 있으며, 이를 극복하기 위해 ‘데이터 정의(data justice)’라는 윤리 프레임 제시한다. Aaronson(2024)도 국가 간 이해관계로 인해 데이터 규범이 단절되고 있다고 보며, AI 시대에 대응하기 위

해서는 윤리와 인권을 기반으로 한 통합적 글로벌 거버넌스 체계 수립이 시급하다고 주장한다. 그는 데이터 주권, 감시 리스크, 기술 편향 등으로부터 시민을 보호할 수 있는 국제적 윤리 기준이 정립되지 않는다면, AI는 민주주의를 위협하는 수단으로 전락할 수 있다고 경고한다.

나. 위기 대응에서의 윤리 실패 사례와 교훈

팬데믹과 같은 위기 상황에서는 데이터 기반 기술의 효율성이 강조되지만, 이는 윤리적 원칙과 충돌할 소지가 있다. Beard & Belchamber(2020)는 코로나19 대응에서 AI 및 데이터 기술이 신속성과 효율성에 치중하면서 시민의 프라이버시 보호와 약자 보호라는 윤리 원칙이 희생되었음을 비판적으로 고찰하였다. 감염자 추적 시스템과 알고리즘 기반 의료 자원 배분 등은 동의 절차 부족, 설명 책임 미흡, 데이터 편향 문제를 야기하여 신뢰를 저해하였다고 분석된다. 이와 유사하게, Arigbabu et al.(2024)은 'Project Nightingale' 사례에서 민간기업 주도의 의료 AI 시스템이 환자의 동의 없이 데이터를 수집·활용하여 공공의 신뢰를 상실한 과정을 분석하였다. 이들은 사전 동의 체계, 시민 감시 메커니즘, 민간 파트너에 대한 규제적 책임이 윤리적 데이터 거버넌스의 필수 요소임을 강조한다.

다. 설명가능성과 책임 원칙의 국제적 정착 필요성

AI 기반 데이터 활용의 확산은 새로운 윤리 원칙의 제도화를 요구하고 있다. Boppiniti(2023)는 책임 있는 데이터 과학을 위한 핵심 원칙으로 설명가능성, 공정성, 책임성, 프라이버시 보호, 다학제 협업 등을 제시하며, 윤리적 기준이 설계와 운영 전 과정에 통합되어야 한다고 주장하였다. 국내 연구자들도 유사한 주장을 제시하고 있다. 선지원(2021)은 데이터 기반 행정이 법치주의와 인간 존엄성, 정보기본권과 충돌할 수 있는 구조적 한계를 지적하며, 공공 알고리즘 등록제와 설명요구권 도입 등을 통해 헌법적 정당성을 확보해야 한다고 주장한다. 김기환(2013)은 정보 비대칭, 목적 외 사용 등 데이터 활용의 위험성을 지적하며, 독립감독기구와 윤리강령 도입을 통해 공공의 신뢰를 확보해야 한다고 본다. 이와 함께 Medhora et al.(2018)는 데이터 거버넌스를 디지털 공공재 기반의 글로벌 구조로 재설계할 것을 제안하며, 시민 데이터 권리, 데이터 빈곤국 지원, 민간 공유 계약 등의 제도화를 통해 윤리적 지속가능성을 확보해야 한다고 강조한다. Winter(2021)와 Winter & Davidson(2019)은 의료 분야 AI 시스템의 확산에 따라 개인 건강정보의 재식별 위험성과

책임소재의 불분명성을 지적하며, '공정성·책임성·투명성(FAT)' 원칙과 다기관 참여형 거버넌스가 대안이 될 수 있음을 제안하였다.

제5절 소 결

1. 이론적 고찰 주요 내용

본 연구의 이론적 고찰에서 데이터 거버넌스의 개념의 변천사와 주요국의 데이터 거버넌스 사례를 살펴보았으며 나아가 데이터 거버넌스의 성공요인과 실행전략 및 데이터 거버넌스의 사회적·윤리적 논의하였다.

먼저, 데이터 거버넌스의 개념은 데이터 거버넌스는 현대 사회의 복잡한 문제를 해결하기 위한 협력적 통치 메커니즘에서 유래한다. 전통적으로 거버넌스는 정부뿐만 아니라 비영리 단체와 민간 부문이 공동으로 공공 문제를 다루는 것을 의미하며, 수직적 명령 체계에서 벗어나 협업, 투명성, 책임성을 핵심 원리로 삼는다(McGuire & Agranoff, 2011; Harrison & Sayogo, 2014). 정보통신기술의 발달로 e-거버넌스로 진화하면서 시민 참여와 데이터 기반 행정이 강화되었으며(Mcnamara, 2012; Misuraca, 2007) 인공지능(AI)의 도입은 데이터 편향성, 프라이버시 문제 등을 야기하며 AI 데이터 거버넌스의 필요성이 부상하였다(Florida et al., 2018). 이러한 개념의 변화에서 한국적 맥락에서 한국의 데이터 주권 강조(유승현, 2022)를 바탕으로 AI 편향성 도전(김운명, 2024)을 해결하기 위해 e-거버넌스 확장과 LLM 관리(Pahune et al., 2025)를 제도화하면 공공 행정 효율성을 높일 수 있다. 또한, 관계적·사회적 관점(Viljoen, 2021; Liu, 2022)을 반영해 데이터화로 인한 불평등(Redden, 2018)을 방지하며 취약계층 참여를 강화하면 사회적 포용을 높여 디지털 격차를 해소할 수 있다는 점을 확인 할 수 있었다.

주요 이론적 프레임워크에 대한 논의에서는 공통적으로 조직 구조와 역할 명확화, 데이터 생애주기 전반의 관리, 위험관리와 유연성, 투명성, 책임성, 윤리적 고려, 협력적 네트워크와 참여 그리고 기술·정책·프로세스의 통합에 대해 강조되었다.

특히 대부분의 프레임워크에서는 최고데이터관리자(CDO)나 데이터 스튜어드 팀의 임명과 같이 역할 분담을 강조하며 Janssen et al.(2020)과 Kanying et al.(2023)의 접근에서처럼 명확한 책임 배분이 조직 내 협력을 촉진하고 AI 프로젝트의 실패 위험을 줄이는 핵심 요소로 간주한다. 또한 데이터 생성, 저장, 활용, 폐기에 이르는 전체 주기를 포괄적으로 다루는 것이 공통적이며 품질, 보안, 추적성을 지속적으로 확보하는 데 초점을 맞추고 있다.

이는 Marcucci et al.(2023)의 글로벌 모델이나 Schoeman(2023)의 AI 거버넌스에서 확인할 수 있듯이 데이터의 장기적 가치와 위험을 고려한 접근으로, 디지털 전환 시대의 데이터 자산 관리를 강화한다. 아울러 위험 기반 접근이 공통적으로 부각되며 AI·디지털 환경의 불확실성에 대응하기 위한 예방적·적응형 전략을 중시한다. Luna-Reyes et al.(2024)의 복합적 체계나 Hausenloy et al.(2024)의 프론티어 거버넌스처럼 이는 데이터 편향성이나 보안 취약점을 사전에 식별하고 대응하는 데 필수적이며, 장기적 변화와 실시간 조정을 균형 있게 지원한다. 더불어 공정성, 설명가능성, 프라이버시 보호 등 윤리적 요소를 통합하며, 사회적 신뢰를 구축하는 데 필수적이라고 강조한다. Alqudah(2021)과 Batool et al.(2025)의 프레임워크에서처럼, 이는 AI 시스템의 블랙박스 문제를 해결하고, 시민 참여를 통해 민주적 가치를 반영하는 방향으로 거버넌스를 확장한다. 나아가 이해관계자(정부, 시민, 민간) 간 협력과 플랫폼 기반 공유를 통해 개방적 생태계를 형성하며, 가치 창출과 균형(공공성·혁신)을 추구한다. Micheli et al.(2020)의 데이터 공유 모델이나 한국지능정보사회진흥원(2024)의 국가 데이터 카탈로그처럼, 이는 조직 간 데이터 파편화를 해소하고, 시민 중심의 참여를 통해 공공 서비스 혁신을 촉진한다. 마지막으로 기술 인프라(클라우드, 블록체인 등)와 정책·프로세스를 결합하여 실시간·지속적 개선을 가능케 한다. Jiong'en & Yingliang(2024)의 클라우드 플랫폼이나 Carlizzi & Quattrone(2022)의 순환모형에서 보듯, 이는 기술적 효율성과 제도적 안정성을 동시에 추구하며 AI 시대의 동적 환경에 적합한 프레임워크를 형성한다. 이러한 주요 이론적 프레임워크의 공통점들은 AI·디지털 전환 시대에 데이터 거버넌스를 신뢰 기반의 전략적 체계로 재정립하는 데 핵심적이며 특히 공공 정책 연구와 디지털 거버넌스 분야에서 다각적 균형을 추구할 필요성을 시사한다.

주요국 사례와 국제기구 차원의 데이터 거버넌스 프레임워크에서는 EU의 GDPR 중심 규제, 미국의 CLOUD Act, 중국의 사이버 주권, 영국과 캐나다의 중앙집중형 AI 거버넌스 등 각국의 정책·제도적 차이를 살펴본 뒤, OECD의 전략·전술·전달 모델과 AI 원칙, 규제 샌드박스, 이해관계자 참여, 민첩 규제, 국제 협력 등의 권고가 글로벌 표준으로 자리 잡고 있음을 지적한다. 이를 통해 한국은 기술·산업 중심이 강하나, 윤리·신뢰 제도화가 부족하므로, 국가 차원 CDO 체계와 데이터 윤리 프레임워크 구축 필요한하며 이는 공공신뢰 확보와 사회적 수용성을 높여 지속 가능한 거버넌스를 실현할 수 있다. 선진국 사례처럼 중앙 조정과 민관 협력을 강화하면 데이터 이타주의와 인간 중심 설계가 한국의 디지털

정부혁신에 적용 가능하다는 점을 시사하였으며 압출 모델(정부·시민·인터랙티브)은 시민 참여와 정부 반응성을 강조하므로 한국의 MERS Map·T맵 사례를 확장해 포괄적 빅데이터 플랫폼을 개발하면 공공정책의 효과성을 제고할 수 있다는 점을 확인하였다. 마지막으로 OECD의 예측적 거버넌스처럼 가치·인텔리전스·참여를 통합하면, AI 불확실성에 대응하면서 혁신을 촉진할 수 있다는 점을 시사하였다.

성공요인과 실행전략에 대한 이론적 검토에서는 성공 요인으로서는 최고 의사결정자의 전략적 리더십과 전담 조직(CDO) 구축, 명확한 정책·법제 기반, 윤리·투명성·책임성 내재화, 데이터 품질·보안·기술 역량 확보, 체계적 모니터링·평가·지속 개선이 꼽혔으며 실행 전략에서는 제도적 기반 정비, 기술 기반 실행 메커니즘 구축, 참여 기반 운영 및 민관·지방 협력 강화, 데이터 책임성과 프라이버시 보호를 위한 제도·기술 결합, 글로벌 윤리 기준과의 정합성 확보로 요약된다.

마지막으로 사회적·윤리적 논의를 통해 먼저 한국은 인공지능 활용 과정에서 발생할 수 있는 윤리적 문제와 사회적 불평등을 최소화하기 위해 정책적 노력을 강화가 필요하다는 점을 시사하였다. 우선, AI 윤리기준에서 강조하는 투명성과 연대성의 원칙을 바탕으로 국민의 설명권과 개인정보 자기결정권을 확대할 필요가 있다. 특히 개인정보보호위원회가 2025년에 제안한 정책 방향을 토대로 AI 자동화 과정에서 발생할 수 있는 규제 공백을 보완한다면 프라이버시 침해 위험을 사전에 차단할 수 있을 것이다.

또한 디지털 격차 해소는 사회적 포용성을 높이는 핵심 과제로 보고있다. 고령층과 같은 취약계층을 대상으로 맞춤형 디지털 교육을 강화하고, 실제 생활 속에서 기술을 실험·적용하는 리빙랩(Living Lab) 방식을 도입한다면 접근성과 활용 격차를 줄이는 데 도움이 된다. 더불어 지방정부의 역량을 강화해 지역 간 불균형을 완화하는 정책이 병행되어야 한다. 이는 단순한 기술 보급을 넘어 사회 전체의 포용적 디지털 전환을 이끄는 기반이 될 수 있다.

국제적 흐름에 부합하는 규범 적용의 중요성을 강조하였다. GDPR과 OECD의 원칙을 국내 법제 및 정책과 연계한다면 초국가적 데이터 흐름에서 발생할 수 있는 문제에 효과적으로 대응할 수 있다. 이는 코로나 팬데믹과 같은 위기 상황에서 드러난 윤리적 실패를 예방하고, 데이터 정의를 실현하는 데 기여할 것이다.

마지막으로 투명성 중심의 거버넌스 체계 구축이 필요하다. 알고리즘 등록제와 영향 평

가 제도를 도입해 인공지능의 비가시성을 해소하고, 시민이 직접 참여할 수 있는 의사결정 구조를 제도화함으로써 신뢰를 확보해야 한다. 이 과정에서 EU AI법에서 제시한 고지 의무 제도는 유용한 참고 사례가 될 수 있다.

2. 이론적 고찰 시사점

이론적 고찰을 통해 데이터 거버넌스가 전통적인 조직 내 데이터 관리 차원을 넘어 AI 시대의 핵심적 사회제도로 진화하고 있음을 보여준다. 특히 AI와 디지털 전환이 데이터의 편향성, 알고리즘 불투명성, 책임소재 불분명성 등 새로운 도전과제를 제기하면서 데이터 거버넌스는 기술적 관리를 넘어 윤리적 책임성과 사회적 신뢰성을 확보하는 포괄적 프레임워크로 재정의되어야 한다는 점이 핵심이다. 성공적인 데이터 거버넌스 구현을 위해서는 최고의사결정자의 전략적 리더십, 명확한 법제도적 기반, 생애주기 기반 품질관리 체계, 시민참여형 운영구조가 통합적으로 작동해야 하며, 동시에 프라이버시 보호와 데이터 활용 간의 윤리적 균형, 디지털 격차 해소, 글로벌 규범과의 연계를 통해 사회적 수용성과 지속가능성을 확보해야 한다는 시사점을 제공한다.

제3장 공공기관 데이터 거버넌스 실제에 대한 탐색

본 연구는 공공부문 데이터 거버넌스의 실태를 진단하고 정책적 실행전략의 타당성을 검증하기 위하여 전문가 및 이해관계자 인터뷰를 실시하였다. 인터뷰는 공공부문 데이터의 수집·관리·활용 과정 전반에서 발생하는 제도적·조직적·기술적 쟁점을 다각도로 파악하고, 현장 경험에 기반한 개선 방향을 도출하는 것을 목적으로 진행되었다.

조사는 2025년 7월 22일부터 8월 25일까지 실시되었으며 총 26명의 전문가 및 이해관계자가 참여하였다. 인터뷰 대상자는 중앙부처, 지방정부, 국책연구기관, 대학, 공기업, 산업계 등 폭넓은 기관에서 선별되었으며, 각 기관의 직위와 담당 업무에 따라 데이터 정책·관리·활용 관련 경험을 가진 실무자와 연구자를 균형있게 포함하였다.

인터뷰는 연구진이 설계한 질문지를 기반으로 한 서면 자가응답 방식으로 이루어졌으며, 일부 응답자는 이메일 또는 전화·온라인 화상회의, 대면 인터뷰 등을 통해 추가 질의가 병행되었다. 질문지는 공공부문 데이터 거버넌스의 다층적 구조를 반영하여 ▲ 개인(직무와 데이터의 관련성), ▲ 데이터(전주기 관리·이슈 등), ▲ 기술과 시스템(AI·플랫폼 활용 등), ▲ 조직과 관리(리더십·조직문화·인력 관리 등), ▲ 법률과 제도(정책·규제 등) 총 5개 영역, 50문항 내외로 구성되었다.

모든 인터뷰는 익명으로 처리되었으며, 연구윤리 준수를 위해 응답 내용은 질적분석을 위해서만 활용되었다. 분석은 문항별 내용 코딩과 핵심키워드 도출, 공통 패턴 및 차이점 비교의 세 단계로 이루어졌으며, 반복적으로 제시된 의견은 제도 개선과 프레임워크 설계 시 핵심 근거로 반영되었다. 본 인터뷰는 공공부문 데이터 거버넌스가 실제 행정·정책 현장에서 작동하는 메커니즘을 실질적으로 점검하고, 향후 데이터·AI 융합 거버넌스 체계의 정책적 타당성과 현장 적용성을 강화하기 위한 근거를 마련하는 과정이었다.

제1절 데이터

1. 데이터 관리와 활용 현황

가. 기관별 운영 방식과 연계 수준

공공기관의 데이터 관리와 연계 수준은 기관별로 큰 차이를 보이며, 이는 데이터 거버넌스의 정착을 어렵게 하는 핵심 요인으로 지적되어 왔다(조세현·차세영, 2022; OECD, 2022). 기존 연구에서도 기관 간 이질적인 관리체계와 표준 부재로 인한 낮은 상호운용성이 가장 큰 문제로 꼽힌다(오윤경·차남준, 2023; OECD, 2022; Harrison et al., 2019). 이번 인터뷰 결과에서도 이러한 지적을 구체적으로 확인할 수 있었다.

“현재 국회에서는 의원실을 포함하여 각 상임위원회, 입법조사처, 예산정책처가 행정부처 및 공공기관에 대한 자료요구권을 바탕으로 행정자료를 취득합니다. 그러나, 그렇게 취득된 데이터가 서로 간에 공유되지는 않습니다. 의원실 상호 간에는 자료 공유가 기본적으로 안되고 있는 것으로 알고 있고, 이는 국회 공무원이 근무하는 상임위, 조사처, 예정처도 마찬가지입니다. 특히 각 실국 및 팀 간에도 데이터 교류는 거의 되지 않고 있어서, 사실상 특정 업무의 담당자가 후임자에게 이관하는 데이터가 유일하다고 보면 될 것 같습니다. 그러나, 전임자로부터 받은 데이터도 체계적으로 정리되지 않아서 후임자가 새로 데이터를 수집하는 경우가 많습니다.”

- 000(입법기관)

“현재 정책 연구를 수행하면서 가장 반복적으로 경험하는 장애요인은 기관 간 데이터 연계 체계의 부재입니다. 이미 존재하는 데이터가 있다고 할지라도 각 기관에서 공유하지 않을 경우, 데이터 활용이 매우 어렵고, 공유한다고 할지라도 매우 제한적입니다. 또한, 코드 매칭 작업을 수작업으로 반복해야 하고, 이 과정에서 발생하는 누락이나 오류는 결과 해석에도 영향을 미칠 수 있습니다. 따라서, 정보 간의 데이터 연계를 제도적으로 설계하고, 이를 기반으로 한 정책 모니터링 체계를 구축하는 것이 필요하다고 생각합니다.”

- 000(B국책연구원)

“가장 큰 애로사항은 이질적인 데이터 간의 통합과 해석의 일관성 유지입니다. 특히 서로 다른 출처에서 수집한 데이터를 연계할 때, 데이터 포맷, 시간 단위, 분류 체계 등이 상이하여 정합성 확보에 많은 시간과 인력이 소요됩니다. 또한, 다른 분야의 연구자들과 융합연구를 하는 경우, 주로 사용하는 라이브러리의 차이나 전처리 방식의 차이가 있다면 같은 주제를 연구하고 있음에도 기대하는 결과물과 상이한 결과를 받기도 합니다. 그래서 이에 대한 대응을 위해 공통 데이터셋 저장소 구축하고, 버전이나 메타데이터 기록을 지속적으로 공유하는 방안을 활용하고 있습니다.”

- 000(A대학교)

“기관 데이터는 여러 부서에 분산되어 관리되고 있습니다. 정보 시스템 관리 부서가 별도로 존재하지만, 관람객 데이터는 전시 운영부서에서, 전시물 관리 데이터는 전시물 유지관리 부서에서 관리하고 있습니다. 또한 각 부서별로 부서 프로그램에 대한 관람 데이터를 수집하여 관리하고 있습니다. 연계가 잘 이루어지지 않고, 데이터에 대한 공유도 잘 이루어지지 않고 있습니다. 수집된 데이터가 타 부서에도 잘 공유되어서 수집된 데이터를 활용하여 사업에 반영될 수 있도록 할 필요가 있습니다.”

- 000(A책임운영기관)

이와 같이 현장의 목소리는 제도적 기반의 미비와 호환성 부족을 공통된 장애요인으로 지적하고 있다. 특히 입법기관 전문가의 인터뷰에서는 국회 내부의 부서 간 데이터 단절 문제가 드러났으며, 정책연구원 연구자는 기관 간 연계 부재로 인한 수작업 반복과 오류 발생을 강조하였다. 또한 학계 전문가의 인터뷰에서는 데이터 포맷 및 분류 체계의 상이함과 전처리 방식의 차이가 일관성 확보를 어렵게 하는 핵심 문제로 제기되었다. 이러한 사례들은 문헌에서 지적된 기관 간 이질적 관리체계와 낮은 상호운용성의 한계를 구체적으로 보여준다(조세현·차세영, 2022; OECD, 2022; Malomo & Sena, 2017). 더불어 데이터 관리와 활용의 지속 가능성을 위협하는 요인으로, 인력 부족 문제 역시 반복적으로 언급되었다.

“행정 인력의 업무 과중으로 인해 데이터 입력의 정확성과 시의성 확보가 어려운 경우가 많습니다. 또한 담당자가 데이터의 현행화를 주기적으로 해줘야 하는 관리의 어려움이 있습니다. 예를 들어 물품 재물조사 시 실물과 시스템 데이터 간 차이가 발

생하고, 이를 맞추기 위해 별도의 점검 및 확인 작업이 반복되는 경우가 있습니다. 인력 지원과 시스템 자동화 기능 확대가 필요합니다.”

- 000(A기초자치단체)

“데이터를 관리할 인력 자원을 구하기가 어렵습니다. 대부분의 인력은 계약직 신분 이어서 지방에서는 지원자가 없습니다. 무기계약직, 또는 특별채용 공무원으로 임명 하지 않는 이상 실제로는 전문가를 구하기도 어렵습니다.”

- 000(E대학교)

“데이터를 관리하는 조직과 전문 인력이 필요합니다. 우리 기관에는 정보담당 부서가 있지만, 주로 홈페이지와 서버를 관리하는 수준이고, 기관에서 발생하는 각종 데이터를 수집하고 관리하는 부서는 아닙니다. 데이터만을 관리하는 조직이 현재 없어 체계적인 관리가 되고 있지 않습니다.”

- 000(A책임운영기관)

다수 응답자는 기관 간 협력의 부재를 데이터 거버넌스 실행을 저해하는 핵심 요인으로 지적하였다. 이는 기술적 연계 부족에 그치지 않고, 정보 공유와 조정 메커니즘의 부재, 그리고 상호 책임성 확보의 미비 등 구조적 한계로 드러났다.

“현재 각 실·국·산하기관의 업무시스템과 BI 대시보드가 분절적으로 운영되고 있습니다. 따라서, 데이터 사전·메타데이터 표준, 품질 규칙, 연계용 공통키가 미흡하고, 외부기관(교육청·경찰·법원·민간위탁기관 등)과는 법적 근거·표준계약 부재로 정기 연계가 어렵습니다.”

- 000(A지방의회)

“주로 애로사항은 유관기관 간 데이터 연계 부족(특히 경찰과의 공동대응)입니다. 이에 대응 방안으로는 소방과 경찰 간 실시간 정보 공유시스템의 구축이 필요하다고 생각합니다. 현재는 주기적으로 소방 경찰 간의 업무 협력관계에 관하여 회의를 하고 있지만 현장에서는 아직도 변한 것이 없습니다.”

- 000(A소방서)

“데이터 거버넌스라고 하면 조직 간의 혹은 정부와 민간 간의 공유 또는 협력을 의미하는 것이라고 생각하는데, 우리나라의 경우는 정부 조직 간의 거버넌스도 어려운 실정입니다. 강력한 리더십을 발휘하지 않는다면 거버넌스는 구축해도 아무 소용이 없을 것으로 봅니다.”

- 000(E대학교)

“데이터 품질 관리 측면에서 가장 큰 어려움은 실적 데이터의 정확성과 일관성 부족입니다. 특히 데이터 분석 시, 00청과 00부의 실적이 불일치하는 사례가 빈번하게 발생합니다. 향후에는 데이터 생성 단계부터 기관 간 실시간 데이터 연계 체계 구축이 함께 이루어져야 할 것으로 생각합니다.”

- 000(B국책연구원)

결국 이러한 결과는 문헌에서 지적된 것처럼(Malomo & Sena, 2017) 내부 관리 역량은 점진적으로 강화되고 있으나, 기관 간 연계와 협력 수준은 여전히 낮으며, 제도적·기술적 한계로 인해 구조적 제약이 심각하다는 사실을 뒷받침한다. 나아가, 지역 간 격차도 중요한 쟁점으로 드러났다. 특히 수도권과 지방 간의 데이터 관리 역량 및 인프라 수준의 차이는 데이터 거버넌스의 균형적 정착을 어렵게 하는 구조적 한계로 반복적으로 제기되어 왔다(김지현·이영주, 2021). 이는 지역 간 데이터 활용 역량의 불균형을 심화시키고, 궁극적으로 국가 차원의 통합적 데이터 거버넌스 체계 구현을 저해하는 요인으로 작용한다(Harrison et al., 2019; 오윤경·차남준, 2023).

“기관에서 중앙으로 데이터를 보낸 후, 미회신되는 경우가 많아 지방의 정책건의 데이터가 소실되는 경우가 있다는 것이 반복적인 장애요인입니다. 이를 해결하기 위해서는 중앙과 지방의 대등한 관계 설정과 정책건의의 내실화 및 중앙의 책임감있는 대응시스템 등이 필요할 것으로 보여집니다.”

- 000(A협의회)

“기관에서 데이터의 중앙관리를 위한 조직을 만든 적이 있습니다. 그런데, 지방에서는 그 업무를 수행할 만한 석사학위 인력을 찾기도 매우 어렵습니다.”

- 000(E대학교)

이와 같은 인터뷰는 중앙 의존적 구조와 지방의 자율성 부족이라는 구조적 문제를 잘 보여준다. 전자는 중앙정부의 회신 부재로 인해 지방의 정책 건의 데이터가 반영되지 못하는 사례를 드러내고 있으며, 후자는 지방 차원의 전문인력 부족으로 인해 중앙에서 마련한 관리체계가 실질적으로 작동하기 어려운 한계를 보여준다. 결국 중앙과 지방 간의 불균형한 권한 배분과 역량 격차가 데이터 거버넌스 실행의 핵심 장애요인으로 반복적으로 확인되고 있다.

나. 데이터 생애주기 과정과 특징

데이터의 생성에서 활용까지 이어지는 과정은 생성·수집, 제공·가공, 분석·활용의 선형적 생애주기로 설명된다(경제·인문사회연구회, 2024). 그러나 실제 현장에서 이 과정은 단계별로 단절적이고 비효율적으로 운영되고 있으며, 데이터 흐름이 원활히 이어지지 못하는 사례가 많았다. 먼저, 데이터 생성 단계에서부터 기관별로 문제점들이 나타났다.

“기관마다 데이터의 생성 기준이 다르기 때문에 데이터 체계가 미비하여 데이터를 공유받는데 소요되는 비용이 큼니다. 향후에는 주요 유관기관 간의 시스템 연계 및 정책연구기관 차원의 데이터 수집 및 관리 프로토콜을 마련할 필요가 있다고 생각합니다.”

- 000(B국책연구원)

“탄약과 관련된 군수 데이터의 경우 생산, 성능 및 수락시험, 야전 배치, 저장, 신뢰성시험, 불출, 폐기 등의 단계에서 데이터들이 생성되는데 모든 데이터가 수치화 되어 있지 않고 각 단계마다 수기, 오기입 등 휴먼에러가 빈번합니다. 탄약의 수명을 예측하여 장기 보관과 교탄 처리, 폐기 등의 의사결정을 위한 데이터 분석을 위해서는 다양하고 복잡한 탄약 수명 주기상의 데이터의 품질이 무엇보다 중요합니다.”

- 000(A연합회)

“어떤 경로로 어떤 목적의 데이터가 생성되는지를 판단하고 그에 맞는 생명주기에 맞는 Governance 정책을 만들 수가 있습니다. 특히, 개발하는 서비스의 목적에 맞는 다양한 경우의 수를 반영할 수 있는 데이터가 필요한데 대부분의 경우 다양하지 못한 품질이 떨어지는 데이터로 서비스를 개발하고, 추후 예상치 못한 데이터 패턴으로 서비스의 품질이 떨어지는 경우가 많이 발생합니다. 그래서 결국 다양한 경우의 수를

반영할 수 있는 데이터가 생성될 수 있도록 서비스를 설계해야 합니다.”

- 000(B대학교)

“개별 데이터 생성·수집 관련 기준이 제각각이기 때문에 데이터 연계 및 통합을 통한 효과적인 성과 창출이 어렵습니다. 예를 들어, 건설 분야에서 계획 도면, 건설 도면, 준공 도면의 구축 기준이 각각 다르기 때문에 건설 도면에 대한 생애주기가 일관적으로 정립되지 않고 있습니다.”

- 000(B공공기관)

이는 문헌에서 강조한 데이터 표준화와 메타데이터 관리의 부재를 실증적으로 보여준다(Koltay, 2016).

수집 단계에서는 데이터 품질 저하 문제가 자주 언급되었다. 이는 데이터가 체계적으로 정제되지 못하고 축적되는 구조적 한계를 반영하며, 장기적으로는 데이터 활용 가능성과 신뢰성을 저해하는 요인으로 작용한다(OECD, 2022; Koltay, 2016).

“데이터의 형식이 통일되어 있지 않아 업로드하는 직원들에 따라 천차만별이고, 이는 다른 직원들에게 데이터 검색 및 활용상의 어려움을 야기합니다. 이로 인해, 직원들이 상기 시스템들을 적극적으로 활용하는 경우가 많지는 않습니다. 또한, 데이터 품질 향상을 위해 업로드하는 데이터의 형식 통일을 강제할 경우 데이터 공유에 더 소극적이 되는 사례들이 있었습니다.”

- 000(입법기관)

“데이터가 저장된 형식이 모두 통일되어 있지 않아서 이를 각각 처리하여 병합해야 한다는 문제가 있습니다. 그래서, 정부나 공공의 데이터는 하나의 기관이나 공개된 플랫폼에서 관리한다면 연구자들의 편의성이나 상호운용성이 높아질 것이라 생각합니다.”

- 000(A대학교)

“체계적인 정보화 시스템과 업무가 잘 연결된 경우라면 개별 시스템과 통합시스템 연계를 통해 잘 관리되고 있으나, 그렇지 않은 업무나 기관의 경우라면 개별 업무 시스템 및 디바이스 로컬에 비정형(메타데이터 관리 안 됨) 형태로 쌓여있는 경우가 많습니다. 00부와 00청 산하 00원의 경우는 비정형 문서와 문서 내 정량적 정보를 엑

셀에 별도로 저장하여 관리되고 있는 경우가 있으며 담당자가 바뀌거나 업무가 중단되었다가 다시 관리되는 경우, 데이터 정리의 기준, 단위, 명칭 등의 일관성이 없이 관리됩니다. OO부의 경우, 관행적으로 측정, 관찰, 관리 등 정보를 수기로 작성하는 곳이 많습니다. 데이터에 개념이 약하고 업무 처리에 집중되어 있어 자료를 정리하거나 데이터화하여 보관하는 등의 업무체계가 잡혀있지 않은데, 이유는 업무규정 등에 관련 내용이 정해져 있지 않을 뿐만 아니라 필요성을 느끼지 못하고 있기 때문입니다.”

- OOO(A연합회)

“데이터의 가치를 효과적으로 구현하기 위해서는 공공-민간 데이터, 정형-비정형 데이터 등을 장기적 관점에서 지속적으로 축적하는 과정이 필요하지만, 단기적 관점에서의 현안 대응 및 즉각적인 성과 창출이 중요하기 때문에 장기적 관점에서의 데이터 축적이 수행되기 어렵습니다.”

- OOO(B공공기관)

“데이터를 활용하고 싶어도 데이터를 수집하지 않는 경우가 많습니다. 데이터를 수집할 수 있는 틀을 만들어 사업마다 관람객 반응에 대한 데이터를 수집한다면, 이를 분석하여 보다 나은 사업으로 개선할 수 있으리라 생각합니다.”

- OOO(A책임운영기관)

이처럼 저장소가 존재하더라도 검색 기능의 한계, 체계적 분류 부재, 관리 인력 부족 등으로 인해 데이터 접근성과 활용성이 현저히 낮아지는 사례가 확인되었다. 이러한 상황은 결국 데이터가 단순 보관에 머무르며, 공공부문 데이터 거버넌스의 핵심 목표인 ‘활용 중심의 데이터 관리’가 실질적으로 달성되지 못하는 구조적 한계를 드러낸다(Harrison et al., 2019).

제공 단계에서는 데이터 접근권한 제한과 비체계적 제공 방식으로 인한 신속성·효율성 저하가 문제로 나타났으며, 가공 단계에서는 표준화 부재와 수작업 결합, 자동화 미비로 인해 재현성과 생산성이 저하되는 한계가 두드러졌다.

“공급망 정책 관련, 데이터 제공 단계에서 정비가 필요하다고 판단됩니다. 관세 데이터는 공급망 관리의 핵심 요소임에도 불구하고, 관세 데이터베이스에 대한 접근권이 타 부처에는 공유되지 않습니다. 대신 기획재정부 등에서 요청할 때마다 관세청

이 해당 데이터를 새로 추출하여 제공하는 방식으로 운영되고 있어, 데이터 제공에 며칠이 소요되는 등 신속성이 떨어지고 효율적인 업무 처리를 제약하는 상황들이 발생하였습니다.”

- 000(A대학교)

“데이터 수집 이후 가공 단계에서는 다양한 출처에서 확보한 데이터를 분석 가능하도록 정렬하고, 정형·비정형 데이터를 융합하거나 시간축을 맞추는 등의 복잡한 처리가 요구됩니다. 특히 기업 보고서와 같은 비정형 텍스트 데이터는 주관적 서술이 많고, 시점이 불분명한 경우도 있어 정형 데이터와의 연계가 쉽지 않습니다. 이러한 데이터를 규격화된 포맷으로 전환하거나, 분석 목적에 맞게 통합하는 과정이 체계화되어 있지 않으면 연구자마다 가공 기준이 달라지고, 이는 연구 결과의 재현성과 비교 가능성을 저해하게 됩니다. 또한 데이터 가공은 반복적이고 시간이 많이 소요되는 작업이지만, 자동화된 도구나 가이드라인이 부족한 경우가 많아 연구 생산성에도 영향을 미칩니다. 따라서 연구자들마다 데이터 가공 기준이 상이해 주관성이 개입될 여지가 크므로, 이를 통일할 수 있는 표준화된 기준과 절차를 마련하는 것이 중요합니다.”

- 000(A대학교)

“학교에서 원자료는 풍부하지만, 이를 의사결정에 적합한 형태로 변환하는 기능이 제한적입니다. 예를 들어, 예산집행 데이터와 시설 점검 결과를 통합 분석하려면 여러 시스템에서 자료를 내려받아 수작업으로 결합해야 합니다. AI 기반 자동 가공·통합 분석 툴이 도입된다면 행정 효율성이 크게 높아질 것입니다.”

- 000(A기초자치단체)

이처럼 제공 단계에서는 접근권한 제한과 비체계적 제공 방식으로 인해 데이터 활용의 신속성과 효율성이 저하되는 문제가 반복적으로 지적되었으며, 가공 단계에서는 표준화 부재, 수작업 결합, 자동화 미비 등으로 인해 연구 결과의 재현성과 비교 가능성이 낮아지고 생산성 또한 크게 떨어지는 한계가 확인되었다. 결국 이러한 단계별 제약은 데이터 활용의 질적 수준을 저해하는 구조적 요인으로 작용하며, 이를 해결하기 위해서는 부처 간 공통 분류체계 구축, 메타데이터 표준화, 그리고 가공 과정의 자동화 체계 마련이 시급한 과제로 제기된다.

분석 단계는 정책 방향성과 의사결정의 정합성을 담보하기 위해 매우 중요한 과정임에도 불구하고, 전문성 부족으로 인한 진입장벽, 그리고 개인 역량에 대한 과도한 의존이라는 한계를 보여주었다.

“분석이 가장 중요하다고 생각합니다. 분석을 객관적으로 하지 않으면 정책의 방향성이 왜곡될 수 있기 때문입니다.”

- 000(B중앙부처)

“기관 임직원의 분석에 대한 수요는 있으나 분석에 대한 진입장벽을 느끼는 경우가 많습니다.”

- 000(B공단)

“국회의 관점에서 데이터의 생애주기 관리에서 가장 중요한 것은 수집과 분석입니다. 행정부처 및 공공기관에 자료요구를 할 경우에는 완벽히 구체화 되지는 않았지만 최소한 특정한 목적을 가지고 합니다. 따라서 그 목적에 부합하는 자료를 수집하여, 수집된 자료를 바탕으로 유의미한 결과를 이끌어내야 하기 때문입니다. 국회 상임위, 입법조사처, 예산정책처의 공무원들은 특정 분야를 오직 한 명이 전담하는 경우가 많기 때문에 개인의 역량에 따라 분석 결과가 커다란 차이를 보이는 경우가 많습니다.”

- 000(입법기관)

위와 같이 분석 단계에서는 정책의 객관성과 정합성을 확보하기 위한 핵심 과정임에도 불구하고 여전히 여러 한계를 내포하고 있음이 확인된다. 특히 전문 인력과 체계적 지원의 부족은 분석 역량을 제약하고, 개인의 경험과 주관적 해석에 과도하게 의존하는 구조를 고착화시켜 정책 결과의 신뢰성을 저하시킬 위험이 크다. 따라서 분석 단계의 효과성을 제고하기 위해서는 전문성 강화를 위한 교육·훈련, 기관 차원의 분석 지원체계 구축, 그리고 데이터 해석의 표준화 방안 마련이 시급한 과제로 제기된다.

마지막으로, 활용 단계에서는 데이터 공유의 부재, 제도적 장치의 미비, 그리고 국민 대상 활용의 한계가 핵심적인 문제로 지적되었다. 특히 기관 간 데이터가 축적·연계되지 못하고 조직 내부에서도 체계적인 인수인계가 부재하여, 데이터가 반복적으로 수집되거나 효과적으로 활용되지 못하는 구조적 문제가 드러났다.

“데이터 생성부터 가공까지는 지속적인 주기를 통해서 반복적으로 이루어지지만, 실제로 활용되는 경우는 소수에 해당합니다. 따라서, 다양한 정책건의를 통한 비정형데이터가 실제로 잘 활용될 수 있도록 하는 제도개선과 정교화 장치가 필요하지 않을까 생각합니다.”

- 000(A협의회)

“데이터 활용 단계가 가장 중요하다고 생각합니다. 국회 내 조직 간 데이터 공유는 거의 이루어지고 있지 않습니다. 또한, 조직 내 데이터 인수인계도 체계적으로 이루어지고 있지 않아서 다시 처음부터 데이터 수집을 하는 경우도 많습니다. 이는 체계적인 데이터 축적을 통한 효과적인 활용을 저해하는 가장 큰 요인으로 보입니다.”

- 000(입법기관)

“활용 단계의 정비가 필요하다고 생각합니다. 기관의 전반적인 업무 특성상 국민을 대상으로 활용이 가능한 데이터가 한정적이기 때문입니다.”

- 000(C공단)

이는 문헌에서 지적하는 바와 같이, 데이터와 정책 간의 실질적 연계 및 활용 부족과 일치한다(OECD, 2022). 특히 데이터가 정책 결정의 근거자료로 체계적으로 반영되지 못하고, 기관 간 단절적 운영으로 인해 데이터가 반복적으로 수집되거나 단순 보관에 머무는 현실은 공공부문 데이터 거버넌스가 지향하는 ‘활용 중심의 관리’ 목표와 괴리됨을 보여준다(Harrison et al., 2019). 결국 이는 데이터 기반 정책결정의 신뢰성과 효과성을 저해하는 핵심 요인으로 작용한다.

2. 주요 문제와 과제

가. 기관별 애로사항과 품질관리 문제

공공부문 데이터 거버넌스의 또 다른 핵심 과제는 기관별로 반복적으로 발생하는 애로사항과 이를 악화시키는 품질관리 문제이다. 문헌에서는 데이터 품질을 정확성, 일관성, 완전성, 적시성 등의 속성으로 정의하며, 이러한 속성이 확보되지 못하면 데이터 기반 행

정은 한계에 부딪힐 수밖에 없다고 강조한다(Loshin, 2011; OECD, 2022). 그러나 인터뷰에서는 이러한 원칙이 충분히 이행되지 못하는 사례들이 다수 제시되었다.

1) 기관 간 공유 부재와 데이터 단절

국회의 경우, 데이터는 행정부처나 공공기관에 자료요구를 통해 확보되지만, 내부 공유는 거의 이루어지지 않고 있었다.

“현재 국회에서는 의원실을 포함하여 각 상임위원회, 입법조사처, 예산정책처가 행정부처 및 공공기관에 대한 자료요구권을 바탕으로 행정자료를 취득합니다. 그러나, 그렇게 취득된 데이터가 서로 간에 공유되지는 않습니다. 의원실 상호 간에는 자료 공유가 기본적으로 안되고 있는 것으로 알고 있고, 이는 국회 공무원이 근무하는 상임위, 조사처, 예정처도 마찬가지입니다. 특히 각 실국 및 팀 간에도 데이터 교류는 거의 되지 않고 있어서, 사실상 특정 업무의 담당자가 후임자에게 이관하는 데이터가 유일하다고 보면 될 것 같습니다. 그러나, 전임자로부터 받은 데이터도 체계적으로 정리되지 않아서 후임자가 새로 데이터를 수집하는 경우가 많습니다.”

- 000(입법기관)

이처럼 전임자-후임자 인수인계 외에는 데이터의 축적·공유 체계가 존재하지 않아, 동일한 자료를 후임자가 새로 수집하는 비효율이 반복되고 있었다.

2) 데이터의 부서별·개인별 관리와 중복 수집

연구기관의 경우에도 데이터는 연구자 개인이 과제별로 개별 수집·관리하는 방식이 일반적이었다.

“연구자가 연구 분야에 따라 필요 데이터를 개별적으로 수집·관리·분석함에 따라 기관 차원에서는 어떠한 데이터를 보유하고 있고, 다른 연구자가 이를 활용할 수 있도록 공유하는 체계가 없습니다. 따라서 유사한 데이터를 중복으로 수집할 수 있는 비효율성이 발생할 수 있다고 생각합니다.”

- 000(A국책연구원)

즉, 기관 차원의 통합 관리 부재로 인해 동일 데이터가 반복 수집되는 구조적 문제가 지적되었다

3) 데이터의 표준화·품질관리 미비

시스템은 존재하지만, 데이터 표준화와 품질 검증이 미비하고, 책임소재 불명확 등 구조적 한계의 문제가 반복적으로 발생하였다.

“많은 공공부문 기관들에서 공공 의사결정 지원 시스템으로서 디지털 트윈 시스템을 활용하고 있지만, 이러한 시스템에 활용되는 데이터 및 시물레이션 모델의 품질(예: 정확성, 신뢰성, 일관성)의 검증이 어렵습니다. 이러한 데이터 및 시물레이션 모델의 품질에 문제가 발생하면 디지털 트윈 시스템의 산출 결과에도 문제가 발생하고 이는 잘못된 의사결정 지원을 초래할 수 있으므로, 이러한 데이터 및 시물레이션 모델의 품질을 사전에 확보하고 사후에 검증할 필요가 있습니다.”

- 000(B공공기관)

“지난 10년 이상 공공데이터 개방 및 활용 정책이 지속적으로 추진되어 왔으나, 데이터 품질 부족, 부처 간 칸막이, 표준 부재, 개인정보 이슈 및 불명확한 책임 소재 등 구조적 한계는 여전히 반복되고 있습니다. 그간 데이터 표준화나 CDO 지정 등 다각적인 대응 전략이 시도되었음에도 현장의 근본적인 문제는 해결되지 못했습니다.

이는 기존의 데이터 거버넌스가 공급자 중심의 형식적이고 분절적인 부처 단위 관리 체계에 머물러 있었기 때문입니다. 데이터 경제와 AI 시대를 맞이하여 과거의 방식을 답습해서는 당면한 문제를 해결할 수 없습니다. 따라서 기존의 틀을 과감히 탈피하여 역할·권한·기술 기반을 전면적으로 재설계하는 새로운 거버넌스 패러다임이 필수적이라고 생각합니다”

- 000(F대학교)

“DBMS에 연계된 사용자 시스템(나이스, K-에듀파인 등)을 통해 사용자의 DB의 CRUD로 관리되고 있으며 사용자의 입장에서 데이터에 접근할 때는 네트워크를 통해서 접근하기 때문에 서버에 사용자의 접근이 물리는 시기(학년도 말 정산시기), 시스템이 서비스를 시작하여 초반 시스템이 불안정한 시기에 데이터 무결성을 유지하기 어려울 때 CRUD를 정상적으로 수행하기 어렵기 때문에 문제가 발생합니다.”

- 000(A기초자치단체)

또한 학교 단위에서는 프로그램별로 별도 관리가 이뤄져 중복 입력과 기준 불일치 문제가 빈번히 발생하였다.

“현재 학교와 관련된 데이터는 돌봄, 방과후, 인력 배치, 프로그램 운영 등 각 분야별로 별도로 관리되고 있습니다. 시도교육청, 교육지원청, 학교 단위로 시스템이 분산돼 있어 유기적 연계가 부족한 편입니다. 특히 타 부서나 외부기관(지자체, 마을교육공동체 등)과의 연계는 수작업으로 이루어져 효율성이 떨어지고, 중복 입력이나 기준 불일치 문제도 발생하고 있습니다.”

- 000(A기초자치단체)

4) 비정형 데이터 관리의 어려움

현장 중심 기관에서는 비정형 데이터의 처리와 연계가 큰 문제였다. 소방기관의 사례는 이를 잘 보여준다.

“출동대원이 작성하는 데이터(일지)는 소방청 통합 시스템으로 자동 전송되어 중앙에서 저장되지만, 외부와의 연계는 현재 거의 없는 것으로 알고 있습니다. 가령 요구조자의 신상 및 출동 정보와 같이 유관기관에서 수집된 데이터는 연계 시스템이 없어서 공유받기가 어렵습니다. 이로 인해서 같은 출동임에도 불구하고 상황이 종료된지도 모른 채 현장 활동을 종료하지 못한 적도 있습니다. 또한, 구급대원의 경우 환자를 이송할 때 매번 병원에 연락을 하여 이송 가능 여부를 확인해야만 요구조자를 병원에 이송할 수 있습니다.”

- 000(A소방서)

이는 현장 데이터가 중앙에 모이더라도 실시간 연계·활용이 어려워 대응 지연을 초래하는 한계를 드러낸다.

5) 관리 주체와 책임성의 불분명

일부 대학·연구기관에서는 데이터 관리 주체조차 명확히 설정되어 있지 않았다.

“우리 조직에서 데이터를 관리하는 주체가 누구인지 정확히 알지 못합니다. 각 학과 및 부서들에서 데이터가 어떻게 수집되는지, 그리고 그 데이터를 어떻게 활용하고 있는지도 알지 못합니다. 데이터에 대한 관심과 중요성을 모르고 있는 것 같기도 합

니다.”

- OOO(E대학교)

이는 데이터 거버넌스의 기본 전제인 관리 책임성(accountability)의 부재를 단적으로 보여준다.

나. 대응 노력과 개선 필요성

기관별 인터뷰를 종합하면, 여러 제약에도 불구하고 데이터 관리·연계 문제를 개선하려는 노력이 일부 나타나고 있으며, 동시에 향후 제도적·기술적 보완의 필요성이 강하게 제기되고 있었다.

1) 내부 전담 조직 신설과 통합 관리 시도

일부 중앙부처는 데이터 관리의 비효율을 개선하기 위해 전담 조직을 설치하거나 부서 차원의 통합 관리를 추진하였다.

“2024년 이전에는 각 부서별로 데이터가 관리되고 이들 간 연계 체계는 부재하였으나, 2024년 이후 기획재정부 부총리 직속 ‘리서치팀’을 신설하면서 기관 차원에서 데이터 관리·공유를 추진해오고 있습니다.”

- OOO(A대학교)

이러한 사례는 부서별 관리에서 기관 차원의 통합 관리로 전환하려는 움직임을 보여주며, 향후 다른 기관으로 확산될 필요성이 있다.

2) 표준화·메타데이터 관리 강화 필요성

데이터 표준 부재 문제를 극복하기 위해 일부 기관은 자체적으로 버전 관리나 포맷 공유를 통해 정합성을 유지하려는 노력을 기울이고 있었다.

“데이터 간 연계성과 정합성을 유지하기 위해 버전 관리와 메타데이터 기록을 통해 체계적으로 관리하고 있습니다. 다만, 연구자 간 데이터 버전 관리 방식이 통일되지 않거나, 메타데이터 작성이 일관되지 않으면 데이터의 추적성과 재현성에 어려움이 발생할 수 있어 주관리자가 저인 경우 미리 사전에 포맷을 공유하여 관리하고 있습니다.”

- OOO(A대학교)

이는 메타데이터 표준화와 품질관리 프로토콜 확립이 시급함을 보여준다.

3) 시스템 연계 및 협력 메커니즘 필요

현장 부처와 연구기관은 수작업 기반 관리의 비효율성을 지적하며, 기관 간 시스템 연계 및 데이터 공유 프로토콜 마련을 요구했다.

“기관마다 데이터의 생성 기준이 다르기 때문에 데이터 체계가 미비하여 데이터를 공유받는데 소요되는 비용이 큼니다. 향후에는 주요 유관기관 간의 시스템 연계 및 정책연구기관 차원의 데이터 수집 및 관리 프로토콜을 마련할 필요가 있다고 생각합니다.”

- OOO(B국책연구원)

이는 데이터 흐름을 제도적으로 정비하고, 기관 간 협력 구조를 제도화해야 한다는 시사점을 제공한다.

4) 비정형 데이터의 체계적 관리 필요

현장에서 발생하는 비정형 데이터를 보다 효과적으로 활용하기 위해 자동화된 연계·정제 시스템 구축이 필요하다는 지적도 있었다. 소방 현장의 사례에서처럼, 단순 저장에 머무르지 않고 실시간 연계와 상호 활용 체계가 마련되어야 한다.

“출동대원들은 출동 시 출동지령단말기로 신고에 대한 정보를 전달받습니다. 그러나 정보가 제대로 제공되지 않는다면 출동에 큰 차질이 발생합니다.”

- OOO(A소방서)

5) 보안과 활용 간 균형 확보

데이터 보안이 강화되면서 오히려 활용이 위축되는 문제를 해소하기 위해, 보안·접근 권한·활용 가이드라인의 균형적 설계가 필요하다.

“학교에서 활용하는 데이터는 대부분의 경우 개인 중요정보가 빠진 데이터를 주로 활용하고, 기업에서는 모네타이제이션(Monetization) 할 수 있는 방안을 함께 고민하기에 정규화, 표준화를 강조하고, 보안상 유출되지 않기위해 암호화 및 데이터베이스에

대한 접근제어가 중요해지는 추세에 있습니다.”

- OOO(B대학교)

이는 보안 중심 접근과 활용 중심 접근 간의 균형을 제도적으로 정립할 필요성을 보여준다.

3. 시사점

가. 해외 사례의 적용 가능성

데이터 거버넌스의 실행 과정에서 해외 사례의 적용 가능성은 한국의 제도적 맥락과 조직적 한계를 고려할 때 특히 중요한 시사점을 제공한다. 다수의 인터뷰 참여자들은 국내 기관의 데이터 관리·연계 한계를 지적하면서도, 해외에서 이미 정착된 제도적·기술적 장치들을 참조할 필요가 있음을 강조하였다.

1) 데이터 관리 체계의 통합

“데이터 공유를 전제로 한 다기관 통합 플랫폼 구축이 필요하다고 생각합니다. 현 대응은 단일 기관이 아닌 다기관 공동대응이기 때문에 실시간 데이터 연계체계를 법제화하고 운영해야 합니다. 예를 들자면 필란드 112 긴급대응 시스템이 있습니다.”

- OOO(A소방서)

2) 민간 데이터의 공익적 활용 제도화

“우리나라에 적용될 필요가 있다고 생각하는 요소는 EU Data Act의 내용을 접목할 수 있도록 ‘재난 공중보건 등 특정 공익 목적에서 공공기관이 민간 데이터 제공을 요청할 수 있는 절차·보호장치를’ 규정할 수 있는 장치를 마련하는 것입니다. 특히, 위기 대응 시 위치, 모빌리티, 유통 등 민간 데이터의 공익 활용이 필요하며, 이는 사전에 지정된 공익적 사유와 절차, 보상, 보안 패키지를 법령에 포함하여 법안을 마련하는 것입니다.”

- OOO(A공단)

3) 싱가포르의 거브테크(GovTech)

“싱가포르 정부는 공공 부문에서 기술 혁신을 주도하기 위해 거브테크를 설립했고, 내부 기술 역량 강화에 중점을 두고, 외주 대신 내부 개발자 육성을 통해 데이터 거버넌스를 구축하고 있습니다. 데이터의 수집, 저장, 분석, 활용 전 과정에서 통일된 기준과 정책을 적용하여 데이터 품질 및 보안을 확보하고 있습니다. 우리나라에서도 정부 주도로 분야별 연구 목적의 표준 데이터셋을 구축하고, 연관성을 계속 넓혀 데이터의 원활한 흐름을 위해 데이터 파이프라인을 연결하고 장애물을 없애는 등의 작업이 필요할 것으로 판단됩니다.”

- 000(B대학교)

“싱가포르는 정부 데이터 표준을 강하게 통일하고, 모든 부처가 공통된 기준으로 데이터를 수집하도록 의무화했습니다. 그 결과 데이터가 부처 간 빠르게 공유되고, *duplicate* 데이터가 크게 감소하였습니다. 이 표준화 덕분에 AI 서비스 개발 속도도 빨라졌다고 할 수 있습니다.”

- 000(C국책연구원)

4) 영국의 데이터 책임체계(Data Stewardship Model)

“OECD 및 주요 선진국의 데이터 거버넌스 사례를 보면, 특히 중앙정부 차원의 통합된 데이터 관리체계 구축과 데이터 책임자의 역할이 중요하다고 생각됩니다. 특히, 영국의 데이터 책임체계(Data Stewardship Model)는 우리나라 공공부문에 도입할 필요성이 크다고 생각합니다. 영국은 Cabinet Office 산하의 Government Digital Service(GDS)를 통해 정부 전 부처의 데이터 표준, API, 메타데이터 체계를 강제하고 있으며, 각 부처에는 Chief Data Officer(CDO)와 데이터 책임관을 명시적으로 지정하고 있습니다. 이는 중복된 데이터 생산이나 정책 연계 실패를 최소화하는 데 핵심적인 역할을 하고 있습니다. 우리나라도 국가데이터정책위원회, 행안부, 기재부, 환경부 등 여러 부처가 각자의 데이터 시스템을 운영하고 있으나, 기관 간 연계가 어렵고, 데이터에 대한 관리 주체와 책임 주체가 불명확하며 데이터 기반 정책 수립이 실질적 의사결정까지 연결되지 않는 경우가 많습니다. 이런 측면에서 우리도 데이터 거버넌스를 단순한 관리가 아닌 ‘정책 실행의 기반’으로 인식을 전환할 필요가 있습니다.”

- 000(B국책연구원)

5) 국세청 데이터를 통한 사회 연구 활용

“스웨덴과 같은 국가에서는 빈곤이나 불평등을 연구할 때 국세청의 납세 데이터를 활용할 수 있도록 합니다. 우리나라는 가계금융복지 조사 등과 같이 서베이를 통해서 불평등과 빈곤에 대해서 연구하는데, 서베이 조사의 한계상 실제 한국의 상태를 왜곡할 여지가 충분하다고 생각합니다. 따라서 국세청 데이터를 활용할 수 있는 방안이 고민되면 좋을 것 같습니다.”

- 000(C대학교)

6) 글로벌 기준의 적용 필요성

“데이터 스튜어드십 · FAIR 원칙 · CDO 체계 · 신뢰할 수 있는 데이터 공간(*secure data room*) 등은 국내 공공부문에도 유효하다고 생각합니다.”

- 000(D대학교)

“안전한 데이터 연결 방식(*PPRL*)이란 서로 다른 기관에 있는 데이터를 개인 정보가 새지 않게 안전하게 연결하는 방법입니다. 예를 들어, 주민번호 같은 민감한 정보를 직접 쓰지 않고 암호화된 코드로 바뀌어서, 같은 사람의 데이터를 여러 기관에서 합쳐서 분석할 수 있게 합니다. 이렇게 하면 개인정보를 보호하면서도 복지, 교육, 보건 데이터를 통합해 더 정확한 정책을 만들 수 있습니다. 또한, 성과 · 영향 평가 체계(*Results Chain*)는 사업이 ‘얼마나 많이 했는지’뿐 아니라 무슨 변화를 만들었는지, 사회에 어떤 장기적인 영향을 줬는지까지 차례대로 평가하는 방식입니다. 예를 들어 직업교육 프로그램이라면, 교육을 몇 명이 받았는지(*Output*), 교육받은 뒤 취업률이 얼마나 올랐는지(*Outcome*), 지역 고용상황이 얼마나 좋아졌는지(*Impact*)까지 보는 것입니다. 이렇게 하면 사업이 단순한 활동량을 넘어, 실질적인 효과와 사회적 가치를 내는지 판단할 수 있습니다.”

- 000(A지방의회)

7) AI 거버넌스 법제화

“EU 인공지능법처럼 국내도 AI 활용이 늘어나는 추세에 대비하여 공공부문 AI 거버넌스 법률 제정 등 검토가 필요합니다.”

- OOO(A공기업)

8) 미국의 데이터 거버넌스 조직 모델

“미국에서는 데이터의 가치를 분석하고 데이터를 활용하여 성과를 창출하는 조직이 대통령 직속 조직으로서 강한 권한과 지위를 보유하고 있을 뿐만 아니라, 데이터 담당자에 대한 대우와 보상의 수준도 높습니다. 한국에서도 데이터 담당 개인 또는 조직이 개별 공공부문 기관 내에서 또는 전체 공공부문 체계 내에서 강한 권한과 지위를 보유하면서 데이터 거버넌스 관련 업무를 수행하도록 할 필요가 있고, 이러한 개인 또는 조직에게 충분한 대우와 보상을 제공할 필요가 있습니다.”

- OOO(B공공기관)

결국 해외 사례의 적용 가능성은 단순한 제도 이식이 아니라, 국내 조직문화와 법적·기술적 환경을 고려한 점진적 제도화 과정으로 접근해야 한다. 인터뷰에서 나타난 바와 같이, 플랫폼 구축, 표준화, 책임성 확보, 전문인력 강화는 해외에서 이미 검증된 방식이며, 이를 한국의 데이터 거버넌스 체계에 적용한다면 데이터 품질 제고와 활용성 확대에 기여할 수 있을 것이다.

나. AI 생애주기 기반 거버넌스 방향

AI 거버넌스는 모델의 개발에서 배포, 운영까지 이어지는 전 생애주기에 걸쳐 데이터의 품질·보안·책임성을 확보하는 체계가 요구된다. 인터뷰 결과, 공공부문에서 우선 고려해야 할 핵심 범주는 크게 데이터 수집·전처리 단계, 모델 개발·검증 단계, 배포·운영 관리 단계로 구분된다.

1) 데이터 수집과 전처리 단계: 품질 관리, 개인정보 보호

“데이터 수집 및 전처리 단계의 품질 관리 단계에서는 개인정보나 민감정보, 유해언어 등에 대한 비식별화, 편향 데이터 제거, 데이터 출처 및 이력 관리가 중요하다고

생각합니다.”

- 000(A대학교)

“데이터 수집과 전처리 단계에서 출처와 품질을 검증하고 개인정보를 비식별화시켜야 합니다.”

- 000(A공단)

“AI의 학습 데이터의 개인정보 및 보안관리가 우선적으로 고려되어야 하는 부분이라고 생각합니다.”

- 000(A중앙부처)

2) 모델 개발 및 검증 단계: 설명 가능성, 윤리성, 편향 점검

“모델 거버넌스에 있어 검증·편향·설명가능성을 고려해야 합니다.”

- 000(D대학교)

“AI 알고리즘 검증 및 윤리 기준 마련: 투명한 기준과 검토 절차를 통해 오작동이나 편향 가능성을 사전에 차단해야 합니다.”

- 000(B기초자치단체)

“AI 모델의 목적은 기관마다 다를 수 있습니다. 예컨대, 예산사업이 많은 기관은 예산사업의 추진현황 및 전망에 대한 측면에 초점을 맞춰 모델을 개발할 필요가 있을 것 같은 반면, 법 적용 및 해석을 주로 하는 규제기관은 과거 규제내역과 함께 규제법령을 구체적인 사례에 적용할 수 있는지에 대한 언어적 모델 개발에 초점을 맞출 필요가 있을 것으로 예상됩니다.”

- 000(입법기관)

“모든 데이터 관리 프레임워크의 범위가 중요한 가운데, 왜곡된 데이터 및 편향된 데이터와 같은 부적절한 데이터가 AI 모델의 개발부터 배포까지의 생애주기에 활용되지 않도록 하기 위해 데이터 품질 검증을 강화할 필요가 있습니다. 특히 공공부문에서 왜곡된 데이터 및 편향된 데이터로 인한 AI 모델 개발 및 배포로 인해 공공성이

훼손되면 넓은 범위의 개인들에게 피해를 초래할 수 있다는 점에서, 데이터 품질 검증이 강화될 필요가 있습니다.”

- 000(B공공기관)

3) 배포와 운영 단계: 지속적인 모니터링, 피드백 체계

“첫째, 데이터 수집 단계의 신뢰성 확보, 둘째, AI 도입 단계의 윤리성 검토 및 위험 평가, 셋째, 운영 이후의 지속적인 모니터링 및 피드백 체계 구축이 중요하다고 생각합니다.”

- 000(A기초자치단체)

“서비스 적용 및 피드백 단계: AI가 실제로 교육 행정에 투입된 이후 현장의 반응과 성과 데이터를 수집·분석해 지속적으로 개선하는 피드백 체계를 구축해야 합니다.”

- 000(B기초자치단체)

4) 공정성과 책임성 확보

“서비스 적용 및 피드백 단계: AI가 실제로 교육 행정에 투입된 이후 현장의 반응과 성과 데이터를 수집·분석해 지속적으로 개선하는 피드백 체계를 구축해야 합니다.”

- 000(B기초자치단체)

“생애주기 기반 AI 거버넌스를 위해 공공부문이 우선적으로 고려해야 할 핵심 범주는 첫째, 데이터 품질과 신뢰성을 확보하는 것입니다. 정책 및 행정에 활용되는 AI의 정확성과 공정성을 위해 수집부터 정제·검증까지 데이터를 체계적으로 관리해야 합니다. 둘째, 투명성을 확보하는 것입니다. AI가 어떤 기준과 근거로 판단했는지를 국민이 이해할 수 있도록 알고리즘 구조와 데이터 특성 등을 공개하고 설명할 수 있어야 합니다. 셋째, 책임성과 위험관리 체계를 구축하는 것입니다. AI의 오류나 편향에 대비한 사전 위험 분석, 운영 중 모니터링, 사후 대응 절차를 마련하고 책임 소재를 명확히 해야 합니다.”

- 000(B국책연구원)

“우선은 신뢰와 책임성 확보라고 생각합니다. 개인정보보호의 설계, AI의 결정을 인간이 이해할 수 있는 방식으로 제공하는 것, AI 모델의 개발에 있어서 전체적인 정확도도 중요하지만 취약계층에 대한 예측 오류 등 공정성지표의 측정, 그리고 어떠한 공공서비스에 AI가 활용되고 있는지를 공개하는 등의 활동들이 필요하다고 생각합니다.”

- 000(C대학교)

5) 데이터 표준화와 제도적 기반

“AI 행정에서 우선 챙겨야 할 세 가지는 ① 데이터 품질과 표준 - 예를 들어, 무상 교복 지원 사업에서 시·군마다 ‘학생 수’ 집계 기준이 달라 사업비 산정이 어긋난 적이 있습니다. 처음부터 입력 형식과 코드 규칙을 통일해야 이런 문제를 막을 수 있습니다. ② 개인정보 보호와 접근 권한 - 다문화가정 지원 정책에서 법무부·교육청·시청 데이터를 연계하려 해도, 접근 권한이 명확하지 않아 몇 달씩 지연된 사례가 있습니다. 권한과 절차를 사전에 명확히 해야 합니다. ③ 성과와 피드백 - 청소년 상담·심리치유 프로그램의 경우, 참여 인원(산출)만 집계하고 실제 재학 유지율(결과)이나 학교 적응도(영향)를 평가하지 않아 예산 효율성을 따지기 어려웠던 적이 있습니다.”

- 000(A지방의회)

6) 실무 차원의 보안과 인프라 문제

“AI 행정에서 우선 챙겨야 할 세 가지는 ① 데이터 품질과 표준 - 예를 들어, 무상 교복 지원 사업에서 시·군마다 ‘학생 수’ 집계 기준이 달라 사업비 산정이 어긋난 적이 있습니다. 처음부터 입력 형식과 코드 규칙을 통일해야 이런 문제를 막을 수 있습니다. ② 개인정보 보호와 접근 권한 - 다문화가정 지원 정책에서 법무부·교육청·시청 데이터를 연계하려 해도, 접근 권한이 명확하지 않아 몇 달씩 지연된 사례가 있습니다. 권한과 절차를 사전에 명확히 해야 합니다. ③ 성과와 피드백 - 청소년 상담·심리치유 프로그램의 경우, 참여 인원(산출)만 집계하고 실제 재학 유지율(결과)이나 학교 적응도(영향)를 평가하지 않아 예산 효율성을 따지기 어려웠던 적이 있습니다.”

- 000(A지방의회)

“비식별 기준·기술·위험평가 방식이 기관마다 달라 재식별 위험이 반복적으로 논란이 되고있습니다. 데이터 활용과 개인정보보호 간의 긴장관계가 제도화되어 있지 않아, 기관들이 ‘안전’을 이유로 활용을 기피하는 문제가 발생합니다.”

- OOO(F대학교)

“첫 번째, 보안, 기관의 주요 데이터를 AI에 학습시켜도 되는지, 학습시킨 데이터의 유출 등은 없을지 검토가 필요합니다. 두 번째, 사용성, 내부 사용자가 손쉽게 사용할 수 있는지 검토가 필요합니다. 세 번째, 지속성, 데이터의 지속적 관리와 담당자 변경 시 변경되지 않을 수 있게 철저한 인수인계가 필요합니다.”

- OOO(B공단)

“데이터 보안, 개인정보 보호, 정확성을 우선적으로 고려해야 한다고 생각합니다.”

- OOO(A공기업)

종합하면, 인터뷰 응답자들은 공공부문의 AI 거버넌스가 단일 단계나 기술적 요인만을 관리하는 차원을 넘어, 데이터 수집에서 모델 개발·검증, 배포와 운영, 사후 모니터링에 이르기까지 전 생애주기를 포괄하는 통합적 관리 체계를 요구하고 있음을 강조하였다. 특히 데이터 품질·보안·개인정보 보호를 기반으로, 모델 검증과 설명가능성, 운영 과정의 지속적인 모니터링과 피드백, 그리고 공정성과 책임성 확보가 반복적으로 지적되었다. 또한 실무 차원에서는 데이터 표준화와 제도적 기반 마련, 보안 및 인프라 문제 해결이 선결 과제로 제시되었다. 구체적으로는 부처별로 데이터 소유권이 분절된 기존의 관리 구조가 데이터의 통합적 활용을 저해하고 유사한 문제를 반복시키는 원인이 되어 왔기에, 이를 해결하기 위한 ‘초부처 통합 구조’와 ‘국가 데이터 플랫폼’ 체계로의 과감한 전환이 필요하다는 의견도 도출되었다. 개별 기관의 자체 양식이 아닌 국가 공통 표준(National Data Standard)에 기반하여 모든 데이터의 생산·관리 체계를 일원화하고, 국가 단위의 통합 데이터 허브 및 공통 API를 구축의 필요성이 제시되었다. 아울러 이러한 통합 체계가 실효성을 갖기 위해서는 부처 간 이해관계를 조정할 수 있는 상위 거버넌스 조직의 실질적 권한을 강화하여, 기존의 분절형 거버넌스를 ‘플랫폼형 국가 데이터 거버넌스’로의 혁신 필요성을 강조하였다.

이러한 의견들은 결국 AI가 행정과 정책 전반에서 활용되기 위해서는 신뢰성·투명성·책임성이라는 거버넌스 핵심 원칙을 생애주기 전반에 내재화해야 함을 시사한다.

제2절 기술과 시스템

공공부문 데이터 거버넌스의 실제 운영 과정에서 기술과 시스템 차원의 논의는 크게 세 가지 영역에서 시사점을 도출할 수 있다. 첫째, 데이터 활용에 있어서 사용되는 기술(AI, 통계패키지, 시각화툴). 둘째, 데이터 활용에 있어서 사용되는 시스템(플랫폼, 대시보드, 클라우드 등). 셋째, 기술과 시스템 차원에 있어서 장애요인과 성공요인이다.

1. 데이터 활용에 있어서 사용되는 기술(AI, 통계패키지, 시각화툴)

공공부문에서 데이터 활용을 위해 도입되고 있는 기술들은 기관별로 다양한 수준과 형태로 나타나고 있다. AI 기반 도구부터 전통적인 통계 패키지까지 폭넓은 스펙트럼을 보이고 있으며, 각 기술의 활용도와 만족도는 기관의 특성과 구성원의 역량에 따라 상당한 차이를 보이고 있다.

인터뷰 결과 최근 공공부문에서는 생성형 AI를 중심으로 한 AI 기반 도구의 도입이 활발히 진행되고 있었다. 특히 ChatGPT, Perplexity와 같은 생성형 AI 도구가 다수 기관에서 활용되고 있는 것으로 나타났으며, 기술의 활용도와 만족도는 기관의 특성과 구성원의 역량에 따라 상당한 차이를 보이고 있었다

“각 부서들에서 생산한 공식 보고서들에 접근할 수 있는 내부 시스템이 있습니다. Perplexity, Deep L의 사용(유료)을 지원하고 있습니다”

- OOO(입법기관)

“기획재정부는 현재 내부망에서 Chat GPT, Perplexity를 운영 중이고 부서별 자료 공유 플랫폼인 웨어박스를 운영하고 있으며, 중장기적으로는 자체적인 생성형 AI 구축을 추진 중입니다”

- OOO(A대학교)

“데이터 활용에 있어서 주로 통계 패키지(STATA, SPSS 등)를 사용합니다. 최근 연구원 차원에서는 Chat GPT를 지원하고 있습니다. AI SW를 추가 지원하고자 원내 구성원에게 수요 조사를 진행하고 있습니다”

- 000(A국책연구원)

일부 기관에서는 긍정적인 효과를 보고하고 있으나, 정확성과 신뢰성에 대한 우려도 함께 제기되었다. 이는 정부에서 AI 활용 시 데이터 관리와 윤리적 고려의 중요성을 실증적으로 뒷받침한다.

“Chat GPT는 연구자들이 활발하게 이용하고 있는데, 아직 정확성에 대한 신뢰도가 떨어지고 한계가 있어 자료 요약이나 정리에 많이 활용하는 것 같습니다. 해당 프로그램을 통해 방대한 양의 자료를 신속하게 정리·요약할 수 있는 장점이 있으나, 연구자가 재검을 통해 정확성이나 신뢰성을 확인해야 하는 추가 절차가 필요하다는 점에서 아직 한계가 있는 것 같습니다”

- 000(A국책연구원)

생성형 AI의 공공부문에서는 여전히 SPSS, STATA, R, Python과 같은 전통적인 통계 패키지와 분석도구가 광범위하게 활용되고 있다. 특히 연구기관과 정책분석을 담당하는 부서에서는 이러한 도구들이 핵심적인 역할을 하고 있는 것으로 나타났다.

“기관 차원보다는 개인적으로 SPSS, AMOS, STATA, HLM 등의 통계패키지를 많이 사용하고 있으며, 필요시 비정형데이터(신문기사, 인터뷰 내용 등)에 대한 워크클라우드, 키워드네트워크이나 특정 결과에 대한 예측이나 추정을 위해서 텍스트, R, 파이썬 등의 일부 기능을 사용하기도 합니다”

- 000(A협의회)

“Python 및 Python 기반 다양한 라이브러리, 구글 드라이브-저장소, colab, Visual Studio Code, Github, Hugging Face, Pytorch, Tensorflow, Aihub를 사용합니다”

- 000(A대학교)

“통계패키지로 SAS를, 시각화툴로 ElasticSearch를, Cloud는 AWS를, 대시보드는 IBM instana를 사용합니다”

- OOO(B대학교)

공공부문에서 활용되는 기술들은 각각 고유한 장점과 한계를 보이고 있다. 특히 오픈소스 기반 도구들의 접근성과 민간 솔루션의 기능성 사이에서 균형점을 찾는 것이 중요한 과제로 나타났다. 이는 Janssen et al.(2020)이 제시한 신뢰할 수 있는 인공지능을 위한 데이터 거버넌스에서 기술적 선택의 중요성과 일치한다.

특히 오픈소스 도구의 접근성과 상용 솔루션의 기능성 사이에서 균형점을 찾는 것이 중요한 과제로 부각되었다. 이는 공공부문의 제한된 예산과 보안 요구사항을 고려할 때, 기술 선택에서 실용적 접근이 필요함을 시사한다.

“Python 및 관련 라이브러리의 장점은 오픈소스 기반으로 이루어져 있기 때문에 누구나 접근이 가능합니다. 또한, 다양한 라이브러리가 지속적으로 개발되고 있으며, 커뮤니티(Github, Papers with Code 등)가 활발하여 오류 추적, 공유 등이 잘 되어 있습니다”

- OOO(A대학교)

“솔루션과 플랫폼에서 제공되는 기능들이 다양하고 많아서 짧은 기간내에 서비스 구축이 가능한 점이 장점이지만, 가지고 있는 데이터를 솔루션을 통해 100% 활용할 수 없는 단점도 있습니다. 예를 들면, 시스템 연동, 데이터 포맷 처리 등”

- OOO(B대학교)

2. 데이터 활용에 있어서 사용되는 시스템(플랫폼, 대시보드, 클라우드 등)

공공부문의 데이터 관련 시스템은 기관별 특성에 따라 다양한 형태로 구축되어 있다. 교육 분야의 통합정보시스템부터 재정 분야의 전문 플랫폼까지, 각 영역별로 특화된 시스템들이 운영되고 있으며, 최근에는 클라우드 기반 서비스의 도입이 확산되고 있다.

교육 분야의 통합정보시스템부터 재정 분야의 전문 플랫폼까지, 각 영역별로 특화된 시스템들이 운영되고 있으며 최근에는 클라우드 기반 서비스의 도입이 확산되고 있다.

"업무상 현재 사용하고 있는 시스템은 나이스, K-에듀파인 두가지 시스템을 사용하여 데이터를 활용하고 있으며, 두가지 시스템은 크게 재정영역(K-에듀파인), 그 외 교무학사, 인사 등 영역(나이스)으로 나뉘져 있으며 단순 사용자의 입장에서 평가하기는 어렵지만, 어느정도 해당분야 관리 시스템으로는 데이터 시스템의 통합이 상당히 이루어져 완성형에 가깝다고 생각함"

- 000(A기초자치단체)

"학교에서는 주로 K-에듀파인, 교육행정정보시스템(NEIS), 기록관리시스템, 교육시설통합정보망 등을 활용하고 있습니다. 장점은 업무별로 기능이 세분화되어 있어 전문성 있는 처리가 가능하다는 점이고, 단점은 각 시스템 간 연계성이 부족하여 동일한 정보를 중복 입력하거나, 자료를 수작업으로 가공해야 하는 불편함이 있다는 점입니다"

- 000(A기초자치단체)

교육 분야에서는 나이스(NEIS), K-에듀파인과 같은 대규모 통합정보시스템이 핵심적인 역할을 하고 있는 것으로 나타났으며 이러한 시스템들은 상당한 수준의 데이터 통합을 달성했으나 여전히 시스템 간 연계와 사용자 편의성 측면에서 개선이 필요한 상황이다.

재정 및 행정 분야에서는 디브레인(dBrain), 열린재정과 같은 전문 플랫폼들이 운영되고 있다. 이러한 시스템들은 특정 업무 영역에 특화되어 있어 해당 분야에서는 높은 효율성을 보이고 있으나, 타 시스템과의 연계에서는 한계를 보이고 있다.

"기획재정부가 대외적으로 운영 중인 데이터 관련 시스템은 열린재정(재정정보시스템), dBrain(디지털예산회계시스템)이 있습니다"

- 000(A대학교)

"예산 업무에 있어서는 디브레인이 있고, 다른 업무로는 국정과제 대시보드 등이 있다고 알고있으나 사용해본적은 없음. 국가기관에서 사용하는 데이터 거버넌스 관점에서는 사용자의 편의성이 부족하고, 다른부처와의 연계성에 있어서는 한계가 있다고 생각함"

- 000(A중앙부처)

최근 공공부문에서는 클라우드 기반 서비스와 협업 플랫폼의 도입이 확산되고 있다. G 드라이브, 웨어박스와 같은 내부 협업 도구부터 외부 클라우드 서비스까지 다양한 형태로 활용되고 있으나, 보안과 활용도 측면에서 여전히 과제가 남아 있다.

“클라우드 시스템 'G드라이브' 사용 중이나 개인자료 백업에만 활용할 뿐 데이터를 공유 및 공동 활용하는 것은 활성화되어 있지 않음”

- 000(B중앙부처)

“자기 자료 보관과 국내 공유를 위한 클라우드 웹하드가 있고, 최근에 <웨어박스>라고 부처 전체에 공유되는 자료를 올리는 곳과 자료의 리서치를 도와주는 리서치팀이 생김!”

- 000(C중앙부처)

전문 연구기관들은 연구 특성에 맞는 특화된 데이터 시스템을 구축하여 운영하고 있다. 특히 DW(데이터웨어하우스) 구축과 대시보드 서비스를 통해 내부 데이터 활용도를 높이려는 노력이 이루어지고 있었다.

“기관 내 DW를 구축하고 주요사업에 대한 일부 데이터를 대시보드 형태로 제공하고 있어서 기관 임직원들이 공단 내 데이터에 대한 접근이 용이하고, 데이터 기반의 의사결정에 기여 가능하다는 장점이 있음”

- 000(B공단)

“저희 기관은 나라장터 시스템을 기반으로 한 내부 통합 데이터 플랫폼을 운용하고 있습니다. 이 플랫폼은 조달계약, 입찰, 구매 실적 등 다양한 정보를 축적하고 있으며, 일부 데이터는 공공데이터포털이나 유관기관과의 연계를 통해 외부로도 공유되고 있습니다”

- 000(B국책연구원)

공공부문의 경우 클라우드 기반 협업 플랫폼의 도입이 확산되고 있으나, 여전히 개인 백업 용도에 머물러 있거나 보안 제약으로 인해 제한적으로 활용되고 있다. 이는 기술적 도입과 실제 활용 사이의 간극이 여전히 크다는 점을 보여준다.

3. 기술과 시스템 차원에 있어서 장애요인과 성공요인

공공부문 데이터 거버넌스에서 기술과 시스템이 성공적으로 작동하기 위해서는 다양한 요인들이 복합적으로 작용한다. 특히 시스템 간 연계성, 데이터 표준화, 사용자 역량, 보안과 활용성의 균형 등이 핵심적인 성공요인으로 작용하는 반면, 표준화 부족, 시스템 분절, 사용자 교육 부족 등이 주요 장애요인으로 나타나고 있다.

“위에서 언급한 두가지 시스템에 대해서 연동 시에 자동연동이 되지 않고 수동으로 버튼을 눌러서 데이터를 연동시켜야 하는 부분(인사, 정보공개 데이터 등)이 많은데, 시스템 구성상 최종적으로 확인된 데이터를 연동하기 위해 수동으로 구현한 것이 아닌지 생각이 됨”

- 000(A기초자치단체)

“예를 들어, 공무원 급여업무 진행 시, NEIS에서 급여 작업 후 K-에듀파인에서 동일한 내용을 다시 등록해야 하는 불편함이 있습니다. 이로 인해 중복 업무, 시간 지연, 오류 발생 가능성이 높아진다고 할 수 있습니다”

- 000(A기초자치단체)

“국가기관에서 사용하는 예산시스템은 디브레인 시스템이고, 지방자치단체에서 사용하는 예산시스템은 E-호조 라는 시스템으로 알고 있는데, 그 두 시스템 간의 연계가 안된다고 알고있음”

- 000(A 중앙부처)

공공부문에서 가장 빈번하게 지적되는 문제는 시스템 간 연계와 통합의 부족이다. 각 기관이나 부서별로 독립적으로 구축된 시스템들이 서로 호환되지 않아 데이터 활용에 상당한 제약을 가하고 있다. 특히 데이터 형식, 용어 체계, 코드값 등의 표준화 부족으로 인해 시스템 간 데이터 교환과 통합에 어려움을 겪고 있으며 이는 데이터의 정확성과 신뢰성에도 영향을 미치고 있다.

“전사차원에서 데이터 포맷 표준화를 추구하지만 서비스되고 있는 시스템이 많은 경우 간혹 전사차원에서 관리되지 못한 영역이 있습니다”

- 000(B대학교)

“같은 군수 정보라고 하더라도 기관마다 동일한 규격을 표시하는 표현방법이 다르고 특히 탄약의 경우 고유한 일련번호를 보관과정에서 알아보기 어렵기 때문에 컷값으로 매칭하거나 추적하기 어려움”

- 000(A연합회)

“시스템 간 데이터 제공 환경이 달라 데이터 가공 시 정리 시간 소요”

- 000(A공기업)

공공부문의 엄격한 보안 기준이 데이터 활용을 제약하는 요인으로 작용하고 있다. 특히 AI 도구나 클라우드 서비스 활용에서 보안과 활용성 사이의 균형점을 찾는 것이 중요한 과제로 부각되고 있다.

“국가기관은 보안기준이 높기 때문에, 데이터 활용 및 운영을 위하여 보안기준을 충족하여야 하는 어려움이 있다고 생각함”

- 000(A중앙부처)

“현재 공공기관의 경우 데이터 활용에 대해 적극 장려하고 있지만, 업무 시스템과 이에 대한 연결은 매우 미흡한 편입니다. 예컨대 AI교육활용에 대한 강좌는 많이 개설되어 있는 편이나, 실제 업무망에는 차단된 AI 사이트들이 많습니다. 또한 통계분석에 활용할 수 있도록 R Studio, Google Drive 등 주요 프로그램과 사이트가 차단되어있는 상황입니다”

- 000(A공단)

데이터 거버넌스 성공을 위한 가장 중요한 기술적 요소로 표준화된 데이터 형식과 시스템 간 상호운용성이 지적되고 있다. 이는 데이터의 품질과 활용도를 동시에 높이는 핵심 요소로 인식되고 있었다.

“표준화된 데이터 형식이 가장 중요하다 생각합니다”

- 000(A대학교)

“상호운용 표준, 권한·접근통제”

- 000(D대학교)

“데이터 생성-수집-가공-분석-활용에 있어 그 주체가 누구이며, 권한은 누가 갖고 최종 책임자는 누구인지에 대한 명확한 기준이 확립되어야 책임소재나 투명성의 문제 등이 최소화될 것이라고 봅니다”

- 000(A국책연구원)

기술적 복잡성을 줄이고 사용자가 쉽게 접근할 수 있는 인터페이스 구축이 성공의 핵심 요소로 강조되고 있다. 특히 비전문가도 활용할 수 있는 직관적인 시스템 설계가 중요하다고 언급되었다.

“① 시스템 간 연계성 확보, ② 사용자 중심의 직관적인 인터페이스, ③ 데이터 정합성 자동 검증 기능이 핵심입니다. 이 세 가지가 갖춰지면 현장에서는 데이터 입력 부담이 줄고, 정확한 정보에 기반한 신속한 의사결정이 가능해집니다”

- 000(A기초자치단체)

“① 시스템 간 연계성 확보, ② 입력 자동화 및 중복 최소화, ③ 사용자 친화적인 인터페이스(UI), ④ 접근 권한의 적절한 분배가 핵심이라고 생각합니다”

- 000(B기초자치단체)

데이터의 정확성과 신뢰성을 확보하기 위한 품질 관리 체계와 지속적인 업데이트 메커니즘이 성공의 필수 요소로 인식되고 있다.

“사용자의 입장에서는 데이터에 대한 신뢰성을 높이기 위해 데이터 무결성을 유지하는 것이 시스템에 요구되는 핵심이라 생각됨”

- 000(A기초자치단체)

“운영 과정에서 나타나는 문제는 지속적인 데이터 품질 관리, 점검 및 업데이트 구축의 미흡이라 생각합니다. 공공 데이터를 공개한 이후, 정기적인 업데이트나 오류 수정이 이루어지지 않는 경우가 많습니다”

- 000(A대학교)

마지막으로 조직 구성원의 역량 강화에 대한 응답도 있었다. 기술과 시스템의 성공적 활용을 위해서는 조직 구성원의 디지털 역량과 데이터 리터러시 강화가 필수적인 요소로 강조되고 있다.

“기술 및 시스템 측면에서 공공부문 종사자의 디지털 전환 교육을 통해 세대 간의 격차를 줄이는 것이 선행되어야 합니다”

- 000(A협회)

“기관 임직원들의 A·데이터 활용에 대한 리터러시 강화를 통한 인식 개선”

- 000(B공단)

“데이터의 중요성은 알지만 조직원들이 이를 실제적으로 사용하는 경우는 극히 드뭅니다. 이에 직원들의 교육과 기관내 생성된 데이터를 사용한 경험들이 많아지고 이것이 공유되어야 한다고 생각합니다”

- 000(A공공기관)

본 연구의 인터뷰 결과 공공부문 데이터 거버넌스의 기술과 시스템 차원에서 도출되는 핵심 시사점은 기술 도입과 실제 활용 간의 구조적 간극으로 나타났다. 생성형 AI를 중심으로 한 신기술 도입이 활발히 진행되고 있으나 정확성과 신뢰성에 대한 우려로 보조적 용도에 머물러 있으며, 교육·재정 등 분야별로 특화된 시스템들은 높은 완성도를 보이지만 시스템 간 연계성 부족으로 인한 중복 업무와 데이터 분절 현상이 심각하다. 특히 표준화 부족, 보안 규제와 활용성 간의 딜레마, 조직 구성원의 역량 격차가 복합적으로 작용하여 기술적 투자 대비 실제 데이터 거버넌스 성과는 제한적인 상황이다.

성공적인 데이터 거버넌스를 위해서는 개별 기술의 우수성보다 시스템적 접근이 더욱 중요한 것으로 나타났다. 표준화된 데이터 형식과 상호운용성 확보, 사용자 친화적 인터페이스

이스 구축, 지속적인 데이터 품질 관리 체계가 핵심 성공요인으로 확인되었으며 이는 기술 도입에 앞서 기존 시스템 간 연계성 강화와 데이터 표준화를 우선시하는 단계적 접근, 보안과 활용성의 균형을 위한 정책 조정, 조직 구성원의 체계적 역량 강화가 선행되어야 함을 의미한다. 현재 대다수 공공기관이 겪고 있는 데이터·AI 전문 인력난은 단발성 교육이나 일회성 컨설팅 등 기존의 소극적 방식으로는 해결이 불가능한 반면, 거버넌스의 성패는 기술 도입이 아닌 이를 운용할 조직과 인력 구조가 가장 중요한 요인임은 분명하다. 따라서 공공기관마다 독립적인 데이터 책임관(CDO) 및 AI 책임관(CAIO) 체계를 공식화하여 실질적 권한을 부여하고, 데이터 스튜어드 및 품질·윤리 관리를 위한 전담 조직을 배치할 필요하며, 나아가 전문 인력 확보의 어려움을 타개하기 위해 순환형 채용 및 범정부 공동 인재풀 제도를 운영하고, 중앙-지방 간 상호 파견을 제도화함으로써 기관 간 역량 격차를 구조적으로 해소가 요구된다.

결국 공공부문 데이터 거버넌스의 성공은 기술, 시스템, 조직, 제도가 유기적으로 결합된 종합적 생태계 구축으로 시사된다.

제3절 조직과 관리

공공부문 데이터 거버넌스의 실제 운영 과정에서 조직과 관리 차원의 논의는 크게 다섯 가지 영역에서 시사점을 도출할 수 있다. 첫째, 조직 구성원의 인식과 참여 수준. 둘째, 리더십과 조직문화. 셋째, 전문 인력과 역량 강화이다.

1. 조직 구성원의 인식과 참여 수준

공공부문 데이터 거버넌스의 성패는 무엇보다도 조직 구성원들의 인식 수준과 참여 의지에 달려 있다. 기존 문헌에서도 이 점은 반복적으로 강조되어 왔다. Dawes(2010)는 데이터 거버넌스가 단순한 기술적 과제가 아니라 조직 전체의 “문화적 전환(cultural shift)”을 필요로 한다고 지적했으며, Kitchin(2014)은 데이터 기반 행정을 가능케 하는 것은 제도적 장치뿐 아니라 “공무원 개개인의 태도와 참여”라고 강조하였다. OECD(2019) 역시 데이터 거버넌스 성공요인 중 첫 번째로 “조직 구성원의 인식 개선과 적극적 참여”를 명시하였다. 그러나 현실에서는 구성원의 낮은 인식 수준과 제한된 참여 의지가 가장 큰 장애물로 나타난다. Janssen과 Helbig(2018)은 데이터 기반 행정 추진이 실패하는 주요 원인 중 하나로 공무원의 낮은 인식과 참여 부족을 꼽았으며, Yang & Wu(2021)는 단순한 교육만으로는 참여가 활성화되지 않으며 보상체계와 가치 전환이 병행되어야 한다고 강조하였다.

본 연구의 인터뷰 결과도 이러한 국제적 논의와 일치한다. OOO(입법기관)는 “구성원들은 데이터를 공유하기 위해 식별·표준화·업로드해야 하는데, 이 일련의 과정이 익숙하지 않은 사람들에게는 상당한 행정적 부담으로 작용한다”고 지적하였다. 그는 또 “자신이 보유한 데이터를 타인과 공유하는 것을 귀찮아하거나 불필요하다고 생각하며, 심지어 자신에게 손해라고 생각하는 경우도 있었다”고 덧붙였다. 이러한 진술은 Burden et al.(2012)이 제시한 ‘행정적 부담(administrative burden)’ 개념이 한국 공공조직에서도 동일하게 관찰될 수 있음을 보여준다.

중앙정부와 더불어 지방정부 실무자들의 응답도 같은 맥락을 보여준다.

“정책과제로 AI기반의 사업을 발전시키려고는 하지만, 기본적인 데이터에 대한 인식수준이 낮다고 생각”

- 000(A중앙부처)

“현재 구성원들 개별적으로는 공공데이터에 대한 CRUD를 수행하고 있지만, 이를 DB관리를 하고 있다고 인지하고 작업하는 것으로 보이지 않음”

- 000(A기초자치단체)

“구성원들의 데이터에 대한 인식은 기록과 증빙 중심으로 한정되어 있으며, 데이터 분석이나 거버넌스의 관점은 낮은편입니다”

- 000(A기초자치단체)

“실무자들 사이에서 데이터의 중요성은 공감대가 있으나, 여전히 “자료제출” 정도로 인식되는 경우가 많습니다”

- 000(B기초자치단체)

이러한 발언들은 데이터 활용이 여전히 형식적 보고 문화에 종속되어 있음을 보여준다. 공공기관 연구자들 또한 비슷한 문제를 지적하였다.

“데이터거버넌스에 대한 인식과 참여도는 낮은 편. 기관의 주요 업무가 아니다보니 사이드 업무라는 인식이 강함”

- 000(B공단)

“조직구성원은 데이터거버넌스가 기존 업무에서 과부하 등 추가 업무로 여겨지는 경향이 있음. 필요성은 인식 하나 업무피로도를 고려하여 생각하고 있기 때문에 적극 적일지는 의문임”

- 000(A연합회)

“현재 기관에서는 데이터거버넌스라는 것은 생소한 의미이며 대부분의 기관이 데이터의 관리를 넘어서 수집과 정제도 제대로 되어있지 않다고 생각”

- 000(A공공기관)

“조직 구성원들은 데이터 거버넌스에 대한 인식과 참여도가 낮은 수준입니다. 데이터를 수집하고, 데이터를 활용하여 객관적인 의사결정을 위한 활용에도 관심이 적습니다. 기존에 해 왔던대로 설문조사를 하고, 설문조사 결과를 통해 결과보고서를 작성하고 끝납니다. 데이터를 활용해 사업에 대한 개선에 활용되는 경우는 많지 않습니다.”

- 000(A책임운영기관)

이와 같은 응답은 Nikiforova & Zuidervijk(2021)이 지적한 공공데이터 개방 과정에서의 저항과 책임 회피 심리, 그리고 Yang & Wu(2021)의 연구에서 확인된 지방정부 공무원의 낮은 데이터 활용 인식과 일맥상통한다. 또한 일부 인터뷰에서는 데이터 거버넌스의 필요성을 인식하면서도, 이를 본인의 업무와 분리된 “전문가의 영역”으로 한정하는 태도도 나타났다.

“많은 구성원들이 데이터의 중요성은 인지하고 있으나, 구체적인 거버넌스 체계나 절차에 대한 이해는 부족하며, 일상 업무에서 적극적으로 참여하거나 책임을 지는 경우는 제한적입니다.”

- 000(B국책연구원)

“공공부문 전체 구성원으로 범위를 확장하면, 데이터 거버넌스의 필요성이나 의의 자체에 대한 인식이 낮고, 실제 업무와 연결된 체감도도 떨어지는 경우가 많습니다. 특히 데이터품질, 표준화, 책임소재 등은 기술적인 문제라기보다는 행정·조직운영과 밀접하게 연관되어 있음에도, 이를 본인의 업무와 분리된 ‘전문가의 영역’으로 간주하는 경향이 존재합니다.”

- 000(A대학교)

그럼에도 불구하고 다수의 응답자들은 인식 개선과 참여 촉진을 위한 방안을 제시하였다. 특히 조직 구성원들이 서로 커뮤니케이션하고 협업하는 과정이 필수적이라고 강조하였으며, 문제 해결 중심 프로젝트 운영과 성과 중심 인사제도와 연계된 데이터 교육을 병행해야 한다고 지적하였다. 또한, 데이터 분석 결과가 정책 결정에 실제 반영되는 성과 환류 구조와 실무자 중심 워크숍, 전담 인력 확보가 필요하다고 언급하였다.

“데이터 기반 행정 추진을 위해서는 혁신적 행동을 장려하는 조직문화 정착이 필요하다.(중략) 데이터 거버넌스 참여를 위한 인센티브는 성과평가·근무성적평정 등에 반영되어야 하며, 모범 사례 발굴 및 포상도 효과적일 수 있다”

- 000(입법기관)

“데이터 기반 행정이 정착되기 위해서는 ‘일 잘하는 조직’이 아니라 ‘데이터로 일하는 조직’으로의 문화 전환이 필요하다”

- 000(A기초자치단체)

“현재 00행정 조직은 여전히 경험 기반 의사결정이 주를 이루고 있어 데이터 중심 문화가 초기 단계에 머물러 있다.(중략) 이를 개선하기 위해 ① 데이터 분석 결과가 정책결정에 실제반영되고있다는 성과 환류 구조 마련, ② 실무자 중심의 데이터 활용 워크숍 운영, ③ 데이터 관리 역량을 갖춘 전담 인력 확보가 필요합니다”

- 000(B기초자치단체)

“데이터 기반 행정이 단기 과제가 아니라 일상적 시스템으로 정착되려면 우수 활용 사례를 포상하거나 인사에 반영하는 인센티브가 뒷받침되어야 한다”

- 000(B국책연구원)

지방의회 차원의 응답도 있었다. 한 의원은 “회의와 보고에서 수치와 근거를 중심으로 말하는 문화”가 필요성을 강조하며, 무상교복 개선 토론회 사례에서 데이터(불량률, 납품 지연 건수 등)을 제시하자 설득력이 크게 높아졌다고 설명하였다(000, A지방의회).

이러한 응답들은 조직 구성원의 인식과 참여 수준을 높이기 위해 교육·성과 관리·문화 전환이 함께 추진되어야 함을 보여준다. 영국의 Data Ethics Framework는 구성원 교육과 데이터 윤리 가이드를 결합하여, 데이터 관리가 단순 의무가 아니라 “더 나은 정책결정에 기여하는 가치 활동”임을 인식시키려는 노력을 제도화하였다. 이는 한국의 공공부문에서도 단순 보고에서 탈피하여 데이터 활용의 성과와 효익을 가시화하는 체계가 필요함을 보여준다.

이처럼 한국의 데이터 거버넌스는 업무 효율성을 제고하기 위한 제도임에도 불구하고, 현장에서는 추가적 과업으로 인식되어 오히려 거부감을 유발할 수 있다는 점이 지적되고 있다. Janssen et al.(2020) 또한 데이터 거버넌스의 성공 조건으로 ‘가시적 가치 창출

(visible value creation)'을 제시하며, 구성원들이 데이터 공유와 활용이 자신의 업무 성과와 직결된다는 경험을 할 때 참여가 촉진된다고 설명한다. 이러한 연구는 인터뷰 응답자들의 경험적 진술을 뒷받침하며, 단순한 규정과 지침만으로는 구성원의 자발적 참여를 확보하기 어렵다는 점을 시사한다.

2. 리더십과 조직문화

데이터 거버넌스가 공공조직에서 정착하기 위해서는 제도적 설계뿐 아니라 리더십과 조직문화의 역할이 핵심적이다. 기존 연구들은 데이터 거버넌스의 성공 요인으로 기술적·법적 기반보다 리더십과 조직문화의 변화를 더 중시한다(Janssen & Van den Hoven, 2015; Klievink & Janssen, 2018). 특히 최고관리자의 전략적 비전과 리더십이 있을 때 데이터 활용은 조직의 의사결정 과정에 제도화되며, 구성원들의 참여를 촉진하는 조직문화가 형성된다(OECD, 2019). Mahanti(2018) 역시 전략적 리더십과 최고경영진의 지원이 데이터 거버넌스 정착의 가장 중요한 요인이라고 보았다.

본 연구의 인터뷰 결과에서도 리더십의 중요성은 일관되게 강조되었다.

“리더십의 가장 중요한 역할은 데이터 거버넌스 개념에 대한 수용성 제고입니다. 조직 내 리더들은 구성원들이 데이터 거버넌스에 적극 참여할 수 있도록 인센티브 등 제도적 개선방안을 강구할 뿐만 아니라, 그들과의 의사소통을 통해 참여를 독려할 필요가 있습니다”

- OOO(입법기관)

“데이터 거버넌스의 필요성과 효용성을 지속적으로 설명하고 인센티브 등을 활용해 독려하며, 관련 미래 비전을 구성원들에게 제시하는 것이 중요하다고 판단됩니다”

- OOO(A대학교)

“리더십은 단순한 승인 권한을 넘어서, 데이터 기반 조직 문화 확산의 촉진자 역할을 해야 한다고 봅니다”

- OOO(B기초자치단체)

“리더십은 데이터의 중요성을 인식하고 방향을 제시하며, 실천을 위한 자원과 지원

을 확보하는 역할”

- OOO(A기초자치단체)

“리더십은 조직이 새로운 데이터 거버넌스 체계를 받아들이고 적응할 수 있도록 변화 관리에 아주 중요한 역할이며, 직접 데이터 거버넌스의 중요성을 강조한다”

- OOO(B대학교)

이러한 발언들은 리더십을 단순 행정 권한이 아니라 조직의 변화를 주도하는 비전 제시자이자 촉진자로 인식하고 있음을 보여준다.

“데이터 거버넌스에서 리더십의 역할은 매우 중요하며, 리더의 의지와 행동이 정책과 지침 등으로 실현될 때 데이터 거버넌스가 작동할 수 있다”

- OOO(A 협의회)

“조직 내부에 데이터 활용에 관한 내용을 명확히 인식시키는 것”

- OOO(A 중앙부처)

“데이터 활용 및 관리를 위한 구성원 역량 강화에 대한 투자는 데이터 거버넌스 리더십의 핵심 역할 중 하나라고 생각합니다”

- OOO(A대학교)

“데이터 거버넌스에서 리더십의 역할은 조직 전반에 데이터의 중요성을 인식시키고, 데이터 활용을 전략적으로 이끌며, 부서 간 협업을 장려하는 것입니다. 특히, 데이터 기반 의사결정을 조직문화로 정착시키기 위해서는 리더가 명확한 방향성을 제시하고, 일관된 메시지를 전달하는 것이 필수적입니다”

- OOO(B국책연구원)

반대로, 리더십 부재가 데이터 거버넌스 추진력 저하로 이어진다는 지적도 나왔다. 이는 기존 문헌에서 Mahanti(2018)가 강조한 ‘전략적 리더십(strategic leadership)’의 부재와 유사하다. Mahanti는 데이터 거버넌스가 조직의 핵심 전략으로 자리잡기 위해서는 최고경영진(C-level)의 강력한 의지와 후원이 필요하다고 보았다. 인터뷰에서 나타난 사례는 리더십의 인식 격차가 제도 실행력과 구성원의 인식과 참여, 그리고 나아가 조직문화에 직접적

인 영향력을 행사한다는 점을 잘 보여준다.

“현재 기관의 리더십은 데이터 관리와 활용의 중요성에 대한 공감대는 형성되어 있으나 일상적인 정책운영에서 데이터 기반 결정이 체계적으로 작동하도록 이끄는 데에는 아직 부족한 부분이 많습니다. 구체적인 성과지표 설정이나 데이터 기반 성과관리 등 리더십 차원에서의 제도적 뒷받침이 강화되어야 한다고 생각합니다”

- 000(B국책연구원)

“리더는 방향과 원칙을 제시하고, 부서 간 갈등을 조정해야 합니다.(중략) 현재 기관의 리더십은 관심은 높지만, 실행 속도와 조율에서 더 적극적인 역할이 필요합니다”

- 000(A지방의회)

“현재 조직의 리더십은 전폭적인 지지를 받고 있지는 않습니다. 기업에서는 데이터 거버넌스 보다 더욱 중요한 것들(예, 수익화, 매출등)이 많기에 우선순위에서 뒤쳐져 있기 때문이죠”

- 000(B대학교)

“현재 우리 기관의 리더십은 데이터 거버넌스의 중요성은 인지하고 있으나, 데이터 관련 조직이나 인력을 배치하는 것에 대해서는 소극적이라고 생각함”

- 000(A중앙부처)

“현재 우리 기관은 데이터에 대한 리더십은 있지만 이를 실질적으로 활용하기 위한 구체적인 방안은 미진하다고 판단됩니다”

- 000(A공공기관)

“기관장이 데이터 활용으로 최종 목표와 조직의 방향성을 설정할 필요가 있습니다.(중략) 현재 우리 기관의 리더십은 데이터 거버넌스에 대해서는 많이 부족해 보입니다. 정부에서 AI가 중요하다고 하니까 AI로 무엇을 해야 할지 구체적인 고민없이 AI를 해야 되는 상황이고, AI와 관련된 데이터에 대한 중요성을 강조하고 있지만, 구체적인 실체가 없는 실정입니다.”

- 000(A책임운영기관)

이처럼 데이터 거버넌스에서 최고관리자의 인식과 비전 제시는 조직 내 실행력으로 연결되는 핵심 동기가 된다. 해외사례는 이러한 시사점을 더욱 강화한다. 영국은 「Government Data Strategy(2020-2025)」를 통해 각 부처 최고책임자에게 데이터 관리 책임을 부여하고 있으며, 미국은 「Foundations for Evidence-Based Policymaking Act(2018)」를 통해 Chief Data Officer(CDO) 제도를 법제화하였다. 싱가포르의 총리실 산하에 'Smart Nation and Digital Government Office'를 설치하여 데이터 기반 행정을 국가 차원에서 주도하였다. 반면 한국의 경우 여전히 기관장 차원에서 데이터 활용을 부차적 사안으로 인식하는 경우가 많아 응답자들의 지적처럼 실질적 추진력이 약화되는 경향을 보이고 있다.

그러나 실제 데이터 거버넌스의 실행과정은 조직문화의 제약에 의해 크게 좌우되는데, 이는 한국의 행정문화에서 크게 두드러진다. 한국의 경우 리더십의 '관심'이 구체적인 권한과 책임, 그리고 자원 배분으로 이어지지 않는다는 점에서 지속성이 결여되어 있다. 인터뷰에 따르면 전담 인력의 부족, 부서 간 칸막이, 책임 불명확, 성과주의적 행정문화 등이 데이터 거버넌스 정착을 가로막는 주요 요인으로 지적되었다.

“국가기관 관리자들에 데이터에 대한 낮은 이해도가 주요 장애요인이라고 생각함”

- 000(A중앙부처)

“주요 장애요인은 부서 간 칸막이식 업무 운영과 데이터 관리의 전문성 부족입니다”

- 000(B기초자치단체)

“조직과 관리 차원에서 데이터 표준이 미흡하고 전담인력이 부족한 상황입니다”

- 000(A공단)

“가장 큰 걸림돌은 부서 간 협력 부족입니다”

- 000(A지방의회)

“책임·소유권 불명확과 사일로 문화임. 데이터 오너·스튜어드 지정과 CDO 중심 통합 의사결정이 시급”

- 000(D대학교)

이는 Dawes(2014)가 강조한 바와 같이 실패를 용인하고 학습을 중시하는 문화가 없는 환경에서는 데이터 활용이 정착되기 어렵다는 점을 잘 보여준다. 또한 Wirtz et al.(2022)은 데이터 기반 조직문화가 없는 상황에서 단순히 리더십만 강조될 경우 거버넌스 체계가 ‘형식적’으로만 남을 수 있다고 경고하였다. 이와 더불어 한국의 성과주의적 행정문화 또한 장애요인으로 지적된다. 이러한 환경 속에서 데이터 거버넌스는 단기성과와 직접 연결되지 않는 한 부차적인 업무로 밀려난다. 이러한 경향은 인터뷰 응답자들의 답변에서 확인할 수 있다.

“조직 내 권한 및 책임 부재(데이터 관련 역할과 책임이 명확히 규정되지 않아 추진에 혼란이 발생할 수 있습니다.) 및 부서 간 협력 부족(여러 부서 간의 협조와 정보 공유가 원활하지 않아 통합 관리가 어려워집니다)”

- 000(B대학교)

“우선, 각 기관이나 부서가 독립적으로 데이터 관리 체계를 운영하면서 협력체계가 미흡해 데이터 공유와 연계가 어려운 부분이 있습니다. 또한, 데이터 관리 및 활용에 대한 책임 주체가 불분명하여 의사결정과 실행이 지연되고, 문제 발생 시 대응이 체계적이지 못합니다”

- 000(B국책연구원)

“기관 간 데이터 표준, 연계의 부재와 데이터 책임 및 관리의 불분명성, 전문 인력 부족이 대표적인 장애요인입니다. 가장 시급한 부분은 전문 인력 부족인데, 데이터를 활용해 본 경험이 있는 사람들이 극히 드문 경우가 많습니다”

- 000(A대학교)

“인력 부족, 전문인력이 아닌 기존 인원이 공공 데이터 업무를 하는 것이 일반적이고 공공 데이터 업무 외 타 업무도 부담이 심하여 데이터에 대한 전문성 강화가 어려움”

- 000(A공공기관)

“조직 관리 차원에서 발생하는 공공 데이터 거버넌스의 장애요인은 조직원 간의 역량과 편차를 어떻게 줄일 것인가? 주축이 될 수 있는 부서 여부와 인력 규모 및 전문

성, 그리고, 현실을 고려한 예산배정 등이지 않을까 판단된다”

- OOO(A협회)

Janssen et al.(2020)이 지적한 바와 같이, 데이터 거버넌스는 규정 차원이 아니라 문화적 내재화(cultural internalization)를 통해서만 안정적으로 작동할 수 있음을 뒷받침한다. 캐나다는 “Digital Standards”를 통해 오류를 학습 기회로 간주하는 문화를 확산시키고 있으며, 호주는 “Data Sharing Principles(2015)”를 통해 위험관리와 데이터 활용을 균형 있게 추진하였다. 이러한 비교는 한국의 조직문화 전환 필요성을 다시금 드러낸다.

한편, 데이터 거버넌스의 지속가능성을 위해서는 기관 간 데이터 공유와 협업이 제도적으로 활성화되어야 한다. 그러나 인터뷰 응답자들은 데이터 공유와 협업과 관련하여 현실의 여러 제약 요인을 지적하였다. 먼저 기술적·제도적 한계와 함께, 참여 동기의 부족이 대표적인 문제로 지적되었다.

“기관 간에 데이터 관리를 각 기관의 시스템을 통해서 수행하는데, 일부는 통합하여 정부24를 통하여 진행하고 있지만, 각 기관 시스템 플랫폼에 최적화 되어 연계가 쉽지 않은 상황. 개인정보 취급 등의 이슈로 공유되었다가도 롤백 되기도 하고 하는 상황에서 활성화는 쉽지 않은 상황이라고 생각됨”

- OOO(A기초자치단체)

“강제적으로 데이터 거버넌스 참여를 유도할 경우에는 할당된 분량만 참여하고, 그 이상의 참여는 기대하기 어렵습니다. 따라서 참여 확대에는 개인의 자발적 참여를 유도하기 위한 어느정도의 인센티브 부여가 필수적이라고 생각합니다”

- OOO(입법기관)

“데이터 공유와 협업활동이 조직과 개인의 평가에 반영될 때, 실질적인 참여 동기가 부여된다 생각합니다.(중략) 시스템 호환성 부족과 절차 복잡성은 협업 장벽이므로, 표준화된 데이터 포맷과 API 기반의 공유 플랫폼을 제공하고 이를 활용하는 정도에 따라 인센티브를 지급하는 방식도 효과적이라 생각합니다”

- OOO(A대학교)

일부 응답자는 성과평가와 연계된 보상체계의 제도화 필요성을 강조하였다.

“기관 간 데이터 공유와 협업을 활성화 하기 위한 효과적인 방안은 성과와 연계하여 실질적으로 인센티브를 제공해야 합니다.(중략) 따라서 데이터 공유실적이 기관평가나 예산 확보, 인사에 반영될 수 있도록 제도화하면 자발적인 참여로 유도할 수 있습니다”

- 000(B국책연구원)

“개인 성과 평가에 데이터 기반 행정 업무를 추가하면 좋겠습니다. 예를 들어 기후 위기 전시 기획전을 개최할 경우 환경부의 기후 변화 데이터를 활용하여 전시물 제작에 사용하는 것을 독려하고, 평가에도 긍정적인 영향을 미칠 수 있게 설계한다면, 데이터에 대한 많은 관심을 가질 수 있을 것입니다.”

- 000(A책임운영기관)

“데이터 제공 실적을 기관의 평가와 예산 배분에 반영하여 인센티브를 제공하고, 데이터 교환 플랫폼을 구축하고 기관별 데이터 중개자를 지정할 필요가 있습니다”

- 000(A공단)

“인센티브로는 공공분야 조직을 움직이기는 쉽지 않고, 의무조항과 인센티브과 함께 들어가야 할 것 같습니다. 필요한 경우 데이터 공유를 의무화하고, 이것을 성과지표로 관리하는 것이 좋을 것 같습니다”

- 000(C대학교)

“성과관리 관련 조치를 통해 현재 데이터 협업을 성실하게 수행해도 유의미한 평가 또는 보상을 받지 못하는 제약을 완화할 필요가 있음.(중략) 데이터 보안 관련 법·제도 완화와 데이터 협업에 대한 유의미한 인센티브 부여가 필요함. 한편으로, 법·제도 관련 조치를 통해 현재 엄격한 데이터 보안 관련 법·제도로 인해 고품질 데이터를 보유한 기관이 다른 기관에게 이러한 데이터를 공유하기 어려운 제약을 완화할 필요가 있음”

- 000(B공공기관)

또한, 협업의 실질적 성과는 담당자 간 신뢰관계와 공동성과 배분 구조를 통해 촉진될 수 있음을 확인할 수 있었다.

“기관 성과평가 등을 활용하는 방안도 일부 효과가 있겠으나 근본적으로는 협업이 필요한 기관 담당자 간 신뢰관계 구축이 효과적이라고 생각합니다”

- 000(A대학교)

“기관 간 데이터 공유와 협업을 활성화 하기 위해서는 윤리적 가치와 신뢰가 우선되어야 한다. 정확한 정보를 제공하는 플랫폼과 공정성 확보 기회가 중요할 것으로 보인다”

- 000(A협회)

“기관 간 공유를 늘리려면 성과를 나누는 구조가 필요합니다. 예를 들어, 경기도교육청과 광명시청이 학부모 교육 데이터를 공유해 중복사업을 줄였더니, 양쪽 모두 예산 절감과 참여율 상승 효과를 봤습니다. 이런 공동성가를 평가·예산에 반영하면 서로 데이터를 더 잘 공유하게 됩니다”

- 000(A지방의회)

이는 OECD(2019)가 강조한 바와 같이, 데이터 공유 활성화의 핵심은 신뢰(trust)와 인센티브라는 점과 맞닿아 있다. 또한 Dawes(2014)는 협업은 강제성보다 상호 신뢰와 공동성 과정에서 강화된다고 설명한다. 해외사례 역시 제도적 장치를 통해 협업을 촉진한다. 호주는 범정부 차원의 통합 데이터 공유 플랫폼을 구축하고, 캐나다는 데이터 공유 성과를 부처 평가에 반영한다. 영국 역시 GDS 체계 하에서 데이터 공유를 의무화하고, 성과지표와 연계한다. 이처럼 신뢰 기반의 협업 구조와 인센티브 연계는 데이터 거버넌스 활성화의 핵심 기제로 작용한다. 한국에서도 이와 마찬가지로 데이터 공유에 대한 필요성에 대한 중요성이 요구되고 있으며, 데이터 공유 플랫폼을 통한 협업과 데이터 공유 활성화에 대한 목소리가 높아지고 있는 추세이다.

“각 기관에서 수집되는 정보를 공유하기 위해서는 공통의 영역이 필요합니다. 모든 공공영역이 데이터를 공유할 필요는 없으며 우선 유사한 기관 간의 공통영역을 찾아 내 먼저 공유하는 것이 필요합니다”

- 000(A공공기관)

“서로 간의 데이터 보유 현황 공개 및 기관 간 자체적 데이터 공유에 대한 평가 실적 인정, 타 기관의 데이터 보유현황을 알 수 없어 필요 데이터를 요청하기 어려우며 기관 간의 협의를 통해 데이터 서로 간 공유 시 평가 실적으로 인정받기가 어려움”

- 000(B공단)

“범정부 통합 플랫폼 구축, 한 플랫폼에 여러 기관이 데이터 공유하는 식으로 플랫폼을 단일화하여 정보 수집 및 관리강화”

- 000(A공기업)

결국 데이터 거버넌스의 성공적 정착은 리더십, 조직문화, 협업 구조가 상호 보완적으로 작동할 때 가능하다. 최고관리자의 비전 제시와 제도적 내재화, 문화적 전환을 통한 구성원의 참여 확대, 그리고 신뢰와 인센티브에 기반한 데이터 공유 활성화가 함께 추진될 때, 데이터 거버넌스는 형식적 규정을 넘어 실질적 행정 성과를 창출하는 제도로 자리매김할 수 있을 것이다.

3. 전문 인력과 역량 강화

데이터 거버넌스의 성패는 제도나 기술보다는 이를 실질적으로 운영하는 사람에게 달려 있다. 공공부문에서 데이터 거버넌스를 추진하기 위해서는 데이터 수집, 관리, 분석, 공유, 활용 등 전 과정에서 적절한 역량을 갖춘 인력이 필요하다. 그러나 현실에서는 전문성의 불균형, 교육·훈련 체계의 부족, 그리고 민간·학계와의 교류 미흡이라는 세 가지 문제가 반복적으로 지적되고 있다. 이러한 문제는 인터뷰 응답자들의 진술을 통해서도 확인되며, 동시에 국내외 문헌과 해외사례가 이를 뒷받침하고 있다.

첫째, 전문성의 불균형 문제는 여러 응답자들이 공통적으로 강조하였다. 한 응답자는 “실무 연계도가 높은 사례 중심 교육이나 현장 맞춤형 코칭 체계는 아직 부족해 이를 보완

할 필요가 있습니다.”고 지적하였다(OOO(B기초자치단체)). 이러한 발언은 공공부문이 단발성 프로젝트 중심의 데이터 관리에 머물러 있으며, 내재적 역량(in-house capacity)이 축적되지 못하고 있음을 보여준다.

“인적자원의 전문성 강화를 위해서는 적절한 역량과 경력을 보유한 전문가를 확보하는 것이 필요하다”

- OOO(A 협의회)

“전문가를 선발하려고 해도 선발할 수가 없습니다. 그리고 선발이 안 되고 있음에도 적극적으로 연봉을 올리지도 못합니다. 다른 부서에서 근무하는 계약직과의 형평성을 맞추기 위해서입니다”

- OOO(E대학교)

“데이터 거버넌스 성공을 위해 해당 기관별로 데이터 관리 전문 담당자가 필요합니다. 전산 업무하는 사람에게 데이터 관리 업무를 한 쪽지로 넣어 주고 하고 있다고 하다면 형식적인 업무가 될 수 있습니다. 데이터 관리를 전담하는 전문 인력이 필요합니다. 상황이 허락된다면 데이터 관리 전담 조직도 필요합니다.”

- OOO(A 책임운영기관)

이와 더불어, “실질적인 전문 인력 확보와 체계적인 역량 개발 체계 구축은 매우 부족한 상황”이라며 “전문직위 신설, 지속적인 경력 개발 경로 설계, 타 부처와의 교류 확대 등이 병행되어야 한다”는 지적도 있었다(OOO, B국책연구원). 실제로 Kim et al.(2022)의 국내 연구는 이러한 상황을 ‘역량 격차(capacity gap)’로 규정하며, 정책부서와 데이터 담당 부서의 협업 부재가 데이터 거버넌스 실패의 주요 원인이라고 지적하였다. Janssen et al.(2020) 또한 AI와 데이터 기반 행정을 위해서는 조직 전체가 일정 수준 이상의 데이터 이해도를 공유해야 하며, 일부 전문가의 고립된 전문성만으로는 정책 실행력을 담보할 수 없다고 분석한다.

해외에서도 유사한 맥락이 발견된다. 영국 중앙정부는 데이터 사이언티스트를 공무원 조직에 직접 채용하는 동시에, 일반 행정직 공무원들에게도 기본적인 데이터분석 훈련을 제공하여 역량 격차를 줄이고 있다. 반면 한국은 데이터 전담인력과 일반 공무원 간의 간

극이 여전히 크며, 이는 데이터 기반 의사결정 문화 정착을 지연시키는 원인이 된다.

전문 인력 확보 문제는 단순 인력 수급 차원을 넘어 역량 체계(capacity framework) 부재와 연결된다. OECD(2019)는 데이터 거버넌스의 핵심 과제 중 하나로 ‘데이터 인력의 역량 강화’를 제시하며, 이를 위해 기술 교육뿐 아니라 데이터 윤리, 개인정보 보호, 정책 적용 능력을 포괄하는 다층적 역량 모델을 강조한다. Janssen et al.(2020) 역시 “데이터 거버넌스는 단순한 기술 훈련으로 해결되지 않으며, 분석 결과를 정책적 맥락에 통합하는 능력이 중요하다”고 지적한 바 있다.

둘째, 공공부문의 데이터 관련 교육과 훈련 체계의 미흡 역시 일관되게 지적되었다. 국내에서는 일부 중앙부처와 지자체, 공공기관 등이 데이터·AI 교육과정을 개설하고 있으나, 여전히 파편적이고 단기적이다. 이는 OECD(2023)가 강조한 데이터 리터러시 교육의 보편화 필요성과 정확히 맞아떨어진다.

“데이터·AI 교육 과정 개설 및 단계별·수준별 교육 시행”

- 000(B공단)

“학교 행정 분야에서는 행정 실무연수, K-에듀파인 연수 등이 있지만, 데이터분석이나 시스템 활용에 특화된 연수는 부족합니다”

- 000(A기초자치단체)

“교육과정 지원 정도에 그치는 것으로 보임”

- 000(D대학교)

OECD(2024)는 한국의 데이터 교육을 “단편적이고 체계성이 부족하다”고 평가하며, 공무원 대상 역량 개발 정책의 일관성 부재를 문제로 지적하였다. 한국의 경우 ‘프로젝트 단위’ 교육에 치중하고, ‘직무 경력 경로(career path)’로서의 제도화는 미흡한 상황이라 할 수 있다.

“전문교육을 이수하도록 교육과정을 개설하고, 자격증 취득을 독려합니다. 데이터 거버넌스는 그 자체로 전문적인 분야이기 때문에, 관련 지식을 체계적으로 배울 수 있는 교육프로그램이나 자격증 취득을 장려하는 것이 중요하다고 생각합니다”

- 000(B대학교)

“데이터 담당자의 직무 전문성을 인정하는 직무경력 관리 제도도 향후 도입이 필요하다고 봅니다”

- 000(B기초자치단체)

“구성원 모두를 대상으로 하는 데이터 리터러시 교육과 실무 중심의 워크숍을 정기적으로 운영하는 방안이 필요합니다”

- 000(A대학교)

반면 해외에서는 전문 인력 강화가 제도화된 방식으로 추진되고 있음을 확인할 수 있다. 영국은 Data Science Campus(ONS 산하)를 설치하여 공무원 대상 데이터 과학 교육을 체계적으로 실시하고, 교육 이수 결과를 인사관리와 승진에 반영한다. 싱가포르의 GovTech Academy를 통해 AI, 데이터 거버넌스, 사이버 보안 등 분야별 맞춤형 커리큘럼을 제공하고 있으며, 부처별 데이터 책임자(CDO)를 지정하여 경력 경로를 제도화했다. 에스토니아 역시 디지털 정부를 이끌어가는 공무원 교육 프로그램을 국가 차원에서 운영하며, 데이터 역량을 공직 핵심 역량의 일부로 통합하였다.

교육 및 훈련과 관련하여 캐나다는 ‘Data Science Academy’를 통해 공무원 교육을 단계적(기초-중급-정책 적용)으로 설계하여, 단순한 지식 전달이 아니라 직무 수행에 직접 적용할 수 있도록 하고 있다. 뉴질랜드는 ‘Data Leadership Network’를 운영해 공무원 간 경험과 노하우를 공유한다. 미국은 ‘Federal Data Strategy - Action Plan」(2019)을 통해 모든 공무원에게 연 1회 데이터 활용 기본교육을 의무화했으며, 호주는 APS Academy를 중심으로 데이터·디지털 역량 교육을 승진 요건과 연결했다. 일본 디지털청은 2021년 설립 이후 민간 전문가와 학계 인사를 강사로 초빙해 정부 내 데이터 역량 교육을 실시하며, 일정 비율 이상의 직원이 심화 과정을 수료하도록 제도화했다.

이들 사례는 모두 데이터 전문성의 제도적 내재화라는 공통점을 지닌다. 국내 응답자들의 문제 제기는 이와 뚜렷한 대조를 이룬다. 이는 교육 자체의 양적 부족뿐만 아니라, 제도적 연계성 부재가 더 근본적인 문제임을 보여준다.

마지막으로 민간 및 학계와의 협력 부족도 중요한 제약 요인으로 확인되었다. 민간의 데이터 분석 기술의 경우 공공부문에 비해 훨씬 앞서 있는 데 반해, 제도적 장치의 미비로 이들과의 교류는 자문 수준에 그치거나 일회성으로 끝나는 경우가 많다는 지적이다.

“조직 내에서는 현재 데이터 거버넌스 관련 전문가가 부재하기 때문에 외부 전문가 및 업체와의 용역사업을 추진하여 이를 보완하고 있습니다”

- 000(입법기관)

“다양한 산업체 및 연구기관과 협력하여 주기적으로 세미나와 워크숍을 개최함으로써, 연구자들이 지속적으로 최신 기술과 지식을 습득하고 역량을 강화할 수 있는 환경을 조성하고 있습니다”

- 000(A대학교)

“산학연 협력 및 외부 전문가 초청이 대표적입니다. 실무 중심의 세미나와 특강을 진행하고 있으며, 이를 통해 최신 동향이나 우수사례를 공유하여 전문성 향상을 도모합니다”

- 000(A대학교)

해외사례는 이러한 한계를 극복하는 다양한 방식을 보여준다. 영국은 GDS(Government Digital Service)에 민간 데이터 전문가를 프로젝트 단위로 채용해 공공-민간 간 학습 효과를 극대화했다. 싱가포르의 'Data Fellowship Programme'을 통해 민간·학계 전문가를 일정 기간 정부 부처에 파견해 공동 프로젝트를 수행하게 했다. 일본 디지털청은 민간 IT기업 전문가를 일정 비율 이상 정규직 공무원으로 채용하여 데이터 정책을 주도하도록 한다. 프랑스는 공공 데이터 혁신을 위해 'Entrepreneur d'Intérêt Général' 제도를 운영, 민간 스타트업 인재가 10~12개월간 정부 부처에서 데이터 프로젝트를 수행하도록 지원한다. 이는 단순한 자문이 아니라, 민간의 전문성을 공공부문에 깊이 이식하는 제도적 장치라는 점에서 참고할 만하다.

문헌에서도 이 같은 필요성이 강조된다. Marcucci et al.(2023)은 글로벌 데이터 거버넌스 메타-프레임워크를 제시하면서, '스튜어드십'과 '책임소재'를 핵심 축으로 제시하며, 민간과 공공 간 경계를 넘어서는 협력이 제도적으로 보장되어야 한다고 주장한다. MacFeely et al.(2022) 역시 데이터의 공공성(public good)과 민간 소유권(proprietary data) 간 균형을 제도적으로 명확히 하지 않을 경우, 정책 체계는 공허해질 수밖에 없다고 지적하였다.

종합하면, 한국 공공부문 데이터 거버넌스의 인적자원 문제는 ▲전문 인력의 절대적 부

죽, ▲단기 순환 보직 구조와 파편적 교육, ▲민간·학계 교류의 제도화 부족이라는 세 가지로 요약된다. 이를 극복하기 위해서는 첫째, 정책기획-데이터분석-성과평가를 연계할 수 있는 복합형 전문가 양성이 필요하다. 둘째, 교육훈련을 일회성 워크숍이 아닌 직무경력관리제도와 연계해야 한다. 셋째, 해외사례처럼 전 직원 대상 데이터 리터러시 교육을 보편화하고, 교육 결과를 인사·승진 체계와 직접 연결할 필요가 있다. 넷째, 민간·학계 전문가와의 지속적 교류를 제도화해 내부 학습효과를 축적해야 한다.

결국, 전문 인력 강화는 데이터 거버넌스의 ‘보조 요건’이 아니라 제도의 성패를 좌우하는 핵심적 전제라 할 수 있다. 이는 기존 연구들이 제시한 “제도의 성패는 결국 사람에 달려 있다”(Ubaldi, 2013; Janssen et al., 2020)는 지적과도 일치한다.

제 4 절 법률과 제도

공공부문 데이터 거버넌스의 운영과정에서 법률과 제도 관련 논의는 네 가지 주요 영역에서 시사점을 도출한다. 첫째, 데이터기반행정법 등 관련 법제도의 영향력, 둘째, 법제도와 실무 간 발생하는 격차, 셋째, 당위론적 방향성과 현실 간 괴리, 넷째, 법제도 차원의 장애요인과 성공요인을 중심으로 살펴본다.

1. 데이터기반행정법 등 관련 법제도의 영향력

공공부문 데이터 거버넌스의 법제도적 영향은 다양한 측면에서 나타난다. 연구들은 이의 복잡성과 실무 현장에 미치는 실질적 함의를 중점적으로 분석해왔다. OECD(2022)는 데이터 거버넌스의 성공적 정착을 위해서는 명확하고 실행 가능한 정책 프레임워크 구축이 필수적이라고 강조하였으며, Marcucci et al.(2023)은 데이터 관리 전 과정에 대한 법적 근거 및 내부 규정 마련과 권한·책임·역할 명확화의 중요성을 제시하였다.

본 연구의 인터뷰 결과 공공부문 데이터 거버넌스의 정착 과정에서 가장 심각한 문제는 관련 법령 간의 목표 충돌과 해석의 혼재로 지적되었다. Kroll(2018)은 개인정보보호와 데이터 활용 간의 균형이 데이터 거버넌스의 핵심 과제라고 지적하였으며, MacFeely et al.(2022)은 데이터의 공공성과 프라이버시 보호 간 균형을 달성하기 위한 정책적·기술적 프레임워크의 필요성을 강조하였다. 본 연구의 인터뷰 응답자들은 이러한 이론적 논의가 현실에서 어떻게 구현되는지 확인할 수 있었다.

“개인정보보호법과 공공데이터법 간의 상호충돌 이슈 존재하며, 관련 법령 간의 상호충돌의 이슈가 존재한다고 봅니다”

- 000(A 협의회)

“개인정보보호법과 공공데이터법, 데이터기반행정법의 목표 간 괴리가 크고 가이드의 구체성이 부족합니다”

- 000(B 대학교)

개인정보보호법과 공공데이터법의 목표 간 괴리가 심각하며, 실무에서는 “보호와 개방이라는 두 가치 사이에서 지속적인 충돌이 발생하고 있다”는 응답은 B대학교, 겸임교수의 지적과 일치한다. 이는 Marcucci et al.(2023)이 강조한 법적 근거의 명확화가 아직 충분히 이루어지지 않았음을 시사한다.

“개인정보보호법은 정보 주체의 프라이버시를 보호하고, 정보 활용에 대한 자기 결정권을 보장하는 것을 목표로 하고, 반면 공공데이터법은 공공데이터를 적극적으로 개방하고 활용하여 새로운 가치를 창출하려는 목적을 가지고 있어서 결국 ‘보호’와 ‘개방’ 두 가치 사이의 충돌이 발생합니다.”

- OOO(B대학교)

특히 데이터기반행정법의 구조적 한계가 반복적으로 지적되었다.

“현재 공공부문의 데이터 거버넌스를 총괄하는 법률로 “데이터기반행정 활성화에 관한 법률” 및 그 하위법령이 있는데, 이 법률은 전형적인 특정산업 진흥법률의 구조 및 내용을 갖고 있어서, 실제로 데이터를 다루는 부처 및 공공기관들에게 도움이 될지는 아직 미지수이다”

- OOO(입법기관)

“데이터기반행정법은 공공기관의 의사결정에 데이터 활용을 강조하지만, 학교 현장에서는 이를 뒷받침할 인력과 시스템이 부족하여 실효성이 낮은 편. 또한 개인정보보호법과 공공데이터법 간의 충돌적 해석이 문제. 예를 들어, 학부모 민원 처리와 관련된 자료 공유 시 어느 법을 우선해야 하는지 현장에서는 혼란이 있음”

- OOO(A기초자치단체)

교육 분야에서의 경험은 법령 간 우선순위 혼란이 어떻게 현장에서 나타나는지 보여준다. Hausenloy et al.(2024)이 제시한 데이터 특성에 맞는 법·제도적 통제장치의 필요성이 실무 현장에서 절실히 요구되고 있음을 확인할 수 있다. 백승민 사무관의 “학부모 민원 처리 관련 자료 공유 시 어느 법을 우선해야 하는지 현장에서 혼란이 있다”는 증언은 Harrison et al.(2019)이 강조한 일관된 적용 체계의 부재를 보여준다.

실제로 정용찬(2018)은 정부의 법제도 기반이 부처 간 데이터 통합 및 공동 활용의 핵심

적 조건임을 강조하였다. 김지현(2023)은 데이터 규제 정책이 보호 중심에서 활용 중심으로 전환되는 글로벌 트렌드를 고려한 국내 정책 조정의 필요성을 제기하였다. 그러나 인터뷰 응답자들의 경험은 이러한 전환이 아직 충분히 이루어지지 않았음을 보여준다.

현재의 법제도는 방향성을 제공하지만 법령 상호간의 목표 상충이나 실행력 확보의 어려움, 그리고 다양한 정책 분야의 현실적 특수성이 반영되지 못하여 충분한 효과를 거두지 못하고 있음을 확인할 수 있다. Marcucci et al.(2023)이 제시한 포괄적 데이터 거버넌스 프레임워크의 구현을 위해서는 법적 정합성 확보, 실행 체계 강화, 분야별 특수성을 고려한 맞춤형 접근이 병행되어야 할 것으로 판단된다.

2. 법제도와 실무 간의 격차

기존 문헌에서 법제도의 이상적 방향성과 실무 현장의 현실 사이에는 상당한 격차가 존재한다는 문제는 반복적으로 지적되어 왔다. 조세현·차세영(2022)은 데이터 거버넌스가 공공부문에서 정착되기 위해서는 제도적 설계와 실무 환경의 조화가 필수적이라고 강조하였으며, 윤건(2023)은 공공기관별로 데이터 관리 및 연계 수준이 상이하며, 정책 수립과 집행 과정에서 실효성 있는 데이터 활용 구조가 충분히 마련되지 못한 상황이라고 진단하였다. Janssen et al.(2020)은 데이터 거버넌스가 업무 효율성을 제고하기 위한 제도임에도 불구하고, 현장에서는 추가적 과업으로 인식되어 거부감을 유발할 수 있다고 분석하였다.

실무 현장에서는 인력과 시스템 측면의 부족, 권한과 역할 분장의 불명확성 등 법제도의 실천력 확보가 어려운 상황이다.

“법과 제도는 이상적인 방향을 제시하지만, 학교 현장은 행정인력 부족과 시스템 미비로 따라가기 어려운 구조 또한 데이터 활용은 권장되지만 개인정보 유출 시 책임은 전적으로 기관에 있음”

- 000(A기초자치단체)

“법률이 현실 업무 환경을 충분히 반영하지 못해 ‘과도한 보안’과 ‘느린 처리’가 일어나고 있음. 현장은 신속한 데이터 활용을 원하지만, 법적 절차와 규제가 복잡하여 민첩성이 떨어짐”

- 000(A기초자치단체)

인터뷰 결과 교육 분야에서의 경험은 법제도와 현장 간의 격차가 어떻게 구체적으로 나타나는지 제시되었다. 개인정보보호위원회(2025)가 제시한 프라이버시 영향평가(PIA) 의무화와 민감정보 처리 시 기술적 보호조치 도입 요구는 이론적으로는 타당하지만, 실무 현장에서는 상당한 제약으로 작용하고 있다. “법과 제도는 이상적인 방향을 제시하지만, 학교 현장은 행정인력 부족과 시스템 미비로 따라가기 어려운 구조”라는 인터뷰의 지적은 윤건·김윤희(2020)가 강조한 데이터 기반 행정의 제약 조건과 일치하게 나타났다.

아울러 인터뷰 과정에서 실무자가 느끼는 법적 책임의 부담을 설명하기도 하였고, 이 과정에서 법적 책임의 부담의 원인을 법제도 해석의 모호함으로 지적되었다.

“데이터 개방 시 법적 검토를 할 수 있는 인원이 기관 내 부재하며, 데이터 개방에 따른 법적 책임이 담당자에게 있어 적극적으로 데이터 개방이 어려움”

- 000(B공단)

“법제도에 명확하게 정해져있지 않고 담당자가 자의적 판단이 필요한 경우가 종종 발생. 이에 업무 담당자들이 부담을 느끼고 업무 추진에 어려움을 겪고 있음”

- 000(B공단)

실무자들이 느끼는 법적 책임의 부담은 데이터 거버넌스 정착을 저해하는 핵심 요인으로 나타났다. Marcucci et al.(2023)이 강조한 권한·책임·역할의 명확화가 현실에서는 충분히 이루어지지 않아, 실무자 개인에게 과도한 책임이 전가되는 구조적 문제가 발생하고 있는 것으로 나타났다.

3. 당위론적 방향성과 실제적 현실 간의 괴리

데이터 거버넌스의 이론적 당위성과 실제 정책 현실 사이에는 구조적 모순이 존재한다. OECD(2019)는 데이터 기반 공공부문으로의 전환을 위한 이상적 프레임워크를 제시하였으나 실제 정책 환경에서는 정치적 영향과 조직적 저항으로 인해 구현에 한계가 있음을 인정하였다. 윤건(2023)은 데이터 거버넌스의 필요성에 대한 이론적 합의에도 불구하고, 실제 정책 수립과 집행 과정에서는 다양한 제약 요인이 작동하여 실효성이 떨어진다고 분석하였다.

특히 정치적 의사결정 구조와 데이터 기반 행정 간의 갈등이 근본적 문제로 지적되고 있다. 인터뷰에서도 데이터 거버넌스의 근본적 한계를 제기하기도 했다.

“가장 큰 문제는, 어쩌면 데이터 거버넌스의 기본 취지와 상충되는 문제일 수도 있는데, 실제로 데이터가 부족하여 또는 데이터 공유가 잘 되지 않아서 효과적인 정책 결정이 이루어지지 않는 사례들이 과연 얼마나 있을까입니다. Evidence-based policymaking의 사례에서 보듯이, 과학적/객관적 증거에 기반한 정책결정은 공공부문에 대한 정치적 영향으로 인해 쉽지 않음”

- OOO(입법기관)

인터뷰의 Evidence-based policymaking의 한계를 지적하는 것으로 “과학적/객관적 증거에 기반한 정책결정은 공공부문에 대한 정치적 영향으로 인해 쉽지 않다” 현실 인식을 반영한다. 이론적으로는 데이터 융합과 활용이 정책 효과성을 높일 수 있음에도 불구하고, 실제로는 조직 간 이해관계, 정치적 고려, 단기성과 압박 등으로 인해 제대로 실현되지 못하고 있다(윤건·김윤희, 2020).

“무엇보다도 과도한 규제라고 생각. 어떻게 발전해나갈지에 대한 것은 전문가들도 모르는 상황에서 정부가 성급하게 규제를 벌이는 것은 산업과 시장의 발전을 막는 기제. 정부 주도의 규제보다는 상향식의 의사결정을 통한 관리가 되어야 할 것이라고 생각하며, 전문가들과 시민사회의 충분한 참여가 필요하다”

- OOO(C대학교)

현재의 데이터 거버넌스 접근 방식이 과도한 규제 중심으로 설계되어 실제 혁신을 저해한다는 비판이 제기되었다. OECD(2022)가 제시한 데이터 활용과 보호 간의 균형이 현실에서는 보호 쪽으로 과도하게 기울어져 있다는 지적이 있었다. Janssen et al.(2020)이 강조한 신뢰할 수 있는 데이터 거버넌스 구축 과정에서의 균형점 찾기의 어려움을 나타냈고, 특히 “정부 주도의 규제보다는 상향식의 의사결정을 통한 관리가 되어야 하며, 전문가들과 시민사회의 충분한 참여가 필요하다”는 제안은 행정안전부(2024)가 추진하는 하향식 데이터 정책 추진 방식과 대조를 이룬다는 점에서 당위론적 참여 거버넌스와 실제 정책 추진 방식 간의 근본적 괴리를 드러낸다.

또한 인터뷰 응답을 통해 실무자들의 데이터 거버넌스에 대한 이해 부족과 형식적 접근 형태 등 공공기관의 현실을 파악할 수 있었다. 행정안전부의 데이터 품질관리 가이드라인과 같은 기술적 지침은 존재하지만 이를 뒷받침할 조직적 이해와 문화적 기반은 부족한 상황이다.

“비교적 최근에 생긴 분야다보니 데이터 제공자, 수집자, 관리자 모두 데이터 거버넌스에 대한 이해 제고 필요하다. 활용도가 뛰어난 데이터는 정보 제공자 미 동의로 활용 불가 등”

- 000(A공기업)

000(A공기업, 관계자)의 인터뷰 내용은 데이터 기반 행정 추진 과정에서의 역량 부족 문제와 직결된다. 데이터 거버넌스가 단순한 기술적 과제가 아닌 조직 전체의 인식 전환을 요구한다는 점에서 현재의 접근 방식은 여전히 표면적 수준에 머물러 있다는 것을 확인할 수 있었다.

“아직까지는 없는 것으로 알고있음. 그만큼 데이터 기반 업무가 활성화 되지 않았기 때문이라고 생각한다”

- 000(A중앙부처)

이론적으로 데이터 기반 행정의 필요성과 효용성이 광범위하게 인정받고 있으나 실제 업무 현장에서는 데이터 기반 의사결정이 충분히 활성화되지 않고 있다. 단순히 시간이 필요한 문제가 아니라 데이터 거버넌스에 대한 근본적 접근 방식의 재검토가 필요하다. 따라서 데이터 거버넌스가 조직과 정책 현장에 내실 있게 뿌리내리기 위해서는 실천적 측면과 제도적 기반을 함께 재검토할 필요가 있다.

4. 법제도 차원의 장애요인과 성공요인

공공부문 데이터 거버넌스의 성공과 실패를 가르는 핵심 변수 중 하나가 법제도적 환경임은 다양한 연구와 현장 실무자 인터뷰를 통해 일관되게 확인되고 있다. Marcucci et al.(2023)의 분석에 따르면 효과적인 데이터 거버넌스 시행을 위해서는 법적인 기반(legal

foundation), 일관된 정책 체계(policy regime), 그리고 현실적으로 작동하는 실행 메커니즘(execution mechanisms)의 세 가지 요소가 긴밀하게 결합되어야 한다. 윤건(2023)은 데이터 거버넌스 구성요소 간 구조적 관계에서 제도적 요인이 얼마나 긍정적 혹은 부정적으로 작용할 수 있는지를 실증적으로 보여주었고, Harrison et al.(2019) 또한 데이터 관리가 정책·프로세스·책임성이 조화를 이룰 때 관리체계의 실효성이 확보될 수 있음을 강조하였다. 실제로 공공 데이터 거버넌스가 안정적으로 정착되기 위해서는 법제도가 데이터의 생산·관리·활용에 대한 명확한 지침을 제공할 뿐 아니라, 조직간 갈등, 해석 불명확, 책임 회피 등 현장의 실질적 쟁점을 해소할 수 있어야 한다. 그렇지만 현실에서는 다수의 법제도적 장애요인이 데이터 거버넌스의 내실 있는 정착을 방해하는 요소로 중첩되어 나타난다.

가장 빈번하게 지적되는 장애요인은 개인정보보호법과 데이터활용 관련 법률(예: 공공데이터법, 데이터기반행정법) 간의 목표 충돌과 해석상의 불명확성이다. 현행 데이터 거버넌스 관련 법령들이 구체적 실행력을 갖추지 못한 채 선언적 수준에 머물러 있으며 인터뷰 결과 현장에서도 동일하게 “각 기관의 데이터 관리 계획 수립이나 CDO 지정 등이 의무화되어 있음에도 이에 대한 실질적인 지원 체계나 평가 기준이 미비해 형식적인 이행에 그치는 경우가 많다”는 응답은 관계부처 합동(2024)이 제시한 기본계획의 실행 전략에 대한 근본적 재검토 필요성을 시사한다.

“공공기관의 데이터 관리·활용에 관한 법령이 지나치게 형식적이거나 선언적인 수준에 머물러 있어 현장에서는 구체적인 실행력을 갖기 어려움. 각 기관의 데이터 관리 계획 수립이나 CDO 지정 등이 의무화되어 있음에도 이에 대한 실질적인 지원 체계나 평가 기준이 미비해 형식적인 이행에 그치는 경우가 많음”

- 000(B국책연구원)

데이터 거버넌스에서 장애요인 중 하나는 데이터 활용 과정에서 발생할 수 있는 문제에 대한 책임 소재가 불분명하다는 점이 언급되었다. OECD(2022)가 강조한 데이터 전 생애주기에 걸친 책임 체계 구축이 현실에서는 충분히 이루어지지 않고 있다.

“데이터 소유권 및 책임 소재의 불명확성이 가장 큰 장애요인. 공공 데이터를 활용하는 과정에서 문제가 발생하더라도, 해당 데이터를 배포한 공공기관이 소유자임에도 불구하고 신속한 대응이 어려운 경우가 많음. 왜냐하면 데이터를 가공하여 제공한 기

관은 기업이나 연구소인 경우가 많기 때문”

- 000(A대학교)

“책임주체가 불분명하여 분쟁과 위협의 요소가 자리잡고 있음. 개인정보가 담겨있는 데이터의 경우 가명정보를 결합하여 운영하면 된다고하지만 이에 대한 결합 책임, 사고 재식별 사고 책임에 대해 규정하기가 쉽지 않음”

- 000(A공단)

특히 가명정보 처리 과정에서 “결합 책임, 사고 재식별 사고 책임에 대해 규정하기가 쉽지 않다”는 문제는 개인정보보호위원회(2025)가 제시한 AI 자동결정에 대한 설명요구권, 디지털 취약계층 보호 등의 정책 방향과 실무 현실 간의 간극을 보여주었다.

“기관에서는 개인정보보호법을 우선으로 하기 때문에 이슈가 발생하지는 않음. 기본적으로 공공데이터법은 진흥에 관련된 법이고 개인정보보호법 제재에 대한 법률이기 때문에 인센티브보다는 패널티에 더욱 민감하게 대응”

- 000(A공공기관)

“데이터 활용시 유출에 대한 제재 등이라고 생각됩니다. 또한 데이터 관련 법률이 제정되어도 강제력이 없어 실제적인 활용이 어려움”

- 000(A공공기관)

“개인정보보호법이나, 보안관련된 규정이 장애요인이 될 수 있으며, 개인정보 보호와 보안에 관한 규제가 엄격한 반면에 데이터를 공유하거나 활용하는데 있어서 이러한 규제가 데이터 개방을 소극적으로 만드는데에 기여”

- 000(A중앙부처)

실제 공공기관에서는 개인정보 유출 및 규제 위반에 따른 패널티와 부담에 더 민감하게 움직이는 경향을 보이고 있었다. 이는 인센티브 중심보다는 패널티 리스크 관리가 현장 조직 전략으로 나타난다는 의미이다.

아울러 법제도의 부정적 측면에도 불구하고 성공요인과 긍정적 효과로는 업무 명확화로 인한 안정감이 언급되었다.

“법제도로 권한과 책임을 명확히 하여 업무에서 안정감을 줌”

- 000(B중앙부처)

반대로, 법제도가 긍정적으로 작동해온 영역도 있다. 대표적으로는 권한과 책임 명확화를 통한 업무 안정성 제고, 표준화와 품질관리의 유도 효과 등이 있다. 함께 공공 이익 실현에 필수적인 정책 및 연구 목적의 데이터 개방 제한을 완화할 필요가 높아지고 있다는 제안도 있었다.

“개방과 공유는 장려하되 품질과 계보 규정은 권고 수준에 그치는 경우가 많음. 따라서 표준 메타데이터와 품질 KPI를 의무화해야 함”

- 000(A공단)

“정보보호 또는 개별 시·군에서의 반발 이슈로 인해 원하는 수준의 데이터를 확보하지 못한 경우가 있음(가공 자료만 공개). 정책 수립과 제안이라는 공익적 목적으로 수행하는 연구에 있어서는 데이터 공유(공개) 제한의 범위를 완화할 필요가 있다고 생각한다”

- 000(A국책연구원)

근본적으로 법제도적 장벽의 이면에는 조직 간 알력, 현실 외면적 지침, 리더십 및 실행 의지 결여 등 사회적·조직적 요인도 깊이 자리하고 있다. 이러한 법제도의 문제점을 조직적 차원으로 접근하기도 하였다.

“법제도적 장애는 결국 법제도를 구현시키는 조직간의 알력다툼이나, 현실을 반영하지 못한 지침, 그리고 법제도를 실현하고자 하는 의지와 실행력이 떨어지는 리더와 직원이라고 보여짐”

- 000(A협의회)

인터뷰 결과 공공부문 데이터 거버넌스의 법제도는 실질 및 선언적 측면에서 모두 조직 행위와 성과에 중요한 영향을 미치지만, ▲이해관계자 간 책임소재 불명, ▲실효성 미비, ▲과도한 패널티와 공유 위축 등 구조적 한계가 혼재한다. 한편, 제도적 법적 기반 강화, 표준화 및 평가체계 확립, 현장 실행 지원 등 입체적인 개선전략이 병행될 경우 법제도는

데이터 거버넌스 혁신의 전제조건이자 안정적 촉진요소로 기능할 수 있다. 앞으로는 데이터 정책 환경 전반을 총괄하는 실행력 있는 법령 체계 확립, 인센티브 중심의 조직문화, 그리고 현장 경험을 반영한 정책 피드백 루프의 제도화가 선행되어야 할 것이다.

제5절 소 결

1. 데이터

본 장의 이론적 고찰과 인터뷰 분석을 종합하면, 공공부문의 데이터 거버넌스는 제도·기술·조직 측면에서 상호 보완적으로 작동해야 함에도 불구하고, 현재는 기관별 상이한 운영 방식과 낮은 연계 수준이 결합되어 구조적 제약을 낳고 있음을 확인할 수 있다. 기존 연구에서도 지적된 기관 간 이질적 관리체계와 표준 부재, 낮은 상호운용성은 인터뷰에서 구체적 사례로 재현되었다(OECD, 2022; Harrison et al., 2019; 오윤경·차남준, 2023). 입법 기관, 정책연구원, 대학, 책임운영기관 등 다양한 주체의 진술은 내부 공유의 부재, 코드·분류·시간 단위의 불일치, 메타데이터 미흡, 수작업 결합의 상시화가 데이터 흐름을 단절시키고 비용을 누적시키는 경로를 뚜렷이 보여준다.

데이터 생애주기 관점에서는 생성·수집, 제공·가공, 분석·활용으로 이어지는 선형 모형이 현장에서는 단속적·비효율적으로 작동한다는 점이 반복적으로 확인되었다. 생성·수집 단계에서는 기관·업무별 구축 기준의 불일치와 체계적 수집·축적의 미비가 이후 단계의 정합성을 훼손하고, 제공·가공 단계에서는 접근 권한의 제한, 비체계적 제공 방식, 표준화 및 자동화 부재로 재현성과 생산성이 저하된다. 분석 단계에서는 객관성·정합성을 담보할 전문 인프라가 부족하여 개인 역량에 대한 과의존이 나타나고, 활용 단계에서는 내부 인수인계의 비체계와 기관 간 공유 부재로 재수집·중복 처리가 반복되는 문제가 확인되었다. 이러한 단계별 문제는 결국 ‘저장’ 중심 관행을 고착화하여 ‘활용’ 중심의 거버넌스 목표와 괴리를 확대한다.

기존의 데이터 정책은 개방한 데이터의 양적 성과에 치중한 나머지, 실질적인 활용 가치를 창출하는 데는 한계가 있었다. 따라서 향후 데이터 거버넌스는 ‘문제 해결 중심(Problem-Solving Oriented)’으로 접근 방식을 근본적으로 전환해야 한다. 이를 위해 교통, 복지, 재난, 환경 등 부처와 지자체의 핵심 정책 과제를 먼저 식별하고, 해결에 필요한 데이터 세트를 역으로 정의하여 수집·정비·연계 체계를 설계가 요구된다.. 아울러 성과 평가 체계 또한 단순 개방 실적이 아닌 ‘활용 사례(Use Case)’ 중심으로 개편하여, 데이터 개방과 공유가 실제 행정 문제 해결로 직결되도록 정책을 재설계가 필요하다.

조직 역량의 측면에서는 전담 조직의 부재, 전문 인력의 부족과 높은 이직·비정규성, 현장 업무 과중으로 인한 데이터 현행화 곤란이 데이터 관리의 지속가능성을 위협하는 공통 요인으로 드러났다. 특히 수도권과 지방 간 인프라·인력의 불균형, 중앙-지방 간 권한 배분과 역량 격차가 결합된 구조적 문제임을 시사한다. 이는 내부 관리 역량이 점진적으로 개선되더라도, 기관 간 연계·협력 메커니즘이 부재한 한 거버넌스의 균형적 정착이 어려움을 의미한다.

이상의 진단은 몇 가지 시사점을 제공한다. 첫째, 부처 간 공통 분류 체계와 코드·시간 단위·메타데이터의 표준화를 통해 연계용 공통키와 품질 규칙을 제도화해야 한다. 둘째, 제공·가공 단계의 자동화와 파이프라인화를 통해 수작업 결합·중복 처리를 최소화하고, 품질 점검과 버전·이력 관리를 상시화해야 한다. 셋째, 분석 단계의 객관성과 재현성을 확보하기 위해 기관 차원의 분석 지원체계와 책임성 있는 검증 절차를 마련하고, 디지털 트윈을 포함한 의사결정 지원 시스템에는 데이터·모델 품질 보증과 사전·사후 검증을 내재화해야 한다. 넷째, 중앙-지방 간 상호책임적 데이터 거버넌스 관계를 구축하여 회신·공유·활용의 선순환을 제도화하고, 지방의 인프라·인력 격차 해소를 위한 표준 데이터셋 보급과 역량 강화 투입을 병행해야 한다.

해외 사례에 대한 인터뷰의 제언은 이러한 개선 방향을 구체화하는 준거 틀을 제공한다. 다기관 통합 플랫폼과 실시간 연계 법제화(핀란드 112), 공익 목적의 민간 데이터 접근 절차·보상·보안 장치의 제도화(EU Data Act), 내부 개발 역량을 축으로 한 표준·보안 일괄 적용(싱가포르 GovTech), 중앙 차원의 데이터 표준·API·메타데이터 강제화 CDO 책임 체계(영국 GDS), 그리고 최고위 권한 하의 데이터 조직과 보상 체계 정비(미국 모델)는 국내의 제도·조직·기술 여건에 맞춘 점진적 이식 가능성을 보여준다. 개인정보 보호를 전제로 한 안전한 연결(PPRL), 성과·영향 중심의 평가 프레임(Results Chain) 등은 활용 단계의 질을 높이는 실무적 수단으로서도 타당하다. 핵심은 단순한 제도 이식이 아니라, 국내 조직 문화와 법·기술 환경을 고려한 단계적 제도화와 운영 규범의 정착이다.

다음 장에서는 본 장에서 도출된 진단과 시사점을 토대로 공공기관 데이터 거버넌스의 전략을 체계화하는 데 초점을 둔다. 표준화와 메타데이터 관리, 제공·가공 단계의 자동화와 파이프라인 구축, 기관 간 공유·연계 프로토콜과 책임성 있는 조정 메커니즘, 중앙과

지방의 균형 있는 역할 분담, 전문 인력과 전담 조직의 강화, 활용 중심의 평가와 환류 체계 등 앞서 확인된 과제를 전략의 핵심 축으로 정교화하고자 한다.

2. 기술과 시스템

공공기관들은 기술적 측면에서 AI·통계·시각화 도구와 플랫폼·대시보드·클라우드 시스템을 빠르게 도입하고 있다. 그러나 표준화와 상호운용성 부족, 품질 관리와 보안-활용 간 균형 문제, 사용자 경험(UX)과 디지털 역량 결핍으로 인해 도입-활용 간 구조적 간극이 발생하고 있음을 확인할 수 있었다.

인터뷰 결과, 다수 기관에서 생성형 AI(ChatGPT, Perplexity 등)를 문서 요약·정리·검색 보조 등에 활용하고 있는 것으로 나타났다. 다만 정확성 검증을 위한 추가 절차가 필요하다는 응답이 다수였으며, 연구·정책 분석의 핵심 업무에서는 여전히 SPSS·STATA·SAS·R·Python 등 전통적 통계·분석 도구가 중심을 이루고 있다. 시각화와 데이터 접근성 강화를 위해 ElasticSearch 기반 시각화나 대시보드 도구를 병용하고 있으나, 사용자 수준과 부서 특성에 따른 활용도 편차가 크다. 오픈소스 도구는 낮은 진입 비용과 풍부한 생태계라는 장점을 제공하지만, 보안·지원·책임성 측면에서 상용 솔루션과의 트레이드오프가 존재한다는 의견이 확인되었다.

현재의 ‘AI 보조 + 전통 통계 코어’ 이중 구조는 AI 시대 데이터 거버넌스의 핵심 범주인 편향 완화, 공정성, 프라이버시, 추적성(라인리지), 안전한 모델 배포를 실무에서 아직 절차화·내재화하지 못했음을 보여준다. 생성형 AI 활용이 문서화·탐색 보조 수준에 머물면서 데이터 품질·표준화·보안·책임분담 등 기반 통제가 약한 상태에서는 알고리즘 투명성과 책임성을 확보하기 어려워진다. 이는 이론적 고찰이 제시하는 AI 데이터 거버넌스의 7대 범주 요구와 일치한다. 기술 스택의 다양성 자체는 긍정적이지만, 현재의 ‘정확성 검증을 수반한 보조적 AI 활용과 전통 통계의 코어 운용, 시각화의 선택적 결합’ 패턴이 전략적 가치로 전환되려면 표준화된 데이터, 품질관리 체계 위에 MLOps/DataOps적 운영통제가 결합되어야 한다.

시스템 측면에서는 교육·재정 분야의 NEIS(나이스), K-에듀파인, dBrain(디브레인), 열린재정 등 업무 특화 시스템들이 높은 완성도를 보인다. 그러나 인터뷰에서는 이들 시스템

간 자동 연계 부재, 중복 입력, 수작업 가공 문제가 반복적으로 지적되었다. 중앙부처와 지자체 간 이기종 예산시스템(E-호조 등) 연계의 한계와 동일 데이터의 부서 간 용어체계 상이함은 전형적인 상호운용성 실패 양상을 보여준다. 클라우드·협업 플랫폼(G드라이브, 웨어박스 등) 도입도 확산되고 있으나, 많은 조직에서 개인 백업용 수준에 머물거나 보안규정으로 인해 실제 협업과 공유가 제한적으로 나타났다.

이러한 병목은 기관별 편차가 크고 연계·통합 부족으로 인한 효율 저하, 보안규제로 인한 AI 분석툴 활용 제약, 포맷 불일치로 인한 품질 저하를 야기한다. 해결책으로는 통합 거버넌스 플랫폼, 사용자 친화적 UI, 자동 검증 시스템, 중앙 데이터 카탈로그 구축이 제시되는데, 이는 인터뷰에서 확인된 현장의 요청(연계 자동화·중복 최소화·정합성 자동검증·권한분배 명료화·직관적 UI)과 일치한다.

도입·활용 간극을 축소하는 경로는 개별 솔루션의 성능 제고가 아닌 전 주기 데이터 거버넌스에 맞춘 통합 인프라와 공통 표준, 중앙 카탈로그·검증·권한체계를 갖춘 플랫폼적 접근이다. 이는 이론적 고찰에서 확인된 OECD 공공데이터 혁신 모델이 제시하는 전략-전술-전달의 3단계 구조와 6대 구성요소와 부합하며 특히 기관 간 협력을 위한 공통 표준과 기술 인프라 확충 권고는 국내 상황에서 지목되고 있는 데이터 연계 부족·품질 저하·표준화 미흡이라는 핵심 병목 해결의 방향성을 제시한다.

3. 조직과 관리

조직과 관리 측면에서 데이터 거버넌스는 단순한 기술적 과제가 아닌, 조직 문화의 근본적인 전환과 리더십의 역할에 달려있음을 확인하였다. 이론적 논의에서 중요하게 다루어지는 ‘전략적 리더십’, ‘전담 조직의 구축’, ‘구성원 참여’의 중요성은 이해관계자 인터뷰를 통해 구체적인 문제점과 한계로 재확인되었다. 즉, 이론적 당위성은 존재하지만, 실제 정책현장에서는 조직 문화와 관행의 제약으로 인해 그 실효성이 낮은 것으로 나타났다. 이러한 문제들을 해결하고 향후 전개될 공공기관 데이터 거버넌스 전략을 체계화하기 위해 다음과 같은 점을 생각해 볼 수 있다.

첫째, 리더십과 실행 의지의 내재화이다. OECD(2022)의 연구에서 강조된 최고 관리자의 전략적 리더십은 이해관계자 인터뷰 결과, 그 부재가 데이터 거버넌스의 실행력을 저해하

는 핵심 요인으로 나타났다. 리더십의 부재는 데이터 정책을 형식적인 과제로 전락시켰고, 조직 구성원들의 참여를 저해하는 핵심 요인으로 작용했다. 따라서 향후 데이터 거버넌스의 실행 전략과 관련하여 최고데이터책임자(CDO)의 역할과 권한을 명확히 하고, 이들의 데이터 기반 의사결정이 실질적인 조직 성과로 이어지는 체계를 포함해야 할 것이다.

둘째, 성과 연계형 인센티브 및 문화로의 전환이 필요하다. OECD(2019)가 제시한 데이터 활용과 관련한 협업 및 인센티브의 중요성이 실제 현장에서는 '추가적인 행정 부담'으로 인식되는 경향이 있었다. 이러한 인식을 극복하기 위해서는 데이터 공유 및 활용 성과가 개인 및 부서의 인사, 예산, 평가에 실질적으로 반영되는 보상 체계가 마련되어야 한다. 이는 단순히 규정과 의무를 강제하는 하향식 접근을 넘어, 구성원들의 자발적 참여를 유도하는 상향식 동기 부여의 기반이 될 것이다. 향후 전략에서는 이러한 인센티브 제도를 구체적으로 설계하여, 데이터 활용이 일상적인 업무 관행으로 정착될 수 있도록 유도해야 한다.

셋째, 데이터 전문 인력 양성 및 협업 체계 구조화가 요구된다. 이해관계자 인터뷰에서 전문 인력의 부족과 부서 간 협력 부재가 반복적으로 지적되었다. 이러한 문제는 이론적 고찰에서 제시된 협력적 거버넌스 모델(Chen & Lee, 2021)의 실현을 가로막고, 데이터 사일로(Data Silo)를 심화시키는 구조적 원인으로 작용하였다. 또한 Schoeman(2023)이 강조한 전문 인력과 역량의 중요성에도 불구하고, 현장에서는 데이터 업무 담당자의 전문성 부족과 데이터 업무 관련 교육의 과편적 교육체계가 확인되었다. 따라서 향후 전략에서는 중앙과 지방, 공공과 민간 간 유기적인 협력 체계가 포함되어야 할 것이다.

아울러, 공공데이터와 민간데이터의 연계는 데이터 기반 행정과 AI 활용 정책의 효과를 극대화하기 위한 핵심 요소로 지속적으로 강조되어 왔다. 특히 AI 기반 정책예측, 위험 분석, 사회문제 해결형 알고리즘 개발 등은 공공부문과 민간부문이 보유한 데이터의 상호 보완적 결합을 통해 가장 큰 효과를 발휘한다. 그러나 실제 정책 환경에서는 법·제도적 충돌, 기술적 한계, 신뢰 부족, 경제적 유인 미비 등 다양한 장애 요소로 인해 공공과 민간 데이터 연계가 충분히 이루어지지 못하고 있는 상황으로 판단된다.

이에 데이터 리터러시(Data Literacy) 교육을 보편화하고, 데이터 담당자의 직무 전문성을 인정하는 경력 관리 제도를 도입함으로써 데이터 거버넌스 역량을 전반적으로 향상시켜야 한다.

결론적으로, 조직과 관리 측면의 문제점들은 단순히 기술이나 제도적 결함이 아닌 조직 내 인간적 요소와 문화적 관행에 깊이 뿌리내리고 있음을 알 수 있다. 따라서 공공기관 데이터 거버넌스 전략은 이러한 문제를 해결하기 위해 리더십의 전략적 역할, 성과 중심의 인센티브, 그리고 협력적 조직문화를 핵심 축으로 삼아야 하며, 이는 데이터 기반 행정이 단기적 정책 과제가 아닌 지속가능한 체계로 자리잡는 기반이 될 것이다.

4. 법률과 제도

공공부문의 데이터 거버넌스는 법제·기술·조직 측면에서 상호 보완적으로 작동해야 함에도 불구하고, 현재는 기관별 상이한 법령 체계와 낮은 제도 연계 수준이 결합되어 구조적 제약을 낳고 있음을 확인할 수 있다. 기존 연구에서도 지적된 기관 간 이질적 법제 운용과 표준 부재, 낮은 상호운용성은 인터뷰에서 구체적 사례로 재현되었다. 중앙부처, 지자체, 공공기관 등 다양한 주체의 진술은 개인정보보호 규범과 활용 법령의 충돌, 법적 근거의 불명확성, 해석 지침의 부재, 기관별 상이한 절차가 데이터 흐름을 단절시키고 비용을 누적시키는 경로를 뚜렷이 보여준다.

인터뷰 결과 공공부문이 직관과 관행 중심의 정책 수립에서 벗어나 과학적·데이터 기반 정책으로 전환하기 위해 거버넌스가 구조적 토대가 되어야 한다는 점이 반복적으로 확인되고 있다. 데이터 거버넌스의 개념 역시 단순한 관리 규범에서 출발해 e-거버넌스를 거쳐, AI 시대에 이르러 통합적·참여적·윤리 지향의 체계로 진화해왔다. 특히 AI의 확산은 편향 완화, 공정성, 프라이버시, 추적성, 안전한 모델 배포 등 핵심 범주를 거버넌스의 필수 의제로 올려놓았고, 알고리즘의 투명성과 책임성이 제도적 과제로 부상했다.

법제적 관점에서 현재 우리나라 데이터 거버넌스의 핵심 쟁점은 개인정보보호법과 공공데이터법 간의 충돌 지점에 집약된다. 인터뷰 참여자들은 개인정보보호법의 엄격한 해석으로 인해 공익적 목적의 데이터 활용조차 위축되는 있는 점과, 법령 간 상충으로 인한 해석의 모호성이 실무진의 위험 회피적 행태를 강화하고 있다고 지적했다. 특히 가명정보 처리에 관한 개인정보보호법 제28조의2와 공공기관의 정보공개에 관한 법률 간 적용 우선순위의 불명확성, 데이터3법 개정 이후에도 여전히 남아있는 규제 공백과 과잉 규제 영역의 혼재가 현장 혼란을 가중시키고 있다.

주요국의 법제 사례는 규범적 명확성과 실용적 균형의 중요성을 보여준다. 미국의 연방 데이터전략법(Federal Data Strategy)은 데이터를 전략적 자산으로 규정하여 연방기관의 데이터 관리·활용 의무를 법제화했고 각 기관의 CDO에게 법적 권한과 책임을 부여했다. 영국의 데이터보호법(Data Protection Act 2018)은 GDPR 준수와 함께 공익적 목적의 데이터 처리에 대한 명확한 예외 조항을 두어 공공 부문의 혁신적 활용을 뒷받침하고 있다. EU의 데이터거버넌스법(Data Governance Act)은 공공 부문 데이터 재사용을 위한 단일 정보 지점(Single Information Point) 설치를 의무화하여 절차적 투명성을 확보했다.

이러한 해외 사례들로부터 도출되는 법제적 시사점은 다음과 같다. 첫째, 데이터 보호와 활용 간 균형을 위한 명확한 법적 기준과 예외 조항의 설정이다. 둘째, 데이터 거버넌스 전담 기구의 법적 지위와 권한의 명문화이다. 셋째, 기관 간 데이터 연계·공유를 위한 법적 근거와 절차의 표준화이다. 넷째, 공공성·투명성 기반의 책임성 확보 메커니즘의 법제화이다.

국내 법제 개선을 위한 구체적 과제를 살펴보면 우선 개인정보보호법상 공익적 목적에 대한 해석 기준을 명확히 하고 공공기관 간 데이터 연계를 위한 법적 근거를 강화해야 한다. 현행 개인정보보호법 제18조(개인정보의 목적외 이용·제공)와 제26조(업무위탁)의 해석·적용 기준을 구체화하여 실무진의 예측가능성을 높이는 것이 급선무이다. 특히 공공의 안전과 복리의 추상적 기준을 세분화하여 교육·보건·안전·환경 등 영역별 구체적 판단 기준을 제시할 필요가 있다.

둘째, 데이터 거버넌스 전담 기구의 법적 근거 마련이다. 대통령령에 근거하는 위원회의 경우 법적 안정성과 지속성이 취약하기에 법률 차원에서 데이터 거버넌스 컨트롤타워의 지위·기능·권한을 명시하고, 각 기관 CDO의 법적 지위와 책임을 규정해야 한다. 이와 함께 기관 간 데이터 연계·활용 분쟁 시 조정 권한과 절차를 법제화하여 실질적 조정 기능을 확보해야 한다.

셋째, 공공데이터 개방·활용을 위한 법제 정비이다. 현행 공공데이터법은 개방 중심으로 설계되어 있어 공공기관 간 내부 활용에 대한 근거가 미흡하다. 공공기관의 보유 데이터를 “공공자산”으로 명시하고 기관 간 공유·활용을 원칙으로 하되 보안·개인정보보호상 예외를 명확히 하는 방향으로 법제를 재편해야 한다. 또한 민간 데이터의 공익적 활용을 위한 법적 근거와 보상 체계, 데이터 신탁 등 중간기구 활용 방안도 법제화가 필요하다.

넷째, AI·빅데이터 시대에 부합하는 새로운 규범의 도입이다. 현재 논의 중인 AI기본법과의 정합성을 고려하여 알고리즘 투명성, 자동화된 의사결정의 책임성, AI 편향 방지 등에 대한 법적 기준을 마련해야 한다. 특히 공공 부문의 AI 활용에서 설명가능성과 인간의 최종 판단권 보장을 법제화하여 시민의 권익을 보호해야 한다. 다섯째, 지방자치단체의 데이터 자치권 확대를 위한 법적 뒷받침이다. 현재 중앙집권적 데이터 거버넌스 구조 하에서 지방의 창의적 데이터 활용이 제약받고 있다. 지방자치법상 자치사무 범위에서 데이터 거버넌스의 지위를 명확히 하고, 지방 특성에 맞는 데이터 정책 수립·시행 권한을 보장해야 한다. 이와 함께 중앙-지방 간 데이터 연계·공유 의무와 지원 체계를 법제화하여 균형 잡힌 데이터 연방주의를 구축해야 한다.

결론적으로 한국의 데이터 거버넌스 법제는 보호 중심에서 균형적 활용으로 분절적 규율에서 통합적 체계로, 중앙집권적 통제에서 협력적 거버넌스로의 전환이 필요하다. 이를 위해서는 개별 법률의 개정을 넘어 데이터 거버넌스 기본법 제정을 통한 종합적·체계적 접근이 요구된다. 데이터를 단순한 정보가 아닌 공공자산이자 정책 도구로 규정하고 이를 책임감 있게 활용할 수 있는 법적 기반을 구축할 때 데이터 기반 행정의 실질적 구현이 가능할 것이다.

제 4 장 공공기관 데이터 거버넌스의 전략

제 1 절 프레임워크 및 전략 도출

1. OECD 기본 프레임의 수정과 고도화

[그림 4-1] 은 앞서 수행된 문헌분석과 면접조사의 결과를 바탕으로 도출된 통합 데이터 거버넌스 모형을 제시한다. 통합 데이터 거버넌스 모형은 크게 수요자(예: 시민, 기업)의 필요와 요구에 관한 요소들로 구성되는 환경 층위와 공급자(예: 행정기관, 공공기관)의 자원과 역량에 관한 요소들로 구성된 제도·조직·기술 층위로 구분된다. 넓은 맥락에서, 환경 층위는 제도·조직·기술 층위에 데이터 및 인공지능 기반 행정에 대한 필요와 요구를 투입하고, 제도·조직·기술 층위는 환경 층위의 필요와 요구에 부응하는 제품 또는 서비스를 산출물로서 제공한다. 좁은 맥락에서, 제도→조직→기술로 연결되는 제도·조직·기술 층위는 데이터와 인공지능을 포함하는 기술의 도입·운영 과정을 보여주고(제도에 의한 기술의 도입·운영 관련 거시 방향·지침 제시, 조직에 의한 필요 기술 도입·운영), 기술→조직→제도로 연결되는 제도·조직·기술 층위는 데이터와 인공지능을 포함하는 기술의 적용·활용 과정을 보여준다(조직에 의한 필요 기술 적용·활용, 필요 기술 적용·활용을 위한 조직의 필요 방향·지침 마련).

통합 데이터 거버넌스 모형의 각 층위는 다음과 같이 구성된다. 첫째, 환경 층위는 크게 순효과 구현과 역효과 차단으로 구성된다. 한편으로, 데이터 거버넌스의 수요자들은 공공 부문 기관들이 데이터 및 인공지능 기반 행정을 바탕으로 내부적으로는 내부업무 효율성을 강화하고, 외부적으로는 공공서비스 품질을 제고하고, 궁극적으로는 시민참여를 촉진하도록 요구한다. 다른 한편으로, 데이터 거버넌스의 수요자들은 데이터 및 인공지능 기반 행정으로 인해 자신이 소외되거나 자신에게 피해가 발생하지 않도록 개인정보 보호, 사이버범죄 차단, 디지털격차 해소를 요구한다. 이러한 환경 층위의 필요와 요구는 제도·조직·기술 층위에 지속적으로 전달되고, 제도·조직·기술 층위는 환경 층위의 필요와 요구에

부응하기 위해 작동한다.

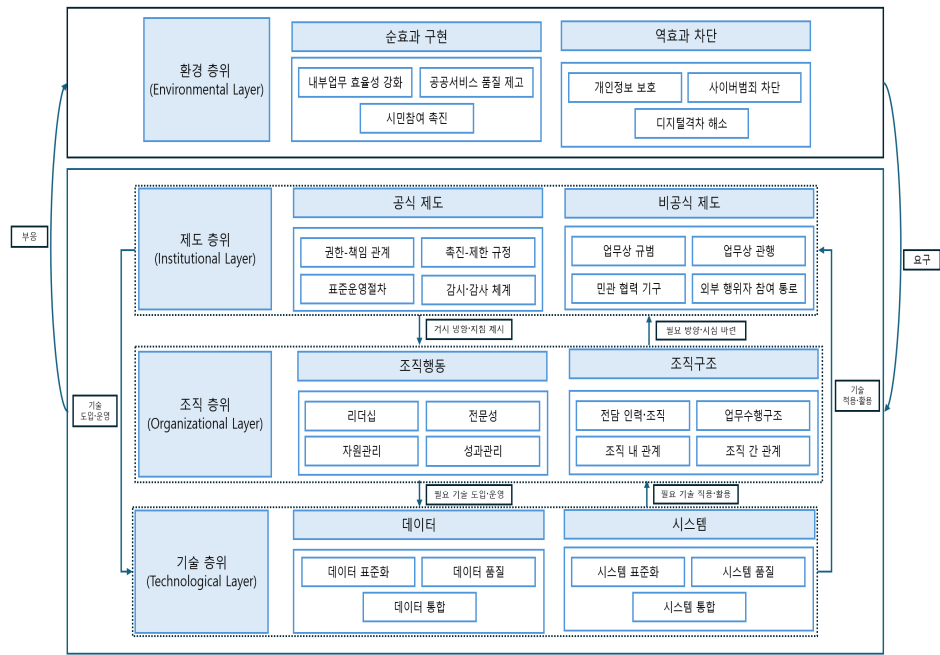
둘째, 제도 층위는 크게 공식 제도와 비공식 제도로 구성된다. 공식제도는 구성원들의 행동방침으로서의 역할을 수행하는 법률, 지침, 규범 등을 의미한다. 데이터·기술 관련 공식 제도는 권한-책임 관계(데이터·기술의 기획-집행-평가 단계에서 각 구성원에게 주어진 권한과 책임의 관계), 촉진-제한 규정(데이터·기술의 활용을 촉진 또는 제한하는 규정), 표준운영절차(데이터·기술 관련 업무에 대한 조직 차원의 공식적인 수행절차), 감시·감사 체계(데이터·기술의 오남용 여부에 대한 감시·감사를 위한 체계)를 포함한다. 비공식 제도는 공식적인 제도는 아니지만 구성원들이 바람직하다고 간주하는 가치 또는 수단을 의미한다. 비공식제도는 업무상 규범(구성원들이 당위적으로 바람직하다고 간주하는 가치), 업무상 관행(구성원들이 비공식적이지만 통상적으로 수행하는 업무 방식), 민관 협력 기구(공공부문과 민간부문의 협력을 촉진하기 위한 기구), 외부 행위자 참여 통로(민간부문의 개인·부서·기관이 공공부문에 참여할 수 있는 방법 또는 수단)를 포함한다.

셋째, 조직 층위는 크게 조직행동과 조직구조로 구성된다. 조직행동은 데이터·기술 관련 업무를 담당하는 조직의 효과성 향상을 위해 필요한 조직 내 개인 특성을 의미한다. 조직행동은 리더십(데이터·기술에 대한 기관장의 관심과 의지 및 관련 자원 동력 능력), 전문성(데이터·기술 관련 관리자·실무자의 전문지식과 경험), 자원관리(데이터·기술 업무에 필요한 인적·물적 자원의 효율적인 관리), 성과관리(데이터·기술 업무의 성과에 대한 적절한 평가와 보상)를 포함한다. 조직구조는 데이터·기술 관련 업무를 담당하는 조직의 효과성 향상을 위해 필요한 조직의 총체적·집합적 특성을 의미한다. 조직구조는 전담 인력·조직(데이터·기술 업무를 고유 업무로 전담하여 담당하는 인력·조직), 업무수행구조(수행구조·수직구조와 같이 데이터·기술 업무를 수행하기 위한 조직의 구조적 특성), 조직 내 관계(공공부문 내 개인·부서·기관 차원의 업무 관계), 조직 간 관계(공공부문과 민간부문 간 개인·부서·기관 차원의 업무 관계)를 포함한다.

넷째, 기술 층위는 크게 데이터와 시스템으로 구성된다. 데이터는 데이터 거버넌스의 운영 및 성과에 필요한 가공되지 않은 자료를 의미한다. 데이터는 데이터 표준화(데이터의 형식과 구조의 일관된 정도), 데이터 품질(데이터의 정확성·적시성·보안성 등의 질적 수준), 데이터 통합(데이터와 데이터 간 연계 및 통합을 통한 확장가능성)을 포함한다. 시스템은 데이터의 생성·수집·관리·활용·폐기 등에 필요한 기술·시설·장비의 체계를 의미

한다. 시스템은 시스템 표준화(시스템의 형식과 구조의 일관된 정도), 시스템 품질(시스템의 정확성·적시성·보안성 등의 질적 수준), 시스템 통합(시스템과 시스템 간 연계 및 호환을 통한 확장가능성)을 포함한다.

[그림 4-1] 통합 데이터 거버넌스 모형

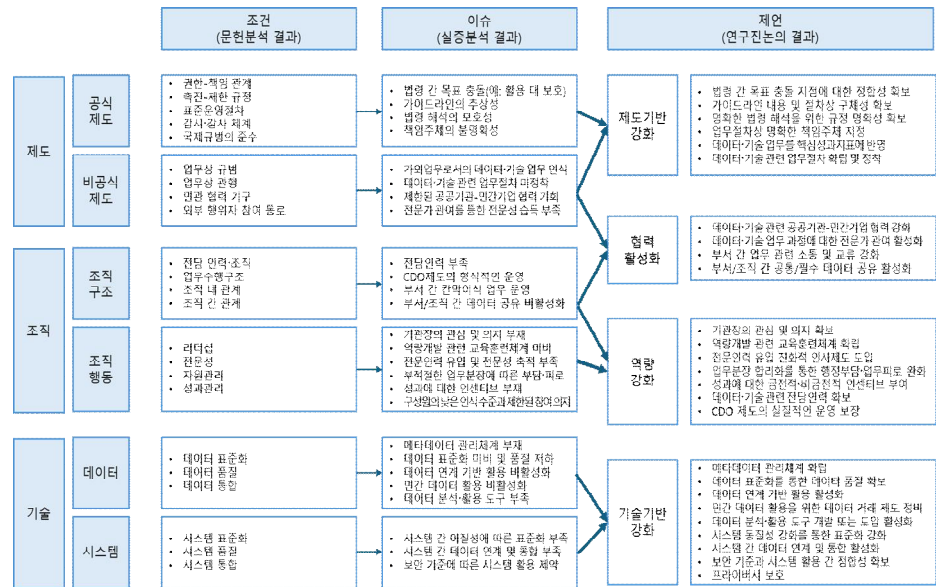


[그림 4-2]는 앞서 제시된 통합 데이터 거버넌스 모형을 구성하는 요소들을 데이터 거버넌스 활성화를 위한 조건들로 간주하고, 각 조건에 대응하는 이슈(문제점 및 제약)와 각 이슈에 대처하기 위한 제언을 연계한 결과를 제시한다. 데이터 거버넌스에 관한 이슈들은 크게 제도, 조직, 기술 차원의 이슈들로 구성된다. 첫째, 제도 차원에서는 구성원의 동기와 사기 관련 제약(예: 가외업무로서의 데이터·기술 업무 인식), 업무 과정의 명확성과 책임성 관련 제약(예: 책임주체의 불명확성), 국제규범 준수 관련 제약(예: 국제규범의 국내 적용 및 이행 제한)이 발생하는 것으로 파악된다. 둘째, 조직 차원에서는 공공부문 기관 내 조직구조 및 조직행동 관련 제약(예: 기관장의 관심 및 의지 부재, 전담인력 부족)과 공공

부문-민간부문 협력 관련 제약(예: 칸막이식 업무 운영, 외부 전문인력 유입 부족)이 발생하는 것으로 파악된다. 셋째, 기술 차원에서는 데이터 품질·연계 관련 제약(예: 데이터 표준화 미비 및 품질 저하)과 시스템 품질·연계 관련 제약(예: 시스템 간 이질성에 따른 표준화 부족)이 발생하는 것으로 파악된다.

데이터 거버넌스를 둘러싼 이슈에 대처하기 위한 제언들은 크게 제도기반 강화, 협력 활성화, 역량 강화, 기술기반 강화로 구분된다. 첫째, 제도기반 강화를 위해 구성원의 동기 와 사기를 고양하기 위한 제도(예: 데이터·기술 업무를 핵심성과지표로 지정), 업무 과정 의 명확성과 책임성을 강화하기 위한 제도(예: 업무절차상 명확한 책임주체 지정)가 확보 되어야 하고, 국제규범의 국내 적용 및 이행이 활성화되어야 한다. 둘째, 협력 활성화를 위 해 공공-공공 협력 활성화 방안(예: 부서·조직 간 공동·필수 데이터 공유 활성화)과 공공-민간 협력 활성화 방안(예: 업무 과정에 대한 전문가 관여 활성화)이 균형적으로 실행되어 야 한다. 셋째, 역량 강화를 위해 관리적 역량 강화 방안(예: CDO 제도의 실질적인 운영 보장)과 실무적 역량 강화 방안(예: 역량개발 관련 교육훈련체계 확립)이 연계적으로 실행

[그림 4-2] 데이터 거버넌스 활성화를 위한 조건-이슈-제언 연계도



되어야 한다. 넷째, 기술기반 강화를 위해 데이터 품질·연계 강화 방안(예: 메타데이터 관리 체계 확립)과 시스템 품질·연계 강화 방안(예: 시스템 동질성 강화를 위한 표준화 강화)이 동시에적으로 실행되어야 한다.

앞서 제시된 데이터 거버넌스 활성화를 위한 제언들은 실행 전략(제도기반 강화, 기술기반 강화, 협력 활성화, 역량 강화) 및 실행 시기(단기, 중기, 장기)에 따라 단기, 중기, 장기 방안들로 재구성될 수 있다([그림 4-3] 참조). 첫째, 단기 방안은 데이터 및 인공지능 기반 행정의 정착에 필요한 기초 역량 및 기반 확보이다. 단기적으로는 모든 공공부문 기관들이 기관 수준 및 지역에 관계없이 데이터 및 인공지능 기반 행정의 안정적인 정착에 필요한 인적·물적 자원과 관계적·제도적 조건을 확보할 수 있도록 지원할 필요가 있다. 역량 강화 차원에서는 데이터 및 인공지능 기반 행정에 대한 기관장의 확고한 리더십 아래 구성원에게 교육·훈련과 성과 인센티브를 부여하여 구성원의 인식 수준과 참여 의지를 고양시켜야 한다. 협력 활성화 차원에서는 내부적으로는 부서 간 업무 관련 소통 및 교류를 활성화하고, 외부적으로는 데이터·기술 업무 과정에 대한 전문가 관여를 활성화해야 한다. 기술기반 강화 차원에서는 데이터·기술의 활용 활성화에 필요한 기초적인 인프라(예: 메타데이터, 데이터 분석 도구, 보안 기준)를 확보해야 한다. 제도기반 차원에서는 구성원의 동기와 사기를 고양하기 위한 명확한 업무 책임주체 지정과 데이터·기술 업무의 핵심 성과지표 설정이 구현되어야 한다.

둘째, 중기 방안은 데이터 및 인공지능 기반 행정의 확산에 필요한 심화 역량 및 기반 강화이다. 중기적으로는 각 공공부문 기관이 데이터 및 인공지능 기반 행정의 안정적인 정착에 필요한 인적·물적 자원과 관계적·제도적 조건의 질적 수준을 고도화할 수 있도록 지원할 필요가 있다. 역량 강화 차원에서는 데이터·기술 전문인력을 확보하고 업무분장 합리화를 통해 구성원의 행정부담과 업무피로를 완화해야 한다. 협력 활성화 차원에서는 기관 내 또는 기관 간 공통·필수 데이터 공유를 활성화해야 한다. 기술기반 강화 차원에서는 데이터 품질 고도화, 데이터 연계·통합 활성화, 시스템 간 연계·통합 활성화가 구현되어야 한다. 제도기반 차원에서는 데이터·기술 관련 가이드라인의 내용·절차 구체성을 확보하고 업무절차를 정착시켜야 한다.

셋째, 장기 방안은 데이터 및 인공지능 기반 행정의 보편화에 필요한 역량 및 기반 공고화이다. 장기적으로는 각 공공부문 기관이 데이터 및 인공지능 기반 행정의 인적·물적 자

원과 관계적·제도적 조건을 각각의 고유 또는 핵심 목표를 달성하기 위한 유용하고 일상적인 수단으로 활용할 수 있도록 지원할 필요가 있다. 역량 강화 차원에서는 전문인력 유입 활성화를 위해 외부 전문인력의 공공부문 진출에 친화적인 인사제도를 도입해야 하고 CDO 제도의 실질적인 운영이 보장되어야 한다. 협력 활성화 차원에서는 데이터·기술 관련 공공기관-민간기업 협력이 내실화 및 실질화되어야 한다. 기술기반 강화 차원에서는 민간 데이터 활용을 위한 데이터 거래 제도 정비와 시스템 동질성 강화를 위한 표준화 강화가 구현되어야 한다. 공공데이터와 민간데이터의 연계에서 가장 큰 장애 요인은 법·제도적 충돌과 이로 인한 활용 제한으로, 공공부문은 개인정보보호법뿐만 아니라 복지, 교육, 보건, 노동, 금융 등 개별법에 의해 데이터 처리 목적이 엄격히 제한되며 목적 외 활용이 법적으로 금지되는 경우가 많다. 반면 민간부문은 데이터 활용 목적이 다양하고 상대적으로 유연하여, 공공의 목적 제한 규정과 결합될 때 데이터 공유가 원천적으로 어려워지는 구조가 발생한다.

특히 데이터기반행정법이 규정하는 데이터 기반 정책분석 기능과 AI기본법이 추진하는 AI 개발·검증·활용 목적 사이에는 법적 해석의 차이가 존재하여, 공공기관은 민간데이터 활용 시 법적 책임과 위반 위험을 부담해야 한다는 부담을 가진다. 또한 데이터 결합을 위해 결합전문기관을 거쳐야 하는 절차는 상당한 시간이 소요되며, 신속성이 요구되는 AI 기반 정책 대응에는 구조적 병목으로 작용하고 있다.

공공과 민간의 데이터 활용 목적이 법령에 의해 상이하게 구분되는 구조를 개선하기 위해, 데이터 활용을 규율하는 법·제도 체계를 통합적으로 정비할 필요가 있다. 목적 제한 원칙을 일률적으로 적용하는 방식에서 벗어나 데이터 활용의 공익성·필요성·적절성을 판단하는 사전심사 체계를 마련하여 보다 유연한 데이터 활용을 가능하게 해야 한다.

또한 데이터기반행정법과 AI기본법, 개인정보보호법 사이의 충돌을 줄이기 위해 '통합 데이터 활용 프레임워크'를 도입하여 공공-민간 영역에서 동일한 규범과 기준 아래서 데이터 처리와 AI 활용을 가능하게 하는 법적 기반을 마련할 필요가 있다.

아울러, 제도기반 강화 차원에서는 법령 간 목표 충돌 지점에 대한 정합성과 명확한 법령 해석을 위한 규정 명확성이 확보되어야 하고, 데이터·기술 관련 국제규범의 국내 준수 및 이행이 활성화될 필요가 있다.

[그림 4-3] 데이터 거버넌스 활성화를 위한 로드맵



데이터 거버넌스 활성화와 함께 공공기관과 민간기업이 안정적으로 협력할 수 있도록 데이터 제공과 활용에 관한 계약 조건을 표준화하고 데이터·AI 모델의 소유권과 책임 범위를 명확하게 규정하는 표준협력계약을 마련해야 한다. 또한 공공-민간 간 위험을 조정하기 위한 공동 책임 모델을 도입하여, 데이터 오류나 AI 알고리즘 문제로 발생하는 법적 위험을 한 기관이 단독으로 부담하지 않도록 해야 한다.

더 나아가 민간데이터의 품질을 객관적으로 인증하는 데이터 신뢰성 검증기관을 설립하여 데이터 안전성에 대한 신뢰를 확보하고, 공공기관이 제공받는 민간데이터의 품질과 활용 가능성을 사전에 판단할 수 있도록 지원이 요구된다.

2. 구체적인 거버넌스 전략

이상 본 연구가 제시한 프레임워크에 기반하여 구체적인 거버넌스 전략을 다음과 같이 재구성할 수 있다. 이는 Janssen et al.(2020)이 제시한 데이터 거버넌스 전략을 본 연구의 프레임워크와 연결하여 재구성한 것이다. 구체적으로 계획 기반 접근, 조직 기반 접근, 위험 기반 접근으로 나누어서 이론적 검토와 질적 분석을 통해 연구진이 검토한 주요 결과를 정리하였다(아래 <표 4-1> 참조).

먼저 계획 기반 접근은 데이터 거버넌스 관련 문제를 해결하고 목표를 달성하기 위해 단기, 중기, 장기 전략 수행계획 수립·실행·감독·평가를 중심으로 접근하는 것이다(Janssen et al., 2020). 예를 들어, 본 연구진이 제도기반 강화의 이슈로 검토한 ‘가이드라인 내용 및 절차상 구체성 확보’와 관련하여 조직 차원에서 1년 단위 또는 3년 이상의 중장기계획으로 과제를 발주할 수 있다(가칭 ‘데이터 거버넌스 가이드라인 연구’). 마찬가지로 본 연구진이 기술기반 강화의 이슈로 검토한 ‘데이터 표준화를 위한 데이터 품질 확보’와 관련해서도 조직 차원에서 “연구 데이터의 표준화·현행화 용역” 과제를 발주할 수 있다. 이는 한국의 조직에서 익숙한 방식이면서 한국에 있는 연구기관, 대학교 등의 전문기관 협업을 통해 데이터 거버넌스 전략을 실행하는 방법이라고 할 수 있다.

특히 공공과 민간 데이터의 연계 측면에서 이러한 기술·표준적 접근은 더욱 중요하다. 현재 공공데이터와 민간데이터는 구축 목적, 데이터 구조, 저장 방식, 코드 체계 등이 서로 다르기 때문에 기술적 결합 과정에서 상당한 비용과 시간이 소요되는 ‘기술적·표준적 불일치’ 문제가 발생하고 있다. 공공데이터는 행정절차와 통계 작성 목적에 따라 정형화된 형태로 구축되는 반면, 민간데이터는 사업 모델과 고객 분석 중심으로 비정형 데이터를 포함하여 다양한 형태로 생성되기 때문이다. 이로 인해 동일 개념을 보더라도 정의나 단위가 서로 달라 연계 시 데이터 손실이나 오류가 발생하며, 공공데이터의 부족한 메타데이터나 긴 갱신 주기, 민간데이터의 품질 검증 체계 부재 등은 효율성을 크게 떨어뜨리는 요인이 된다.

따라서 향후 전략은 이러한 한계를 해소하기 위해 ‘기술적·표준적 기반 강화’에 집중해야 한다. 구체적으로 국가 단위의 데이터 표준을 공공과 민간에 함께 적용하는 공통 표준 체계를 구축하여 데이터 속성, 메타데이터, 라벨링, 품질 기준 등을 공통화함으로써 연계

비용을 획기적으로 절감해야 한다. 또한, 안전한 데이터 결합을 위한 가명데이터 공간과 폐쇄형 분석 환경(Secure Data Room)을 확대하고, 민간 원본 데이터 없이도 공동 분석이 가능한 연합학습(Federated Learning) 기술을 적극 도입해야 한다. 아울러 기관 간 협력 체계를 통해 공공데이터의 정확성과 갱신 주기를 민간의 고빈도 데이터와 결합 가능한 수준으로 개선하는 것이 필수적이다.

둘째, 조직 기반 접근은 데이터 거버넌스에 부합하는 조직구조를 설계하고 조직 내 권한을 부여하는 것을 강조하는 접근이다(Janssen et al., 2020). 이 접근은 데이터 거버넌스 전략에 맞게 조직구조를 재설계하고 데이터 거버넌스로 필요한 직무를 수행하는 보직(예: CDO)을 신설하고 그에 맞는 전문인력을 채용·훈련하는 것을 중요하게 여긴다. 이 접근을 통해 본 연구진이 검토한 역량 강화와 관련된 주요 방안들을 수행할 수 있다.

나아가 조직 차원을 넘어 공공과 민간이 함께하는 생태계 기반의 강화 전략이 병행되어야 한다. 현재 민간기업은 데이터 제공에 따른 비용과 관리 부담은 큰 반면, 직접적인 경제적 이득을 얻기 어려워 공공부문의 지속적인 데이터 공유 동기가 부족한 실정이다. 이를 해결하기 위해 데이터 제공 기업에 대한 세액 공제, 기술 상용화 지원, 공동 AI 모델 개발 수익 배분 등 실질적인 경제적 보상 체계 마련이 필요하다.

또한, 공공과 민간이 공동으로 참여하는 AI 실증 프로젝트를 추진하여 공공은 정책 효과를, 민간은 비즈니스 기회를 창출하는 상생 구조를 만들어야 한다. 예를 들어, 지역별로 공공기관, 기업, 대학 등이 참여하는 ‘지역 데이터 얼라이언스’와 같은 협력형 조직 체계를 구축한다면, 지속 가능한 데이터 연계 기반과 지역 단위 데이터 생태계를 활성화할 수 있을 것이다.

〈표 4-1〉 데이터 거버넌스 수행전략

구분	도식화	내용
계획 기반 접근	<pre> graph TD objectives --> plan plan --> implement implement --> evaluate evaluate --> objectives </pre>	문제해결을 위하여 단기, 중기, 장기 전략 수행계획을 수립·실행·감독·평가함 예: “데이터 거버넌스 가이드라인 연구”를 과제발주/입찰 방식으로 수행(제도기반 강화) 예: “연구 데이터의 표준화·현행화 용역”을 과제발주/입찰 방식으로 수행 중장기계획으로 데이터 거래 제도 정비
조직 기반 접근	<pre> graph TD CIO[CIO] --> B1[] CIO --> CDO[CDO] CIO --> B2[] CIO --> B3[] CDO --> B4[] CDO --> B5[] </pre>	데이터/알고리즘 전문인력 관리부서 신설 (전문인력 채용, 교육훈련 등 담당) 데이터·기술 관련 전담인력 확보 CDO 보직의 신설 성과에 대한 금전적·비금전적 인센티브 부여
위험 기반 접근	<pre> graph LR collect --> dots[...] dots --> destroy </pre>	보안 기준과 시스템 활용 사이 부정합성 발생 경고 시스템 도입 프라이버시 침해 등 보안 문제 식별 시스템 간 데이터 오류 발생 식별 데이터 표준화 관련 체계적 오류 발생 식별

출처: Janssen et al.(2020)을 참조하여 연구진이 재구성

셋째, 위험 기반 접근은 데이터 거버넌스의 빅데이터 알고리즘 시스템에서 발생할 수 있는 위험을 식별하고 대응하는 것에 초점을 맞춘 것이다(Janssen et al., 2020). AI와 데이터 관련 위험들을 정리하고(예: 데이터 편향 위험, 개인정보 유출 위험 등), 각 위험을 단위로 대응방법을 마련하는 방식이다. 본 연구진이 정리한 기술 기반 강화와 관련된 주요 이슈들은 위험 기반 접근으로 다루어지기 적합한 것들이다. 예를 들어, 프라이버시 문제, 데이터 정합성 문제, 데이터 오류 점검 등의 이슈 위험의 종류로 분류하여 다룰 수 있다.

따라서 성공적인 거버넌스를 위해서는 이러한 위험 요인을 선제적으로 해소하는 신뢰 기반 협력 거버넌스가 구축 필요성이 강조된다. 법·제도적 정비를 통해 책임 소재와 권한

관계를 명확히 하고, 데이터 제공자와 활용자 간의 신뢰를 담보할 수 있는 투명한 절차와 안전한 활용 환경을 조성함으로써, 상호 간의 위험 인식을 낮추고 적극적인 협력을 유도
가 필요하다.

3. 데이터 거버넌스-AI 거버넌스 연계 프레임워크

가. 데이터 거버넌스-AI 거버넌스 연계 구조

본 연구에서는 데이터 거버넌스와 AI 거버넌스의 공통점과 차이점을 체계적으로 분석하여 두 거버넌스 간 연계 프레임워크를 제시하였다. <표 4-2>에 나타난 바와 같이 데이터 거버넌스와 AI 거버넌스는 네 가지 핵심 영역에서 상호보완적 관계를 형성하고 있다.

데이터 거버넌스는 조직이 보유하고 활용하는 데이터를 정확하고, 안전하고, 일관되게 관리하기 위한 체계를 말한다. 즉, 데이터를 수집·저장·가공·공유·폐기하는 전 과정에서 품질, 표준, 보안, 접근 권한, 책임 소재를 명확히 해 데이터가 신뢰할 수 있는 자산으로 활용되도록 만드는 것이 핵심이라고 할 수 있다.

반면, AI 거버넌스는 이러한 데이터를 활용해 만들어진 AI 모델과 시스템이 공정하고, 투명하며, 책임 있게 운용되도록 관리하는 체계에 집중한다. 이는 알고리즘의 편향, 설명 가능성, 안전성, 개인정보 침해 위험, AI 결과물에 대한 책임 등을 포함한다. 다시 말해, 데이터 거버넌스가 ‘재료’를 관리한다면 AI 거버넌스는 그 재료로 만든 ‘도구’가 안전하고 윤리적으로 작동하도록 보장하는 관리 체계라고 할 수 있다.

이를 세부적으로 살펴보면 다음과 같다. 정의 측면에서 데이터 거버넌스는 데이터 생애 주기 전반(수집·저장·활용·공유·폐기)에 걸친 품질·보안·표준·정책·보안·위험 관리를 통해 체계적이고 신뢰할 수 있는 데이터 활용 촉진을 목표로 하는 반면, AI 거버넌스는 시스템 모델·알고리즘의 개발·운영 전 과정에서 정책과 윤리적 원칙을 준수하면서 공정성·투명성·책임성을 확보하고 위험성을 관리하는 것으로 정의된다. 이는 디지털 자원의 신뢰성 확보를 위한 제도적 토대와 조직적 활동을 의미한다는 공통점을 가지면서도, 데이터는 정보자원 중심, AI는 알고리즘과 의사결정 시스템 중심이라는 차이점을 보인다.

목표 영역에서 데이터 거버넌스는 데이터 품질 보장, 상호운용성 개선, 보안 강화 등에 초점을 맞추는 반면, AI 거버넌스는 알고리즘의 공정성 실현, 가능성 확보, 사회적 책임 수

행을 중시한다. 양자 모두 공공 서비스 품질 향상과 시민 신뢰 제고라는 궁극적 목표를 공유하지만, 데이터 거버넌스는 정합성과 보안, AI 거버넌스는 공정성과 윤리에 각각 중점을 둔다.

관리 대상 측면에서 데이터 거버넌스는 원천 데이터, 메타데이터, 공공/민간 데이터를 포괄하며, AI 거버넌스는 모델, 알고리즘, 학습데이터, 예측결과 등을 포함한다. 공통점으로는 두 거버넌스 모두 디지털 자원을 관리대상으로 포괄한다는 점이다. AI 거버넌스는 생애주기 관리(Lifecycle governance) 측면에서 전 생애주기를 관리한다는 공통점이 있으나, 데이터 거버넌스는 데이터 자체, AI는 모델과 의사결정 중심이라는 차이를 보인다.

주요 쟁점에서 데이터 거버넌스는 개인정보보호와 데이터 활용 간 균형, 시스템 간 상호운용성 확보, 데이터 품질 및 표준 부재 등 주로 데이터 유통·활용 기반의 제도적 인프라를 둘러싼 이슈가 중심이 된다. 반면 AI 거버넌스는 알고리즘 편향과 차별 문제, 위험 기반 접근(Risk-based Framework), 책임소재 불명확성, 블랙박스 이슈 등 의사결정 주체로서의 AI에 대한 사회적 수용성과 윤리·안전성 문제가 핵심 쟁점이다. 특히 데이터는 ‘표준화, 품질’이 이슈라면 AI는 ‘윤리, 공정성’이 핵심이라는 차이가 존재한다.

정책 및 제도 기반 측면에서 데이터 거버넌스는 개인정보보호법, 공공데이터법, 데이터 기반행정법 등 데이터 개방·보호·거버넌스를 각각 분절적으로 규율하는 다(多) 법령 체계 위에 구축되어 있으며, 국제적으로는 GDPR 및 OECD의 데이터 프레임워크 등이 참고 기준이 된다. AI 거버넌스는 EU AI Act, OECD AI 원칙, UNESCO AI 윤리 원칙 등 위험기반 규제와 책임성 중심의 국제 규범을 제도화하는 흐름 속에 있다. 이처럼 데이터 거버넌스가 국내 법제 중심이라면, AI 거버넌스는 국제 윤리·규제 중심이라는 특징이 있다.

이러한 차이점에도 불구하고 두 거버넌스 모두 신뢰성 확보, 위험 관리, 책임성 명확화, 프로세스 표준화, 조직 전체 차원의 관리체계 구축을 목표로 한다. 또한, 데이터와 AI 모두 기술적 측면만으로 관리할 수 없고, 법·제도·조직 문화·윤리 기준 등 비기술적 요소까지 포괄해야 한다는 공통점을 가진다. 무엇보다 두 체계는 상호 의존적이라고 할 수 있다. 데이터가 부실하면 어떤 AI도 신뢰할 수 없고, AI가 불투명하게 운영되면 데이터 관리의 목적도 달성할 수 없다. 따라서 데이터 거버넌스 기반 위에 AI 거버넌스가 확장적으로 구축되는 구조라고 볼 수 있을 것이다.

〈표 4-2〉 데이터 거버넌스와 AI 거버넌스 간 공통점/차이점 정리

구분	데이터 거버넌스	AI 거버넌스	공통점	차이점
정의	• 데이터의 생성·수집·저장·공유·활용 전 과정에서 품질·보안·표준·책임을 확보하기 위한 관리·의사결정 체계	• AI시스템(모델·알고리즘)의 개발·검증·운영 전 과정에서 공정성·투명성·책임을 확보하기 위한 관리·의사결정 체계	• 디지털 자원의 신뢰·책임성·투명성 확보를 위한 제도적·조직적 체계	• (데이터) 정보관리 중심 • (AI) 알고리즘/의사결정 관리 중심
목표	• 데이터 품질 보장, 상호운용성, 개인 정보보호, 데이터 활용 촉진	• 알고리즘 공정성, 설명 가능성, 안전성, 사회적 수용성 확보	• 공공 신뢰성 제고, 투명성, 사회적 가치 창출	• (데이터) 정합성, 보안 중심 • (AI) 윤리, 편향성 제거 중심
관리 대상	• 원천 데이터, 메타 데이터, 가명정보, 공공/민간 데이터 자산	• 모델, 알고리즘, 학습데이터, 예측결과물	• 디지털 자원을 포괄	• (데이터) 데이터 자산 중심 • (AI) 모델과 의사결정 중심
관리 범위	• (데이터 생애주기) 생성 → 저장 → 제공 → 분석 → 활용/폐기	• (AI 생애주기) 데이터 → 학습 → 모델 개발 → 배포 → 운영·모니터링	• 전 생애주기 관리 (Lifecycle governance) 목표	• (데이터) 정보 관리 중심 • (AI) 의사결정 관리 중심
주요 쟁점	• 표준화 부족, 시스템 간 연계 부재, 책임소재 불명, 개인 정보보호 vs. 활용 갈등	• 알고리즘 편향, 설명 가능성 한계, 책임 주체 불명확, 윤리적 리스크	• 법·제도·윤리적 문제 모두 포함	• (데이터) 품질·표준화 중심 • (AI) 윤리·공정성 중심
정책/제도 기반	• 공공데이터법, 데이터기반행정법, 개인정보보호법	• EU AI Act, OECD AI 원칙, AI 기본법/기본법 시행령 (한국)	• 국제적 규범과 표준화 필요	• (데이터) 국내 데이터법 중심 • (AI) 국제 윤리·규제 중심
조직/관리 구조	• CDO, 공공데이터 전략위원회 (한국) 등	• CAIO, 국가AI 전략위(한국) 등	• 전담조직 및 책임자 필요	• (데이터) 정보 관리 중심 • (AI) 윤리/리스크 관리 중심

다음 〈표 4-3〉은 인터뷰 결과(데이터 거버넌스의 실제)를 바탕으로 데이터 거버넌스와 AI 거버넌스 간 연계점을 리더십, 법·제도, 운영체계, 기술 인프라 관점에서 체계적으로

구조화한 것이다. 특히 두 거버넌스의 차별성과 연계성을 동시에 부각하면서, 단순 비교가 아니라 디지털 전환 시대에 통합된 거버넌스 설계가 어떻게 확장·심화되어야 하는지를 보여준다.

리더십 영역에서 데이터 거버넌스는 최고데이터책임관(CDO) 설치 및 전담 조직 운영, 데이터 기반 행정 문화 정착, 부서 간 조정 및 책임·권한 명확화 등에 중점을 둘 수 있다. AI 거버넌스 역시 전담 조직을 기반으로 운영되며, 특히 CAIO와 같은 전문 직위를 통해 CDO와의 연계 협업 체계를 강화할 수 있다. CDO와 CAIO는 각각 데이터 관리와 AI 운영을 총괄하는 핵심 리더십으로, 데이터-AI 거버넌스가 실제로 작동하기 위해 두 역할 간 협업 구조를 명확히 설계하는 것이 중요하다. CDO는 데이터 품질·표준·보안 등 데이터 기반을 구축하고, CAIO는 이를 바탕으로 AI 모델의 윤리·공정성·위험관리 체계를 담당한다. 조직 유형에 따라 분리 또는 통합 운영이 가능하지만, 공공부문에서는 투명성과 책임성을 고려하여 두 직위를 분리하되 협의체를 통해 조정·연계를 강화하는 방식이 바람직하다. 이러한 리더십 배치는 데이터에서 AI로 이어지는 전 주기적 거버넌스 체계를 안정적으로 연계하는 데 필수적이다.

법·제도 측면에서 데이터 거버넌스는 데이터 관리 전 생애주기를 법제화하고 책임 및 소유권 규정을 명확화하는 등의 역할을 수행할 수 있다. 데이터 거버넌스를 통해 확립된 데이터 법·제도는 AI 거버넌스를 통해 AI Act, OECD 원칙 등 글로벌 규범과의 정합성을 확보할 수 있다. 또한, AI 학습과정에서 개인정보 활용시 가명처리/비식별화 기준 등을 데이터 거버넌스에서 제공한다면, AI 거버넌스에서는 데이터 편향성·윤리 검증 등으로 확장할 수 있는 연계점이 존재한다.

표준화 측면에서는 데이터 거버넌스가 데이터 표준/메타데이터 체계를 확립, 상호운용성 확보, 품질관리 등의 역할을 수행할 수 있다. 데이터 거버넌스에서 데이터 품질관리 및 표준화 등의 역할을 수행한다면 AI 거버넌스는 이를 AI 학습 데이터셋 품질 검증 지표로 연결할 수 있는 연계점이 존재한다.

기술 인프라 측면에서 데이터 거버넌스는 클라우드 기반 통합 인프라 구축, 데이터 리터러시 교육 및 인력양성 등의 역할을 수행할 수 있다. 이 중 데이터 리터러시 교육은 AI 거버넌스를 통해 AI 리터러시 교육으로 확장될 수 있다는 연계점이 존재한다.

〈표 4-3〉 데이터 거버넌스-AI 거버넌스 연계 프레임워크

차원	데이터 거버넌스의 실제	데이터 거버넌스의 역할	AI 거버넌스와 연계 포인트
리더십	• 최고관리자의 전략적 리더십 및 비전 부족 • 데이터 업무를 부차 업무로 인식 • 사일로, 책임 불명확	• CDO 임명 및 전담 조직 운영 • 데이터 기반 행정 문화 정착 • 부서간 조정 및 책임/권한 명확화	• AI 책임자 (Chief AI Officer) 지정, 거버넌스 협업 구조 연계 – AI 프로젝트는 데이터 기반 이므로, 데이터 리더십을 CAIO 등과 연계 필요 – CDO가 데이터 품질 관리, CAIO는 이를 바탕으로 모델 개발, 윤리, 성과 관리까지 책임
법·제도	• 개인정보보호법(보호) vs. 공공데이터법/데이터 기반 행정법(개방) 충돌 • 데이터기반행정법 실행력 낮음(진흥법적 구조에 기인) • 책임·소유권 모호, 패널티 중심	• 법령 간 정합성 확보(보호와 활용 간 균형 추구) • 데이터 관리 전 생애주기 법제화 • 책임 및 소유권 규정 명확화	• AI Act, OECD 원칙 등 글로벌 규범과의 정합성 확보 필요 • AI 학습 과정에서 개인정보 활용 시 가명 처리/비식별화 기준 등을 데이터 거버넌스에서 제공 → AI 거버넌스에서 데이터 편향성·윤리 검증으로 확장 • 데이터 거버넌스는 데이터 전 생애주기 법제화를 추구, AI 거버넌스는 이를 모델 전 생애주기 규제로 연결
표준화	• 기관별 관리체계 상이 → 기관별 포맷·코드 불일치 • 메타데이터·품질관리 부족 • 데이터 코드·포맷 불일치 • 데이터분석·활용도구 부족	• 데이터 표준/메타데이터 체계 확립 • 상호운용성 확보 • 품질관리 및 정기 업데이트 체계화	• AI 학습데이터 품질 및 출처 검증 체계 마련 • 데이터 표준화 → AI 학습 데이터셋 품질 검증 지표로 연결

차원	데이터 거버넌스의 실제	데이터 거버넌스의 역할	AI 거버넌스와 연계 포인트
기술 인프라 (인적/ 물적)	• 시스템 간 연계부재 (수작업, 이중입력) • 전담인력 및 역량 부족 (특히 지방의 경우) • 보안 규제로 AI/클라우드 활용 제약	• 클라우드 기반 통합 인프라 구축 • 데이터 리터러시 교육 및 전문 인력의 양성 및 지방 파견	• 데이터 리터러시 교육 → AI 리터러시 교육으로 확장

나. 데이터-AI 법률 정합성 확보 전략

데이터기반행정법과 AI기본법은 각각 데이터 활용 촉진과 인공지능의 안전·신뢰 확보라는 목표를 갖고 있으나, 공공 행정에서 데이터 기반 업무와 AI 기술이 결합되는 과정에서는 두 법률의 목적·규율·적용 범위가 교차하면서 여러 수준의 긴장과 충돌이 발생하고 있다. 특히 개인정보 보호 수준, 데이터 활용 범위, AI 투명성 규정, 책임 귀속, 규제 강도 등에서 법령 간 정합성이 충분히 확보되지 못한 부분이 존재하며, 이에 대한 체계적 조정이 필요하다.

첫째, 데이터 공유와 활용을 촉진하는 데이터법의 규정과 고위험 AI·프로파일링·생성형 AI 등에 대해 엄격한 안전조치를 요구하는 AI법 규정 사이에서는 개인정보 처리 및 활용범위를 둘러싼 해석 차이가 나타날 수 있다. 둘째, 공공서비스에 적용되는 AI의 설명가능성과 투명성을 요구하는 규정은 민간 AI 모델의 영업비밀 및 보안 요소와 충돌할 가능성이 있으며, 이에 따라 공공 부문에서 적용 가능한 ‘핵심 정보 공개’ 기준의 정립이 필요하다. 셋째, 데이터 기반 자동화 과정에서 AI가 개입하는 행정 결정의 오류·차별·왜곡 등이 발생할 경우 책임 귀속의 불명확성이 존재하여, 데이터 제공자·모델 개발자·운영기관 간 단계별 책임을 명확히 규정하는 체계가 요구된다. 넷째, 데이터법의 활용 촉진 기조와 AI법의 위험기반 규제 기조가 조정되지 않을 경우 규제 공백 또는 과도한 규제 부담이 발생할 수 있다. 마지막으로, 데이터 관리 표준과 AI 모델 검증 기준이 분리된 체계로 운영될 경우 감독 및 품질관리의 일관성이 약화될 우려가 있다.

이러한 문제를 해소하기 위해서는 첫째, 법 목적 간 정합성을 확보하고, 데이터 민감도 및 AI 위험수준에 따라 규제 강도를 조정하는 위험기반·비례성 원칙을 적용할 필요가 있다. 둘째, 데이터 생애주기와 AI 모델 생애주기 전 과정을 아우르는 단계별 책임 배분체계를 마련하여 오류 발생 시 책임주체를 명확히 해야 한다. 셋째, 데이터 분류·비식별 기준,

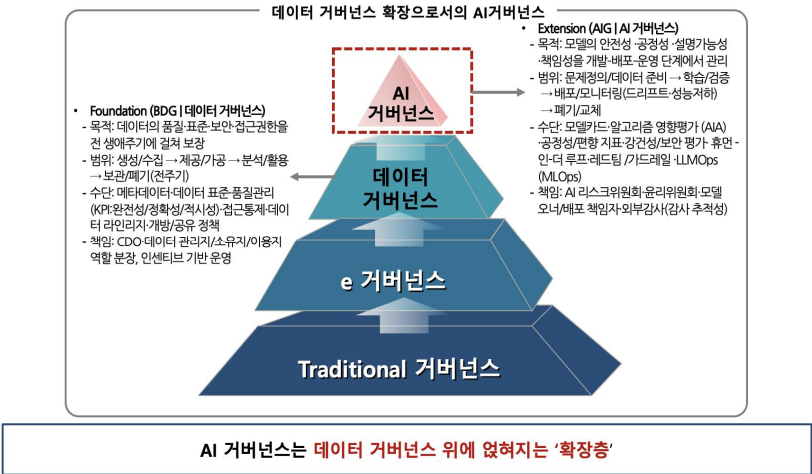
투명성 요구 수준, 개인정보 처리 기준 등 공통적 기반 규정을 통합적으로 정비하고, 알고리즘 영향평가(AIA), 규제 샌드박스, 제3자 검증 제도 등 실무적 조정 메커니즘을 제도화하여 법령 간 충돌 위험을 최소화해야 한다. 마지막으로, 범정부 차원의 ‘데이터·AI 조정기구’를 설치하여 표준·가이드라인·법령 해석을 일원화하고, 사례 중심의 분쟁·충돌 조정 기능을 수행하도록 하는 것이 필요하다.

이와 같은 통합성 확보 전략은 데이터 거버넌스가 제공하는 표준·품질 기반을 AI 거버넌스의 윤리·책임성 체계와 연계하는 역할을 수행하며, 공공부문에서 데이터와 AI가 결합된 서비스의 신뢰성과 안전성을 확보하기 위한 제도적 기반을 마련하는 데 기여할 것이다.

4. 거버넌스의 층위화 multi-layered 프레임워크

앞서 살펴본 바와 같이, 거버넌스의 발전은 사회적 난제(Wicked Problem)에 대응하기 위한 제도적 장치에서 출발하여 정보통신기술(ICT), 데이터, 그리고 인공지능(AI)의 도입에 따라 새로운 층위를 더해왔다. 본 연구는 이를 전통적 거버넌스(Traditional Governance) → e-Governance → 데이터 거버넌스(Data Governance) → 인공지능 거버넌스(AI Governance)로 이어지는 층위적 진화 과정으로 이해하고자 한다.

[그림 4-4] 데이터 거버넌스의 확장층으로서의 AI 거버넌스



우선, 전통적 거버넌스는 사회적 난제를 해결하기 위한 기본 개념으로 출발하였으며 (McGuire & Agranoff, 2011), 복잡한 공공문제에 대응하기 위해 제도적 협력과 이해관계자 간 조정, 정책 네트워크의 작동에 중점을 두었다. 그러나 이는 정보통신기술을 전제로 하지 않았다는 점에서 한계를 지녔다. 이러한 한계는 ICT의 발달과 함께 e-거버넌스로 진화하게 되었다. e-거버넌스는 전자정부 시스템을 기반으로 대규모 공공데이터를 축적하면서 정부의 반응성과 시민참여를 높이는 데 기여했지만(송경재 외, 2018), 동시에 데이터 활용과 보호를 둘러싼 새로운 규범과 조정의 필요성을 불러일으켰다.

이러한 맥락에서 등장한 것이 데이터 거버넌스이다. 데이터 거버넌스는 단순히 데이터를 수집하고 저장하는 차원을 넘어, 품질·표준·보안·접근을 전 주기에 걸쳐 관리하는 것을 목적으로 한다(Harrison et al., 2019; Janssen et al., 2020; Nielsen, 2017). 데이터 주기는 생성과 수집, 제공과 가공, 분석과 활용, 보관과 폐기로 이어지며, 각 단계마다 품질 관리와 접근 통제, 개방 정책이 정책 수단으로 작동한다. 동시에 데이터의 관리·소유·이용 주체 간 역할 분담과 책임 소재가 핵심 요인으로 부각된다. 그러나 데이터는 그 자체로 완결된 정보와 동일하지 않으며, 추론·예측·판단·의사결정 과정을 거쳐야만 의미 있는 정보로 전환된다.

이 지점에서 인공지능의 도입은 데이터 활용 방식에 근본적 변화를 가져왔다. AI는 데이터의 정보화 과정에서 새로운 대상과 수단을 제공하며, 데이터 거버넌스의 방향성과 외연을 재정의해야 할 필요성을 제기하였다. AI 거버넌스는 본질적으로 데이터 거버넌스의 확장층으로 이해될 수 있다. 신뢰할 수 있는 대규모 데이터의 확보가 없이는 AI가 작동할 수 없으며(Janssen et al., 2020), AI 거버넌스는 데이터 분석과 활용의 단계에서 출발했으나 점차 생성과 수집, 제공과 가공, 나아가 보관과 폐기 단계까지 융합하는 방향으로 확장되고 있다(Pahune et al., 2025). 전 주기적 데이터 거버넌스 속에서 제도화된 AI 활용이 바로 AI 거버넌스라 할 수 있다. 따라서 AI 거버넌스는 빅데이터 거버넌스의 확장이자 진화의 방향으로 규정할 수 있다.

그럼에도 AI 거버넌스는 데이터 거버넌스와 구별되는 차별성을 지닌다. 데이터 거버넌스가 데이터 관리와 품질 보장을 핵심으로 한다면, AI 거버넌스는 모델의 운영과 관리에 초점을 맞춘다. 구체적으로, 모델카드(Model Card), 알고리즘 평가, 데이터 편향 지표와 같은 도구는 투명성과 공정성, 책임성을 강화하는 수단으로 작동하며(Pahune et al., 2025;

Janssen et al., 2020), 이는 데이터 거버넌스의 본래 의도를 AI 환경에서 구현하는 방식이다. 결국 두 거버넌스는 상호 보완적이면서도 특화 영역이 다르기 때문에 독립적 관리가 필요하다(Schoeman, 2023). 데이터 거버넌스는 정책적 토대를 제공하고, AI 거버넌스는 이를 기반으로 확장되는 구조적 관계를 가진다.

이상의 논의를 종합하면, 거버넌스는 전통적 거버넌스에서 출발하여 e-거버넌스, 데이터 거버넌스, 그리고 인공지능 거버넌스로 이어지는 층위적 진화를 보여준다. 각 단계는 이전 단계의 성과와 한계를 토대로 확장되며, 특히 AI 거버넌스는 데이터 거버넌스를 기반으로 한 새로운 진화의 방향을 제시한다. 따라서 두 체계는 독립적이면서도 상호 보완적인 관계를 형성하며, 현대 공공 거버넌스 체계의 핵심 축으로 자리매김한다.

5. 중앙-지방 역량 격차 해소 전략

데이터 거버넌스와 AI 거버넌스는 국가 전반의 디지털 전환을 추진하는 핵심 기반으로, 중앙정부와 지방자치단체 간의 역량 차이는 거버넌스 체계의 실질적 작동을 저해하는 주요 요인으로 지적되고 있다. 특히 지방자치단체는 전문 인력 부족, 예산 제약, 기술 활용 능력의 한계 등으로 인해 중앙정부가 제시하는 정책 방향을 현장에서 실제로 구현하는 데 어려움을 겪고 있다. 중앙정부와 지방자치단체 간 데이터 및 AI 거버넌스 역량 차이는 실질적인 거버넌스 작동을 막는 주된 장애물이다. 이러한 상황에서 효과적이고 지속 가능한 거버넌스를 구축하기 위해서는 중앙과 지방의 역할을 재정립하고, 역량 차이를 구조적으로 완화할 수 있는 전략이 필요하다. 또한 이 격차를 해소하고 국가 차원의 거버넌스 효율성을 높이려면 다음과 같은 다층적이고 실질적인 전략이 필요하다.

가. 국가 차원의 공통 기반(Common Infrastructure) 구축

첫째, 중앙정부는 데이터·AI 활용에 필요한 기술적·제도적 기반을 표준화하여 지방자치단체가 동일한 품질 수준의 서비스를 제공할 수 있도록 해야 한다. 지방자치단체가 독자적으로 고도화된 시스템을 구축하기에는 인력 및 재정 측면에서 어려움이 존재하므로, 중앙정부가 전국 단위로 통일된 기반을 제공하는 방식이 가장 효율적이다.

이를 위해 중앙정부는 전국적으로 활용 가능한 데이터 표준과 메타데이터 관리 체계를 마련하여 지자체 간 데이터 구조의 이질성을 최소화해야 한다. 아울러 AI 분석 도구, 빅식

별화 도구, 클라우드 기반 AI 검증환경 등 공통 AI 플랫폼을 제공함으로써 지자체가 기술적 역량에 관계없이 동일한 수준의 도구를 사용할 수 있도록 지원해야 한다. 또한 국가 데이터 허브와 공통 API 체계를 구축하여 지자체는 데이터·AI 서비스를 개발하기보다 활용과 응용에 집중할 수 있는 구조를 만드는 것이 중요하다.

이러한 공통 기반 구축은 지자체 간 데이터 품질과 활용 수준의 불균형을 해소하고, 행정 서비스의 일관성을 높이는 데 기여한다.

나. 수준별 역할 분담을 통한 책임·기능 재설계

둘째, 중앙과 지방이 동일한 역할을 수행하는 구조는 비효율을 초래하며, 역량 차이를 더욱 심화시킨다. 따라서 데이터·AI 거버넌스의 실질적 작동을 위해서는 중앙정부, 광역자치단체, 기초지방자치단체 간의 역할을 명확히 구분하고, 각 기관이 수행해야 할 책임 범위를 재설계해야 한다.

중앙정부는 정책 방향 설정, 법·제도 정비, 데이터·AI 표준·보안 규칙 수립, 고난도 AI 모델 개발 등 국가적 수준의 고도화된 기능을 담당해야 한다. 반면 광역자치단체는 지역단위의 데이터 통합·조정 기능을 수행하며, 기초지방자치단체는 데이터 생산과 현장 중심의 행정서비스 적용을 담당하도록 역할을 구분할 필요가 있다.

아울러 민간 기업, 지역 대학, 연구기관 등을 포함한 지역 협력체계를 구축하여 지자체의 기술 역량을 보완하는 ‘분권형 거버넌스’ 모델을 도입하는 것도 중요하다. 이러한 역할 분담 모델을 통해 역량이 부족한 지자체도 복잡한 업무를 중앙과 광역의 지원 하에 수행할 수 있으며, 중복 업무를 최소화하여 효율성을 확보할 수 있다.

다. 지방자치단체의 인력·조직 역량 강화

세 번째 전략은 지방자치단체의 역량 부족 문제를 구조적으로 해결하기 위한 인력 및 조직 역량 강화이다. 많은 지자체는 데이터·AI 분야의 전문 인력을 안정적으로 확보하기 어렵기 때문에, 중앙과 지방 간의 인재 순환 및 지원 체계가 필요하다.

우선 중앙정부-광역자치단체-기초지자체 간 데이터·AI 전문 인력을 순환 배치하는 ‘순환형 인재제도’를 도입하여 인력의 편중 문제를 완화할 수 있다. 또한 지방 공무원을 대상으로 실무 중심의 데이터 분석, 품질관리, AI 리스크 관리 교육을 제공하는 지역 AI 아카데미를 운영할 필요가 있다. 여기에 민간 전문가를 일정 기간 지자체에 파견하거나 공동 프

로젝트를 수행하는 ‘민관 인재교류 프로그램’도 활용된다면 지자체의 기술 역량 향상에 큰 도움이 될 것이다.

조직 차원에서는 모든 지자체가 데이터 책임관(CDO)과 AI 책임관(CAIO)을 지정하여 데이터·AI 업무의 책임성과 전문성을 확보하고, 전담 조직을 단계적으로 강화할 필요가 있다.

라. 현장 기반의 활용 사례(Use Case) 중심 확산 전략

마지막으로, 데이터·AI 거버넌스가 실제로 작동하기 위해서는 현장에서 즉시 적용 가능한 실증 사례를 중심으로 거버넌스를 확산하는 전략이 중요하다. 제도·기반만 있는 거버넌스는 종종 문서화된 체계에 머물기 때문에, 실제 행정 문제를 해결하는 사례를 기반으로 거버넌스를 활성화해야 한다.

중앙정부는 지자체가 바로 적용할 수 있는 데이터·AI 활용 사례 패키지를 개발하여 제공할 필요가 있다. 예를 들어 민원 분류 자동화, 도로 파손 탐지, 복지 사각지대 예측, 재난 대응 지원 등 공통적으로 수요가 높은 분야의 AI 활용 모델을 구축해 지자체에 제공하면, 지자체는 이를 자체 데이터로 재학습(Fine-tuning)하여 지역 특성에 맞게 사용할 수 있다. 이러한 실증 중심의 접근은 지자체가 스스로 역량을 축적하는 기반이 되며, 성공 사례가 축적될수록 전국 확산도 자연스럽게 이뤄질 수 있다.

나아가 앞서 제시한 공통 기반 구축, 역할 분담 재설계, 인력·조직 역량 강화 전략과 함께 추진될 때, 중앙~지방 간 역량 격차를 실질적으로 완화하고 국가 전반의 데이터·AI 거버넌스 체계를 안정적으로 작동시키는 기반을 마련할 수 있다.

제2절 프레임워크 전략에 대한 전문가 검증 및 보완

1. 대상자 및 실시 내용

본 연구는 공공기관 데이터 거버넌스의 전략과 프레임워크의 정책적 타당성과 실효성을 검증하기 위해 전문가 자문을 실시하였다. 자문은 연구진이 도출한 프레임워크와 실행 전략이 실제 행정·정책 현장에서 작동 가능한지를 평가하고, 제도적·기술적·윤리적 측면에서의 보완 방향을 도출하는 것을 목적으로 진행되었다.

2025년 10월 14일부터 10월 17일까지 전문가 대상으로 자문이 이루어졌으며 총 21명의 전문가가 참여하였다. 참여자는 데이터 거버넌스, 인공지능(AI) 거버넌스, 공공정책, 기술 표준, 법·제도 등 다양한 분야를 대표하는 전문가로 구성되었다. 구성의 다양성은 거버넌스 프레임워크의 실질성과 포괄성을 확보를 위해 고려하였다.

자문은 서면 및 온라인 병행 방식으로 이루어졌으며 연구진이 사전에 설계한 검증 문항에 대한 서면 응답과 심층 토의 방식의 온라인 미팅을 결합하였다. 자문에서는 프레임워크의 기본 구조, 개념적 정합성, 적용 가능성, 실행 전략, 단계별 추진체계, 데이터-AI 연계 구조 등 구체적 개선 방향을 심층적으로 논의하였다.

검증 문항은 총 5개 대분류 및 8개 세부 문항(A-H)으로 구성되었으며, 각 문항은 연구진이 제시한 프레임워크의 타당성, 현실 적용성, 보완 필요성을 다각도로 평가할 수 있도록 설계되었다. 세부 문항의 구성은 다음과 같다.

[OECD 기본 프레임의 수정과 고도화]

① 통합 데이터 거버넌스 모형

A. 위와 같이 수정된 OECD 통합 데이터 거버넌스 모형의 구성요소에 대해 어떻게 생각하십니까?

② 데이터 거버넌스 활성화를 위한 조건-이슈-제언 연계도

B. 위와 같이 통합 데이터 거버넌스 활성화를 위한 조건-이슈-제언 연계도에 대해 어떻게 생각하십니까?

③ 데이터 거버넌스 활성화를 위한 로드맵

C. 위와 같이 데이터 거버넌스 활성화를 위한 로드맵에 대해 어떻게 생각하십니까?

[구체적인 거버넌스 전략]

D. 위와 같이 데이터 거버넌스 수행전략에 대해 어떻게 생각하십니까?

[데이터 거버넌스-AI 거버넌스 연계 프레임워크]

① 데이터 거버넌스와 AI 거버넌스 간 공통점/차이점

E. 위와 같이 <표 2>의 공통점과 차이점에 대해 어떻게 생각하십니까?

② 데이터 거버넌스-AI 거버넌스 연계 프레임워크

F. 위와 같이 <표 3>의 데이터 거버넌스-AI 거버넌스 연계 프레임워크에 대해 어떻게 생각하십니까?

[거버넌스의 층위화 multi-layered 프레임]

G. 데이터 거버넌스 층위화 다층 프레임워크 제안에 대해 어떻게 생각하십니까?

[종합 평가 및 제언]

H. 전반적인 종합 의견 및 보완 사항에 대해 말씀해 주세요.

각 문항은 OECD의 통합 데이터 거버넌스 프레임을 한국적 맥락에 맞게 수정한 틀의 타당성을 검증하고 공공기관의 데이터 관리·활용 정책이 인공지능 거버넌스와 어떻게 연계될 수 있는지를 검토하기 위한 것이다. 문항 설계 단계에서는 기존 연구에서 제시된 거버넌스의 핵심 요인을 포괄하도록 구조화되었으며, 각 문항은 개방형 질문 형태로 제시되어 전문가의 서술형 의견을 수집할 수 있도록 구성되었다.

자문 과정은 ① 문항별 서면검토 → ② 의견 수집 및 분석 → ③ 보완 논의 및 종합 의견 정리의 3단계로 진행되었다. 전문가들의 의견은 내용 분석 방식을 통해 항목별로 분류·요약되었으며, 반복적으로 제시된 주요 키워드와 쟁점을 중심으로 피드백을 정리하였다.

종합적으로, 본 자문은 데이터 거버넌스의 개념적 정합성과 실행전략의 현실성을 검증함과 동시에, 데이터-AI 거버넌스 간 연계 구조의 완성도를 점검하는 과정이었다. 이를 통해 본 연구는 기술적 프레임을 넘어 정책적·제도적·조직적 실행체계로 발전할 수 있는 근거를 확보하였으며, 전문가 검증 결과는 향후 공공부문 디지털 전환 정책의 실효성 제고에 중요한 기초자료로 활용될 것이다.

2. 전문가 자문 내용

가. OECD 기본 프레임의 수정과 고도화

전문가 자문 결과, OECD 통합 데이터 거버넌스 기본 프레임을 국내 행정환경에 맞게 수정·보완하는 과정은 크게 세 가지 측면에서 논의되었다. 첫째, 프레임의 구성요소 및 체계의 타당성을 검토하고, 둘째, 조인-이슈-제언 연계도를 중심으로 정책 실행과정의 논리

구조를 점검하였으며, 셋째, 이를 바탕으로 통합 데이터 거버넌스 프레임워크의 고도화 방향을 제시하였다. 세 자문 결과는 상호 연계되어 있으며, 모형의 구조적 타당성 검증에서 출발해 실행 연계 및 전략적 확장으로 이어지는 단계적 발전 구조를 이루고 있다.

1) 통합 데이터 거버넌스 모형의 구성요소

전문가 자문 결과, 수정된 OECD 통합 데이터 거버넌스 모형은 기존의 기술 중심적 접근을 보완하고, 행정체계의 제도·조직·기술 요소를 통합적으로 반영하고 있다는 점에서 대체로 긍정적인 평가를 받았다. 다수의 전문가는 본 모형이 OECD 기본 틀의 원칙과 방향성을 유지하면서도 한국적 행정·조직 환경을 고려한 실질적 보완이 이루어졌다는 점에 주목하였다. 특히, 제도적 기반과 조직적 실행 구조를 분리해 구성한 점, 그리고 수요자(시민·기업)와 공급자(행정기관·공공기관) 간 상호작용을 중심에 둔 점이 실무 적용 가능성을 높였다는 평가가 제시되었다.

여러 전문가는 수정된 모형이 단순한 관리 도식이 아니라 데이터의 수요자와 공급자 간 연결 구조를 시각적으로 명확히 제시했다는 점을 긍정적으로 언급하였다. 한 전문가는 “단순한 기술 도입 요인 나열을 넘어 행정 현장에서 실제로 발생하는 ‘효과와 역효과’, ‘공식적 요인과 비공식적 요인’과 같은 복잡다단한 측면을 포괄하려 한 점이 매우 인상적”이라고 평가하였으며 또 다른 전문가는 “OECD의 통합 데이터 거버넌스 모형은 환경, 제도, 조직, 기술의 네 층위를 유기적으로 연결해 데이터 행정을 체계적으로 설명한 완결성 높은 구조”라고 언급하였다.

또한, 전문가들은 본 모형이 제도적, 조직적, 기술적 층위 간의 수직적·수평적 관계를 균형 있게 표현하고 있다는 점을 공통적으로 지적하였다. 다수의 의견에 따르면, 기존 OECD 프레임은 기술·데이터 관리 중심의 정태적 성격이 강했으나 수정된 모형은 제도와 조직의 관계를 포함해 실제 거버넌스가 작동하는 구조를 역동적으로 반영했다는 것이다. 일부 전문가는 “제도적으로 규정되어도 조직이 움직이지 않으면 실질적 거버넌스는 작동하지 않는다”며, 이번 모형이 이러한 구조적 불균형을 해소하기 위한 방향성을 보여준다고 평가하였다.

한편, 모형 내 ‘환경 층위(environmental layer)’를 추가한 점에 대해 많은 전문가들이 공통적으로 긍정적 의견을 보였다. 이 층위는 사회·경제·기술적 변동 요인을 포괄함으로써,

데이터 거버넌스의 외부적 조건과 정책적 맥락을 반영할 수 있게 한다는 평가였다. 일부는 “환경 층위가 데이터 생태계를 둘러싼 맥락적 요인을 시각화함으로써, 데이터 정책을 고립된 기술 행정이 아닌 사회적 시스템으로 인식하게 한다”고 지적하였다. 다만 몇몇 전문가는 환경 층위와 제도·조직·기술 층위 간의 관계를 좀 더 정교하게 설명할 필요가 있다고 제안하였다.

일부 전문가들은 제도·조직 간 상호작용의 실질적 구현 방안이 부족하다는 점을 지적하였다. 한 전문가는 “제도적 장치와 조직 간 협력의 연계 구조는 도식화되어 있으나 이를 구체적으로 운영할 주체와 절차가 제시되지 않았다”고 평가하였고 또 다른 전문가는 “협업에 대한 공식적/비공식적 인센티브/디센티브가 조직구조 뿐 아니라 제도층위에 포함되어야 할 수 있다”고 덧붙였다. 이는 모형이 제도적 구조를 명확히 하였음에도 불구하고 실제 거버넌스의 실행 가능성 측면에서는 추가 보완이 필요하다는 견해로 해석된다.

다수의 전문가들은 또한 수정된 모형이 기존의 기술 중심적 접근을 넘어 시민 참여와 사회적 책임을 강조하는 거버넌스 전환의 방향성을 제시했다는 점을 높이 평가하였다. 특히 데이터 관리와 활용의 중심을 ‘국가-행정-시민’의 삼각 구조로 확장함으로써, 공공데이터의 사회적 가치와 신뢰성 제고에 기여할 수 있다는 의견이 제시되었다. 한 전문가는 “기술결정론적 관점을 지양하고, 공공부문 데이터 거버넌스 실행의 사회·기술적(socio-technical) 복잡성을 균형 있게 고려한 접근”이라고 언급하였다. 또 다른 전문가는 “환경 층위와 제도·조직·기술 층위로 구분한 것이 적절하다고 판단되나, 두 층위의 작용이 일방향적으로 요구와 부응으로 표시되는 것보다 적극적인 시민참여(시민의 서비스 및 재화 공급), 적극 행정(맞춤형, 찾아가는 행정서비스 공급-능동형) 등을 반영하여 양방향 형태를 반영한 용어로 변경하는 것을 제안”하였다.

마지막으로, 일부 전문가들은 OECD 프레임의 개념적 보편성과 국내 행정 현실 간의 간극을 지적하면서 수정된 모형이 그 간극을 일정 부분 해소했다고 평가하였다. 그러나 여전히 구체적 운영 단계에서의 조직 간 조정·협력 메커니즘, 책임 구조의 명확화, 데이터 관리주체 간 역할 구분이 부족하다는 의견이 있었다. 이에 따라 전문가들은 향후 거버넌스 모형의 발전 방향으로, “개념적 프레임워크를 넘어 제도 설계와 운영 지침 수준까지 구체화되어야 한다”고 공통적으로 강조하였다.

종합하면, 전문가 자문은 수정된 OECD 통합 데이터 거버넌스 모형이 기존의 한계를 보

완하며 한국형 데이터 거버넌스 체계로 발전할 수 있는 기초를 제시했다는 점에서 긍정적으로 평가하였다. 동시에, 제도와 조직 간의 실질적 작동 구조를 명확히 하고, 환경 변화에 대응할 수 있는 유연성을 보강해야 한다는 현실적 과제를 함께 지적하였다.

2) 조언-이슈-제언 연계도 분석 및 보완 방향

전문가 자문 결과, 통합 데이터 거버넌스 활성화를 위한 ‘조언-이슈-제언 연계 구조’는 전체적으로 논리적 일관성과 구조적 체계성을 갖추고 있다는 평가가 다수였다. 전문가들은 이 연계도가 데이터 거버넌스의 실행과정에서 발생할 수 있는 문제를 단순히 나열하는 수준을 넘어 문제의 원인(조건)-현상(이슈)-해결(제언) 간의 인과적 관계를 시각적으로 정리했다는 점에서 의미가 크다고 보았다. 특히, 제도적 조건과 조직적 여건을 선행 요인으로 설정하고, 이를 중심으로 실행상의 이슈와 개선 제언이 연결된 구성은 현실적인 접근이라고 평가되었다.

여러 전문가들은 본 연계도가 데이터 거버넌스의 복합성을 반영했다는 점에서 긍정적인 의견을 보였다. 한 전문가는 “데이터 거버넌스의 복합적 성격을 제도·조직·기술 차원에서 구조적으로 진단하고, 실행 가능한 개선 방향을 단계적으로 제시한 체계적 접근”이라고 평가하였다. 또 다른 전문가는 “이론적 모형(조건)을 바탕으로 현장의 실질적인 문제점(이슈)과 구체적인 정책적 대응(제언)을 논리적으로 연계한 점은 매우 실용적이며 가치 있는 접근”이라고 평가하였다.

다만 일부 전문가는 연계 구조가 지나치게 포괄적이어서 각 요소 간 우선순위와 영향도 분석이 부족하다는 점을 지적하였다. “모든 이슈가 동등하게 다뤄지는 듯한 인상을 주는데, 정책 설계 측면에서는 중요도에 따라 핵심 이슈를 선별해야 한다”는 의견이 제시되었다. 또한, 조언-이슈-제언 간의 연결선이 개념적 차원에 머물러 있어, 실제 정책 집행 단계에서 어떤 주체가 어떤 시점에 대응해야 하는지 명시되어야 한다는 구체적 지적도 있었다.

다수의 전문가는 연계도의 기본 틀은 타당하나 제도적 기반과 조직 역량의 현실적 제약이 병행 고려되어야 한다고 강조하였다. “데이터 거버넌스는 법과 제도의 문제만이 아니라, 공공조직 내부의 인식과 행태가 중요한 변수”라며, “제도적 정비가 이루어져도 현장의 수용성과 실행력 확보 없이는 제언이 작동하기 어렵다”는 의견이 반복되었다. 특히, 데이터 관리 담당자에 대한 인식 제고, 인력·예산의 구조적 제약 완화, 책임성 있는 리더십 확

보 등이 연계도 내에서 선행조건으로 구체화될 필요가 있다는 제안이 있었다.

또한, 전문가들은 연계도에서 제시된 ‘조언’과 ‘제언’의 개념 구분을 명확히 해야 한다는 점을 공통적으로 언급하였다. 일부는 “조언이 분석적 근거라면 제언은 정책적 해법이어야 하는데, 현재는 두 개념이 혼재되어 있다”고 지적하며, 두 단계 간의 논리적 위계를 명확히 구분할 필요가 있다고 하였다. 이러한 구분이 명확해질 경우, 각 단계별로 담당 조직과 추진 방안이 자연스럽게 연계될 수 있을 것이라는 의견도 제시되었다.

한편, 몇몇 전문가는 연계도의 형식적 구조는 우수하지만, 정책 실행 단계에서의 피드백 메커니즘이 부족하다고 평가하였다. 조언-이슈-제언의 연계가 선형적으로만 표현되어 있어, 실제 행정 과정에서의 상호작용과 반복적 환류 구조가 반영되지 않았다는 것이다. “데이터 거버넌스는 시행착오와 피드백을 통해 성숙하는 체계이므로, 순환형 구조로 재설계할 필요가 있다”는 제안이 다수였다.

또한 일부 전문가는 “조건-이슈-제언의 논리적 흐름은 타당하나, 정책 추진 과정에서 나타나는 정치적·조직문화적 변수를 반영하지 못하고 있다”고 지적하였다. 데이터 거버넌스 활성화의 실질적 장애요인이 제도 미비보다는 이해관계 충돌, 부처 간 경쟁, 현장 인식 부족 등에 있다는 것이다. 따라서 연계 구조가 단순한 절차나 프로세스가 아닌, 이러한 거버넌스 내부의 동학과 저항 구조를 설명할 수 있는 동태적 모델로 발전해야 한다는 의견이 제시되었다.

반면, 일부 전문가는 연계도 자체의 단순성과 시각적 명료성이 거버넌스 전략 수립에 유용하다고 평가하였다. “이 구조는 복잡한 데이터 정책 영역을 요약적으로 이해하는 데 효과적이며, 정책결정자나 실무자가 쉽게 활용할 수 있다”는 의견도 있었다. 그러나 그 단순성이 오히려 구체적 실행 전략의 부재로 이어질 수 있다는 점에서 “요소 간 관계를 단계별로 확장하거나, 세부 실행 도구를 병기하는 형태로 보완할 필요가 있다”는 제언도 함께 제시되었다. 예를 들어, “기술적 인프라(hardware/software) 구축이라는 물질 투자가 조직 구성원의 데이터 리터러시 및 분석 역량 강화라는 인적 자본 투자와 어떻게 선순환을 이루어 실질적인 성과로 이어지는지, 그 구체적인 경로를 보장한다면 연계도가 더욱 설득력을 가질 것”이라는 제언이 있었다.

결론적으로, 전문가 자문은 통합 데이터 거버넌스의 조언-이슈-제언 연계 구조가 복합적 거버넌스 문제를 분석하는 유용한 틀이라는 점을 인정하면서도, 각 요소 간 우선순위

설정, 실행 주체의 명확화, 피드백 체계의 제도화 등 현실적 보완이 요구된다는 점에서 일관된 의견을 보였다. 이 구조는 개념적 타당성에서 높은 평가를 받았으나, 이슈의 복잡성, 층위 간 개념의 명확성, 제언의 구체성 측면에서 보완이 병행되어야 한다는 것이 전문가들의 공통된 의견이었다.

3) 통합 데이터 거버넌스 프레임워크의 고도화 방향

전문가 자문 결과, 데이터 거버넌스 활성화를 위한 로드맵은 제도적 기반, 조직 간 협력, 역량 강화, 기술 고도화 등 4개 축을 중심으로 구성된 점에서 대체로 타당하다는 평가를 받았다. 다수의 전문가는 단계별 접근(단기-중기-장기)을 통해 정책 추진의 실효성과 연속성을 확보한 구조를 긍정적으로 평가하였다. 한 전문가는 “데이터 거버넌스 활성화를 위한 로드맵은 단기·중기·장기의 시간적 구획을 두어 정책의 지속성과 실행 가능성을 높인 체계적인 구성”이라고 언급하였다. 또 다른 전문가는 “데이터 기반 행정의 복잡한 구조를 단계적으로 구체화하고 있어, 정책 실행의 현실성을 높이는 설계로 볼 수 있다”고 평가하였다.

여러 전문가는 특히 책임주체의 명확화 필요성을 강조하였다. 한 전문가는 “업무 절차상 명확히 책임 주체를 지정하지 않으면 데이터 공개나 관리 과정에서 책임 소재가 불분명해질 수 있다”고 지적하였으며, 다른 전문가는 “제도기반·협력활성화·역량강화·기술기반의 4개 축 구성은 적절하나, 각 단계별 주체의 역할이 구체화되지 않으면 실행이 지연될 가능성이 있다”고 언급하였다.

로드맵의 세부 구성에 대해서는 다양한 보완 의견도 제시되었다. 한 전문가는 “제도 기반 강화, 기술 기반 강화, 협력 활성화, 역량 강화 등 네 가지 축이 잘 구성되어 있으나, 기술적 측면의 비중이 상대적으로 높아 제도적 안정성과 정책 일관성을 확보할 장치가 병행되어야 한다”고 평가하였다. 또 다른 전문가는 “로드맵이 행정 효율성과 기술적 확장성을 모두 포괄하고 있으나, 데이터 표준화와 품질 관리에 대한 구체적 실행계획이 필요하다”고 강조하였다. 한 전문가는 “단기에는 법·제도 정비를 통해 기반을 마련하고, 중기에는 조직 역량 강화와 협력 네트워크 확충, 장기에는 기술 기반의 자율적 데이터 거버넌스 체계로 발전해야 한다”고 제언하였다. 또 다른 전문가는 “AI와 빅데이터를 포함한 기술적 혁신이 행정 내재화로 이어지기 위해서는 인력 역량 강화와 지속적인 교육 체계가 병행되어야 한다”고 언급하였다.

여러 전문가는 협력 구조의 제도화 필요성을 공통적으로 강조하였다. “데이터 거버넌스

는 법적·제도적 장치만으로 작동하기 어렵고, 협업을 촉진할 수 있는 조직문화적 기반이 필요하다”는 의견이 다수였으며, 일부는 “부처 간 협업 인센티브를 제도적으로 설계해야 한다”고 제안하였다.

한편, 일부 전문가는 정책의 피드백 메커니즘 부재를 지적하였다. “정책의 추진 단계가 일방향으로 설정되어 있어, 실행 이후의 평가와 환류가 어렵다”며 “주기적 점검과 피드백 구조를 내재화해야 한다”는 의견이 제시되었다. 한 전문가는 “데이터 거버넌스 로드맵은 일회성 계획이 아니라, 정기적인 평가와 수정이 가능한 순환형 구조로 운영되어야 한다”고 언급하였다. 또 다른 전문가는 “로드맵을 실행 중심의 운영 매뉴얼로 전환하고, 각 단계별 세부 과업과 책임 주체를 명시해야 실효성을 확보할 수 있다”고 제안하였다.

일부 전문가는 사회적 신뢰 확보의 중요성을 언급하였다. 한 전문가는 “기술 인프라의 확충만으로는 데이터 거버넌스가 작동하지 않으며, 데이터 관리의 투명성과 시민 참여의 제도화가 병행되어야 한다”고 지적하였다. 다른 전문가는 “데이터 공유와 활용 과정에서 개인정보 보호, 보안, 윤리 문제를 균형 있게 반영해야 한다”고 강조하였다.

마지막으로, 다수의 전문가는 민관 협력과 시민 참여의 확대를 고도화 방향의 핵심 과제로 제시하였다. “공공부문 중심의 접근을 넘어 민간·시민사회와의 데이터 협력 체계를 제도적으로 확보해야 한다”는 의견과 “데이터 거버넌스의 성공은 기술이 아니라 신뢰와 협력의 제도에 달려 있다”는 의견이 공통적으로 제시되었다.

종합하면, 전문가 자문 결과 데이터 거버넌스 로드맵은 정책 추진의 단계적 구조와 체계적 방향성 면에서 긍정적 평가를 받았으나, 제도적 기반 강화, 협력 구조의 내재화, 인적 역량 강화, 피드백 체계의 구축, 시민 참여의 제도화가 향후 고도화의 필수 조건으로 제시되었다.

나. 구체적 거버넌스 전략

전문가 자문 결과, 본 연구가 제시한 데이터 거버넌스 전략은 거시적 방향성과 정책적 타당성을 충분히 갖추고 있으나, 실제 행정현장에서 적용하기 위해서는 보다 구체적인 실행체계와 단계별 추진 로드맵이 필요하다는 점이 지적되었다. 특히 데이터 거버넌스가 단순한 기술적 관리체계에 그치는 것이 아니라, 조직의 구조·인사·문화·제도·위험 관리가 통합적으로 작동하는 정책 프레임워크로 발전해야 한다는 점에서 전문가들의 의견이 일치

하였다. 자문 결과는 Janssen et al.(2020)이 제시한 세 가지 접근법인 계획 기반, 조직 기반, 위험 기반 접근의 논리 구조를 중심으로 정리될 수 있다.

먼저, 계획 기반 접근의 측면에서 전문가들은 데이터 거버넌스 전략이 단기적 대응 중심의 정책에서 벗어나 중·장기적 발전 전략으로 전환되어야 한다고 강조하였다. 공공기관이 보유한 방대한 데이터가 지속적이고 체계적으로 활용되기 위해서는, 단일 과제 수행이나 시범사업에 머무는 수준이 아니라, 정책 주기 전반에 걸친 전략적 계획이 요구된다는 것이다. 이를 종합해 데이터 거버넌스의 실행 구조를 단기-중기-장기 전략으로 구분하고, 각 단계별로 상호 연계되는 과제 수행체계를 설계하였다. 단기적으로는 기관별 데이터 거버넌스 가이드라인 마련과 실무 매뉴얼 개발을 추진하고, 중기적으로는 ‘데이터 표준화 및 품질관리 현행화 용역’을 통해 관리체계를 내재화하며, 장기적으로는 데이터 거래 제도 정비와 국제규범 이행체계를 확립하는 방향으로 계획이 구체화되었다.

이 과정에서 전문가들은 데이터 거버넌스 정책의 ‘연속성’과 ‘일관성’이 핵심이라고 지적하였다. 단발성 정책이 아니라 정책-조직-기술이 서로 피드백을 주고받는 선순환적 구조가 필요하다는 것이다. 따라서 본 연구는 데이터 정책을 ‘가이드라인 수립 → 과제 수행 → 모니터링 및 평가 → 개선’으로 이어지는 순환형 구조로 제시하였으며, 이를 통해 각 기관의 자율성을 보장하면서도 국가적 수준의 통합 거버넌스를 실현할 수 있도록 방향성을 설정하였다. 또한 공공부문과 학계·연구기관 간 협력을 활성화하여, 전문성과 실행력을 동시에 확보할 수 있도록 한 점도 중요한 보완 사항으로 제시하였다.

다음으로, 조직 기반 접근은 데이터 거버넌스의 제도적 정착을 위한 핵심 요소로 강조되었다. 자문 참여자들은 데이터 거버넌스가 조직 내 실질적으로 작동하기 위해서는 전담 조직의 신설과 인적 자원의 체계적 관리가 필수적이라고 지적하였다. 특히 최고데이터책임자(CDO)의 제도적 위상 강화와, 데이터·알고리즘 관련 전문인력을 관리·훈련하는 전담 부서의 신설이 필요하다는 의견이 다수 제시되었고, 공공부문 내 데이터 전담조직을 설계할 때 단순한 직무 지정이 아닌 권한 위임과 책임의 명확화를 병행해야 한다는 점이 강조되었다.

전문가들은 또한 데이터 거버넌스의 지속가능성을 위해 공공부문 내 외부 전문인력의 유입이 활성화되어야 한다고 제언하였다. 이를 위해 민간 전문가가 공공부문 프로젝트에 참여할 수 있는 유연한 인사제도를 도입하고, 데이터 분야 경력직의 공공부문 진출을 위

한 제도적 장벽을 완화해야 한다는 의견이 제시되었다. 더불어, 성과에 따라 금전적·비금전적 인센티브를 제공하는 성과 기반 보상체계의 필요성도 언급되었다.

한편, 협력 체계의 강화도 조직 기반 접근의 중요한 축으로 제시되었다. 자문 결과, 공공기관 간 데이터 공유·활용이 여전히 부처 간 칸막이 구조에 의해 제약받고 있으며, 민간기업 및 연구기관과의 협업도 제한적으로 이루어지고 있다는 점이 확인되었다. 따라서 향후 거버넌스 전략에서 공공-민간 간 데이터 협력의 범위와 역할을 명확히 정의하고, 기술적 표준화와 법적 근거를 보완함으로써 개방형 협력 생태계를 구축해야 한다는 점을 강조한다.

마지막으로, 위험 기반 접근에 대한 논의는 데이터 거버넌스의 신뢰성과 안전성을 확보하는 차원에서 가장 중점적으로 다루어졌다. 전문가들은 데이터 활용이 확대될수록 개인정보 보호, 데이터 품질 저하, 알고리즘 편향 등 새로운 형태의 위험이 증가할 수 있다는 점을 지적하였다. 특히 공공기관 내에서는 보안 기준과 실제 시스템 활용 간의 불일치로 인한 오류나 위험이 빈번하게 발생할 수 있으므로, 이를 예방할 수 있는 상시 모니터링 체계가 필요하다는 의견이 제시되었다.

다. 데이터 거버넌스-AI 거버넌스 연계 프레임워크

1) 데이터 거버넌스와 AI 거버넌스 간 공통점/차이점

전문가 자문을 통해 확인된 주요 의견은 데이터 거버넌스와 인공지능(AI) 거버넌스가 공통적으로 신뢰(trust), 투명성(transparency), 책임성(accountability)을 핵심 가치로 지향하고 있음에도 불구하고, 실제 정책 설계와 운영 단계에서는 여전히 이원적이고 단절된 체계로 존재하고 있다는 점이었다. 특히 데이터 거버넌스가 정보자원의 품질·표준화·보안·활용 촉진 등 기술적·관리적 측면에 초점을 맞춘 관리 프레임이라면, AI 거버넌스는 알고리즘의 공정성·설명가능성·윤리성 등 사회적·윤리적 통제 프레임으로 확장된다는 점에서 두 체계는 관리의 초점과 규율의 방식에서 뚜렷한 차이를 보인다.

그러나 이러한 차이는 대립적인 구분이라기보다 발전 단계의 차이와 상호 보완의 구조로 이해되어야 한다는 점이 전문가들 사이의 공통된 견해였다. 다수의 자문은 데이터 거버넌스가 AI 거버넌스를 가능하게 하는 기반 인프라이며, AI 거버넌스는 데이터 거버넌스가 지향하는 신뢰와 품질관리의 원칙을 사회적 책임성과 윤리성으로 확장하는 과정이라고

보았다. 다시 말해, 데이터 거버넌스가 ‘무엇을 관리할 것인가’에 초점을 둔다면, AI 거버넌스는 ‘그 데이터를 어떻게 관리하고 판단할 것인가’에 초점을 둔 체계로, 두 거버넌스는 동일한 가치 체계 위에서 상호 의존적으로 작동해야 한다는 것이다.

특히 전문가들은 공공부문에서 이 두 체계가 단절된 채 병렬적으로 운영될 경우, 데이터 관리의 규범과 AI 관리의 규범이 서로 불일치하여 정책적 일관성과 법적 적합성이 훼손될 위험을 지적하였다. AI의 학습데이터는 데이터 거버넌스의 품질기준을 충족해야 하지만, 실제 운영에서는 데이터팀과 AI팀이 서로 다른 기준으로 작업을 수행하는 경우가 빈번하다는 것이다. 이러한 분리 구조는 데이터 품질관리와 AI 윤리관리 간의 공백을 초래하고, 결과적으로 공공데이터 기반 의사결정의 신뢰성을 저하시킬 수 있다. 따라서 전문가들은 데이터 거버넌스와 AI 거버넌스를 위계적 상하관계로 설정하기보다, 순환적이고 상호 보완적인 구조로 재설계해야 한다고 제안하였다. 즉, 데이터 관리의 원칙이 AI의 윤리원칙으로 확장되고, AI의 운영성과가 다시 데이터의 품질 향상으로 환류되는 선순환적 구조를 구축해야 한다는 것이다.

법·제도 측면에서 전문가들은 데이터 거버넌스가 「공공데이터법」과 「데이터기반행정부」 등 국내 법체계 중심으로 발전해온 반면, AI 거버넌스는 「EU AI Act」나 OECD의 AI 원칙 등 국제 규범에 기초하여 구성되고 있다는 점을 지적하였다. 이러한 구조적 불균형은 법적 체계의 적합성을 저해할 수 있으며, 보호와 활용 간의 균형이 무너지거나 개인정보 처리의 불명확성이 발생할 수 있다는 점에서 보완이 필요하다고 보았다.

조직적 측면에서도, 두 거버넌스의 관리주체를 단절적으로 운영하기보다는 협력적 리더십 구조로 전환할 필요성이 강조되었다. 데이터 거버넌스의 최고책임자인 CDO(Chief Data Officer)가 데이터의 품질과 표준을 관리하고, AI 거버넌스의 책임자인 CAIO(Chief AI Officer)가 이를 기반으로 알고리즘 개발과 윤리성, 사회적 수용성을 관리하는 이중 리더십 구조를 확립해야 한다는 의견이 제시되었다. 이는 공공기관 내에서 데이터와 AI 관리의 일관성을 확보하고, 조직 간 칸막이로 인한 중복업무나 책임 불분명 문제를 해소하는 효과적인 방안으로 평가되었다. 나아가 이러한 협력체계를 제도화하여, CDO와 CAIO가 공동으로 정책 설계와 성과관리, 위험 대응을 수행하는 체계를 구축할 필요가 있다는 구체적 제안도 제시되었다.

기술적 차원에서는 데이터 거버넌스에서 확립된 표준화·품질관리 체계를 AI 거버넌스

의 윤리적 관리 체계와 연계함으로써, 데이터 표준화가 곧 AI 학습데이터의 품질 검증 지표로 작동할 수 있도록 연동하는 구조가 필요하다고 보았다. 특히 데이터 거버넌스에서 구축된 품질검증과 메타데이터 관리, 보안 및 상호운용성 체계가 AI의 학습데이터 품질·출처 검증 단계에 직접 반영될 수 있도록 하는 것이 핵심 과제로 제시되었다. 또한, AI 거버넌스는 단순히 알고리즘 단계에 그치지 않고 배포 이후의 모니터링 및 사고 대응체계를 포함하는 전주기 관리가 이루어져야 한다는 점도 강조되었다. 이는 데이터 관리와 AI 윤리 관리의 경계가 점점 모호해지는 디지털 행정 환경에서, 전 과정에 걸친 통합 거버넌스가 신뢰 확보의 핵심이라는 점을 보여준다.

일부 전문가들은 연구진이 제시한 공통점과 차이점의 구분이 기술적 수준의 차이에 지나치게 집중되어 있다고 평가하면서, 단순한 병렬 비교보다는 공통된 가치 위에서의 차이, 즉 각 체계가 동일한 원칙을 어떻게 다르게 해석하고 적용하는가의 관점에서 재구성할 필요가 있다고 조언하였다. 이는 두 거버넌스가 공통된 원칙을 기반으로 하되, 데이터는 품질과 표준화의 관점에서, AI는 윤리와 사회적 책임의 관점에서 같은 원칙을 서로 다른 층위에서 구현한다는 의미이다.

2) 데이터 거버넌스-AI 거버넌스 연계 프레임워크

데이터 거버넌스와 인공지능(AI) 거버넌스 간의 연계 프레임워크에 대한 전문가 검증 결과, 전반적인 방향성과 구성은 적절하다는 평가가 이루어졌다. 특히 데이터 거버넌스가 공공부문에서 신뢰할 수 있는 데이터의 품질·표준·보안을 확보하는 기반적 체계로서 기능하고, AI 거버넌스가 이를 토대로 투명성·공정성·책임성을 관리하는 구조로 제시된 점이 긍정적으로 평가되었다. 다만 일부 전문가들은 두 거버넌스 간의 관계가 단순한 병렬 비교로 읽힐 우려가 있다고 지적하였으며, 이를 보완하기 위해 관계의 해석을 위계적 구조가 아닌 연속적·순환적 구조로 명확히 제시할 필요가 있음을 제안하였다.

법·제도 영역에서는 데이터 관련 법제와 AI 규범 간의 정합성을 확보할 필요성이 강조되었다. 일부 전문가는 데이터법과 AI법이 각각 독립적으로 추진될 경우 중복과 충돌이 발생할 수 있다고 지적하며, 데이터 처리 단계에서 마련된 가명처리·비식별화 기준이 AI 학습과정의 윤리 검증으로 자연스럽게 이어질 수 있도록 법적 연속성을 강화해야 한다고 제안하였다. 이를 반영해 보호와 활용의 균형, 국내 법제와 국제 규범의 연계성 확보를 핵심

원칙으로 추가하고, 데이터와 AI 생애주기 전반에서의 상호준거성을 확보하는 방향으로 내용을 수정하였다.

조직 및 리더십 체계에 대해서는 CDO(Chief Data Officer)와 CAIO(Chief AI Officer)의 역할을 명확히 구분하면서도, 실질적인 의사결정이 협업적으로 이루어질 수 있도록 설계해야 한다는 의견이 제시되었다. 일각에서는 두 직위가 분리될 경우 불필요한 의사결정 지점이 늘어날 수 있다는 우려도 있었으나, 데이터 품질과 표준을 관리하는 CDO와 모델의 윤리와 투명성을 책임지는 CAIO가 공동으로 의사결정하고 상호검증하는 구조가 필요하다는 의견이 우세하였다.

표준화와 품질관리 측면에서는 기관별 데이터 포맷 불일치, 메타데이터 관리 미흡 등으로 인해 AI 학습데이터의 품질 저하가 발생할 수 있다는 지적이 있었다. 이를 반영하여, 데이터 표준화와 메타데이터 관리가 AI 학습데이터 품질 검증 단계로 자연스럽게 연결되는 기술적 연계 구조를 보다 구체적으로 보완하였다. 또한, 데이터 표준화의 강화가 AI 단계의 성능 및 신뢰성 확보로 직결됨을 명시하여 연계 구조의 실효성을 높였다.

기술 인프라 측면에서는 기존의 클라우드 기반 통합체계와 데이터·AI 리더러시 확장 방향이 타당하다는 평가가 있었으나, 지방 및 중소기업의 인프라 취약성을 고려한 인력 보강과 교육체계 확충이 필요하다는 의견이 제시되었다.

마지막으로, 일부 전문가들은 데이터 거버넌스와 AI 거버넌스의 관계를 확장과 포섭의 관점에서 접근할 필요가 있다고 제안하였다. 즉, AI 거버넌스가 데이터 거버넌스를 포함하여 상위 개념으로 작동하는 측면을 반영하고, 두 체계가 독립적이면서도 상호보완적으로 기능하는 관계를 보다 분명히 제시해야 한다는 것이다.

라. 거버넌스의 층위화 multi-layered 프레임

1) 데이터 거버넌스 층위와 다층 프레임워크 제안에 대한 의견

대부분의 전문가들은 ‘전통적 거버넌스 → e-거버넌스 → 데이터 거버넌스 → AI 거버넌스’로 이어지는 진화적 구조의 논리성과 타당성을 인정했으며, 특히 AI 거버넌스를 데이터 거버넌스의 상위 대체개념이 아닌 확장층으로 설정한 점이 이론적 정합성과 실천적 통찰을 제시하고 거버넌스 제도의 연속성을 보장한다고 평가하였다. 일부 전문가들은 AI 거버넌스가 데이터 거버넌스의 ‘수단’인지 ‘목적’인지 명확히 구분할 필요성을 제기하며, 단순

한 위계적 순서보다는 상호작용하며 피드백을 주고받는 네트워크형 거버넌스로의 전환을 제안했고, 데이터 거버넌스는 ‘무엇을 관리할 것인가(객체 중심)’에, AI 거버넌스는 ‘어떻게 관리할 것인가(프로세스 중심)’에 초점을 둔다는 기능적 구분도 제시되었다.

시각적 표현에 대해서는 여러 전문가들이 피라미드 형태가 AI 거버넌스를 종착점으로 인식하게 할 우려를 제기하며 수평적 확장이나 순환적 구조, AI 거버넌스가 데이터 거버넌스를 감싸는 형태, 각 단계별 feedback loop 표시 등 다양한 개선안을 제안했다.

이론과 실무의 균형 측면에서 일부 전문가들은 층위 세분화가 학문적으로는 의미가 있지만 현장에서는 혼란을 가중시킬 수 있어 이론적으로는 분리하되 현실적으로는 통합 운영하는 방향을 제안했으며, 데이터 품질-모델 품질-정책 품질 간 3단계 품질 연계체계 구축과 층위별 통합 전략의 필요성도 강조되었다.

또한 일부 전문가들은 기술 진화만을 강조하여 기술결정론적 시각으로 비춰질 우려를 제기하며 제도 기반에 따른 기술 구현 흐름의 보완과 선행적 진화보다 ‘융합적 확장’ 개념으로의 설명을 제안했고, 국제 표준과의 정합성 측면에서 AI 거버넌스가 데이터 거버넌스를 내부 구성요소로 포함하는 구조로 재정의할 필요성도 지적되었으며, 인적 인프라 강화와 전문분야에서의 민간 협력형 거버넌스 모델 도입의 중요성도 언급되었다.

마. 종합 평가 및 제언

전문가들의 종합 평가 및 제언을 정리하면 다음과 같다. 대부분의 전문가들은 본 연구가 데이터 거버넌스와 AI 거버넌스의 관계를 체계적으로 정립하고 디지털 행정의 핵심 인프라 구축 과정에서 제도·조직·기술의 상호작용을 다층적으로 이해한 점에서 높은 완성도와 시의성을 갖췄다고 평가하였다. 특히 전통적 거버넌스에서 AI 거버넌스로 이어지는 층위적 진화 구조를 통해 기술 발전을 거버넌스 가치의 확장 과정으로 해석하고, 데이터 거버넌스를 AI 거버넌스의 기반이자 확장층으로 정의하여 정책의 연속성과 기술의 윤리성을 동시에 확보했다는 점이 높이 평가되었다. 그러나 여러 전문가들은 데이터 거버넌스와 AI 거버넌스의 구분이 현실적으로 어렵고 실무상 통합적으로 작동해야 한다는 점을 지적하면서, 두 거버넌스를 구분하는 것의 실질적 의미에 대한 근본적 의문을 제기하였다.

특히 중소 규모 기관에서는 데이터와 AI 관련 기능이 통합 수행되는 경우가 많고, AI의 핵심 쟁점인 편향·공정성·성능이 근본적으로 데이터 품질에 기인한다는 점에서 긴밀한

연계가 필수적이라고 강조되었다. 향후 보완 방향으로서는 통합적 프레임워크로의 전환, 성과지표와 성숙도 지표 구축을 통한 피드백형 관리체계 마련, 데이터-AI 연계 법제 가이드라인 개발, CDO-CAIO 리더십 제도화, 시민 참여형 거버넌스 메커니즘 확립, 중앙과 지방간 불균형 해소를 위한 다방향 네트워크형 거버넌스 구조 구축 등이 제시되었으며, 궁극적으로 데이터-AI-환경이 순환적으로 연결되는 책임 있는 디지털 전환 거버넌스 모델로 발전되어야 한다고 제안하였다.

또한 일부 전문가들은 공공부문 거버넌스의 비합리적 측면과 조직적·정치적 모호성에 대한 고려, AI 활용 목적별 데이터 관리체계의 차별화, 행위자(수요자) 측면의 추가 보완, 그리고 AI 디지털 전환의 빠른 속도를 고려한 거버넌스의 신속한 대응 필요성을 강조하였다.

제5장 결 론

제1절 연구 결과

본 연구는 인공지능과 디지털 기술의 급속한 발전에 따른 공공부문의 데이터 거버넌스 체계 구축 방안을 모색하기 위해 수행되었다. 정부의 데이터 기반 행정 활성화라는 정책 목표를 효과적으로 달성하기 위해서는 체계적인 데이터 거버넌스가 필수적이라는 인식 아래, 국내외 이론적 논의와 실제 운영 현황을 종합적으로 분석하고 한국형 데이터 거버넌스 프레임워크, 구체적인 전략 및 로드맵을 제시하였다.

1. AI 시대 데이터 거버넌스의 진화와 핵심 성공 요인

먼저, 문헌 분석을 통해 데이터 거버넌스는 전통적인 관리 개념에서 출발하여 전자정부를 거쳐 AI 시대에는 통합적이고 참여적인 관리 체계로 진화하고 있다는 점을 확인하였다. 초기에는 데이터 품질, 표준화, 보안, 책임 분담 등 관리 요소 중심이었으나, 최근에는 네트워크형, 플랫폼형, 참여형 모델로 발전하며 사회적·윤리적 관점까지 포괄하고 있다.

특히 AI 확산은 편향 완화, 공정성, 프라이버시, 추적성, 안전한 모델 배포 등 7대 핵심 범주를 요구하며 알고리즘 투명성과 책임성이 새로운 과제로 부각되고 있다.

또한, 주요국의 사례에서 미국은 연방 차원에서 CDO 제도를 도입하여 기관별 책임성을 강화하고 데이터 개방과 활용을 촉진하였다. 영국은 국가 데이터 전략을 수립해 민간과 공공의 협력 구조를 제도화하고 투명성과 신뢰성을 확보하였다. 독일은 엄격한 데이터 보호 규제를 바탕으로 산업 데이터 공유 플랫폼인 Gaia-X를 추진하여 보안과 활용의 균형을 추구하였다. 캐나다는 디지털 정부 전략을 통해 국민 참여형 데이터 서비스 확대와 기관 간 데이터 연계를 강화하였다.

이와 같은 문헌 및 주요국 사례 분석을 종합한 결과 문헌과 주요국 사례를 통해 데이터 거버넌스의 성공 요인들을 도출하였다.

첫째, 데이터 관리 전담 조직과 리더십을 통한 책임성 확보하는 것이 필요하다. 둘째, 데이터 활용을 뒷받침하는 법·제도의 정비가 필수적이다. 셋째, 상호운용성을 확보하기 위한 기술 인프라와 표준화 체계 구축이 요구된다. 넷째, 공공성과 투명성 확보를 통한 국민 신뢰 형성이 핵심 요인으로 확인되었다.

2. 공공데이터 거버넌스 실태 진단과 주요 문제점

공공데이터 거버넌스 실태 진단을 통해 국내 공공기관의 데이터 거버넌스 실태를 데이터, 기술·시스템, 조직·관리, 법률·제도의 네 가지 차원에서 진단한 결과, 다음과 같은 주요 문제점이 발견되었다.

데이터 차원에서는 기관 간 데이터 공유 부족, 표준화 및 호환성 미비, 중복 수집, 비정형 데이터 처리 한계가 공통 문제로 나타났다. 수도권과 지방 간 역량 격차, 전문 인력 부족, 수작업 오류도 심각한 수준이었다. 일부 기관은 메타데이터 관리, 통합 관리로 개선을 시도하고 있으나 보안과 활용의 균형 및 제도적 기반이 여전히 미흡한 것으로 나타났다.

기술·시스템 차원에서는 기관별 편차가 크며 연계 및 통합 부족으로 효율성이 저하되고 있다. 보안 규제로 AI 분석 도구 활용이 제한되고 데이터 포맷 불일치로 품질 저하가 발생하고 있다.

조직·관리 차원에서는 구성원의 낮은 인식과 소극적 참여가 한계로 지적되었다. 가시적 성과와 보상 체계가 마련되어야 참여가 활성화될 수 있으며, 리더십 부재 시 제도화가 성과와 연결되지 못하는 것으로 나타났다. 또한 전문 인력 부족과 교육 한계가 문제로 지적되어 CDO 제도화, 전담 조직 강화, 실무 역량 개발이 필수적인 것으로 확인되었다.

법률·제도 차원에서는 개인정보보호법과 활용 법령 간 충돌로 현장에서 소극적 대응이 이루어지고 있다. 법령 간 정합성 부족, 데이터 소유권 불명확, 지역 간 제도 확산 한계가 문제점으로 도출되었다.

3. 통합 데이터 거버넌스 프레임워크 제시

본 연구는 OECD 기본 프레임워크를 수정·고도화하여 통합 데이터 거버넌스 모형을 제시하였다. 이 모형은 환경 층위와 제도·조직·기술 층위로 구분되며, 환경 층위는 수요자

(시민, 기업)의 필요와 요구로 구성되고, 제도·조직·기술 층위는 공급자(행정기관, 공공기관)의 자원과 역량으로 구성된다.

제도 층위는 공식 제도(권한-책임 관계, 촉진-제한 규정, 표준 운영 절차, 감시·감사 체계)와 비공식 제도(업무상 규범, 업무상 관행, 민관 협력 기구, 외부 행위자 참여 통로)로 구성된다.

조직 층위는 조직 행동(리더십, 전문성, 자원관리, 성과관리)과 조직 구조(전담 인력·조직, 업무 수행 구조, 조직 내 관계, 조직 간 관계)로 구성된다.

기술 층위는 데이터(데이터 표준화, 데이터 품질, 데이터 통합)와 시스템(시스템 표준화, 시스템 품질, 시스템 통합)으로 구성된다.

전문가 자문 결과, 수정된 OECD 통합 데이터 거버넌스 모형은 기존의 기술 중심적 접근을 보완하고, 행정체계의 제도·조직·기술 요소를 통합적으로 반영하고 있다는 점에서 긍정적인 평가를 받았다. 특히 제도와 조직의 관계를 포함해 실제 거버넌스가 작동하는 구조를 역동적으로 반영했다는 점이 높이 평가되었다. 다만 제도적 장치와 조직 간 협력의 연계 구조는 도식화되어 있으나, 실질적 구현 방안이 부족하다는 점과 환경 층위와 제도·조직·기술 층위 간의 관계를 좀 더 정교하게 설명할 필요가 있다는 점이 지적되었다.

4. 데이터-AI 거버넌스 연계 체계

본 연구는 데이터 거버넌스와 AI 거버넌스의 공통점과 차이점을 체계적으로 분석하여 두 거버넌스 간 연계 프레임워크를 제시하였다. 두 거버넌스는 디지털 자원의 신뢰·책임성·투명성 확보를 위한 제도적·조직적 체계이며, 공공 신뢰성 제고와 사회적 가치 창출이라는 궁극적 목표를 공유한다. 그러나 데이터 거버넌스는 정보자원 중심으로 데이터 품질 보장, 상호운용성 개선, 보안 강화에 초점을 맞추며 원천 데이터·메타데이터·공공/민간 데이터를 관리 대상으로 하는 반면, AI 거버넌스는 알고리즘과 의사결정 시스템 중심으로 알고리즘의 공정성 실현, 설명가능성 확보, 사회적 책임 수행을 중시하며 모델·알고리즘·학습데이터·예측결과물을 관리 대상으로 한다. AI 거버넌스는 본질적으로 데이터 거버넌스의 확장층으로 이해될 수 있으며, 신뢰할 수 있는 대규모 데이터의 확보가 없이는 AI가 작동할 수 없다는 점에서 데이터 거버넌스를 기반으로 하지만, 데이터 분석과 활용의

단계에서 출발하여 점차 생성과 수집, 제공과 가공, 나아가 보관과 폐기 단계까지 융합하는 방향으로 확장되고 있다. 그럼에도 AI 거버넌스는 데이터 거버넌스와 구별되는 차별성을 지니며, 데이터 거버넌스가 데이터 관리와 품질 보장을 핵심으로 한다면 AI 거버넌스는 모델의 운영과 관리에 초점을 맞추기 때문에 두 거버넌스는 상호 보완적이면서도 특화 영역이 다르므로 독립적 관리가 필요하다. 전문가 자문에서는 데이터 거버넌스와 AI 거버넌스가 공통적으로 신뢰, 투명성, 책임성을 핵심 가치로 지향하고 있음에도 불구하고 실제 정책 설계와 운영 단계에서는 여전히 이원적이고 단절된 체계로 존재하고 있다는 점이 지적되었으며, 데이터 거버넌스가 ‘무엇을 관리할 것인가’에 초점을 둔다면 AI 거버넌스는 ‘그 데이터를 어떻게 관리하고 판단할 것인가’에 초점을 둔 체계로서 두 거버넌스는 동일한 가치 체계 위에서 상호 의존적으로 작동해야 한다는 점이 강조되었다. 따라서 CDO(Chief Data Officer)와 CAIO(Chief AI Officer)의 역할을 명확히 구분하면서도 협력적 관계를 제도화하고, 데이터-AI 거버넌스 통합 운영 체계를 구축하여 전 생애주기에 걸친 연계성을 강화하는 것이 필요하다.

이러한 배경하에, 데이터 거버넌스와 AI 거버넌스는 공공 및 민간부문에서 각각의 역할과 책임 영역을 명확히 하며 운영될 필요가 있다.

먼저, 정부와 공공기관은 데이터와 AI를 공공 서비스 혁신과 정책 결정의 핵심 도구로 사용한다. 따라서 데이터 거버넌스는 행정 전반의 공공 데이터를 정확하고 표준화된 형태로 확보해 정책의 근거를 강화하는 기반으로 이해해야 한다. 또한, 공공기관은 국민의 데이터를 다루기 때문에 개인정보 보호, 보안, 투명성 같은 원칙을 더욱 엄격하게 적용해야 한다. AI 거버넌스 측면에서는 공공 영역에서 AI가 의사결정에 활용되는 경우 사회적 영향력이 매우 크므로, 알고리즘의 편향 방지, 결과에 대한 설명 가능성 확보, 국민 신뢰 유지가 중요하다. 즉, 정부는 AI를 단순히 효율성을 높이는 도구로만 볼 것이 아니라 사회적 책임을 동반한 기술 시스템으로 이해하고 운영해야 한다. AI 기반 행정 서비스가 확대될수록 국민의 권익 보호를 위한 감시·감독 체계를 함께 강화해야 한다.

다음으로, 또 다른 주요 주체라고 할 수 있는 민간기업 역시 이 두 거버넌스를 통해 경쟁력을 확보해야 한다. 민간기업은 데이터를 경쟁력의 핵심 자산으로 삼아, 데이터 거버넌스를 통해 데이터의 일관성·품질·보안을 확보해 비즈니스 성과를 강화해야 한다. 예를 들어, 고객 데이터를 여러 시스템에 분산 보유하면 정확한 분석이 불가능해지는데, 데이터

거버넌스는 이러한 비효율을 제거하고 데이터를 기반으로 한 의사결정의 신뢰도를 높인다. 또한 데이터 규제 환경이 강화되는 상황에서 데이터 거버넌스는 법적 리스크를 낮추는 역할도 한다. AI 거버넌스 측면에서는 AI 활용이 곧 조직의 리스크 관리 문제로 연결되기 때문에, 기업은 AI 알고리즘의 윤리성, 설명 가능성, 투명성, 품질 관리 체계를 내재화해야 한다. AI가 고객 서비스·마케팅·운영 자동화·채용 등 다양한 분야에 확대됨에 따라, 기업은 기술 개발자뿐 아니라 경영진·법무·인사·보안팀이 함께 참여하는 전사적 AI 관리 체계를 구축할 필요가 있다.

정부·공공기관·기업 모두 데이터 거버넌스를 기초 인프라, AI 거버넌스를 고도화된 통제·관리 체계로 받아들이고, 두 체계를 따로가 아니라 연속적인 하나의 체계로 운영해야 할 필요가 있다. 데이터가 잘 정비되어 있을수록 AI는 더 책임 있게 활용되고, AI가 안전하게 운영될수록 데이터 활용의 가치는 더 높아지기 때문이다. 결국 두 거버넌스는 조직의 디지털 역량을 강화하는 핵심 구조이며, 신뢰 기반의 기술 활용이라는 공통된 목적을 지닌다고 볼 수 있다.

이와 더불어, 본 연구는 데이터 거버넌스와 AI 거버넌스의 연계 프레임워크를 제시하여, AI 거버넌스가 데이터 거버넌스의 확장층으로서 기능함을 명확히 하였다. 거버넌스의 층위화 관점에서 전통적 거버넌스 → e-거버넌스 → 데이터 거버넌스 → AI 거버넌스로 이어지는 진화 과정을 제시하고, 각 단계가 이전 단계의 성과와 한계를 토대로 확장되는 층위적 구조임을 밝혔다. 각 단계는 이전 단계의 성과와 한계를 토대로 확장되며, 특히 AI 거버넌스는 데이터 거버넌스를 기반으로 한 새로운 진화의 방향을 제시한다. 따라서 두 체계는 독립적이면서도 상호 보완적인 관계를 형성하며, 현대 공공 거버넌스 체계의 핵심 축으로 자리매김한다.

제2절 정책적 시사점

1. 제도적 차원의 시사점

공공부문 데이터 거버넌스의 제도적 기반을 강화하기 위해서는 우선 개인정보보호법과 공공데이터법 간의 정합성을 확보할 필요가 있다. 이를 위해 현행 개인정보보호법 제18조(개인정보의 목적 외 이용·제공)와 제26조(업무 위탁)에 대한 해석 및 적용 기준을 구체화함으로써, 실무 현장에서의 예측 가능성과 법적 안정성을 제고해야 한다. 또한 ‘공공의 안전과 복리’와 같은 추상적 기준을 교육, 보건, 안전, 환경 등 정책 영역별로 세분화하여 구체적인 판단 기준을 제시할 필요가 있다.

나아가 개별 법률의 단편적 개정을 넘어 데이터를 공공자산이자 정책 도구로 명시하는 데이터 거버넌스 기본법을 제정할 필요가 있다. 이를 통해 기존의 ‘보호 중심’ 접근에서 ‘보호와 활용의 균형’으로, ‘분절적 규율’에서 ‘통합적 체계’로, 그리고 ‘중앙집권적 통제’에서 ‘협력적 거버넌스’로 전환하는 제도적 기반을 마련해야 한다.

법적 안정성 강화를 위해 법률 차원에서 데이터 거버넌스 컨트롤타워의 지위, 기능, 권한을 명확히 규정하여야 한다. 또한, 각 기관 최고데이터책임자(CDO)의 법적 지위와 책임을 명시해 현재 형식적으로 운영되고 있는 최고데이터책임자(CDO) 제도의 개선이 필요하다.

더불어 기관 간 데이터 연계 및 활용 과정에서 발생하는 분쟁에 대해 조정 권한과 절차를 법제화하여 실질적인 조정 기능을 확보할 필요가 있다.

아울러 현행 공공데이터법이 개방 중심으로 설계되어 공공기관 간 내부 활용 근거가 미흡한 한계를 보완해야 한다. 공공기관이 보유한 데이터를 ‘공공자산’으로 규정하고, 기관 간 공유와 활용을 원칙으로 하되, 보안 및 개인정보보호와 관련한 예외를 명확히 규정해야 한다. 또한 민간 데이터의 공익적 활용을 위한 법적 근거와 보상체계, 그리고 데이터 신탁 등 중간기구를 활용하는 방안을 법제화함으로써 민관 간 데이터 협력의 제도적 기반을 마련할 필요가 있다.

특히 AI 시대에 부합하는 새로운 규범 체계를 구축하기 위해 논의 중인 「인공지능 기본법」과의 정합성을 고려해야 한다. 이에 따라 알고리즘의 투명성, 자동화된 의사결정의 책임성, 인공지능 편향 방지 등에 관한 법적 기준을 정립하고, 공공 부문의 AI 활용 과정에서

설명가능성 확보와 인간의 최종 판단권 보장을 법제화함으로써 시민의 권익을 보호해야 한다.

마지막으로, 중앙집권적 데이터 거버넌스 구조로 인해 제약받고 있는 지방의 자율적 데이터 활용 기반을 확대할 필요가 있다. 이를 위해 지방자치법 상 자치사무 범위 내에서 데이터 거버넌스의 법적 지위를 명확히 하고, 지방정부가 지역 특성에 맞는 데이터 정책을 수립·시행할 수 있도록 법적 권한을 보장해야 한다. 아울러 중앙과 지방 간 데이터 연계·공유 의무 및 지원 체계를 법제화하여, 균형 잡힌 데이터 연방주의를 구축하는 방향으로 제도 개선이 이루어져야 한다.

2. 조직적 차원의 시사점

조직 역량 강화를 위한 전문 인력 양성은 정책기획-데이터분석-성과평가를 연계할 수 있는 복합형 전문가의 육성을 목표로 해야 한다. 이는 단순한 데이터 분석 역량을 보유한 인력을 양성하는 데 그치지 않고, 정책 수립부터 성과 평가까지의 전 과정을 이해하고 관여할 수 있는 통합형 인재를 확보하는 데 중점을 두어야 한다.

교육훈련은 단기적·일회성 워크숍에 머물러서는 안 되며, 직무경력관리제도와 연계한 장기적·지속적 인력개발체제로 전환해야 한다. 또한 해외 주요국의 사례처럼 전 직원 대상 데이터 리터러시 교육을 제도적으로 보편화하고, 교육 성과를 인사·승진 체계와 직접 연계함으로써 조직 전반에 데이터 기반 행정 문화를 정착시켜야 한다. 더불어 민간 및 학계 전문가와의 지속적인 교류를 제도화하여 내부 학습효과와 혁신 역량을 축적할 필요가 있다.

한편, 최고데이터책임자(CDO) 제도는 다수 기관에서 형식적으로 운영되고 있는 한계를 보완해야 한다. 이를 위해 CDO의 역할과 권한을 명확히 규정하고, 데이터 기반 의사결정이 실질적인 조직 성과로 연결되는 관리체계를 구축해야 한다. 또한 기관장의 전략적 리더십이 데이터 거버넌스의 실행력을 좌우한다는 점에서, 기관장이 데이터 정책을 단순한 행정 과제가 아닌 핵심 전략으로 인식할 때 구성원들의 참여와 협력이 촉진된다는 점을 강조할 필요가 있다.

마지막으로, OECD(2019)가 제시한 바와 같이 데이터 활용에 있어 협업 및 인센티브 체

계의 중요성이 현장에서는 종종 ‘추가적인 행정 부담’으로 인식되는 문제를 극복해야 한다. 이를 위해 데이터 공유 및 활용 성과를 인사, 예산, 평가 등 조직의 보상체계에 실질적으로 반영함으로써, 구성원들이 데이터 기반 행정을 조직성과의 핵심 요소로 인식하도록 유도할 필요가 있다.

3. 데이터 관리 시사점

데이터 관리에 있어서 무엇보다 데이터 표준화와 자동화 기반의 상호운용성 제고가 필수적이다. 우선 부처 간 공통 분류체계와 코드, 시간 단위, 메타데이터의 표준화를 제도화함으로써, 연계용 공통키와 품질 관리 규칙을 정립해야 한다. 이를 통해 현장에서 지속적으로 지적되어 온 기관 간 이질적 관리체계, 표준 부재, 낮은 상호운용성의 문제를 근본적으로 개선할 수 있다.

또한 데이터의 제공·가공 단계에서 자동화 및 파이프라인을 구현하여 수작업 결합과 중복 처리를 최소화해야 한다. 아울러 품질 점검, 버전 관리, 이력 관리의 상시화를 통해 현재 많은 기관에서 발생하고 있는 중복 입력 및 수작업 결합 문제를 구조적으로 해소할 필요가 있다.

더불어 데이터 활용의 효율성을 제고하기 위해 사용자 친화적 UI, 자동 검증 시스템, 중앙 데이터 카탈로그를 갖춘 통합 거버넌스 플랫폼을 구축해야 한다. 이를 통해 분산된 시스템 간 자동 연계 부재로 인한 데이터 흐름의 병목 현상을 해소하고, 데이터 도입과 활용 간의 시간적·기술적 간극을 축소할 수 있을 것이다.

데이터의 생애주기 전 단계에 걸친 품질관리 체계를 구축하는 것 또한 중요하다. 특히 분석 단계에서의 객관성과 재현성 확보를 위해 기관 차원의 분석 지원체계와 검증 절차를 제도화해야 하며, 개인의 역량에 과도하게 의존하는 현재의 분석 구조를 개선하기 위해 조직 단위의 분석 인프라를 확충할 필요가 있다.

마지막으로, 디지털 트윈을 포함한 의사결정 지원 시스템에는 데이터 및 모델의 품질 보증과 사전·사후 검증 절차를 내재화해야 한다. 이를 통해 AI 거버넌스의 핵심 원칙인 편향 완화, 공정성, 설명가능성을 보장하는 기술적 기반을 마련함으로써, 공공부문 데이터 활용의 신뢰성과 책임성을 제고할 수 있을 것이다.

4. 데이터 거버넌스 체계 구축 전략

공공기관 데이터 거버넌스의 성공적 구현을 위해서는 단기(1년), 중기(2~3년), 장기(3년 이상)로 구분된 단계적 접근 전략이 필요하다. 단계별로 추진 목표와 실행과제를 명확히 설정함으로써 정책 추진의 연속성과 체계성을 확보해야 한다.

먼저, 단기적으로는 기관장의 확고한 리더십을 기반으로, 데이터 및 인공지능 기반 행정에 대한 조직 구성원의 인식 수준과 참여 의지를 제고하는 것이 우선 과제이다. 이를 위해 교육·훈련 프로그램 및 성과 인센티브 제도를 운영하여 구성원의 참여를 유도하고, 특히 전 직원 대상 데이터 리터러시 교육의 보편화를 추진해야 한다.

내부적으로는 부서 간 업무 관련 소통과 교류를 활성화하여 협업 문화를 정착시키고, 외부적으로는 데이터 및 기술 업무 과정에 전문가의 참여를 제도화함으로써 실무 전문성을 강화해야 한다. 또한 메타데이터, 데이터 분석 도구, 보안 기준 등 기초 기술 인프라를 확보하고, 명확한 업무 책임주체 지정 및 핵심성과지표(KPI) 설정을 통해 데이터·기술 업무의 관리체계를 구체화해야 한다.

중기적으로는 데이터·기술 전문인력의 확보와 함께 업무분장 합리화를 통해 행정부담과 업무피로를 완화해야 한다. 단순한 인원 확충이 아니라 정책·분석·기술 역량을 겸비한 복합형 전문가 양성이 핵심 과제이다. 기관 내·기관 간 공통 및 필수 데이터의 공유를 활성화하고, 특히 중앙-지방 간 상호책임적 데이터 거버넌스 관계를 구축하여 데이터의 회신-공유-활용이 선순환적으로 이루어지도록 제도화해야 한다. 또한 데이터 품질 고도화, 데이터 연계 및 통합 활성화, 시스템 간 상호운용성 강화를 통해 기술적 효율성을 제고해야 한다. 아울러 데이터 및 기술 관련 가이드라인의 내용과 절차를 구체화하고, 이를 기반으로 업무 절차의 표준화와 제도적 정착을 추진해야 한다.

마지막으로 장기적으로는 공공부문에 대한 외부 전문인력 진출을 촉진할 수 있는 인사제도 개선이 필요하다. 또한 최고데이터책임자(CDO) 제도의 실질적 운영을 보장함으로써 조직 내 데이터 거버넌스의 지속가능성을 확보해야 한다.

데이터·기술 관련 공공기관-민간기업 간 협력체계를 내실화하고, 개방형 협력 생태계를 조성하기 위해 양 부문 간 협력 범위와 역할을 명확히 규정해야 한다. 이를 뒷받침하기 위해 기술적 표준화와 법적 근거를 보완하고, 민간 데이터 활용을 위한 거래제도 정비 및 시

시스템 동질성 강화를 위한 표준화 정책을 추진해야 한다.

아울러 법령 간 목표 충돌 지점에 대한 정합성 검토와 명확한 법령 해석 기준을 마련하여 규제 일관성을 확보해야 하며, 데이터 및 기술 관련 국제 규범의 국내 준수와 이행을 활성화함으로써 글로벌 데이터 거버넌스 환경에 대응할 수 있는 제도적 기반을 강화해야 한다.

참 고 문 헌

[국내 문헌]

- 강정목·송효진·김현성(2014), “스마트시대의 디지털 리터러시 측정을 위한 진단도구의 개발과 적용,” 《한국지역정보학회지》, 17(3), pp. 143-173.
- 개인정보보호위원회(2023), 『인공지능 시대 안전한 개인정보 활용 정책방향』.
- _____ (2025), 『2025년 주요업무 추진계획: 안전한 개인정보, 신뢰받는 인공지능 시대』.
- 경제·인문사회연구회(2024), 『분권형 균형발전정책연구』.
- 고학수·임용·박상철(2021), 『유럽연합 인공지능법안의 개요 및 대응방안. 서울대학교 인공지능정책 이니셔티브』.
- 과학기술정보통신부·한국지능정보사회진흥원(2022), 2022 디지털 정보격차 실태조사.
- 과학기술정보통신부·한국지능정보사회진흥원·한국정보통신기술협회(2024), 『인공지능 학습용 데이터 품질관리 가이드라인』.
- 관계부처 합동(2020), 사람이 중심이 되는 「인공지능(AI) 윤리기준」.
- _____ (2022), “제4차(’23~’25) 공공데이터 제공 및 이용 활성화에 관한 기본계획”.
- _____ (2024), “제2차 데이터기반행정 활성화 기본계획”.
- 김경훈·이준배·윤성욱(2021), 『EU 데이터거버넌스 법안(Data Governance Act) 주요 내용 및 시사점』, 정보통신정책연구원.
- 김기환(2013), “공공부문 빅데이터의 활용성과 위험성”, 《정책분석평가학회보》, 23(2), pp. 1-27.
- 김문조·김종길(2002), “정보격차(Digital Divide)의 이론적·정책적 제고,” 《한국사회학》, 36(4), pp. 123-155.
- 김정남·방은찬·장희영·송선영(2023), “인공지능(AI)의 학습용 데이터 윤리 프레임워크 개발에 관한 연구,” 《윤리연구》, 140, pp. 283-309.

- 김준형·이사빈·염지선·김민영(2020), 『지속가능한 스마트 시티 운영을 위한 정부혁신 방안에 관한 연구』, 한국행정연구원 기본연구과제, 2020(0), pp.1-473.
- 김정곤·이효진·강반디(2024), 『인도의 데이터 거버넌스 분석과 한 인도 협력에 대한 시사점』, 대외경제정책연구원(KIEP) 중장기통상전략연구.
- 김지영(2021), “데이터 전담 직렬제도의 필요성과 제도화 방안.” 《정부학연구》, 27(2), pp. 89-112.
- 김지현(2023), 『디지털 정책과 규제 변화 분석: Digital Policy Alert 통계를 중심으로』, 대외경제정책연구원.
- 김윤명(2024), “AI 안전을 위한 규제와 거버넌스.” 《GRI 연구논총》, (기획), pp. 25-53.
- 명승환(2020), 스마트 전자정부론. 윤성사.
- 방정미(2021), “공공데이터 활용 실재화를 위한 데이터기반행정.” 《비교법연구》, 21(1), pp. 87-126.
- 배금일·이현규(2021), 『인공지능 시대의 데이터 공급망 관리』, 한국연구재단.
- 성옥준(2017), “데이터 생애주기 관점에서 본 공공부문 빅데이터 활성화 방안.” 《한국지역정보화학회지》, 20(2), pp. 25-41.
- 성지은·송위진·박인용(2014), “사용자 주도형 혁신모델로서 리빙랩 사례 분석과 적용 가능성 탐색.” 《기술혁신학회지》, 17(2), pp. 309-333.
- 선지원(2021), “데이터 기반 행정의 법적 문제,” 《행정법연구》, 66, pp. 107-135.
- 손형섭(2021), “디지털 전환(Digital Transformation)에 의한 지능정보화 사회의 거버넌스 연구,” 《공법연구》, 49(3), pp. 199-230.
- 송경재·장우영·조인호(2018), “빅데이터 거버넌스의 가능성과 과제에 관한 탐색,” 《사회이론》, 53, pp. 153-186.
- 신재형(2024), “인공지능 학습데이터의 수집 및 이용에 대한 규제”. 서울대학교 대학원 법학과 박사학위논문.
- 엄석진·고윤석·박정은·박건철(2025), 『한국의 데이터 정책과 AI정부』, 한올아카데미.
- 엄석진·김준형·최낙범·황한찬·정희정·서교리(2020), 『정부의 디지털 혁신』. 문우사.
- 유승현(2022), “디지털 전환 시대의 데이터 정책과 이용자 권리 보호: 데이터 독점과 데이터 주권을 중심으로,” 《방송통신연구》, 6(3), pp. 38-73.

- 윤건·김윤희(2020), “공공데이터 융합의 영향요인과 매개효과에 관한 탐색적 연구: 데이터 기반행정의 맥락을 중심으로,” 《한국지역정보학회지》, 23(3), pp. 1-22.
- 윤건(2023a), “공공분야 데이터 융합에 관한 연구: 심층면접을 통한 제약조건 탐색을 중심으로,” 《한국사회와 행정연구》, 33(4), pp. 151-158.
- _____(2023a), “공공분야 데이터 융합에 관한 연구: 심층면접을 통한 제약조건 탐색을 중심으로,” 《한국사회와 행정연구》, 33(4), pp. 151-158.
- _____(2023b), “데이터 거버넌스 실증연구: 구성요소 간 구조적 관계와 영향을 중심으로,” 《정보화정책》, 30(3), pp. 29-48.
- 정용찬(2018), 『포스트 코로나19 시대의 데이터 주권과 데이터 거버넌스』, KISDI 정책보고서, 20-10.
- 정원섭(2020), “인공지능 알고리즘의 편향성과 공정성,” 《인간·환경·미래》, 25, pp. 55-73.
- 조세현·차세영(2022), 『클라우드 전환 시대 데이터기반행정 추진 전략: 데이터 분석·활용 강화를 중심으로』, 한국행정연구원 기본연구과제, 2022(0), pp. 1-326.
- 차오신·오세환(2024), “데이터 플랫폼 활용 영향요인에 대한 탐색적 연구: 한국 공공데이터 포털 사례를 중심으로,” 《정보화정책》, 31(4), pp. 87-103.
- 최정묵(2025), 『이타적이고 유능한 AI 정부』, 푸른나무.
- 한국지능정보사회진흥원(2024), 『데이터 신질서와 생태계 조성을 위한 “국가 데이터 인프라” 개념과 추진 전략』, 한국지능정보사회진흥원.
- 행정안전부(2019), “편리한 활용과 융·복합 촉진을 위한 공공데이터 품질관리 중장기계획”.
- _____(2024), “2024년 데이터기반행정 활성화 시행계획(안)”.
- _____(2025), “2024년 데이터기반행정 실태점검 결과”.
- _____(2025), “2025년 데이터기반행정 활성화 시행계획”.

[국외 문헌]

- Aaronson, S. A.(2020). Data governance, AI, and trade: Asia as a case study. *IDEAS Working Paper Series from RePEc*, 6, 2020-6.
- _____(2024). Data governance is not ready for AI. In *The Role of Governance*

- in *Unleashing the Value of Data*(p. 15).
- Aaronson, S., Struett, T., & Zable, A.(2021). DataGovHub Paradigm for a Comprehensive Approach to Data Governance.
- Al-Badi, A., Tarhini, A., & Khan, A. I.(2018). Exploring big data governance frameworks. *Procedia Computer Science*, 141, 271-277.
- Alqudah, M. A.(2021). Towards the governance of government data using artificial intelligence. *SSRN Working Paper*, 3992303.
- Arigbabu, A. T., Olaniyi, O. O., Adigwe, C. S., Adebisi, O. O., & Ajayi, S. A.(2024). Data governance in AI-enabled healthcare systems: A case of the project nightingale. *Asian Journal of Research in Computer Science*, 17(5), 85-107.
- Arner, D. W., Castellano, G. G., & Selga, E. K.(2022). The transnational data governance problem. *Berkeley Technology Law Journal*, 37, 623-700.
- Asian Development Bank.(2022). *Data Management Policies and Practices in Government*. Asian Development Bank.
- Astana Civil Service Hub & United Nations Development Programme.(2022). *Open Government Data Policies and Practices: Select Country Cases*. Astana: Astana Civil Service Hub & UNDP.
- Batool, A., Zowghi, D., & Bano, M.(2025). AI governance: A systematic literature review. *AI and Ethics*, 1-15.
- Beard, S., & Belchamber, J.(2020). *AI and data governance issues in responding to covid-19: a briefing*.
- Benfeldt, O., Persson, J. S., & Madsen, S.(2020). Data governance as a collective action problem. *Information Systems Frontiers*, 22, 299-313.
- Bhouri, H.(2025). Navigating Data Governance: A Critical Analysis of European Regulatory Framework for Artificial Intelligence. *Recent Advances in Public Sector Management*, 171.
- Boppiniti, S. T.(2023). Data ethics in ai: Addressing challenges in machine learning and data governance for responsible data science. *International Scientific Journal for*

- Research*, 35), 1-29.
- Burden, B. C., Canon, D. T., Mayer, K. R., & Moynihan, D. P.(2012). The effect of administrative burden on bureaucratic behavior: Evidence from health care policy. *Public Administration Review*, 72(5), 741-751.
- Carlizzi, D. N., & Quattrone, A.(2022). Artificial intelligence and data governance for precision ePolicy cycle. In *Artificial Intelligence and Economics: the key to the Future*(pp. 67-84). Cham: Springer International Publishing.
- Cazzaniga, M., Pizzinelli, C., Rockall, E., & Tavares, M. M.(2024). Exposure to artificial intelligence and occupational mobility: *A cross-country analysis*. International Monetary Fund.
- Cerrillo-Martinez, A., & Casadesús-de-Mingo, A.(2021). Data governance for public transparency. *Profesional de la Información*, 30(4), 1-13.
- Chamola, V., Hassija, V., Sulthana, A. R., Ghosh, D., Dhingra, D., & Sikdar, B.(2023). A review of trustworthy and explainable artificial intelligence(XAI). *IEEE Access*, 11, 78994-79015.
- Charles, V., Rana, N. P., & Carter, L.(2022). Artificial intelligence for data-driven decision-making and governance in public affairs. *Government Information Quarterly*, 39(4), 101742.
- Chen, Y. C., & Lee, J.(2021). Collaborative data networks for public service: governance, management, and performance. In *Digital Government and Public Management*(pp. 40-58). Routledge.
- Choenni, S., Bargh, M. S., Busker, T., & Netten, N.(2022). Data governance in smart cities: Challenges and solution directions. *Journal of Smart Cities and Society*, 1(1), 31-51.
- Davis, F. D.(1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- Dawes, S. S.(2010). Stewardship and usefulness: Policy principles for information-based transparency. *Government Information Quarterly*, 27(4), 377-383.
- _____(2014). Information policy and governance. *Journal of the Association for*

Information Science and Technology, 65(5), 856-875.

Dunleavy, P., & Margetts, H.(2025). Data science, artificial intelligence and the third wave of digital era governance. *Public Policy and Administration*, 40(2), 185-214.

European Commission(EC).(2020). *A European Strategy for Data*. Brussels: European Commission.

Effoduh, J. O., Akpudo, U. E., & Kong, J. D.(2024). Toward a trustworthy and inclusive data governance policy for the use of artificial intelligence in Africa. *Data & Policy*, 6, e34.

Fernandes, L. M.(2022). *Trajectories and Challenges for Data Governance Implementation in the Public Sector: A systematic review*.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E.(2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689-707.

Fumega, S.(2024). The Global Landscape of Data Governance. *In The Role of Governance in Unleashing the Value of Data*, 5.

Giron, J., et al.(2023). A comparative analysis with machine learning of public data governance and AI policies in the European Union, United States, and China.

Harrison, T. M., ...-(2019). The data firehose and AI in government: Why data management is a key to value and ethics. In *Proceedings of the 20th Annual International Conference on Digital Government Research* (pp. 171-176).

Harrison, T., F. Luna-Reyes, L., Pardo, T., De Paula, N., Najafabadi, M., & Palmer, J.(2019, June). The data firehose and AI in government: Why data management is a key to value and ethics. In *Proceedings of the 20th annual international conference on digital government research* (pp. 171-176).

Harrison, T., F. Luna-Reyes, L., Pardo, T., De Paula, N., Najafabadi, M., & Palmer, J.(2019, June). The data firehose and AI in government: Why data management is a key to value and ethics. *In Proceedings of the 20th annual international conference on digital government research*, 171-176.

- Hausenloy, J., McClements, D., & Thakur, M.(2024). Towards Data Governance of Frontier AI Models. *arXiv preprint* arXiv:2412.03824.
- Heeks, R.(2001). *Understanding e-governance for development*. Manchester: Institute for Development Policy and Management.
- Houser, K., & Bagby, J. W.(2023). Next-generation data governance. *Duke Law & Technology Review*.
- Janssen, M., & Van den Hoven, J.(2015). Big and open linked data(BOLD) in government: A challenge to transparency and privacy? *Government Information Quarterly*, 32(4), 363-368.
- Janssen, M., & Helbig, N.(2018). Innovative and open government: The role of public values and trust. *Information Polity*, 23(1), 79-95.
- Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T.(2020). Data governance: Organizing data for trustworthy Artificial Intelligence. *Government information quarterly*, 37(3), 101493.
- Janssen, M., Konopnicki, D., Snowdon, J. L., & Ojo, A.(2020). Data governance: Organizing data for trustworthy artificial intelligence. *Government Information Quarterly*, 37(3), 1-8.
- Ji, M., Gu, X., Guo, Q., & Ding, X.(2024, August). Research on Government Data Governance in the Era of Large Language Model. In *Proceedings of the 9th IEEE International Conference on Data Science in Cyberspace(DSC)* (pp. 668-671). IEEE.
- Jiong'en, X., & Yingliang, W.(n.d.). AI-enhanced government data governance in the era of big data: Exploring sharing mechanisms and management systems. *Innovative Applications of AI*, 1(2), 22-30.
- Kanying, T., Thammaboosadee, S., & Chuckpaiwong, R.(2023, December). Formulating Analytical Governance Frameworks: An Integration of Data and AI Governance Approaches. In *Proceedings of the 13th International Conference on Advances in Information Technology* (pp. 1-9).
- Kitchin, R.(2014). *The data revolution: Big data, open data, data infrastructures and their*

- consequences*. Sage.
- Klievink, B., Van Der Voort, H., & Veeneman, W.(2018). Creating value through data collaboratives: Balancing innovation and control. *Information Polity*, 23(4), 379-397.
- Koltay, T.(2016). Data governance, data literacy and the management of data quality. *IFLA Journal*, 42(4): 303-312.
- Kroll, J. A.(2018). Data science data governance [AI ethics]. *IEEE Security & Privacy*, 16(6), 61-70.
- Kuziemski, M., & Misuraca, G.(2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Government Information Quarterly*, 37(3), 101456.
- Kuzio, J., Ahmadi, M., Kim, K. C., Migaud, M. R., Wang, Y. F., & Bullock, J.(2022). Building better global data governance. *Data & Policy*, 4, e25.
- Kim, S., Andersen, K. N., & Lee, J.(2022). Platform government in the era of smart technology. *Public Administration Review*, 82(2), 362-368.
<https://doi.org/10.1111/puar.13422>.
- Liu, J.(2022). Social data governance: Towards a definition and model. *Big Data & Society*, 9(2), 1-14.
- Luna-Reyes, L. F., & Harrison, T. M.(2024). Enterprise data governance for artificial intelligence: implications from algorithmic jobseeker profiling applications in government. In *Research Handbook on Public Management and Artificial Intelligence* (pp. 187-206). Edward Elgar Publishing.
- MacFeely, S., Me, A., Fu, H., Veerappan, M., Hereward, M., Passarelli, D., & Schüür, F.(2022). Towards an international data governance framework. *Statistical Journal of the IAOS*, 38(3), 703-710.
- MacFeely, S., Nastav, B., & Stangl, A.(2022). Governing data as a public good: National and international perspectives. *Statistical Journal of the IAOS*, 38(1), 1-20.
- Mahanti, R.(2018). Data Governance Implementation: Critical Success Factors. *Software Quality Professional*, 20(4).

- Mahanti, R.(2018). Data governance: Perspectives and practices. *Computer*, 51(2), 94-97.
- Malomo, F., & Sena, V.(2017). Data intelligence for local government? Assessing the benefits and barriers to use of big data in the public sector. *Policy & Internet*, 9(1), 7-27.
- Mao, Z., Wu, J., Qiao, Y., & Yao, H.(2022). Government data governance framework based on a data middle platform. *Aslib Journal of Information Management*, 74(2), 289-310.
- Marcucci, S., Alarcón, N. G., Verhulst, S. G., & Wüllhorst, E.(2023). Informing the global data future: benchmarking data governance frameworks. *Data & Policy*, 5, e30.
- Marcucci, E., Gatta, V., Le Pira, M., & Yamoah, F.(2023). A global meta-framework for data governance: Principles and practices. *Government Information Quarterly*, 40(1), 101755.
- McGuire, M.(2006). Collaborative public management: Assessing what we know and how we know it. *Public administration review*, 66, 33-43.
- McNamara, M.(2012). Starting to untangle the web of cooperation, coordination, and collaboration: A framework for public managers. *International Journal of Public Administration*, 35(6), 389-401.
- Medhora, R. P., Aaronson, S. A., Aggarwal, S., Breznitz, D., Ciuriak, D., Clement, A., ... & Wylie, B.(2018). *Data governance in the digital age*. Centre for International Governance.
- Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A.(2020). Emerging models of data governance in the age of datafication. *Big Data & Society*, 7(2), 1-15.
- Misuraca, G.(2007). *E -governance in Africa, from theory to action: A handbook on ICTs for local governance*. Idrc.
- Nielsen, O. B.(2017). A comprehensive review of data governance literature. *Selected Papers of the IRIS*, 8, 120-133.
- Nikiforova, A., & Zuiderwijk, A.(2021). Barriers and enablers of open data provision and usage. *Government Information Quarterly*, 38(2), 101566.

- Nookala, G.(2024). Adaptive Data Governance Frameworks for Data-Driven Digital Transformations. *Journal of Computational Innovation*, 4(1).
- Pahune, S., Akhtar, Z., Mandapati, V., & Siddique, K.(2025). The importance of AI data governance in large language models. *Big Data and Cognitive Computing*, 9(6).
- Park, J.(2018). A comparative study of big data, open data, and my data. *The journal of Bigdata*, 3(1), 41-46.
- Pitoura, E.(2020). Social-minded measures of data quality: fairness, diversity, and lack of bias. *Journal of Data and Information Quality(JDIQ)*, 12(3), 1-8.
- Redden, J.(2018). Democratic governance in an age of datafication: Lessons from mapping government discourses and practices. *Big Data & Society*, 5(2), 2053951718809145.
- Robinson, Diane, Cleary, Shay, Gibson, Sarah, & Miller, A. L.(2024). Data Governance and AI in State Courts. *Trends in State Courts*, 2024, 9-20.
- Rollings, M., Duncan, A. D., & Logan, V.(2019). *10 Ways CDOs Can Succeed in Forging a Data-Driven Organization*. Gartner, Inc.
- Safarov, I.(2019). Institutional dimensions of open government data implementation: Evidence from the Netherlands, Sweden, and the UK. *Public Performance & Management Review*, 42(2), 305-328.
- Schoeman, P.(2023). *Bridging the AI / Data Governance Gap* [Master's thesis, Open Universiteit]. Open Universiteit research portal.
- Sugureddy, A. R.(2024). Generative AI for data stewards: Enhancing accuracy and efficiency in data governance. *International Journal of Artificial Intelligence Research and Development*, 2, 189-202.
- Taeihagh, A.(2025). Governance of Generative AI. *Policy and Society*, 44(1), 1-22.
- Tan, E.(2021). A conceptual model of the use of AI and blockchain for open government data governance in the public sector. *DIG4FED Consortium*.
- Tennison, J.(2024). Why we need inclusive data governance in the age of AI. *The Role of Governance in Unleashing the Value of Data*, 74-82.
- Tukia, A.(2022). *Data governance for sustainable artificial intelligence* [Master's thesis,

- Tampere University]. Tampere University repository.
- Ubaldi, B.(2013). Open government data: Towards empirical analysis of open government data initiatives. *OECD Working Papers on Public Governance, No. 22*. OECD Publishing.
- Van Dijk, J.(2020). *The digital divide*. John Wiley & Sons.
- Van Zoonen, L.(2020). Data governance and citizen participation in the digital welfare state. *Data & Policy, 2*, e10.
- Verma, R., Shrivastava, P., & Merla, N.(2024). Tracing the path: Data lineage and its impact on data governance. *International Journal of Global Innovations and Solutions, 1*-10.
- Viljoen, S.(2021). A relational theory of data governance. *The Yale Law Journal, 131*(2), 573-654.
- Wang, A. X., Simpson, C. R., & Nguyen, B. P.(2024, March). Enhancing Data Governance through Data-Centric AI: Case Study in New Zealand Government Sector. In *2024 16th International Conference on Computer and Automation Engineering*, 64-69.
- Watanabe, T.(2024). *AI & data governance*.
- Winter, J. S.(2021). AI in healthcare: Data governance challenges. *Journal of Hospital Management and Health Policy, 5*(8).
- Winter, J. S., & Davidson, E.(2019). Governance of artificial intelligence and personal health information. *Digital Policy, Regulation and Governance, 21*(3), 280-290.
- Wirtz, B. W., Weyerer, J. C., & Kehl, I.(2022). The governance of artificial intelligence: Ethical, legal and societal challenges. *Information Polity, 27*(1), 37-51.
- Yang, T. M., & Wu, Y. J.(2021). Exploring the determinants of public employees' data sharing intention: The roles of motivation, pressure, and trust. *Government Information Quarterly, 38*(2), 101566.
- Yukhno, A.(2024). Digital transformation: Exploring big data governance in public administration. *Public Organization Review, 24*(1), 335-349.

[기타 자료]

김학래(2020), “데이터기반행정 활성화법 통과에 따른 데이터 거버넌스 재정립”, 제3회 전자정부의 날 기념식 발표자료.

머니투데이(2019. 7. 1), “데이터경제 활성화, 쪼개진 ‘정부 거버넌스’ 통합부텨”,

<https://news.mt.co.kr/mtview.php?no=2019062914103777780>.

전자신문(2022. 8. 31), “공공데이터 거버넌스”, <https://www.etnews.com/20220831000201>.

투이컨설팅(2020. 12. 22), “공공부문의 데이터 거버넌스”,

<https://www.2e.co.kr/news/articleView.html?idxno=301052>.

Cabinet Office.(2020). UK government data strategy 2020-2025. HM Government.

Digital Agency of Japan.(2021). Digital governance strategy. Government of Japan.

French Government.(2017). Entrepreneur d'intérêt général programme. Etalab, Prime Minister's Office.

Government Digital Service(GDS).(2018). Data ethics framework. UK Cabinet Office.

Government of Australia.(2015a). Data sharing principles. Department of the Prime Minister and Cabinet.

_____ (2015b). Public data policy statement. Department of the Prime Minister and Cabinet.

Government of Canada.(2019). Digital standards. Treasury Board Secretariat.

_____ (2020). Data stewardship program. Treasury Board Secretariat.

Government of the United States.(2018). Foundations for Evidence-Based Policymaking Act of 2018. Public Law No. 115-435.

_____ (2019). Federal data strategy – 2020 action plan. Office of Management and Budget.

GOVNET TECHNOLOGY(2023. 12. 6), “A comprehensive guide to public sector data governance framework templates”,

<https://blog.govnet.co.uk/technology/a-comprehensive-guide-to-public-sector-data-governance-framework-templates>.

GovTech Singapore.(2021). GovTech academy programs. Government Technology Agency.

kiteworks(2024. 9. 20), “Data governance and digital transformation in the public sector,
<https://www.kiteworks.com/cybersecurity-risk-management/data-governance-in-the-public-sector/>,

New Zealand Government.(2020). Data leadership network: Annual report. Government Chief Data Steward.

IMD.(2024). IMD World Digital Competitiveness Ranking 2024. Lausanne: International Institute for Management Development.

Republic of Estonia.(2018). Digital agenda 2020 for Estonia. Ministry of Economic Affairs and Communications.

Smart Nation and Digital Government Office.(2018). Smart nation strategy. Prime Minister’s Office, Singapore.

Stanford University Human-Centered Artificial Intelligence.(2023). Artificial Intelligence Index Report 2023. Stanford, CA: Stanford University.

sunlightfoundation, “Open data policy guidelines”,
<https://sunlightfoundation.com/opendataguidelines/>.

OECD(2019). The path to becoming a data-driven public sector. OECD Digital Government Studies. OECD Publishing.

_____(2022), Going digital guide to data governance policy making, OECD Publishing, Paris.

_____(2023). Government at a glance 2023. OECD Publishing.

_____(2024). 2023 digital government index: Results and key findings. OECD Public Governance Policy Papers, No. 44, OECD Publishing.
<https://doi.org/10.1787/1a89ed5e-en>.

thepublicnews(2019. 1. 21), “공공 분야 빅데이터 어떻게 활용해야 하나”,
<https://www.thepublicnews.co.kr/news/articleView.html?idxno=20132>.

ZDNET(2025. 3. 28), “소버린 클라우드가 AI 시대 데이터 거버넌스 구도를 재편하는 방법”,
<https://zdnet.co.kr/view/?no=20250328163508>.

[부록 1] 공공부문 데이터거버넌스 실제에 대한 서면 인터뷰 프로토콜

안녕하세요.

(사)한국정책학회의 연구진은 정보통신정책연구원이 매년 수행하는 메가트렌드 연구의 일환으로 <AI 디지털 전환에 따른 공공부문 데이터거버넌스 전략>을 연구하고 있습니다.

공공부문 데이터거버넌스에 대하여 대학과 연구기관 전문가의 의견을 수렴할 뿐만 아니라, 중앙부처, 지방자치단체, 공공기관 등 공공부문에서 데이터가 어떻게 실제로 활용되고 있고 원활한 활용을 위해 어떤 전략이 필요한지에 대해 전문가와 실무자의 의견을 모으고자 합니다.

공공부문 데이터거버넌스에 관련하여, 데이터의 수집과 이용의 차원, 조직이 이용가능한 기술과 시스템의 차원, 인적자원과 조직관리의 차원, 법률과 제도의 차원 등 다양한 관점에서 의견을 주시면 고맙겠습니다.

본 인터뷰는 개별 인터뷰, Focus Group 인터뷰, 대면 인터뷰, 전화 인터뷰, 온라인 인터뷰, 서면 인터뷰 등 다양한 방식으로 진행될 예정입니다. 본 인터뷰는 연구 윤리를 철저히 준수하여, 인터뷰에 응해주시는 전문가의 성명과 소속 등을 익명으로 처리하고 인터뷰의 내용도 학술적으로 일반화할 수 있는 전략을 도출하는 목적으로만 사용될 것입니다.

공공부문 데이터거버넌스의 혁신을 위해서는 이론적 전문가와 실무적 전문가분들의 경험, 관찰, 통찰이 매우 중요하기 때문에, 고견을 청취하여 데이터거버넌스 전략 수립에 보탬이 되도록 하겠습니다.

감사합니다.

연구진

- 책임연구원: 남태우(성균관대 국정전문대학원)
- 공동연구원: 기남훈(성균관대 국정전문대학원), 김준형(국립순천대학교), 황한찬(국립순천대학교), 신승윤(소프트웨어정책연구소), 김기동(중부대학교)
- 연구보조원: 조수향, 하선우, 김정원(성균관대 국정전문대학원 박사과정)

인터뷰 내용(범주)

1. 개인

- 본인의 직무와 경력(공공조직 관련련)
- 본인이 다루는 데이터의 속성(정책영역, 정형/비정형)
- 데이터를 직접적으로 다루지 않는다면 데이터와 직무와의 관련성

2. 데이터

- 데이터가 생성 혹은 수집되어 활용하기까지의 소과정
- 데이터 소주기에 있어서의 이슈와 쟁점
- 데이터 소주기 관리에 있어서의 장애요인과 성공요인

3. 기술과 시스템

- 데이터 활용에 있어서 사용되는 기술(AI, 통계패키지, 시각화툴)
- 데이터 활용에 있어서 사용되는 시스템(플랫폼, 대시보드, 클라우드 등)
- 기술과 시스템 차원에 있어서 장애요인과 성공요인

4. 조직과 관리

- 데이터 거버넌스에 있어서 리더십의 역할
- 조직내 계획의 체계성
- 인적자원의 전문성 강화
- 아웃소싱 관리 역량
- 데이터기반행정 관련 조직문화
- 조직 관리 차원에 있어서 장애요인과 성공요인

5. 법률과 제도

- 데이터기반행정법 등 관련 법제도의 영향력
- 법제도와 실무 간의 격차
- 당위론적 방향성과 실제적 현실 간의 괴리
- 법제도 차원의 장애요인과 성공요인

인터뷰 계획

연구진은 2~3인이 조를 이루어 개인면접이나 FGI를 대면, 서면, 온라인 등의 방식으로 진행합니다. 인터뷰 시간은 1~2시간 소요됩니다.

본인의 경험과 관찰에 비추어 응답할 수 있는 질문에 응답해주시고, 별지에 편하게 의견을 말씀해주시는 등 다양한 의견을 주시면 감사하겠습니다.

인터뷰 질문

1. 개인

(A1) [개인] 현재 맡고 있는 직무와 역할에 대해 간략히 설명 부탁드립니다.

(A2) [개인] 현재 다루고 있는 데이터의 특성(정책영역, 데이터 유형(정형/비정형) 등)에 대해 설명 부탁드립니다.

(A3) [개인] 본인이 직접 데이터를 다루지 않으신다면, 본인의 직무와 데이터 간의 관련성은 어떻게 인식하고 계십니까?

(A4) [개인] 본인의 업무에서 데이터가 중요하다고 느끼는 부분 또는 데이터 활용에 있어 어려움을 경험했던 사례가 있으시다면 설명 부탁드립니다.

2. 데이터

(B1) [데이터] 귀 기관의 데이터 관리 및 연계 체계는 현재 어떻게 운영되고 있는지 설명해 주십시오.(운영 방식, 연계 수준, 문제점 등을 포함)

(B2) [데이터] 귀하의 기관에서 데이터가 생성되어 활용되기까지의 과정을 단계별로 설명 부탁드립니다.

(B3) [데이터] 기관의 데이터 생명주기 관리에서 가장 중요하다고 생각하는 단계는 무엇인지, 그 이유와 함께 설명 부탁드립니다.

(B4) [데이터] 데이터의 생성 → 수집 → 제공 → 가공 → 분석 → 활용의 전 생애주기 중 귀 기관(또는 귀하의 정책영역)에서 가장 정비가 필요하다고 생각하는 단계와 그 이유를 말씀해 주십시오.

(B5) [데이터] 귀 기관 또는 부문에서 데이터 연계·활용 또는 전주기 관리 과정에서 겪는 주요 애로사항이나 반복적인 장애요인이 있다면 설명해 주시고, 이를 해결하기 위한 기관의 대응 방안이 있다면 함께 설명 부탁드립니다.

(B6) [데이터] 데이터 품질 관리 측면에서는 어떤 어려움을 겪고 계시는지 구체적인 사례를 들어 설명 부탁드립니다.

(B7) [데이터] 데이터 관리가 성공적으로 이루어지기 위해 꼭 필요한 조건이나 요소가 있다면 무엇이라고 생각하십니까?

(B8) [데이터] 해외(OECD 등) 데이터 거버넌스 사례 중, 우리나라 공공부문에 적용할 필요가 있다고 생각하는 요소와 그 이유를 말씀해 주십시오.

3. 기술과 시스템

(C1) [기술과 시스템] 귀 기관이 데이터 활용 과정에서 사용하고 있는 주요 기술, 시스템 또는 플랫폼이 있다면 작성해 주십시오.(예: AI 분석툴, 통계 패키지, 시각화 툴, 클라우드, 대시보드 등) 이러한 기술이 기관의 데이터 거버넌스 측면에서 어떤 장점과 한계를 가진다고 생각하십니까?

(C2) [기술과 시스템] 기관에서 운영하는 데이터 관련 시스템(플랫폼, 클라우드 등)의 현황에 대해 설명 부탁드립니다.

(C3) [기술과 시스템] 기관의 시스템 간 연계 및 통합 측면에서 불편함을 겪은 경험이 있으시다면 설명 부탁드립니다.

(C4) [기술과 시스템] AI 기반 데이터 분석 및 활용 행정을 위해 반드시 도입되거나 개선이 필요하다고 생각하시는 기술·시스템 요소가 있다면 말씀해 주십시오.

(C5) [기술과 시스템] 기술 및 시스템 측면에서 데이터 거버넌스 성공을 위한 핵심 요소는 무엇이라고 생각하십니까?

(C6) [기술과 시스템] 데이터 활용 기술이나 시스템 운영 과정에서 가장 큰 어려움은 무엇인지, 그리고 이를 개선하기 위한 방안을 제안해 주십시오.

4. 조직과 관리

(D1) [조직과 관리] 생성AI 대응 정책: ChatGPT 등 생성AI의 공공부문 도입에서 데이터 보안, 환각 현상, 저작권 등의 이슈를 어떻게 거버넌스할 것인가?

(D2) [법률과 제도] OECD AI 원칙, ISO/IEC AI 표준, EU AI Act 등 국제 표준과의 정합성 확보 방안은?

(D3) [데이터] 생애주기 기반 AI 거버넌스: AI 모델의 개발부터 배포까지 전체 생애주기에 걸친 데이터 관리 프레임워크에서 공공부문이 우선적으로 고려해야 할 핵심 범위는?(최소 세가지 범주)

(D4) [조직과 관리] 공공 데이터 거버넌스 추진 과정에서 조직·관리 차원에서 발생하는 주요 장애요인은 무엇인지, 그 중 가장 시급히 해결해야 할 과제와 이유를 설명해 주십시오.

(D5) [조직과 관리] 조직 구성원들의 데이터 거버넌스에 대한 인식과 참여도는 어떤 수준이며, 개선이 필요하다면 어떠한 방안이 효과적인 것이라 생각하십니까?

(D6) [조직과 관리] 기관 간 데이터 공유와 협업을 활성화하기 위한 인센티브 또는 운영 방안으로 효과적이라고 생각하는 방법과 그 이유를 말씀해 주십시오.

(D7) [조직과 관리] 데이터 기반 행정 추진을 위한 조직문화의 특징, 구성원 인식과 참여, 그리고 이를 높이기 위한 조직 운영 및 인력관리 방안에 대해 말씀해 주십시오.

(D8) [조직과 관리] 데이터 거버넌스에서 리더십의 역할은 무엇이라고 생각하시며, 현재 귀하 기관의 리더십은 어떻게 평가하십니까?

(D9) [조직과 관리] 데이터 거버넌스 관련 인적 자원의 전문성 강화를 위해 기관에서는 어떤 노력을 기울이고 있습니까?

(D10) [조직과 관리] 귀하의 기관 내에서 데이터 관련 계획이나 정책이 얼마나 체계적으로 수립되어 진행되고 있다고 생각하십니까?

(D11) [조직과 관리] 성공적인 데이터 관리 조직의 사례 또는 데이터 거버넌스 성공을 위한 핵심 요소가 있다면 무엇이라고 생각하십니까?

5. 법률과 제도

(E1) [법률과 제도] 공공 데이터 거버넌스 정착을 저해하는 법·제도적 장애요인은 무엇이라고 보십니까? 현행 제도의 한계나 실무와의 괴리 사례도 함께 설명해 주십시오.

(E2) [법률과 제도] 귀하의 기관에서 데이터 기반 행정 관련 법제도(예: 데이터기반행정법 등)가 실무에 미치는 영향이 있습니까? 또는 개인정보보호법, 공공데이터법 등 관련

법령들 간의 상호작용에서 이슈가 있습니까?

(E3) [법률과 제도] 현행 관련 법제도와 실제 업무 현장 간에 어떤 격차가 존재한다고 생각하십니까? 이러한 격차를 해소하기 위해서는 어떠한 개선 방안이 필요하다고 생각하십니까?

(E4) [법률과 제도] 데이터 기반 행정 활성화를 위해 가장 시급히 개선해야 할 법·제도적 영역은 무엇이라고 생각하십니까? 그 이유와 함께 설명해 주십시오.

(E5) [법률과 제도] 향후 정부의 데이터 거버넌스 정책에서 반드시 보완하거나 강화해야 한다고 보시는 기술적·제도적 지원 요소가 있다면 말씀해 주십시오.

(E6) [법률과 제도] 이상적인 데이터 거버넌스 구현을 위해 법제도 측면에서 우선적으로 개선이 필요한 점은 무엇이라고 생각하십니까? 또한, 법제도 차원에서 데이터 거버넌스 발전을 위한 핵심 성공 요인은 무엇이라고 생각하십니까?

[부록 2] 공공부문 데이터거버넌스 프레임워크에 대한 서면 인터뷰 프로토콜

안녕하세요.

(사)한국정책학회 연구진은 정보통신정책연구원이 매년 수행하는 메가트렌드 연구의 일환으로 <AI 디지털 전환에 따른 공공부문 데이터거버넌스 전략>을 연구하고 있습니다.

연구진이 제시한 공공부문 데이터거버넌스 관련 프레임워크와 전략에 대하여 대학과 연구기관 전문가의 의견을 수렴할 뿐만 아니라, 중앙부처, 지방자치단체, 공공기관 등 공공부문에서 데이터가 어떻게 실제로 활용되고 있고 원활한 활용을 위해 어떤 전략이 필요한지에 대해 전문가와 실무자의 의견을 모으고자 합니다.

본 인터뷰는 개별 인터뷰, Focus Group 인터뷰, 대면 인터뷰, 전화 인터뷰, 온라인 인터뷰, 서면 인터뷰 등 다양한 방식으로 진행될 예정입니다. 본 인터뷰는 연구 윤리를 철저히 준수하여, 인터뷰에 응해주시는 전문가의 성명과 소속 등을 익명으로 처리하고 인터뷰의 내용도 학술적으로 일반화할 수 있는 전략을 도출하는 목적으로만 사용될 것입니다.

공공부문 데이터거버넌스의 혁신을 위해서는 이론적 전문가와 실무적 전문가분들의 경험, 관찰, 통찰이 매우 중요하기 때문에, 고견을 청취하여 데이터거버넌스 전략 수립에 보탬이 되도록 하겠습니다.

감사합니다.

연구진

- 책임연구원: 남태우(성균관대 국정전문대학원)
- 공동연구원: 기남훈(성균관대 국정전문대학원), 김준형(국립순천대학교),
황한찬(국립순천대학교), 신승윤(소프트웨어정책연구소), 김기동(중부대학교)
- 연구보조원: 조수향, 하선우, 김정원(성균관대 국정전문대학원 박사과정)

인터뷰 계획

연구진은 2~3인이 조를 이루어 개인면접이나 FGI를 대면, 서면, 온라인 등의 방식으로 진행합니다. 인터뷰 시간은 1~2시간 소요됩니다.

본인의 경험과 관찰에 비추어 응답할 수 있는 질문에 응답해주시고, 별지에 편하게 의견을 말씀해주시는 등 다양한 의견을 주시면 감사하겠습니다.

1. OECD 기본 프레임의 수정과 고도화

1) 통합 데이터 거버넌스 모형

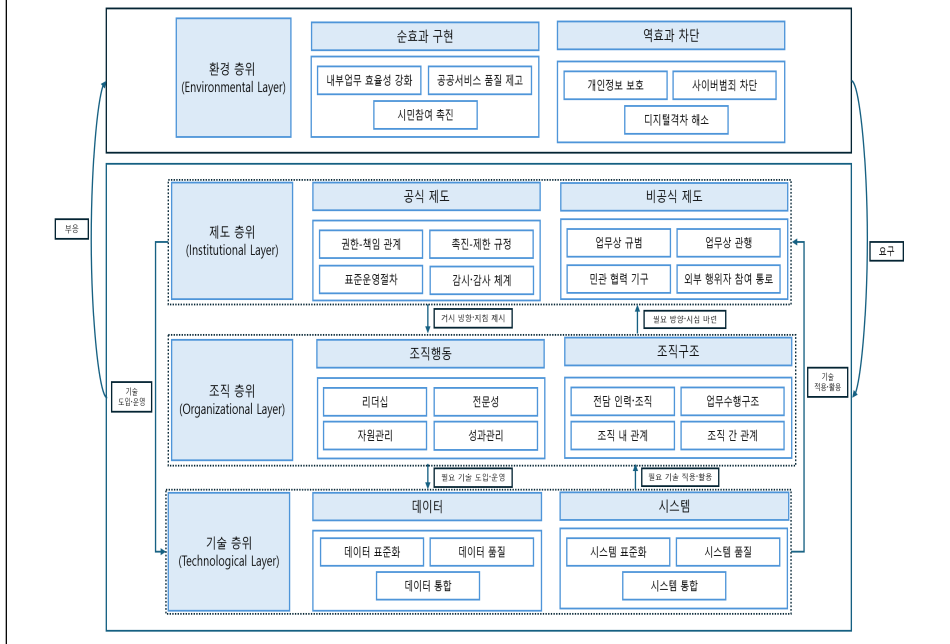
통합 데이터 거버넌스 모형 설명

통합 데이터 거버넌스 모형은 크게 수요자(예: 시민, 기업)의 필요와 요구에 관한 요소들로 구성되는 환경 층위와 공급자(예: 행정기관, 공공기관)의 자원과 역량에 관한 요소들로 구성된 제도·조직·기술 층위로 구분됨

(넓은 맥락) 환경 층위는 제도·조직·기술 층위에 데이터 및 인공지능 기반 행정에 대한 필요와 요구를 투입하고, 제도·조직·기술 층위는 환경 층위의 필요와 요구에 부응하는 제품 또는 서비스를 산출물로서 제공함

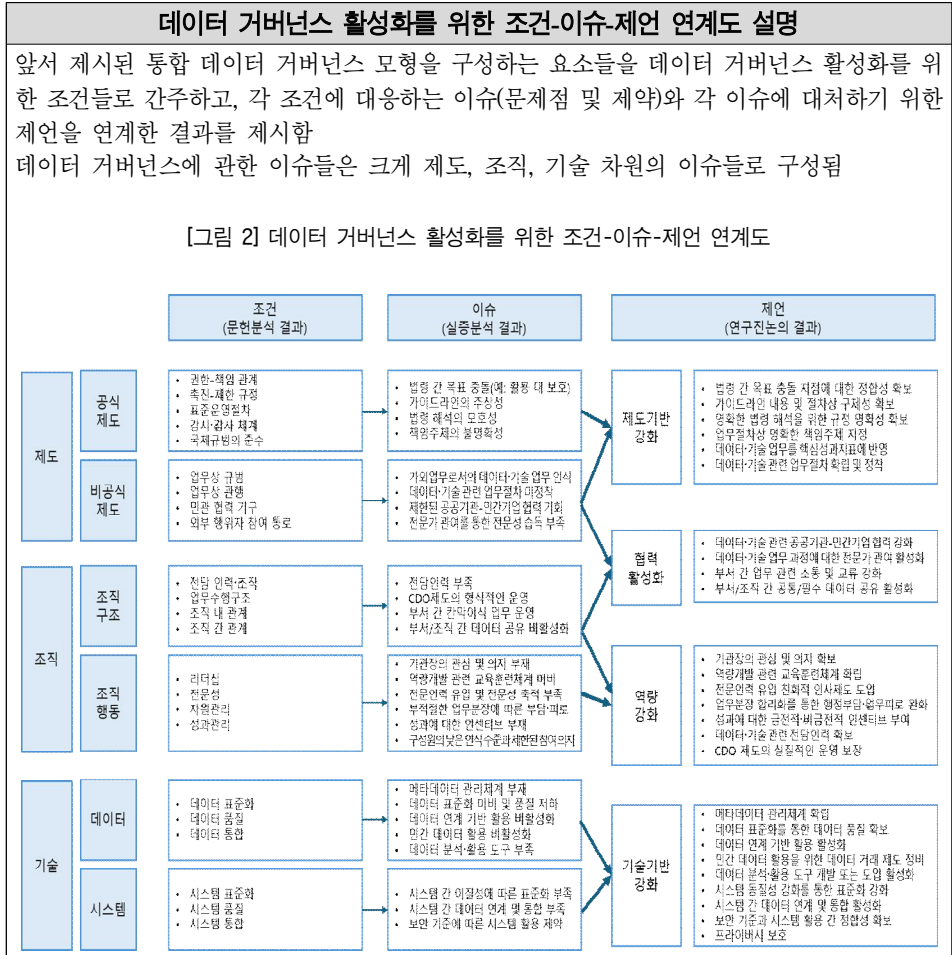
(좁은 맥락) 제도→조직→기술로 연결되는 제도·조직·기술 층위는 데이터와 인공지능을 포함하는 기술의 도입·운영 과정을 보여주고(제도에 의한 기술의 도입·운영 관련 거시 방향·지침 제시, 조직에 의한 필요 기술 도입·운영), 기술→조직→제도로 연결되는 제도·조직·기술 층위는 데이터와 인공지능을 포함하는 기술의 적용·활용 과정을 보여줌(조직에 의한 필요 기술 적용·활용, 필요 기술 적용·활용을 위한 조직의 필요 방향·지침 마련).

[그림 1] 통합 데이터 거버넌스 모형



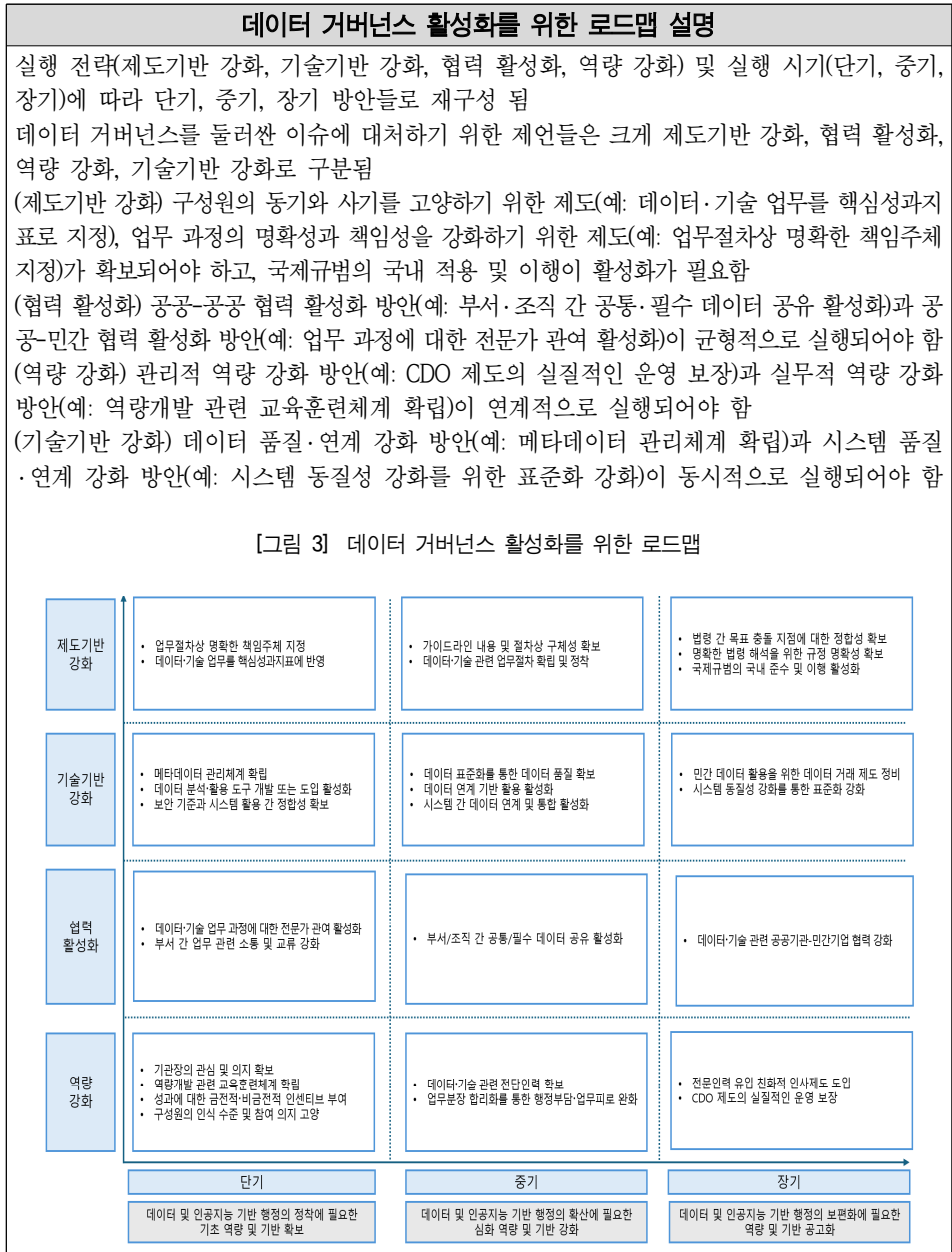
A. 위와 같이 수정된 OECD 통합 데이터 거버넌스 모형의 구성요소에 대해 어떻게 생각하십니까?

2) 데이터 거버넌스 활성화를 위한 조건-이슈-제언 연계도



B. 위와 같이 통합 데이터 거버넌스 활성화를 위한 조건-이슈-제언 연계도에 대해 어떻게 생각하십니까?

3) 데이터 거버넌스 활성화를 위한 로드맵



C. 위와 같이 데이터 거버넌스 활성화를 위한 로드맵에 대해 어떻게 생각하십니까?

2. 구체적인 거버넌스 전략

데이터 거버넌스 전략 설명		
본 연구가 제시한 프레임워크에 기반하여 구체적인 거버넌스 전략을 다음과 같이 재구성할 수 있음. 이는 Janssen et al.(2020)이 제시한 데이터 거버넌스 전략을 본 연구의 프레임워크와 연결하여 재구성함 구체적으로 계획 기반 접근, 조직 기반 접근, 위험 기반 접근으로 나누어서 이론적 검토와 질적 분석을 통해 연구진이 검토한 주요 결과를 정리함 〈표 2〉 데이터 거버넌스 수행전략		
구분	도식화	내용
계획기반 접근		• 문제해결을 위하여 단기, 중기, 장기 전략 수행 계획 수립·실행·감독·평가 • 예: “데이터 거버넌스 가이드라인 연구”를 과제발주/입찰 방식으로 수행(제도기반 강화) • 예: “연구 데이터의 표준화·현행화 용역”을 과제발주/입찰 방식으로 수행 • 중장기계획으로 데이터 거래 제도 정비
조직기반 접근		• 데이터/알고리즘 전문인력 관리부서 신설(전문인력 채용, 교육훈련 등 담당) • 데이터·기술 관련 전담인력 확보 • CDO 보직의 신설 • 성과에 대한 금전적·비금전적 인센티브 부여
위험기반 접근		• 보안 기준과 시스템 활용 사이 부정합성 발생 경고 시스템 도입 • 프라이버시 침해 등 보안 문제 식별 • 시스템 간 데이터 오류 발생 식별 • 데이터 표준화 관련 체계적 오류 발생 식별

D. 위와 같이 데이터 거버넌스 수행전략에 대해 어떻게 생각하십니까?

3. 데이터 거버넌스-AI 거버넌스 연계 프레임워크

1) 데이터 거버넌스와 AI 거버넌스 간 공통점/차이점

데이터 거버넌스와 AI 거버넌스 간 공통점/차이점 설명				
데이터 거버넌스는 데이터의 표준/책임/품질기반을 마련, AI 거버넌스는 이를 투명성/윤리/책임성 관리로 확장하는 구조				
〈표 2〉 데이터 거버넌스와 AI 거버넌스 간 공통점/차이점 정리				
구분	데이터 거버넌스	AI 거버넌스	공통점	차이점
정의	데이터의 생성·수집·저장·공유·활용 전 과정에서 품질·보안·표준·책임 확보 위한 관리·의사결정 체계	AI 시스템(모델·알고리즘)의 개발·검증·운영 전 과정에 공정성·투명성·책임성 확보 위한 관리·의사결정 체계	디지털 자원의 신뢰·책임성·투명성 확보를 위한 제도적·조직적 체계	(데이터) 정보관리 중심 (AI) 알고리즘/의사결정 관리 중심
목표	데이터 품질 보장, 상호운용성, 개인정보보호, 데이터 활용 촉진	알고리즘 공정성, 설명가능성, 안전성, 사회성 수용성 확보	공공 신뢰성 제고, 투명성, 사회적 가치 창출	(데이터) 정합성, 보안 중심 (AI) 윤리, 편향성 제거 중심
관리 대상	원천 데이터, 메타데이터, 가명정보, 공공/민간 데이터 자산	모델, 알고리즘, 학습 데이터, 예측결과물	디지털 자원 전체를 포괄	(데이터) 데이터 자산 중심 (AI) 모델과 의사결정 중심
관리 범위	(데이터생애주기) 생성→저장→제공→분석→활용/폐기	(AI생애주기) 데이터→학습→모델개발→배포→운영·모니터링	전 생애주기 관리 (Lifecycle governance) 목표	(데이터) 정보 관리 중심 (AI) 의사결정 관리 중심
주요 쟁점	표준화 부족, 시스템 간 연계 부재, 책임소재 불명, 개인정보보호 vs. 활용 갈등	알고리즘 편향, 설명가능성 한계, 책임주체 불명확, 윤리적 리스크	법·제도·윤리적 문제 모두 포함	(데이터) 품질·표준화 중심 (AI) 윤리·공정성 중심
정책/제도 기반	공공데이터법, 데이터기반행정법, 개인정보보호법	EU AI Act, OECD AI 원칙, AI 기본법/기본법 시행령(한국)	국제적 규범과 표준화 필요	(데이터) 국내 데이터법 중심 (AI) 국제 윤리·규제 중심
조직/관리 구조	CDO, 공공데이터전략위원회(한국) 등	CAIO, 국가AI전략위(한국) 등	전담조직 및 책임자 필요	(데이터) 정보 관리 중심 (AI) 윤리/리스크 관리 중심

E. 위와 같이 <표 2>의 공통점과 차이점에 대해 어떻게 생각하십니까?

2) 데이터 거버넌스-AI 거버넌스 연계 프레임워크

데이터 거버넌스-AI 거버넌스 연계 프레임워크 설명			
인터뷰 결과를 바탕으로 데이터 거버넌스-AI 거버넌스 연계 프레임워크를 제시함			
〈표 3〉 데이터 거버넌스-AI 거버넌스 연계 프레임워크			
차원	데이터 거버넌스의 실제	데이터 거버넌스의 역할	AI 거버넌스와 연계 포인트
리더십	최고관리자의 전략적 리더십 및 비전 부족 데이터 업무를 부차업무로 인식 사일로, 책임 불명확	CDO 임명 및 전담 조직 운영 데이터 기반 행정 문화 정착 부서간 조정 및 책임/권한 명확화	AI 책임자(Chief AI Officer) 지정, 거버넌스 협업 구조 연계 AI 프로젝트는 데이터 기반 이므로, 데이터 리더십을 CAIO 등과 연계 필요 CDO가 데이터 품질을 관리, CAIO는 이를 바탕으로 모델 개발, 윤리, 성과 관리까지 책임
법·제도	개인정보보호법(보호) vs. 공데이터법/ 데이터 기반 행정법(개방) 충돌 데이터기반행정법 실행력 낮음(진흥법적 구조에 기인) 책임·소유권 모호, 패널리즘심	법령 간 정합성 확보(보호와 활용 간 균형 추구) 데이터 관리 전 생애주기 법제화 책임 및 소유권 규정 명확화	AI Act, OECD 원칙 등 글로벌 규범과의 정합성 확보 필요 AI 학습과정에서 개인정보 활용 시 가명처리/비식별화 기준 등을 데이터 거버넌스에서 제공 → AI 거버넌스에서 데이터 편향성·윤리 검증으로 확장 데이터 거버넌스는 데이터 전 생애주기 법제화를 추구, AI 거버넌스는 이를 모델 전 생애주기 규제로 연결
표준화	기관별 관리체계 상이 → 기관별 포맷·코드 불일치 메타데이터·품질관리 부족 데이터 코드·포맷 불일치 데이터분석·활용도구 부족	데이터 표준/ 메타데이터 체계 확립 상호운용성 확보 품질관리 및 정기 업데이트 체계화	AI 학습데이터 품질 및 출처 검증 체계 마련 데이터 표준화→AI 학습데이터셋 품질 검증 지표로 연결
기술 인프라 (인적/물적)	시스템 간 연계 부재(수작업, 이중입력) 전담인력 및 역량 부족(특히 지방) 보안 규제로 AI/ 클라우드 활용 제약	클라우드 기반 통합 인프라 구축 데이터 리터러시교육 및 전문인력의 양성 및 지방 파견 보안-활용 균형 설계	데이터 리터러시교육 → AI 리터러시 교육으로 확장

F. 위와 같이 <표 3>의 데이터 거버넌스-AI 거버넌스 연계 프레임워크에 대해 어떻게 생각하십니까?

4. 거버넌스의 층위화 multi-layered 프레임

데이터 거버넌스 및 AI 거버넌스 연계 프레임워크 설명
본 연구는 전통적 거버넌스(Traditional Governance) → e-Governance → 데이터 거버넌스(Data Governance) → 인공지능 거버넌스(AI Governance)로 이어지는 층위적 진화 과정으로 이해하고자 함 (전통적 거버넌스) 사회적 난제를 해결하기 위한 기본 개념으로 출발하였으나 이는 정보통신 기술을 전제로 하지 않았다는 점에서 한계를 지님 (e-거버넌스) ICT의 발달과 함께 e-거버넌스는 전자정부 시스템을 기반으로 대규모 공공데이터를 축적하면서 정부의 반응성과 시민참여를 높이는 데 기여했지만 동시에 데이터 활용과 보호를 둘러싼 새로운 규범과 조정의 필요성을 불러일으킴 (데이터 거버넌스) 데이터 거버넌스는 단순히 데이터를 수집하고 저장하는 차원을 넘어, 품질·표준·보안·접근을 전 주기에 걸쳐 관리하는 것을 목적으로 함 데이터 주기는 생성과 수집, 제공과 가공, 분석과 활용, 보관과 폐기로 이어지며, 각 단계마다 품질 관리와 접근 통제, 개방 정책이 정책 수단으로 작동함 동시에 데이터의 관리·소유·이용 주체 간 역할 분담과 책임 소재가 핵심 요인으로 부각되지만 데이터는 그 자체로 완결된 정보와 동일하지 않으며, 추론·예측·판단·의사결정 과정을 거쳐야만 의미 있는 정보로 전환됨 (AI 거버넌스) AI는 데이터의 정보화 과정에서 새로운 대상과 수단을 제공하며, 데이터 거버넌스의 방향성과 외연을 재정의해야 할 필요성을 제기함 AI 거버넌스는 본질적으로 데이터 거버넌스의 확장층으로 이해될 수 있음 신뢰할 수 있는 대규모 데이터의 확보가 없이는 AI가 작동할 수 없으며 AI 거버넌스는 데이터 분석과 활용의 단계에서 출발했으나 점차 생성과 수집, 제공과 가공, 나아가 보관과 폐기 단계까지 융합하는 방향으로 확장되고 있음 전 주기적 데이터 거버넌스 속에서 제도화된 AI 활용이 바로 AI 거버넌스라 할 수 있기에 AI 거버넌스는 빅데이터 거버넌스의 확장이자 진화의 방향으로 규정할 수 있음 AI 거버넌스와 데이터 거버넌스 차별성 데이터 거버넌스가 데이터 관리와 품질 보장을 핵심으로 한다면, AI 거버넌스는 모델의 운영과 관리에 초점을 맞춤 구체적으로, 모델카드(Model Card), 알고리즘 평가, 데이터 편향 지표와 같은 도구는 투명성과 공정성, 책임성을 강화하는 수단으로 작동하며 이는 데이터 거버넌스의 본래 의도를 AI 환경에서 구현하는 방식임 결국 두 거버넌스는 상호 보완적이면서도 특화 영역이 다르기 때문에 독립적 관리가 필요함

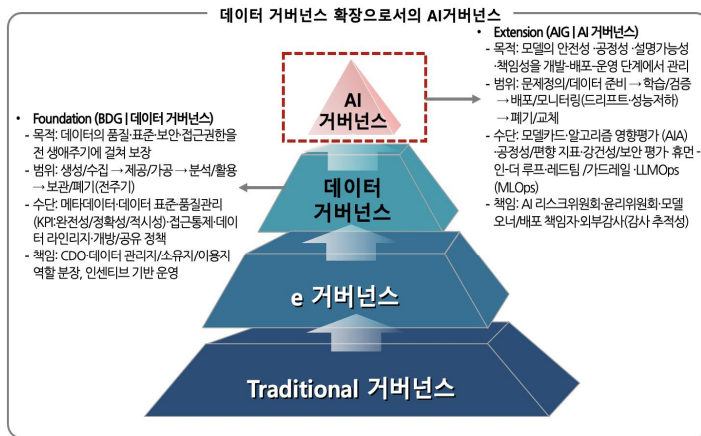
데이터 거버넌스는 정책적 토대를 제공하고, AI 거버넌스는 이를 기반으로 확장되는 구조적 관계를 가짐

이상의 논의를 종합하면, 거버넌스는 전통적 거버넌스에서 출발하여 e-거버넌스, 데이터 거버넌스, 그리고 인공지능 거버넌스로 이어지는 층위적 진화를 보여줌.

각 단계는 이전 단계의 성과와 한계를 토대로 확장되며, 특히 AI 거버넌스는 데이터 거버넌스를 기반으로 한 새로운 진화의 방향을 제시함

따라서 두 체계는 독립적이면서도 상호 보완적인 관계를 형성하며, 현대 공공 거버넌스 체계의 핵심 축으로 자리매김함

[그림 4] 데이터 거버넌스의 확장층으로서의 AI 거버넌스



AI 거버넌스는 데이터 거버넌스 위에 얹혀지는 '확장층'

G. 데이터 거버넌스 층위와 다층 프레임워크 제안에 대해 어떻게 생각하십니까?

5. 종합 평가 및 제언

H. 전반적인 종합 의견 및 보완 사항에 대해 말씀해 주세요.

디지털 대전환 메가트렌드 연구 시리즈

- 25-28-01 디지털 대전환 메가트렌드 연구 총괄보고서 V: 2035 AI·디지털 대전환 메가트렌드
(문아람 외, 정보통신정책연구원)
- 25-28-02 2035 AI·디지털 대전환 미래전략 (문아람 외, 정보통신정책연구원)
- 25-28-03 디지털 전환 기반 구축을 위한 차세대 통신기술 및 양자기술 활용 방안 (이인규 외, 한국통신학회)
- 25-28-04 디지털 자산 시장 발전에 따른 혁신과 공정의 기업 생태계 조성 (양성병 외, 한국경영학회)
- 25-28-05 디지털 전환기 기업 동태 분석 기반 혁신성장 정책 방향 (이경원 외, 정보통신정책학회)
- 25-28-06 AI·디지털 전환기 노동 구조의 재편과 노동권 보호를 위한 안전망 설계 (임운택 외, 한국사회학회)
- 25-28-07 디지털 전환기 민주주의 회복력 강화를 위한 정치제도의 변화 (김범수 외, 한국정치학회)
- 25-28-08 디지털 전환에 대응하는 AI·데이터 기반 디지털 복지서비스 개발 연구 (윤건 외, 한국행정학회)
- 25-28-09 AI·디지털 전환에 따른 공공부문 데이터 거버넌스 전략 (남태우 외, 한국정책학회)
- 25-28-10 디지털 저탄소 전환을 위한 기술적·제도적 방안 (김현 외, 대한전자공학회)
- 25-28-11 디지털 전환기 AGI 기술 경쟁력 확보를 위한 중장기 전략 (송길태 외, 한국정보과학회)

남 태 우

- 고려대학교 행정학 학사
- 인디애나대학교 행정학 석사
- 뉴욕 주립대학교 행정학 박사
- 현 성균관대학교 행정학과·
국정전문대학원 교수

김 준 형

- 위스콘신대학교 경제학 학사
- 서울대학교 행정대학원 행정학 석사
- 서울대학교 행정대학원 행정학 박사
- 현 국립순천대학교 행정학전공 조교수

신 승 윤

- 중앙대학교 공공인재학부 행정학 학사
- 서울대학교 행정대학원 정책학 석사
- 서울대학교 행정대학원 정책학 박사
- 소프트웨어정책연구소 선임연구원

하 선 우

- 이화여자대학교 국어국문학 학사
- 서울대학교 행정대학원 행정학 석사
- 성균관대학교 국정전문대학원 행정학
박사수료
- 현 성균관대학교 국제정보정책·
전자정부연구소 연구원

김 정 원

- 선문대학교 글로벌 경영학,
스페인어중남미학 학사
- 성균관대학교 경영전문대학원 EMBA
석사
- 성균관대학교 국정전문대학원 행정학
박사수료
- 현 성균관대학교 국정전문대학원
플랫폼 거버넌스 교육연구단 연구원

기 남 훈

- 고려대학교 행정학 학사
- 고려대학교 행정학 석사
- 플로리다 주립대학교 행정학 박사
- 현 성균관대학교 행정학과·
국정전문대학원 부교수

황 한 찬

- 서울대학교 사회학 학사
- 서울대학교 행정대학원 행정학 석사
- 서울대학교 행정대학원 행정학 박사
- 현 국립순천대학교 행정학전공 조교수

김 기 동

- 영남대학교 행정학 학사
- 성균관대학교 국정전문대학원 행정학
석사
- 성균관대학교 국정전문대학원 행정학
박사
- 현 중부대학교 학생성장교양학부 조교수

조 수 향

- 중앙대학교 한국음악학 학사
- 성균관대학교 국정전문대학원 행정학
석사
- 성균관대학교 국정전문대학원 행정학
박사과정
- 현 성균관대학교 국제정보정책·
전자정부연구소 연구원

ICT 진흥 및 혁신기반 조성사업 RS-2025-02173110
정책연구 25-28-09

AI·디지털 전환에 따른 공공부문 데이터
거버넌스 전략
(A Study on Strategies for Public Sector Data
Governance in the Era of AI and Digital
Transformation)

2025년 12월 일 인쇄

2025년 12월 일 발행

발행인 과학기술정보통신부 장관

발행처 과학기술정보통신부

세종특별자치시 갈매로 477 정부세종청사

Homepage: www.msit.go.kr

인쇄 인성문화
