

AI친화 · 고가치 데이터

발굴·활용·개방·공유 가이드

(‘26.06.23 / 항만정보화실)

1. AI친화 데이터 정의

□ AI친화 데이터의 개념

- AI 친화 데이터란 인공지능(AI) 학습 및 활용에 최적화된 데이터를 의미하며, 이는 단순히 데이터가 존재하는 수준을 넘어 구조적 일관성, 품질 관리, 표준화, 메타데이터 충실도, 개인정보 비식별화가 확보된 상태를 포함
즉, AI 모델이 효과적으로 학습할 수 있도록 오류가 최소화되고, 데이터의 의미가 명확하며, 안전하게 공유·활용할 수 있는 데이터를 의미함

구조화 (Structured)	• 데이터가 정형화된 스키마(테이블, 필드명, 자료형)를 갖추고 있어, 기계가 쉽게 인식·처리 가능함. • JSON, CSV, DB 테이블 등 표준 포맷을 활용하여 AI 알고리즘 적용 용이.
품질 관리 (Quality Management)	• 결측치(Null), 이상치(Outlier), 중복 데이터가 최소화됨. • 데이터 정합성(코드 일치율, 참조 무결성) 확보. • 최신성 보장(정기적 갱신 및 버전 관리).
메타데이터 충실도 (Metadata Completeness)	• 데이터명, 설명, 출처, 단위, 코드표, 수집주기, 라이선스 등 부가정보가 완비됨. • 데이터 계보(Lineage)가 명확하여 출처와 변경 이력을 추적 가능
비식별화 및 보안 (Safe & Secure)	• 개인정보 및 민감정보는 가명처리, 범주화, 노이즈 주입 등 비식별 처리 완료. • 재식별 위험도 평가 및 합성데이터 적용으로 안전성 강화.
활용성 (Usability & Accessibility)	• 라벨링(Labeling)되어 있어 지도학습 등에 즉시 활용 가능. • API, 데이터 포털, 카탈로그를 통해 접근성 보장. • 머신러닝·딥러닝 모델 적용 시 추가 정제 과정이 최소화됨.

- AI 친화 데이터 후보 선정은 단순히 데이터의 존재 여부가 아니라, AI 분석·예측

모델로 연결 가능한 속성을 갖추었는가에 초점을 맞추고 있음

- 항만 운영·물류 데이터는 스마트항만 서비스 고도화의 핵심 자원이 되고, 안전·환경 데이터는 국민 안전과 친환경 정책 실현에 직접 활용 가능함
- 경영지원 데이터는 내부 행정 효율성·투명성 확보에 기여하며, 기관의 공공데이터 평가 우수기관 달성과도 직결됨
- 앞서 정의한 AI 친화 데이터의 주요 특징(구조화, 품질 관리, 메타데이터 충실도, 비식별화, 활용성)을 실제 인천항만공사 보유 데이터에 적용한 사례는 다음과 같음

구분	AI 친화 데이터 후보	의미	주요 활용 방법
항만 운영	선석 배정 이력 데이터	선박의 입항 시간, 선석 배정 기록, 처리 시간 등을 포함	- AI 기반 선석 배정 최적화- 선박 대기시간 단축 예측- 항만 혼잡도 분석
물류	컨테이너 물동량 통계	부두별·시간대별 컨테이너 처리량, 체류시간 등	- 물동량 예측 모델 학습- 물류 병목 구간 탐지- 운영 효율성 평가
선박 관리	입출항 선박 제원 및 항로 데이터	선박 길이, 너비, 톤수, IMO 번호, 입출항 항로 정보	- 대형 선박 수용성 분석- 항로별 안전성 평가- 최적 항로 추천 서비스
안전	사고 및 근접사고 이력	사고 발생 일시, 장소, 원인, 피해 규모	- 안전사고 발생 확률 예측- 고위험 구간 AI 기반 모니터링- 예방 조치 가이드 마련
환경	항만 환경 센서 데이터(미세먼지, 수질, 소음 등)	항만 주변의 실시간 환경 데이터	- 환경오염 예측 및 대응 정책 수립- 탄소배출 관리 및 보고- ESG(환경·사회·지배구조) 경영 지원
경영 지원	예산 집행 및 계약 데이터	부서별 예산 배정·집행 현황, 계약 건수 및 금액	- 지출 패턴 분석- 비용 예측 및 절감 모델- 투명한 경영 관리 지표 제공

2. 고가치 데이터 정의

□ 고가치 데이터의 개념

- 고가치 데이터란 국민·기업·연구기관의 활용 수요가 높고, 사회적·경제적 파급력이 큰 데이터를 의미함. 이는 단순한 정보 제공을 넘어, 새로운 서비스 창출·산업 혁신·정책 개선 등 광범위한 분야에서 활용될 수 있는 잠재력이 큰 데이터를 의미함. 특히 공공기관에서 보유한 고가치 데이터는 다음과 같은 특성을 가짐

수요성	• 국민과 기업이 필요로 하는 정보로, 개방 시 높은 활용률을 보이는 데이터
파급성	• 사회·경제적 효과가 크고, 산업 혁신과 정책 효과 제고에 기여할 수 있는 데이터
정책성	• 정부의 중점 추진 과제(스마트항만, 디지털 전환, 친환경, 안전 강화 등)와 직접적으로 연계되는 데이터
지속성	• 주기적으로 생산·갱신되어 시계열 분석과 장기적 정책 수립에 기여할 수 있는 데이터

□ 고가치 데이터의 활용 효과

산업적 파급효과	• 물류기업, 해운사, 스타트업이 데이터를 활용해 최적화 모델, 예측 서비스, 신사업을 개발할 수 있다.
정책적 효과	• 환경·안전·교통·탄소중립 등 국가정책과 직접적으로 연계되어 정책 신뢰성과 실효성을 강화한다.
행정 혁신 효과	• 내부적으로는 데이터 기반 의사결정을 강화하여 예산 효율성과 운영 효율성을 높인다.
국민 체감 효과	• 항만 혼잡 완화, 환경 질 개선, 안전사고 예방 등으로 국민의 삶의 질 향상에 기여한다.

3. 시친화·고가치 데이터 발굴 절차

□ 단계별 발굴 절차

① 데이터 현황 조사

- 기관 내 보유 데이터 전수조사: 정보시스템, 엑셀, 수기 자료 포함
- 메타데이터 기본 등록 : 데이터명, 담당부서, 보안등급, 주기, 형식(상세 메타 정의는 4장에서 상세 기술)

구분	조사 항목	세부 내용	예시
기본 정보	데이터명	공식 명칭	선박 입출항 통계, CCTV 영상
	관리 부서	소유 부서명	운영본부, 안전관리팀
	담당자	책임자 및 연락처	홍길동 / 032-000-0000
데이터 유형	유형	정형 / 비정형	정형(통계), 비정형(영상)
	포맷	CSV, XLSX, DB, JSON, PDF, 이미지(JPG/PNG), 동영상(MP4), 텍스트(HWP/DOCX)	DB 테이블, CCTV 영상
기술 정보	규모	건수(정형), 용량(비정형)	10만 레코드, 5TB 영상
	컬럼/속성 수	정형: 필드 개수, 비정형: 메타 속성(촬영일시, 위치 등)	25개 필드, 영상 해상도·촬영시간
	생성 주기	일간 / 월간 / 이벤트 기반	일간 수집, 사고 발생 시 기록
품질 상태	결측치율	정형: % 단위, 비정형: 누락 파일/프레임	2% 결측, 영상 3건 손상
	이상치 발생	정형: 극단값, 비정형: 화질 저하·노이즈	선박 톤수 오류, 센서 노이즈
	표준코드 사용 여부	표준 코드·단위 준수 여부	IMO 코드, WGS84 좌표계
보안/법적	개인정보 포함 여부	이름, 얼굴 영상, 차량번호판 포함 여부	CCTV 영상: 얼굴 포함
	법적 제한	공개 제한 사유, 관련 법률	개인정보보호법, 해양안전법
메타데이터	설명	데이터 정의·활용 목적	선박 입출항 현황 분석용
	계보(Lineage)	생성·수정·활용 이력	운영시스템→DB→분석용 추출
	라이선스/권한	공개/부분공개/내부전용	부분공개(API만 제공)

[표] 메타데이터 관리 예시

② 가치 평가

- 공공·산업 수요도, 사회·경제적 파급력, 정책 연계성 평가
- 고가치 데이터 평가지표 활용 (수요성·파급성·정책성·지속성 점수화)
- 국민 체감 효과가 큰 데이터 우선 검토

평가항목	평가유형	설명	평가점수 (예시)
고가치성	민간 활용성	민간 서비스나 제품 개발에 직접 활용 가능한 주제/내용인지 여부	80
	정책 연계성	정부 정책, 과제, 지표 등과의 연계 가능성 여부	70
	공익성	사회문제 해결 또는 행정서비스 개선에 기여 가능한지 여부	80
	수요도	국민 또는 기업의 수요가 실제로 존재하거나 잠재 수요가 큰지 여부	90
	개방 가능성	법적·기술적으로 개방이 가능한지 여부 (비공개 사유 미해당 등)	80

[표] 평가 유형 정의

③ AI친화성 평가

- 데이터 구조화 정도, 품질 상태(결측·이상치·중복), 표준화 수준 확인
- 개인정보 포함 여부 및 비식별 가능성 평가

평가항목	평가유형	설명	평가점수 (예시)
AI 친화성	시계열성	데이터가 일정 주기로 축적되어 시계열 분석이 가능한지 여부	60
	완결성	항목별 결측이나 누락 없이 데이터가 충분히 수집·정리되어 있는지 여부	70
	구조화 수준	데이터의 표준화 및 정형화 정도 (예: 정형 테이블 구조 여부)	50
	결측치 유무	결측치가 과도하지 않고, 보간 또는 대체 가능한 수준인지 여부	50
	AI 학습 활용 가능성	실제로 AI 모델에 학습용으로 사용 가능한 구조 또는 컬럼을 포함하고 있는지 인지 여부	50

[표] 평가 유형 정의

④ 우선순위 결정

- “가치 평가”와 “AI 친화성 평가”를 교차 분석
- 고가치성과 AI 친화성이 모두 높은 데이터 → 1순위 후보군 선정
- 평가 결과는 데이터 전략팀 주관 하에 위원회 심의

⑤ 개선 및 가공 (Improvement & Processing)

- 품질 개선 : 결측치 보완, 코드 정합성 확보, 이상치 제거
- 표준화 : 단위·코드·시간 포맷 통일
- 비식별화 : 개인정보·민감정보 제거 및 합성데이터 적용

□ 시·친화·고가치 데이터 개방 현황

우선 순위	분야	목록명	형태	개방시점	개방URL
①	교통물류	인천항 국제여객터미널 실시간 운항정보 데이터	■ 정형	2016-04-19	https://opendata.icpa.or.kr/OpenAPI/service/ipaFerryNavigatInfo/getIntrNvgList
②	교통물류	컨테이너터미널 반출입예정 정보	■ 정형	2015-07-18	https://opendata.icpa.or.kr/OpenAPI/service/ipaTrmnlIntegratInfo/getSinGateInfo
③	환경기상	인천항 풍향 풍속 정보	■ 정형	2015-08-02	https://opendata.icpa.or.kr/OpenAPI/service/ipaFldgt/getWdWs
④	환경기상	인천항 실시간 내외항수위 정보	■ 정형	2015-08-02	https://opendata.icpa.or.kr/OpenAPI/service/ipaFldgt/getRltmInoutprtWal
⑤	환경기상	인천항 월별 예상 수위 정보	■ 정형	2015-08-02	https://opendata.icpa.or.kr/OpenAPI/service/ipaFldgt/getMnbyExpectWal
⑥	환경기상	인천항 갑문기상 정보	■ 정형	2020-01-06	https://opendata.icpa.or.kr/OpenAPI/service/ipaLmInfo/getPortWeatherInfo
⑦	교통물류	인천항 갑문 조석표 정보	■ 정형	2021-09-30	https://opendata.icpa.or.kr/OpenAPI/service/ipaLmInfo/getPortWeatherInfo
⑧	환경기상	인천항 자연담수 정보	■ 정형	2021-09-30	https://opendata.icpa.or.kr/OpenAPI/service/ipaLmWaterInfo/getLmLgNaturalFswt
⑨	교통물류	인천항 터미널 혼잡도 정보	■ 정형	2015-08-02	https://opendata.icpa.or.kr/OpenAPI/service/ipaTrmnlCnf/getTrmnlCnf
⑩	교통물류	컨테이너터미널 본선작업정보 조회	■ 정형	2016-04-19	https://opendata.icpa.or.kr/OpenAPI/service/ipaTrmnlIntegratInfo/getSinShipWorkInfo
⑪	교통물류	컨테이너터미널 선석배정정보 조회	■ 정형	2016-04-19	https://opendata.icpa.or.kr/OpenAPI/service/ipaTrmnlIntegratInfo/getSinVescallinfo
⑫	-	연안여객터미널 주차혼잡정보	■ 정형	예정	예정

[표] 데이터 개방 현황

4. 시친화 · 고가치 데이터 정의 방법

□ 메타데이터 관리개요

- 인공지능 친화적 공공데이터란 인공지능 모델의 학습·추론·평가에 쉽게 활용될 수 있도록 위험 최소화 및 기본권 보호를 전제로 기계 이해성 및 품질이 다음과 같이 확보된 수준의 공공데이터를 의미
- 인공지능 친화적 요건은 공공데이터가 AI 활용을 전제로 구조·메타데이터·문서화 품질 정보를 갖추도록 준비된 관리 수준을 의미
- ‘인공지능 친화적’ 상태는 접근성과 상호운용성(표준 포맷, API, 메타데이터), 문서화와 거버넌스(출 처·처리 과정·이용 권한 범위), 데이터 품질, 특정 AI 활용 목적에 대한 적합성 등을 갖춘 수준을 의미
- 인공지능 친화적 공공데이터는 다음과 같은 핵심 특성을 포함

특성	내용
기계가공 및 상호운용 가능 기술 구조	• CSV, Parquet, JSON, GeoJSON 등 기계 이해가 가능한 개방형 포맷으로 제공하고, 명시적인 스키마(필드 정의, 단위, 코드값)를 제공 • API를 통해 대용량 데이터에 안정적으로 접근할 수 있는 체계를 확보하고, 버전·배포 정보 및 식별자(영구식별자, URI등) 부여를 통해 데이터의 연속성과 상호운용성 보장
메타데이터와 문서화	• 데이터셋 및 속성 단위의 상세 설명, 생성·수집·정제·가공 과정, 어노테이션 기준, 품질 진단 방법과 지표를 메타데이터 및 부속 문서로 제공 • AI시스템 개발 시 해당 데이터의 적합성을 사전에 진단 할 수 있도록 활용 목적, 기술적 제한사항, 권장·비권장 활용 사례 명시
내용 품질과 대표성	• 정책·서비스에서 대상이 되는 인구·지역·현상을 균형 있게 반영하여 데이터의 대표성을 확보하고, 이와 동시에 원천 데이터와의 정보 일치성 및 데이터의 최신성, 무결성을 확보 • 데이터의 한계·불균형·잠재적 편향을 메타데이터와 문서에 명시하여 AI모델의 편향·차별·유해 결과를 최소화
권리·윤리·보안·책임 거버넌스	• 라이선스, 저작권, 개인정보·민감정보 처리, 재사용 조건 등을 명확히 표시하고, • 관련 법령 및 AI윤리 원칙에 부합하도록 설계 • 데이터 변경이력과 책임 주체의 추적 가능성을 확보하고, 데이터 오남용·위험 발생 시 대응 절차와 연락 창구를 지정

[표 1] 인공지능 친화적 공공데이터 특성

□ 메타데이터 정의

- 2026년 공공데이터 제공 및 데이터기반행정 평가는 AI 정책, 민간 활용 성과, 고가치 공유데이터 등 실질 성과 중심으로 개편되었다.
- 메타데이터 제공 : 데이터 출처, 변수 정의, 변환 이력 등을 포함한 메타데이터를 제공하여 기계의 가독성 과 이해도를 높임으로써 정확한 데이터 처리 및 모델 학습을 지원하고, 데이터 오해석에 따른 위험을 최소화
- 데이터 제공자는 AI 활용 역량 강화를 위해 포괄적이고 구조화된 메타데이터를 데이터와 연계하여 제공
- 국제 표준 기반의 메타데이터 체계를 적용하여 국가기관·플랫폼 간 데이터 구조·의미의 동일성을 보장하고, 시스템 간 자동화된 데이터 교환·통합·검색을 지원 함으로써 글로벌 상호운용성을 확보
- 글로벌 상호운용성 및 AI 학습 도구와의 호환성 확보를 위해 DCAT, Dublin Core, Croissant 등 국제 표준 메타데이터 규격을 준용하여 적용 필요

특성	내용
DCAT (Data Catalog Vocabulary)	플랫폼 간 데이터 카탈로그 공유 및 데이터 검색 용이성 확보를 위해 데이터셋 기술 방식을 규정하는 W3C 제정 표준 어휘
Dublin Core	자원 식별에 필요한 15개 핵심 요소를 정의하여 범용적 데이터 기술 및 관리를 지원하는 국제 표준(ISO) 규격
Croissant	AI 모델과 학습 도구가 데이터를 즉시 처리할 수 있도록 데이터 구조와 활용법을 정의한 AI 데이터 전용 기술 규격

[표] 국제 표준 메타데이터 규격

□ 데이터관리 메타데이터 정의

- 데이터의 식별, 검색, 접근성 확보 및 시스템 간 유기적 연계·활용을 위해 관리해야 할 핵심 속성들로 구성 되었으며, 우선순위에 따라 필수, 권장, 선택으로 구분하여 데이터 관리 메타데이터 작성 필요

항목명		속성		설명	예시
기 본	데이터명	dct:title	필수	데이터셋의 공식 명칭	서울시 대기질 측정 데이터_2025
	데이터 설명	dct:description	필수	데이터셋의 내용, 보유 목적, 수집 대상 등 설명	서울시 내 구 대상 대기질 수집, 대기 환경 분석을 위함
	관련법령	dct:references	필수	법령 참조, 법적 근거	○○○법 제 ○○조 (URL로 연결)
	소관기관 (운영부서명)	dct:creator	필수	데이터셋을 생산·관리하는 주체	데이터 전략과
	제공기관	dct:publisher	필수	데이터셋을 대외적으로 배포·게재하는 기관	서울특별시
	제공시스템	dcat:landingPage	필수	배포 또는 추가 정보에 대한 액세스를 제공하는 웹 페이지	https://data.seoul.go.kr/data~
	갱신주기	dct:accrualPeriodicity	필수	데이터셋 업데이트 빈도	http://purl.org/cld/freq/annual
	업무/주제분류체계	dcat:theme	필수	데이터셋의 카테고리를 참조. 데이터셋은 여러 테마와 연결 가능	환경
	키워드	dcat:keyword	필수	검색 및 분류를 위한 핵심 키워드	환경, 대기질, 미세먼지
추 가	고유식별자	dct:identifier	필수	중복되지 않는 영구 식별자	10.1234/seoul.air.pm.2024.v1
	담당자연락처	dcat:contactPoint	필수	데이터셋 담당자에 대한 연락처 정보 이름,전화번호,이메일을하위로제공	홍길동
	언어	dct:language	필수	데이터셋의 언어를 참조하며 여러 언어가 있는 경우 이 속성을 반복 가능	http://id.loc.gov/vocabulary/iso639-1/ko
	배포본	dcat:distribution	필수	데이터셋에 배포본이 존재함을 의미	Distribution 1
	접속 URL	dcat:accessURL	필수	데이터셋의 배포본 URL	https://data.seoul.go.kr/data~
	데이터용량	dcat:byteSize	필수	데이터의 크기(바이트)	2000
	데이터셋 유형	dcat:mediaType	필수	수치, 이미지, 음성, 영상 등 구분	image/jpeg
	시간범위	dct:temporal	권장	데이터셋에서 다루는 시간적 기간	2025.01.01.~ 2025.12.31
	공간범위	dct:spatial	권장	데이터셋에서 다루는 지리적 영역	대한민국 서울특별시 (25개 자치구)
	연계데이터셋	dct:relation	선택	연계 데이터셋이나 참조 리소스	서울시 대기질 측정 데이터_2024

[표] 데이터 관리 메타데이터

○ 데이터 관리 메타데이터 작성 예시

항목	작성 예시
데이터명	"dct:title": "2025 서울시 연령별 인구통계 정보"
제공범위	"dct:description": "서울시 거주 인구를 연령대별·구별로 분류하여 인구구조 분석, 정책 수립 및 행정 서비스 개선을 위해 제공되는 통계 데이터."
관련법령	"dct:references": "통계법 제3조 및 국가통계작성지침 준수"
소관기관	"dct:creator": "서울특별시 빅데이터담당관"
제공기관	"dct:publisher": "서울특별시"
표준 데이터셋 제공시스템	"dcat:landingPage": "https://data.seoul.go.kr/dataset/2025-pop-age"
접속 URL	"dcat:accessURL": "https://data.seoul.go.kr/datafile/POP_2025_AGE.csv"
갱신주기	"dct:accrualPeriodicity": "http://purl.org/cld/freq/annual"
고유식별자	"dct:identifier": "seoul-pop-age-2025"
담당자 연락처	"dcat:contactPoint": "fn": "서울특별시 빅데이터담당관 인구통계 담당자", "hasEmail": "mailto:data@seoul.go.kr"
업무/주제분류 체계	"dcat:theme": "인구·사회"
키워드	"dcat:keyword": "인구", "연령별", "서울시", "2025", "인구통계", "demographics"
언어	"dct:language": "http://id.loc.gov/vocabulary/iso639-1/ko"
데이터 용량	"dcat:byteSize": "2400000"
데이터셋 유형	"dcat:mediaType": "text/csv"
관련 데이터셋	"dct:relation": "2024 서울시 연령별 인구통계 정보", "서울시 구별 인구 동향 데이터"
시간범위	"dct:temporal" "time:hasBeginning": "2025-01-01", "time:hasEnd": "2025-12-31"
공간범위	"dct:spatial": "서울특별시"

[표] 데이터 관리 메타데이터 작성 예시 - 2025 서울시 연령별 인구통계 정보

□ 데이터 계보 메타데이터

- 데이터셋 배포 이후의 변경 이력을 투명하게 관리하고 최신성을 보장하기 위해 수정·업데이트 관련 속성 을 정의하며, 데이터 신뢰도 및 재현성 제고를 위해 필수항목으로 구성

항목명		속성		설명	예시
추가 확장	버전	owl:versionInfo	필수	현재 버전 정보	2025.01.v1
	등록일시	dct:issued	필수	시스템 최초 등록 날짜 (초기 등록 후 변경금지)	YYYY-MM-DD
	수정일시	dct:modified	필수	마지막 수정 날짜	YYYY-MM-DD
	버전 노트	adms:versionNotes	필수	이전 버전 대비 변경 사항	데이터셋 최초 공개 결측치 처리 방식 변경 측정지역 코드화 및 적용

[표 5] 데이터 계보 메타데이터

○ 데이터 계보 메타데이터 작성 예시

항목	작성 예시
버전	"owl:versionInfo": "v1.2"
등록일시	"dct:issued": "2025-01-15T09:00:00+09:00"
수정일시	"dct:modified": "2025-03-02T14:30:00+09:00"
버전 노트	"adms:versionNotes": "센서 오류값 제거 및 생육지표(엽장·착과수) 추가. 영상데이터 결측치 보정 알고리즘 개선."

[표 6] 데이터 계보 메타데이터 작성 예시 - 딸기 생육과정

□ 데이터 이용 메타데이터

- 데이터 이용 시 사용 범위와 제약 사항을 정의하여 법적 불확실성을 해소하고, 데이터의 투명한 이용 및 권리 보호를 위한 권한 속성으로 구성

항목명		속성		설명	예시
추가 확장	라이선스	dct:license	필수	데이터셋의 권리 관련 공식적인 법적 문서를 기술하는 속성 -특정라이선스문서(예:CCBY,GPL등)의URI를직접참조하는용도	https://creativecommons.org/licenses/by-nc-sa/4.0/
	저작권	dct:rights	필수	데이터셋 관련 저작권 정보 전반에 관한 정보를 기술하는 속성 -저작권표시,비영리여부,변경/금지여부,변경허락여부	상업적·비상업적 이용 모두 허용

[표 7] 데이터 이용 메타데이터

○ 데이터 이용 메타데이터 작성 예시

항목	작성 예시
라이선스	“dct:license”: “https://.terms/daegu-dataset-usage-policy” (이용조건 URL)
권한	“dct:rights”: “본 데이터셋의 저작권은 대구광역시에 있으며, 상업적 이용 및 제3자 제공은 내부 승인 절차에 따른다.”

[표 8] 데이터 이용 메타데이터 작성 예시 - 대구시 업종별 사업자

□ 데이터 품질 메타데이터

- 데이터 품질 수준 및 한계점을 사전에 인지하여 AI 모델의 신뢰성을 확보할 수 있는 속성으로 구성

항목명		속성		설명	예시
추가 확장	데이터 편향성	rai:dataBiases	권장	확인된 편향 유형	“도시 지역(서울·수도권)에 측정소가 집중되어 있어 농촌 지역의 대표성이 낮음”
	품질 검증 정보	dqv:hasQualityAnnotation	선택	수행된 품질 검토 및 평가 내용	“월 1회 수동 검수 완료 (환경전문가 검토)”
	데이터 한계	rai:knownLimitations	선택	데이터셋의 구조적·내용적 제한사항 및 해석 시 유의 사항	“일부 측정소는 장비 유지보수 기간동안 측정값 누락 (측정값 누락 지역 ○○○, ○○○, ○○○)”
	결측치 정보	rai:dataCollectionMissingData	선택	결측치 비율 및 처리방식	“결측 사유: 센서 점검(2.1%), 통신 오류(1.1%)”

[표 9] 데이터 품질 메타데이터

○ 데이터 품질 메타데이터 작성 예시

항목	작성 예시
알려진 편향 명세	“rai:dataBiases”: “일부 출신국가는 귀화 허가 대상자 규모가 매우 작아 통계적 변동성 이 큼.”, “여성 비율이 높은 특정 국가 편중 현상이 존재하여 전체 귀화자 성비를 과대/ 과소 추정할 가능성 있음.”, “경제활동 연령대(20 - 40대)가 주로 귀화하는 경향으로 인해 고령층 데이터가 부족함.”
데이터셋 품질 검증 여부	“dqv:hasQualityAnnotation”: “@type”: “dqv:QualityAnnotation”, “dqv:hasQualityMeasurement”: “법무부 연례 검증 절차를 통해 원시 귀화 허가 데이터의 정확성·일관성 점검 완료(2025년 검증 기준).”
데이터 한계	“rai:knownLimitations”: “비공식 체류자의 국적 변경 이력은 포함되지 않아 귀화자 전체 규모를 완전히 대변하지 못함.”, “귀화 심사 탈락자·보류자의 통계는 포함되어 있지 않아 국가별 전환율 분석이 제한됨.”, “표본 크기가 작은 국가의 경우 단일 관측치 변화 가 통계에 큰 영향을 미침.”
결측치 정보	“rai:dataCollectionMissingData” : “센서 점검(2.1%), 통신 오류(1.1%)로 결측이 발생할 수 있음.”

[표 8] 데이터 품질 메타데이터 작성 예시 - 국내귀화자 출신 국가

○ DCAT 기반 메타데이터 작성 예시

구분	문서명	핵심내용
데이터 관리 메타데 이터	데이터명	“dct:title”: “2025 서울시 연령별 인구통계 멀티모달 데이터셋”
	제공범위	“dct:description”: “2025년 기준 서울특별시 거주 인구를 연령대별·구별로 집계한 수치 데이터와 함께, 인구 분포 시각화 이미지, 정책 설명 음성 자료, 인구 구조 분석 영상 콘텐츠를 통합 제공하는 멀티모달 통계 데이터셋”
	관련법령	“dct:references”: “통계법 제3조, 국가통계작성지침, 공공데이터 제공 및 이용 활성화에 관한 법률”
	소관기관	“dct:creator”: “서울특별시 빅데이터담당관”
	제공기관	“dct:publisher”: “서울특별시”
	표준데이터 셋 제공시스템	“dcat:landingPage”: “https://data.seoul.go.kr/dataset/2025-population- multimodal”
	접속 URL	“dcat:accessURL”: “https://data.seoul.go.kr/datafile/POP_2025_ MULTIMODAL”
	갱신주기	“dct:accrualPeriodicity”: “http://purl.org/cld/freq/annual”
	고유식별자	“dct:identifier”: “seoul-population-multimodal-2025”
	담당자명	“dcat:contactPoint”: “fn”: “서울특별시 빅데이터담당관 인구통계 담당자”, “hasEmail”: “mailto:data@seoul.go.kr”
	업무/주제분 류체계	“dcat:theme”: “인구·사회”
	키워드	“dcat:keyword”: “인구”, “연령별”, “서울시”, “2025”, “인구통계”, “demographics”, “multimodal”
	언어	“dct:language”: “http://id.loc.gov/vocabulary/iso639-1/ko”
	데이터 용량	“dcat:byteSize”: “1850000000”
	데이터셋 유형	“dcat:mediaType”: “text/csv”
	관련 데이터셋	“dct:relation”: “2024 서울시 연령별 인구통계 정보”, “서울시 구별 인구 동향 데이터”
	시간범위	“dct:temporal”: “time:hasBeginning”: “2025-01-01”, “time:hasEnd”: “2025-12-31”
	공간범위	“dct:spatial”: “서울특별시”
데이터 계보 메타데 이터	버전	“owl:versionInfo”: “v1.2”
	배 포일시	“dct:issued”: “2025-01-15T09:00:00+09:00”
	최 종 업데이트/ 수정일시	“dct:modified”: “2025-12-15T14:30:00+09:00”

	버전 노트	“adms:versionNotes”: “일부 구 단위 인구 누락값 보정, 연령 구간 세분화(5세 단위), 시각화 이미지 및 정책 설명 영상 신규 추가.”
데이터 이용 메타데 이터	라이선스	“dct:license”: “https://.terms/daegu-dataset-usage-policy”
	권한	“dct:rights”: “본 데이터셋은 서울특별시가 저작권을 보유하며, 행정·정책 수립 및 비상업적 연구 목적에 한하여 이용 가능합니다.”
데이터 품질 메타데 이터	알려진 편향 명세	“rai:dataBiases”: “외국인 단기 체류자는 통계에 포함되지 않아 실제 생활 인구와 차이가 있을 수 있음.”, “고령층의 주소 미변경으로 일부 자치구 인구가 과소 추정될 가능성 있음.”
	데이터셋 품질검증 여부	“dqv:hasQualityAnnotation”: “@type”: “dqv:QualityAnnotation”, “dqv:hasQualityMeasurement”: “법무부 연례 검증 절차를 통해 원시 귀화 허가 데이터의 정확성·일관성 점검 완료(2025년 검증 기준).”
	데이터 한계	“rai:knownLimitations”: “주민등록 기준 통계로 실제 상주 인구를 완전히 반영하지 못함.”, “연령대별 이동 인구 분석에는 부적합함.”
	결측치 정보	“rai:dataCollectionMissingData”: “센서 점검(2.1%), 통신 오류(1.1%)로 결측이 발생할 수 있음.”

5. 활용 전략

- 인천항만공사가 보유한 AI 친화 고가치 데이터는 내부 행정 효율화뿐 아니라, 외부 민간·연구 분야에서도 광범위하게 활용될 수 있음
- 활용 전략은 내부 활용과 외부 활용으로 나누어 추진하며, 각 사례는 기관의 스마트항만·안전·친환경 정책과 직접 연계됨

□ 행정기관 내부업무 효율화 적용사례¹⁾

- (특허) 특허문서 등을 학습한 특허전용 초거대 AI 언어모델을 구축하고, 이를 특허 문서 검색·분류 등 심사 업무에 활용(특허청)
 - AI·OCR(광학문자인식) 기술로 특허심판 서류 내용을 추출하여, 방식심사 자동화 및 첨부서류 재분류에 활용



< AI·OCR을 활용한 특허심판 방식심사 자동화 >

□ 공공기관 활용 사례²⁾

NO	구 분		과 제 명	주관·수요 기관 등	사업 유형
	(유형)	(기능)			
< 사회문제 해결 >					
1	돌봄	장애인	장애인 소통지원 초거대AI 멀티모달 기반 서비스	장애인특수학교, 에어패스 등	초거대AI 서비스('24)
2	돌봄	아동	초거대 AI 기반 느린 학습자 조기발견 지원 서비스 개발	전남대어린이병원, 자란다 등	초거대AI 서비스('24)
3	돌봄	아동	아동 건강정책 기반의 AI 챗봇	김해시	초거대AI 서비스('24)
4	돌봄	노인	심탄도 센서 기반 시니어 라이프로그 관제 시스템, Clomental (클로멘탈)	(재)경남사회서비스원 , 돌봄드림	GovTech 창업지원('24)
5	저출생	임신	난임시술 빅데이터 기반 가임력 관리서비스 플랫폼 구축	대구대학교 산학협력단	DPG인프라 ('23,'24)
6	저출생	임신	지역 맞춤형 임신 전, 후 스마트 통합 관리 서비스	주식회사 디에이블	GovTech 창업지원('24)

1) 전국민 인공지능 일상화 실행계획 참조

2) 공공부문 초거대 AI도입·활용 가이드라인 참조

7	복지혜택	취약계층	인공지능 활용 초기상담 등으로 복지 사각지대 해소	보건복지부, 한국 사회보장정보원 등	국민드림프로젝트('22, '23)
8	복지혜택	청년	AI 기반 맞춤형 청년정책 추천·알림	과기정통부, (주)웰로 등	국민드림프로젝트('23)
9	복지혜택	장애인	장애인 보행안전 정보제공 서비스	주식회사 엘비에스테크	GovTech 창업지원('24)
10	복지혜택	장애인	배리어프리 여행 지원 서비스	코나투스 등	DPG 인프라('23)
11	복지혜택	장애인	장애인, 고령자 등을 위한 생성형 AI 기반 배리어프리 콘텐츠 제작 기술 개발·실증	와이즈넷컨소시엄	플래그쉽 프로젝트('24~'25)
12	범죄	아동·청소년	AI기반 아동·청소년 온라인 성착취 선제적 대응시스템 지원	여성가족부	DPG-col lab('25)
13	범죄	아동, 청소년, 장애인 등	성범죄 위기대응(수사·안전·지원) 플랫폼 개발	광주광역시경찰청	DPG-col lab('25)
14	범죄	빈집 정비	빈집 정비 통합 지원 시스템 구축	한국부동산원	DPG-col lab('25)
15	경제활동	노동약자	AI 노동법 상담 및 AI 근로감독관 지원시스템 구축	고용노동부, 마음AI	초거대AI 서비스('24)
16	경제활동	청년	청년 농업인 정착지원 등 특화서비스 개발	농진청, 네이버클라우드	초거대AI 서비스('24)
17	경제활동	사업자	AI·데이터 기반 맞춤형 공공입찰 추천	과기정통부, (주)웰로	국민드림프로젝트('23)
18	경제활동	소상공인	소상공인 빅데이터 플랫폼	중기부, 소상공인 시장진흥공단	국민드림프로젝트('22)
19	경제활동	사업자	인공지능 기반 산재·고용보험 사업종류 추천 솔루션 개발	고용노동부, 이지메이아이	부처협업 기반('23~'25)
20	경제활동	구직자	AI기반 구인·구직 통합지원 솔루션 개발 및 실증	고용노동부, 원티드랩	부처협업 기반('24~'26)
21	경제활동	기업	AI기반 마른김 품질 등급 판별 솔루션 개발 및 실증	해양수산부, 티맥스티베로	부처협업 기반('24~'26)
22	경제활동	기업	해외국가·인증정보 기반 무역투자 GPT 구축	대한무역투자진흥공사 등	DPG 인프라('24)
23	경제활동	구직자	월드잡플러스 시민 응대 및 업무지원을 위한 AI 서비스	한국산업인력관리공단	초거대AI 서비스('23)
24	경제활동	근로자	근로자의 산재보험 결정을 위한 민원 질의 실증	근로복지공단	초거대AI 서비스('23)
25	경제활동	사업자	기업 맞춤형 정보 제공을 위한 AI 모델 및 맞춤형 지원 플랫폼 구축	주식회사 위티	GovTech 창업지원('24)
26	경제활동	소상공인	외식통계조회 서비스	한국농수산식품공사, 나이스지니데이터	DPG인프라 ('23)
27	경제활동	소상공인	소규모 사업자를 위한 통합 Mall관리 플랫폼 구축	제주도 등	DPG인프라 ('23)
28	경제활동	기업	AI 챗봇을 활용한 도내 중소기업 수출 활성화	경기도 경제 과학진흥원	초거대AI 서비스('24)
29	경제활동	기업	관세 분쟁과 관련한 판례 검색 챗봇 서비스	관세청	초거대AI 서비스('24)

30	경제활동	근로자	취업규칙에 대한 적합/부적합 판단 서비스 실증	고용노동부	초거대AI 서비스('23)
31	경제활동	해외직구, 개인무역	민원 데이터와 관세행정, 법규 등을 활용한 민원 응대 실증	관세청	초거대AI 서비스('23)
32	경제활동	사업자	홈택스 종합소득세 분야 챗봇 실증	국세청	초거대AI 서비스('23)
33	의료	응급환자	실시간 의료자원정보 플랫폼 구축	고대의료원, 에이아이티스토리	국민드림 프로젝트('24)
34	의료	응급환자	개인 의료정보 공유체계 구축 및 지역 실증	아이티아이즈	국민드림 프로젝트('24)
35	의료	개인의료	진료기록 조회와 관리, 집에서도 가능	보건복지부	국민드림 프로젝트('22)
36	의료	임상 참여 환자	LLM 챗봇 기반 임상시험 참여 정보 질의 시스템	주식회사 메디아이프러스	GovTech 창업지원('24)
37	재난·재해	인파관리	데이터 기반 K-인파관리지원시스템 구축·운영	행정안전부	국민드림 프로젝트('23)
38	재난·재해	산림해충	AI 기반 산림병해충 방제 지원	엘컴텍	부처협업 기반('22~'24)
39	재난·재해	화학물질	AI 융합 유해 화학물질 판독시스템	바이브컴퍼니	부처협업 기반('23)
40	재난·재해	농작물	e-바이오 리액토믹스 기반의 농작물 불량환경 AI 관제 서비스	경기도농업기술원	DPG-col lab('25)
41	재난·재해	결로	AI기반 결로 예방 자동화 시스템	한국과학기술원 등	DPG 인프라('24)
42	재난·재해	군중밀집	ICT기술 기반 군중 밀집 공공데이터 활용 관리기술 개발	주식회사 애나	GovTech 창업지원('24)
43	재난·재해	지반침하	건물 신축시 AI기반 지반침하 탐지 시스템 개발	한국국토정보공사, 뉴롭 등	GovTech 창업지원('24)
44	재난·재해	항공안전	항공안전 이슈 분석 자동화 초거대 AI PoC	항공안전기술원	초거대AI 서비스('24)
45	재난·재해	산사태	강수량을 활용한 산사태 위험도 예측 프로그램	라온소프트	GovTech 창업지원('24)
46	재난·재해	화재	소방 원격관제 서비스 비디앱 기반의 화재 진입로·대피로 안내 시스템 개발	이엠시티 주식회사	GovTech 창업지원('24)
47	재난·재해	미세먼지	다중이용시설의 미세먼지 저감 AI예측 및 제어솔루션 개발	(주)애플스	GovTech 창업지원('24)
48	교통안전	도시철도	초거대 AI 기반의 도시철도 교통안전 GPT 서비스	서울교통공사	초거대AI 서비스('23)
49	교통안전	경찰	AI기반 경찰 교통안전시스템	경찰청, 주식회사 엠티데이터 등	DPG 인프라('24)
50	환경	빛공해	AI 빛공해 이미지 분석 솔루션 개발 및 실증	환경부	부처협업 기반('24~'26)
51	환경	국토변화 탐지	AI기반 국토 변화탐지 솔루션 개발 및 실증	국토교통부	부처협업 기반('24~'26)

52	환경	수질	수질 시분석 지원 시스템 'AME' 개발	(주)워터아이즈	GovTech 창업지원('24)
53	환경	강 유역	미호강 유역 환경 관리를 위한 민관 협력 통합환경정보 플랫폼 구축	충북도청, (주)데이터센트릭	GovTech 창업지원('24)
54	환경	공공안전	AI 기반 불법 현수막 탐지기술 개발 및 서비스 실증	(주)스마트뱅크	GovTech 창업지원('24)
32	경제활동	사업자	홈택스 종합소득세 분야 챗봇 실증	국세청	초거대AI 서비스('23)
33	의료	응급환자	실시간 의료자원정보 플랫폼 구축	고대의료원, 에이아이티스토리	국민드림 프로젝트('24)
34	의료	응급환자	개인 의료정보 공유체계 구축 및 지역 실증	아이티아이즈	국민드림 프로젝트('24)
35	의료	개인의료	진료기록 조회와 관리, 집에서도 가능	보건복지부	국민드림 프로젝트('22)
36	의료	임상 참여 환자	LLM 챗봇 기반 임상시험 참여 정보 질의 시스템	주식회사 메디아이플러스	GovTech 창업지원('24)
37	재난·재해	인파관리	데이터 기반 K-인파관리지원시스템 구축·운영	행정안전부	국민드림 프로젝트('23)
38	재난·재해	산림해충	AI 기반 산림병해충 방제 지원	엘컴텍	부처협업 기반('22~'24)
39	재난·재해	화학물질	AI 융합 유해 화학물질 판독시스템	바이브컴퍼니	부처협업 기반('23)
40	재난·재해	농작물	e-바이오 리액토믹스 기반의 농작물 불량환경 AI 관제 서비스	경기도농업기술원	DPG-col lab('25)
41	재난·재해	결로	AI 기반 결로 예방 자동화 시스템	한국과학기술원 등	DPG 인프라('24)
42	재난·재해	군중밀집	ICT기술 기반 군중 밀집 공공데이터 활용 관리기술 개발	주식회사 애나	GovTech 창업지원('24)
43	재난·재해	지반침하	건물 신축시 AI 기반 지반침하 탐지 시스템 개발	한국국토정보공사, 뉴롭 등	GovTech 창업지원('24)
44	재난·재해	항공안전	항공안전 이슈 분석 자동화 초거대 AI PoC	항공안전기술원	초거대AI 서비스('24)
45	재난·재해	산사태	강수량을 활용한 산사태 위험도 예측 프로그램	라온소프트	GovTech 창업지원('24)
46	재난·재해	화재	소방 원격관제 서비스 비디앱 기반의 화재 진입로·대피로 안내 시스템 개발	이엠시티 주식회사	GovTech 창업지원('24)
47	재난·재해	미세먼지	다중이용시설의 미세먼지 저감 AI 예측 및 제어솔루션 개발	(주)애플스	GovTech 창업지원('24)
48	교통안전	도시철도	초거대 AI 기반의 도시철도 교통안전 GPT 서비스	서울교통공사	초거대AI 서비스('23)
49	교통안전	경찰	AI 기반 경찰 교통안전시스템	경찰청, 주식회사 엠티데이터 등	DPG 인프라('24)
50	환경	빛공해	AI 빛공해 이미지 분석 솔루션 개발 및 실증	환경부	부처협업 기반('24~'26)
51	환경	국토변화 탐	AI 기반 국토 변화탐지 솔루션 개발 및 실증	국토교통부	부처협업 기반('24~'26)

		지			
52	환경	수질	수질 AI분석 지원 시스템 'AME' 개발	(주)워터아이즈	GovTech 창업지원('24)
53	환경	강 유역	미호강 유역 환경 관리를 위한 민관 협력 통합환경정보 플랫폼 구축	충북도청, (주)데이터센트릭	GovTech 창업지원('24)
54	환경	공공안전	AI기반 불법 현수막 탐지기술 개발 및 서비스 실증	(주)스마트뱅크	GovTech 창업지원('24)

< 대국민 서비스 혁신 >

1	정보제공	질의, 검색	대국민 공공정책 자동신청서비스	주식회사 웰로	GovTech 창업지원('24)
2	정보제공	질의, 검색	시간별 주요 뉴스를 분석·요약하는 데일리 리포트 실증	국회도서관	초거대AI 서비스('23)
3	정보제공	질의, 검색	도서관 발간 자료를 학습한 질의응답 챗봇 실증	국립중앙도서관	초거대AI 서비스('23)
4	정보제공	질의, 검색	수요자 맞춤형 과학기술 분야 특화 챗봇 서비스 실증	한국과학기술기획 평가원	초거대AI 서비스('23)
5	정보제공	질의, 검색	광주시 버스노선도 기반의 교통안내 챗봇 서비스 PoC	광주광역시	초거대AI 서비스('24)
6	정보제공	질의, 검색	시민을 위한 지능형 맞춤형 정보 제공 서비스	화성시, 유알피	DPG 인프라('23)
7	정보제공	질의, 검색	범죄와 형사사법 통계분석 및 RAG기반의 민원 응대 서비스 제공	한국형사법무정책 연구원	초거대AI 서비스('24)
8	정보제공	질의, 검색	부동산 건축 분야 스마트검색 및 챗봇 실증	한국부동산원	초거대AI 서비스('23)
9	정보제공	질의, 검색	방송광고 관련 문서 학습 및 질의응답 지원 챗봇 서비스	한국방송광고 진흥공사	초거대AI 서비스('24)
10	정보제공	질의, 검색	국가 디지털 사업 관련 챗봇 고도화 실증	한국지능정보 사회진흥원	초거대AI 서비스('23)
11	정보제공	질의, 검색	홈페이지 내 콘텐츠 안내 등 대국민 정보제공을 위한 생성형AI 기반 챗봇 구축	한국양성평등 교육진흥원	초거대AI 서비스('24)
12	정보제공	질의, 검색	학생 개별 맞춤형 정보지원, 홈페이지 챗봇 서비스	경기도교육청	초거대AI 서비스('24)
13	상담지원	민원	초거대 AI 기반의 공공기관 민원상담 어시스턴트 서비스	화성시청	초거대AI 서비스('23)
14	상담지원	민원	국민신문고 민원데이터 AI 기반 분석 및 분류 고도화 실증	국민권익위원회	초거대AI 서비스('23,'24)
15	상담지원	민원	공항 보유 시설정보 관련 문의 대응이 가능한 챗봇 서비스	한국공항공사	초거대AI 서비스('23)
16	상담지원	민원	양산시민 대상 민원 상담정보 제공 서비스 실증	양산시	초거대AI 서비스('23)
17	상담지원	민원	통신서비스 분야 전문 상담 챗봇 서비스	한국통신사업자연합 회	초거대AI 서비스('23)
18	상담지원	민원	초거대 AI를 적용한 병무청 대민 서비스 개선	병무청	초거대AI

					서비스('24)
19	상담지원	민원	전력 관련 약관, 규정, 민원 이력 기반 챗봇 서비스 고도화	한국전력공사	초거대AI 서비스('23)
20	상담지원	민원	국민연금 일시금 분야 민원 대응 AI 휴먼 실증	국민연금공단	초거대AI 서비스('23)
21	상담지원	민원	초거대 AI를 활용한 외국인 민원상담 지원 서비스	안산시	초거대AI 서비스('24)
22	상담지원	민원	수원시 복지 민원 서비스 개선	수원시청, (주)라이프로그	DPG 인프라('24)
23	이용편의 등	학습·교육	국가대표 AI코칭 솔루션 실증	문화체육관광부, 심플랫폼	부처협업 기반('23~'25)
24	이용편의 등	학습·교육	과학기술인 학습플랫폼(알파캠퍼스) AI 튜터	국가과학기술인력개발원	초거대AI 서비스('24)
25	이용편의 등	법률·규제	AI융합 약관심사플랫폼 개발 및 실증	공정거래위원회	부처협업 기반('24~'26)
26	이용편의 등	법률·규제	생성형 AI 에이전트 기반 법률 사건 분석 지원 서비스 개발	로앤컴퍼니	플래그십 프로젝트('24~'25)
27	이용편의 등	법률·규제	LLM 기반 규제 모니터링·해설 서비스	(주)씨지인사이드	GovTech 창업지원('24)
28	이용편의 등	법률·규제	규제정보 검색시스템 고도화	한국행정연구원 등	DPG 인프라('24)
29	이용편의 등	건강관리	맞춤형 건강기능식품 정보 제공 플랫폼	주식회사 원스글로벌	GovTech 창업지원('24)
30	이용편의 등	건강관리	AI 금연정보 검색요약 서비스 고도화	한국건강증진개발원	초거대AI 서비스('24)
31	이용편의 등	건강관리	에듀테크 체력측정 및 건강관리 서비스	인천시설공단, 더좋은운동으로	GovTech 창업지원('24)
32	이용편의 등	문화·관광	AI 페르소나 기반 문화매칭 서비스	(주)미디어그룹 사람과숲 등	DPG 인프라('24)
33	이용편의 등	문화·관광	국립현대미술관 맞춤형 전시환경 서비스 구현	(주)미디어그룹 사람과숲 등	DPG 인프라('23)
34	이용편의 등	문화·관광	감성분석 모듈을 적용한 사용자 친화적 관광지 정보 제공 서비스 실증	한국관광공사	초거대AI 서비스('23)
35	이용편의 등	기타	멀티스펙트럼 시각지능 온디바이스 AI기반 공공문서 판독기능의 출입국심사 솔루션	로지체인	GovTech 창업지원('24)
36	이용편의 등	기타	민간클라우드 기반 DPG 통합테스트베드	과기정통부, 주식회사 클로잇 등	DPG 인프라('24)
37	이용편의 등	기타	지능형 수도행정(자가검침 및 예측모델 등) 서비스	양평군청, 나이스지니데이터(주)	DPG 인프라('24)

< 일하는 방식 개선 >

1	행정	R&D	초거대 AI를 이용한 통합 연구지원 서비스 개발	경제인문사회연구회, 네이버클라우드	초거대AI 서비스('24)
2	행정	R&D	국가 R&D 사업비 Risk 관리 솔루션 개발	주식회사 루모스	GovTech 창업지원('24)
3	행정	R&D	정책 제안 지원을 위한 데이터 분석 서비스	경제인문사회연구회	초거대AI

			실증		서비스('23)
4	행정	R&D	기관 보유 연구자료를 활용한 정책 제안 업무 지원 서비스 실증	중소기업중앙회	초거대AI 서비스('23)
5	행정	특허심사	AI기반 특허심사 업무지원 서비스 개발	특허청, 솔트룩스	초거대AI 서비스('24)
6	행정	소방행정	초거대 AI기반 스마트소방 행정 서비스 구축	서울소방재난본부, 아일리스프런티어	초거대AI 서비스('24)
7	행정	건축행정	생성형 AI 기반 국방시설 건축 행정 지원 서비스 개발	국방부, SQI소프트	초거대AI 서비스('24)
8	행정	통관	AI융합 통관 영상 관리 솔루션 실증	관세청	부처협업 기반('23~'25)
9	행정	업무처리	경량형 디지털트윈 기반 공공업무 지원환경	디아지트 주식회사	GovTech 창업지원('24)
10	행정	물자관리	단종품목·조달수량 군수물자 예측서비스	육군군수사령부, (주)지뉴소프트 등	DPG 인프라('24)
11	행정	입지분석	공공복합시설 입지분석 플랫폼 구축	중앙대산학협력단, (주)선도소프트	DPG 인프라('24)
12	행정	감사	공공기관 직원의 비위행위에 대한 기사 바탕으로 유사 사례 도출 AI감사 서비스	에스알	초거대AI 서비스('24)
13	행정	제안요청	e-제안요청도움 서비스	조달청	초거대AI 서비스('24)
14	행정	업무지원	내부 직원을 위한 복지데이터 검색 및 채팅 서비스	서울시복지재단	초거대AI 서비스('24)
15	행정	업무지원	초거대 AI 도입을 통한 교육 행정 응대 PoC	경기도교육청	초거대AI 서비스('24)
16	행정	계약업무	맞춤형 학습을 통한 학교 및 교육기관의 계약 업무 지원 서비스	서울시교육청	초거대AI 서비스('24)
17	행정	검색, 문서작성	문화재 기록물 검색 및 사업계획서 작성 지원 서비스	한국문화재단	초거대AI 서비스('23)
18	행정	계약업무	계약관리 업무방법서를 활용할 수 있는 챗봇 서비스 실증	소상공인진흥공단	초거대AI 서비스('23)
19	행정	문서작성	초거대 AI를 이용한 문서 자동 생성 기능	서울시설공단	초거대AI 서비스('24)

□ 외부 활용 전략

적용 주체	주요 활용 시나리오	기대 효과
민간 기업	- 물류기업: 컨테이너 체류시간·물동량 데이터로 AI 기반 최적 물류 루트 설계 - 해운사: 선박 입출항 통계로 스케줄 최적화 - 스타트업: 항만 데이터를 활용한 신서비스(예: 항만 모니터링 플랫폼) 개발	- 물류비 절감 - 항만 서비스 혁신 - 신산업 창출

연구기관·대학	- 교통·물류 연구: 항만 운영 및 물류 효율성 분석 - 환경 연구: 항만 미세먼지·수질 데이터로 환경 영향 연구 - 안전공학 연구: 사고 데이터 기반 위험도 모델 개발	- 연구 성과 창출 - 정책 근거 자료 제공 - 글로벌 학술 협력 확대
정부·공공기관	- 국가 물류정책·스마트항만 정책 지원 - 탄소중립 이행 모니터링 및 보고 - 공공데이터 포털을 통한 국민 활용 확대	- 정부 정책 효과 제고 - 국가 데이터 전략과 정합성 확보
국민·지역사회	- 개방 데이터를 활용한 항만 모니터링 서비스 - 항만 환경·안전 정보를 시민에게 제공	- 대국민 알권리 보장 - 지역사회 안전·환경 개선

6. 데이터 개방 전략

□ 데이터 개방 전략

- 데이터 개방은 단순히 파일을 공개하는 것을 넘어, 내부 활용 → 파일 개방 → API 제공 → 실시간 스트리밍의 단계적 접근이 필요함
- 또한, 개방 데이터는 반드시 품질·설명·이력 관리가 보장되어야 국민·기업·연구자가 안심하고 활용할 수 있음

□ 단계적 개방 확대

1. 파일 개방 단계

- 목적 : 기본적인 데이터 활용을 위해 파일 단위로 개방
- 권장 포맷 : CSV(UTF-8), XLSX, JSON(line-delimited)
- 필수 제공 항목: 데이터 설명서(Definition), 품질 지표, 코드표, 변경 이력, 샘플 데이터
- 예시 : 선박 입출항 통계, 항만 환경 센서 월별 집계

2. API 제공 단계

- 목적 : 이용자가 최신 데이터를 손쉽게 활용할 수 있도록 API 제공
- 구현 원칙 : REST/JSON 기반, 표준 응답 스키마, 인증·권한 제어
- 필수 기능 : 필터링(기간·지역·항목), 페이징, 오류 코드 정의
- 예시 : 갑문 운영 스케줄 API, 컨테이너 물동량 실시간 조회 API

3. 실시간 데이터 개방 단계

- 목적 : IoT·AI 활용을 위한 실시간 데이터 제공
- 방식 : MQTT, WebSocket, Kafka 등 스트리밍 기술 적용
- 활용 사례 : 실시간 환경 센서 데이터, 항만 장비 가동 현황, 선박 위치 추적
- 보안 고려 : 실시간 데이터는 개인정보·안전정보와 연계될 수 있으므로 권한 기반 차등 제공

□ 개방 시 필수 제공 요소

구분	필수 제공 항목	설명
데이터 설명서	데이터 정의, 범위, 관리 부서, 수집 주기	활용자 이해도 제고
품질 지표	결측률, 중복률, 코드 정합률, 최신성	데이터 신뢰성 확보
변경 이력	버전, 수정 일자, 주요 변경 내용	재현성 및 추적 가능성 보장
샘플 데이터	실제 데이터 일부 (민감정보 제거)	활용자 사전 검증 지원

□ 개방 우선순위 원칙

- 수요 기반 : 국민·기업 요청이 많고 활용도가 높은 데이터 우선
- 정책 연계 : 스마트항만, 디지털 트윈, 친환경·안전 정책과 연계된 데이터 우선
- 안전성 확보 : 개인정보·민감정보 비식별화가 완료된 데이터만 개방

붙임1

인공지능 친화적 공공데이터 관리 체크리스트(상세)

항목	구분	점검 내용	체크
형식 및 구조	필수	데이터를 특정 소프트웨어에 종속되지 않는 개방형 표준 포맷으로 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	필수	데이터를 기계 이해 및 자동 처리가 가능하도록 구조화된 형태로 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	필수	데이터 구조(스키마)는 필드(키)·자료형·관계·중첩 규칙등을 포함하여 명확히 정의되어 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	필수	메타데이터를 표준 스키마(예: DCAT)또는 기관·국가 적용 프로파일에 따라 기계 이해 가능한 형태로 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	필수	메타데이터 작성 기준에서 정의한 필수 항목을 누락 없이 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	데이터 검증 기준(예: 허용값, 필수값, 형식, 관계 제약 등)을 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	대용량 데이터는 압축 제공 또는 대용량 처리에 적합한 포맷(예: Parquet 등) 으로 제공하여 이용자가 용이하게 다운로드·활용할 수 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
의미 및 표준	필수	변수(항목)별 정의·단위·산정기준·코드값 의미를 포함하는 데이터 사전 (또는 이에 준하는 문서)을 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	필수	날짜·시간·코드 등 주요 값의 표기 형식은 표준 포맷에 따라 일관되게 적용되고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	필수	동일한 개념은 동일한 라벨·용어로 일관되게 사용하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	적용한 표준 포맷 및 적용 기준(예: 날짜·시간 형식, 코드 체계, 누락 값 표기 등)을 문서화하여 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	도메인 표준 용어체계 또는 온톨로지가 존재하는 경우, 해당 표준과의 정합 성 점검 기준(매핑·검증 절차 포함)을 마련·적용하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
품질 및 이력	필수	결측값 및 특수값 표기 규칙을 통일하여 일관되게 적용하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	필수	데이터 품질평가의 수행 여부 및 결과 내용을 메타데이터에 명시하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	필수	인공적으로 생성된 합성 데이터나 기계가 자동으로 단 주석이 포함된 경우 이를 명확히 구분하여 표기했는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	결측값 및 특수값의 정의, 발생 사유, 처리 원칙(예: 제외, 대체, 별도 코드 처리 등) 등의 정보를 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음

	권장	데이터의 원천·수집경로·가공/전처리(정제·비식별) 이력을 기록·관리하고 공개 가능한 범위에서 정보를 제공하는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	데이터 정제 및 변환에 사용된 소스 코드는 라이선스를 명시하여, 공개 가능한 범위 내 오픈소스 형태로 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	데이터, 메타데이터 및 핵심 문서(스키마/데이터사전·코드북/접근안내)는 이용자가 동일한 배포 경로에서 함께 확인할 수 있도록 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	데이터 또는 문서가 변경되는 경우, 관련 메타데이터 및 핵심 문서를 동시에 개정·현행화하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	데이터 특성(용도·품질·리스크 등)을 요약한 데이터 카드(또는 이에 준하는 설명 문서)를 함께 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
관리 및 보안	필수	개인정보(또는 민감정보)가 포함된 경우, 관련 법령·지침에 따라 비식별/가명처리 등 보호조치를 적용하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	필수	이용허락 범위(재사용, 재배포, 상업적 이용, 2차 저작물 작성 등)를 라이선스로 명확히 표시하고, 해당 라이선스 정보를 메타데이터로 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	재식별 위험에 대한 평가·검토를 수행하고, 그 결과 및 조치 사항을 문서화하여 관리하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	데이터에 영구 식별자(PID 또는 이에 준하는 지속 식별자)를 부여하여 관리 하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
데이터 제공	권장	데이터는 이용자의 자동 수집 및 연계를 지원할 수 있도록 웹 기반 표준 API(예: RESTful API) 형태로 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	데이터가 API로 제공되는 경우, 이용 편의 및 장애 대응을 위해 파일 다운로드 등 최소 1개 이상의 추가 제공 경로를 함께 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	API 제공 시 데이터 응답에는 필요한 메타데이터를 함께 포함하거나, 메타데이터를 확인할 수 있는 표준화된 참조 링크를 함께 제공하도록 설계되어 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	대용량 데이터 처리 및 자동 수집을 위해 필터링, 조건 검색, 분할 조회(페이지네이션) 등 효율적인 조회 기능을 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
관리 유지	필수	품질·보안·윤리·업데이트에 대한 책임 주체(역할 기준)를 명확히 하고 이용자 문의/오류 신고 연락 창구를 명시하고 운영하는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	필수	이용자 오류 신고 및 개선 요청을 접수·처리할 수 있는 절차를 마련하고, 피드백 체계를 운영하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	데이터 변경 사항은 버전별로 추적 가능하도록 변경 이력 정보를 제공하는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	권장	데이터의 변조 방지 및 신뢰성 확인을 위한 전자서명이나 해시(Hash)값 제공 등 기술적 검증 방식 운영하는가?	☐ 적정 ☐ 미흡 ☐ 해당없음

**인공지능 친화적(AI-Ready) 공공데이터 평가용 체크리스트
[편람기준]**

분류	점검 내용	체크
공통	◦ [필수] (포맷) 특정 프로그램이 없어도 열리는 개방형 표준 포맷 (CSV, JSON, XML 등)으로 제공하고 있는가?	☐ 적정 ☐ 미흡
	◦ [필수] (구조) 기계 판독이 가능한 형태(데이터가 항목별로 나누어 정리된)로 제공하고 있는가?	☐ 적정 ☐ 미흡
품질	◦ [필수] 여러 개의 파일데이터를 제공하는 경우, 데이터의 항목명과 순서는 동일하게 관리하고 있는가?	☐ 적정 ☐ 미흡
	◦ [필수] 데이터 내 빈값, 오류(오타, 깨진글자 등), 중복값 등이 없는가?	☐ 적정 ☐ 미흡
	◦ [필수] 데이터 내 같은 항목은 표기 형식을 동일하게 관리하고 있는가? 예) 등록일을 표기하는데 2026-03-20, 260320, 26-03_20 등으로 혼용하고 경우 x	☐ 적정 ☐ 미흡
	◦ 파일명이나 라벨링 규칙은 일정한 형식으로 관리하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	◦ 비정형 데이터의 경우, 화질 저하가 없는 형태(PNG, JPEG 등)인가? 예) 이미지 확대 시 글자 주변 깨끗, 글자 가독성(선명함), 도표의 선과 배경 구분 명확	☐ 적정 ☐ 미흡 ☐ 해당없음
	◦ 인공적으로 생성된 합성데이터나 기계가 자동으로 단 주석이 포함된 경우 이를 명확히 구분하여 표기했는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
데이터	◦ 개방한 데이터가 최소 1,000개 이상의 누적 데이터인가? 예) 정형(파일) : 1,000행 이상, 비정형(이미지, 문서 등) : 1,000장/건 이상	☐ 적정 ☐ 미흡
	◦ 개방한 데이터가 AI가 학습하기에, 충분한 기간의 내용이 포함되어 있는가? 예) 시간대별, 요일별 월별 등 데이터의 포함 여부	☐ 적정 ☐ 미흡 ☐ 해당없음
	◦ 데이터가 다양한 분야, 상황 등의 유형을 골고루 포함하고 있는가? 예) AI가 상황별로 구분할 수 있는 다양한 사례(지역·성별·연령 등 일부 집단이 아닌 전국 단위 등) 포함 여부	☐ 적정 ☐ 미흡 ☐ 해당없음
	◦ 데이터가 가공·시각화된 결과 데이터가 아닌 가공 전의 원천데이터를 제공하고 있는가? 예) 지도에 시각화한 데이터는 원천 데이터 개방에 해당하지 않으며, 지도에 포함된 각 항목에 관한 원천데이터, 시각화 파일데이터(shp 등) 등 개방 필요	☐ 적정 ☐ 미흡
	◦ 데이터 항목별 정의·단위·산정기준·코드값 의미를 포함하는 데이터 사전(또는 이에 준하는 문서)을 제공하고 있는가?	☐ 적정 ☐ 미흡 ☐ 해당없음
	◦ 비정형 데이터(이미지, 음성 등)의 경우, 해당 데이터를 해석할 수 있는 추가 설명데이터를 제공하고 있는가? 예) 사진 : 장소/일시/이미지 속 주요객체이름, 음성 : 화자정보, 자막(텍스트)데이터 등	☐ 적정 ☐ 미흡 ☐ 해당없음

붙임3

인공지능 친화적(AI-Ready) 공공데이터 데이터 카드 상세

구분	문서명	핵심내용
데이터 셋 개요	데이터셋명/ID	(공식 명칭) / (고유 식별자)
	발행기관/담당부서/연락처	(기관명, 부서명, 이메일/전화)
	데이터셋 요약	데이터 내용, 어떤 행정·정책 목적에 활용되는지 3~5문장으로 기술
	지원 태스크(활용 목적)	예: 분류 / 예측 / 추천 / 검색 / 요약 / 시계열 예측 / 이상탐지 등(해당 시)
	언어 정보	ko / en / 다국어 여부, 텍스트 데이터의 언어 비율(가능 시)
	대상 범위	기간(시작~종료), 공간(행정구역 단위), 대상(기관/업무/서비스)
	갱신 정책	갱신주기(일/주/월), 갱신 기준일(전일/실시간), 갱신 지연 허용 범위
	제공 형태	파일 다운로드 / API 제공 여부(엔드포인트 유무)
데이터 셋 구조	데이터 인스턴스(샘플)	1~3개 샘플 레코드를 예시로 제시(비식별/마스킹 적용)
	데이터 필드 정의	필드 정의 표로 제공
	키/식별자 규칙	기본키/복합키, 중복 허용 여부, 참조키 정의
	결측·특수값 규칙	결측 표기 방법(NULL/빈값/NA), 0과 결측 구분, 마스킹 값 규칙
	분할 정보(Splits)	학습/검증/평가 분할 제공 여부(해당 시), 분할 기준(시간순/무작위)
	포맷·인코딩·표준	데이터포맷, UTF-8, 날짜·시간 ISO 8601, 표준 코드체계(행정표준코드 등)
	기계 이해 스키마 제공	스키마 파일 또는 메타데이터 파일 제공 위치
데이터 셋 생성	구축 배경	정책/행정 목적, 구축 필요성, 제공 범위 결정 근거
	데이터 소스(출처)	원천 시스템 / 기관, 수집 방식, 수집 기간
	정제·전처리 과정	중복 제거, 오류 정정, 표준화(코드/단위/형식), 이상치 처리 원칙
	어노테이션 과정(해당 시)	라벨 정의, 작업자(내부/외부), 가이드라인, 품질검수 등
	비식별·보호 조치	개인정보 포함 여부, 비식별 처리 방식, 재식별 위험 완화(소량 셀 마스킹 등)
	품질관리 절차	생성 단계별 검증 항목(스키마 검증/코드값 검증/샘플링 대조 등)과 책임 주체
사용 시 고려사항	권장 사용 범위	가능한 사용 목적(정책분석/서비스 개선/AI 학습·평가 등) 명시
	비권장/금지 사용	감시/차별/재식별 시도 등 오용 방지 목적의 제한 명시

	사회적 영향	데이터 활용이 시민 / 취약계층 / 정책결정에 미칠 수 있는 영향 요약
	편향성 및 한계	대표성 한계(지역/기간/집단/업무유형 편중), 측정오차, 누락 가능성
	기술적 제약	해상도(시간/공간), 자연시간, 표본 크기, 변경 이력 영향(버전 간 비교 주의)
	리스크 완화 권고	모델링 시 주의사항(시간순 분할, 편향 점검, 품질 플래그 활용 등)
추가 정보 및 라이선스	라이선스	적용 라이선스(공공누리 등), 허용 / 제한 조건(출처표시/변경금지 등)
	인용	데이터셋 인용 문구(기관명, 데이터명, 버전, 발행일, 접속일)
	작성자/관리자 정보	작성 / 검토 / 승인 책임부서, 문의 채널
	버전/변경 이력	버전 규칙, 주요 변경사항, 하위호환성 안내
	관련 문서	데이터사전, 코드북, 품질보고서, 활용사례, API 문서 링크/문서번호
기술적 사양	재현성	동일 조건에서 동일 결과를 얻기 위한 전처리·생성 파이프라인 설명
	환경 설정	전처리 / 검증 / 평가 실행 환경(OS, 컨테이너/가상환경 사용 여부)
	매개변수/규칙	집계 기준, 필터링 기준, 마스킹 임계값, 표준화 규칙 등 핵심 파라미터
	코드 및 라이브러리 버전	전처리 / 검증 / 평가 코드 제공 위치, 의존성 목록 및 버전
	무결성 검증 정보	체크섬 제공 여부, 파일 크기 / 레코드 수 검증 기준
	접근 사양(해당 시)	API 버전, 인증 방식, 호출 제한, 응답 스키마, 오류코드 정의

참고

분야별 SI 친화 · 고가치 데이터 예시

분 야	데이터명
국토관리	도시재생구역, 위성/항공/드론 데이터, 건축도면, 입지분석, 주택공급 등
교통물류	교통사고 상세정보, 물류시설, 수출입통관, 교통약자 시설정보, 자율주행 데이터, 철도교통망, 해상운송망 등
환경기상	탄소배출 데이터, 폐기물 처리정보 등
산업고용	사업자등록정보, 고용 및 기업 정보, 소비 데이터, 발전 시설 정보 등
보건의료	의료영상, 질환별 통계, 반려동물 의료데이터, 의약품 등
농축수산	농업방역, 수산물유통, 축산농가현황 등
식품건강	건강기능식품 정보, 음식 사진, 레시피DB, 식품원재료, 의료기기 등
공공행정	인구통계, 선거데이터, 공문서(회의록), 주소정보 등
과학기술	3D지도데이터, 드론비행, 로봇, R&D 등
재난안전	인구 밀집 지역 CCTV, 실시간 지진 예측 정보, 비상사태 및 재난 구호 관련 정보 등
문화관광	국악, 대중음악 등 음성·음악 데이터, 문화재 3D데이터 등
사회복지	장애인, 독거노인 등 소외계층 지원, 저출산·고령화 대책 등
교육	학교 및 학원 분포, 교육사업, 교육수준 및 역량 통계 등
법률	저작권 분쟁 사례 등
재정금융	정부 세부 지출, 주식정보, 신용정보 등
통일외교안보	방위사업, 북한정보, 병역현황, 보훈데이터 등

참고

AI 친화 · 고가치 데이터 평가표

평가항목	평가유형	설명	평가점수 (예시)
AI 친화성	시계열성	데이터가 일정 주기로 축적되어 시계열 분석이 가능한지 여부	60
	완결성	항목별 결측이나 누락 없이 데이터가 충분히 수집·정리되어 있는지 여부	70
	구조화 수준	데이터의 표준화 및 정형화 정도 (예: 정형 테이블 구조 여부)	50
	결측치 유무	결측치가 과도하지 않고, 보간 또는 대체 가능한 수준인지 여부	50
	AI 학습 활용 가능성	실제로 AI 모델에 학습용으로 사용 가능한 구조 또는 컬럼을 포함하고 있는지 인지 여부	50
고가치성	민간 활용성	민간 서비스나 제품 개발에 직접 활용 가능한 주제/내용인지 여부	80
	정책 연계성	정부 정책, 과제, 지표 등과의 연계 가능성 여부	70
	공익성	사회문제 해결 또는 행정서비스 개선에 기여 가능한지 여부	80
	수요도	국민 또는 기업의 수요가 실제로 존재하거나 잠재 수요가 큰지 여부	90
	개방 가능성	법적·기술적으로 개방이 가능한지 여부 (비공개 사유 미해당 등)	80