



# 데이터 품질 확보방안

[국가 공인 데이터 품질인증]

**2024.06.12**





*Garbage In, Garbage Out*

，

*Data 환경의 변화*

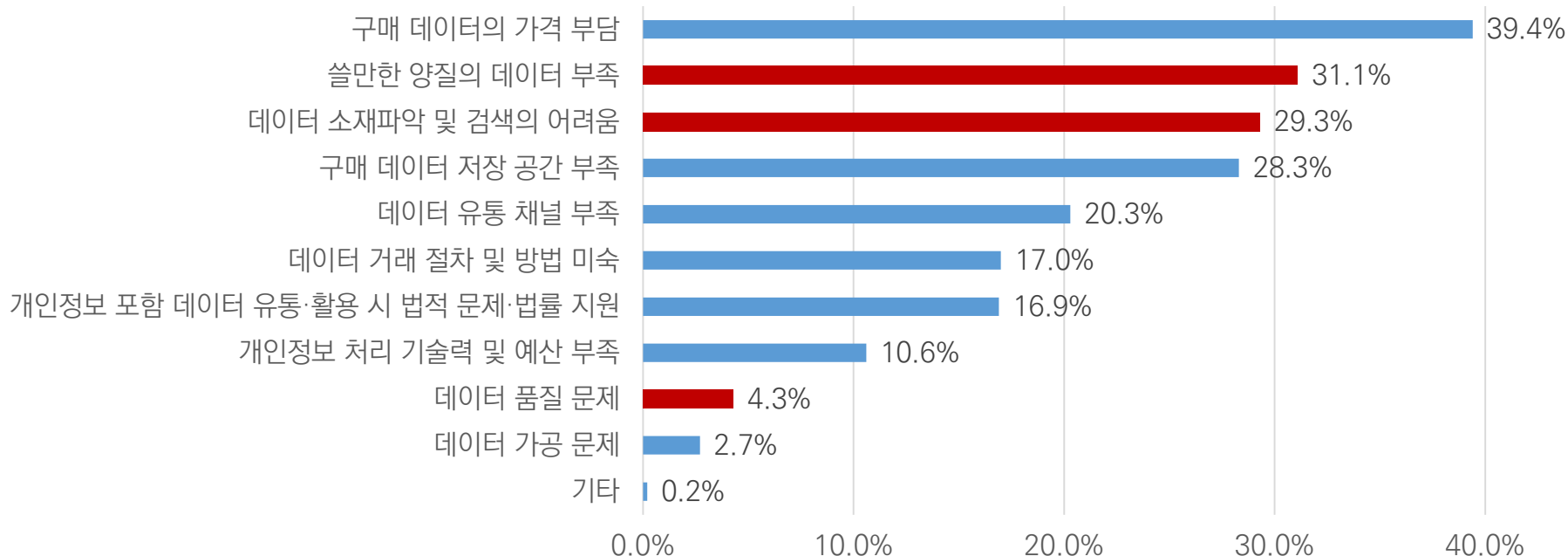
*(DataOps, Data Downtime...)*



출처 : tvN

# 현실

- 데이터 거래 시 애로사항



※ 2023 데이터산업 현황조사, 국가승인통계 제 127004호, '24.5

- 인공지능 사업 운영 상 기업이 느끼는 애로사항에 대해 'AI 인력 부족' 다음으로 '데이터 품질 문제'가 높게 나타남

※ 2023 인공지능 산업 실태조사, 국가승인통계 제127016호, '24.4

- 인공지능분야 기업은 사업 수행의 주요 애로사항이 데이터 품질(44.6%) 응답

※ 국내 AI 연구기관의 R&D 관련 특성과 시사점, 한국전자통신연구원, '22.6

# 국내 데이터 품질 관련 정책 및 법제도 현황

<p>대한민국 디지털 전략 (2022.09.)</p>	<p>거래 : 데이터가 가치를 인정받고 유통되는 시대 본격화</p>	<ul style="list-style-type: none"> <li>• (품질인증) 데이터의 품질향상을 위한 품질 대상 및 기준 등 확립: 데이터 품질인증 제도 시행</li> <li>• (가치평가) 데이터의 객관적 가치평가를 위한 평가기법 및 체계 확립: 데이터 가치평가 제도 시행</li> <li>• (데이터 거래사) 수요자와 공급자 간 데이터 거래를 알선할 전문가 양성: 데이터 거래사 등록제 및 관련 교육 등 시행</li> <li>• (유통질서) 데이터의 합리적 유통 및 공정한 거래 질서 확립 : 데이터 분쟁조정위원회, 표준계약서 등 제도 시행</li> <li>• (과금·보상) 데이터 전송 과금 및 보상 체계 마련</li> <li>• (자산 보호) 데이터 자산 보호 및 부정사용 방지</li> </ul>
<p>데이터 법제도 현황</p>	<p>데이터 산업진흥</p>	<ul style="list-style-type: none"> <li>• 데이터 산업진흥 및 이용촉진에 관한 기본법 (약칭: 데이터 산업법)</li> <li>• 산업 디지털 전환 촉진법</li> </ul>

# 데이터 품질인증 제도란?

---

과학기술정보통신부가 지정한 **데이터품질인증기관**이

「데이터 산업진흥 및 이용촉진에 관한 기본법」 제20조에 따라

**①데이터의 내용**과 **②데이터 관리체계**를 심사하고

데이터의 수준을 평가하여 **데이터품질을 인증하는 제도**

## 관련 법령 및 지침

---

「데이터 산업진흥 및 이용촉진에 관한 기본법」 제20조

「데이터 산업진흥 및 이용촉진에 관한 기본법」 시행령 제20조의3,4,5

「데이터 품질인증기관 지정 및 운영에 관한 지침」(이하 '데이터품질인증기관 지침')

---

# 데이터 품질인증 제도 목적

고품질 데이터 **유통 및 거래 활성화**를 위해  
공신력 있는 **데이터 품질인증 제도**를 시행

공신력 확보

인증 자문단 운영

인증기관지정

인증기관 관리감독

국제표준기반 심사

인증심의위원회

 WISESTONE

## 인증 대상에 따라 데이터 내용 인증, 데이터 관리체계 인증

### 1 데이터 내용

#### 데이터 형태에 따라

정형 데이터  
반정형 데이터  
비정형 데이터

#### 데이터 용도에 따라

거래용 데이터  
시스템 데이터  
학습용 데이터

### 2 데이터 관리체계

데이터 관리 원칙 정의

데이터 관리에 대한 업무 기능

데이터 조직 및 프로세스 체계적 정립 및 수행

# 인증 기준

## 관련 법령 및 지침에 따른 데이터품질기준

완전성	물리적 구조, 논리적 설계, 필수 항목의 누락 여부를 진단하는 기준
유효성	논리 관계의 준수 여부, 정해진 기준의 충족 여부 등을 진단하는 기준
일관성	데이터가 지켜야할 속성이 일관되게 정의되고 사용되는지를 진단하는 기준
정확성	업무 규칙에 따른 정확한 값인지 여부를 진단하는 기준
보안성	보안관리체계의 구축 여부 및 관리체계 수준을 진단하는 기준
유용성	사용자가 요구하는 사항의 충족 수준을 진단하는 기준
접근성	데이터를 용이하게 사용할 수 있는 지를 진단하는 기준
적시성	사용자가 원하는 시점에 적합한 데이터를 제공하기 위한 유지관리활동 여부 및 수준을 진단하는 기준
다양성	사용자의 이용 목적에 부합하는 다양한 특성 정보를 포괄하는지를 진단하는 기준
유일성	중복되지 않고 유일한 값인지를 진단하는 기준



## 인증 기준을 기반으로 국제표준을 참조하여 심사 항목 선정

### 1 데이터 내용

---

정형 데이터	ISO/IEC 25024	데이터 품질 측정에 관한 국제 표준
--------	---------------	---------------------

---

비정형 데이터	ISO/IEC 5259-2	AI 데이터 품질 측정에 관한 국제 표준
---------	----------------	------------------------

---

### 2 데이터 관리체계

---

데이터 관리체계	ISO 8000-62	데이터 품질 관리에 관한 국제 표준
----------	-------------	---------------------

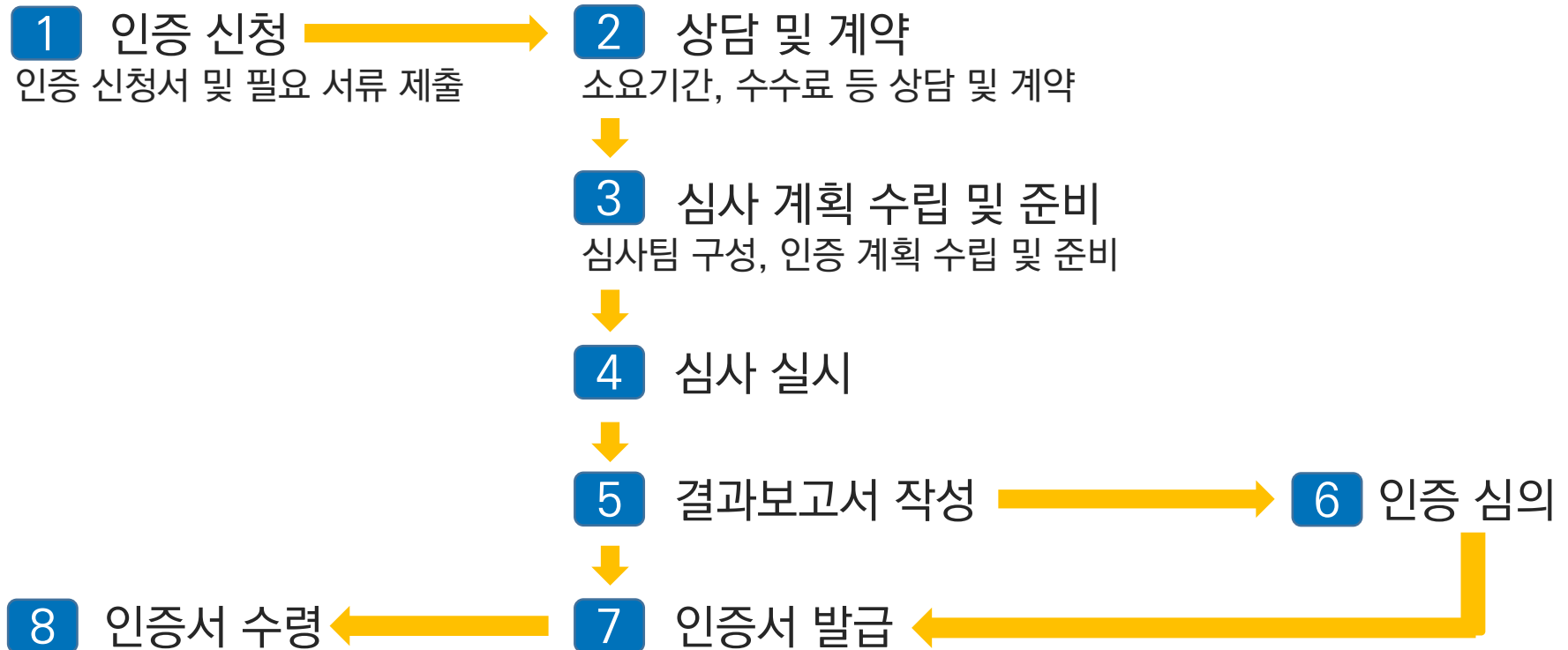
---

# 인증 절차

인증 신청인

인증기관   
WISESTONE

인증심의위원회



# 인증 신청 시 필요 서류

## 정형 데이터 내용

인증 신청서	신청인, 인증 대상 등 인증 신청 정보를 작성한 문서
테이블 정의서	테이블명, 엔티티 정보 등을 정의한 문서
컬럼 정의서	컬럼 정보를 정의한 문서 (컬럼에 적용된 도메인 규칙 및 업무 규칙)
코드 정의서	컬럼에 기재된 각 코드 정보를 정의한 문서
ERD (객체관계도)	각 엔티티 간의 관계도, 각 컬럼 별 연관성을 분석할 수 있는 문서

# 인증 심사 절차

---

- 1 심사 대상 분석 및 심사 계획
- 2 심사 수행
- 3 심사 결과 분석 및 오류 추정 보고서 작성
- 4 심사 결과 공유 회의 진행
- 5 심사 대상 수정 보완(신청 기관)
- 6 최종 심사 수행
- 7 심사결과보고서 작성

# 데이터 내용(정형 데이터) 인증 심사 항목

---

## 데이터 구조 · 특성에 따라 적용 가능한 필수 항목, 선택 항목으로 구분

---

필수 항목 ( 7개 )	완전성 데이터 값 완전성, 데이터 파일 레코드 완전성 유효성 구문 유효성, 의미 유효성, 범위 유효성, 관계 유효성 일관성 참조 무결 일관성
선택 항목 ( 7개 )	유효성 데이터 값 정밀성 일관성 데이터 포맷 일관성, 공통 어휘 일관성 정확성 메타 데이터 정확성, 데이터 값 정확성 접근성 표준 기반 데이터 접근성 유일성 데이터 값 유일성

---

## 인증 등급 및 기준

데이터의 특성과 심사 범위에 따라 **A등급, B등급, C등급**으로 구분  
심사결과 **평균 점수와 개별심사항목점수**로 인증 결정

A등급	필수 항목 모두 심사, 선택 항목 3개 이상 심사	<ul style="list-style-type: none"><li>· 심사 항목 평균 점수 99% 이상</li><li>· 개별 심사 항목 점수 95% 이상</li></ul>
B등급	필수 항목 모두 심사	<ul style="list-style-type: none"><li>· 심사 항목 평균 점수 97% 이상</li><li>· 개별 심사 항목 점수 95% 이상</li></ul>
C등급	필수 항목 일부 심사 (데이터 특성에 따라 일부 항목 진행 불가 시)	<ul style="list-style-type: none"><li>· 개별 심사 항목 점수 95% 이상</li></ul>

# 인증 심사 필수 항목

정형 데이터 내용  
인증 심사 필수 항목

데이터 구조 · 특성에 따라 적용 가능한  
**기본 중요 지표**

완전성

**데이터 값 완전성** 데이터 파일의 데이터 아이템에 대한 값의 완전한 비율  
**데이터 파일 레코드 완전성** 데이터 파일 내의 레코드 수가 완전한 비율

유효성

**구문 유효성** 구문적으로 정확한 값을 갖는 데이터 비율  
**의미 유효성** 의미적으로 정확한 데이터 값의 비율  
**범위 유효성** 지정된 범위 안에 포함된 데이터 비율  
**관계 유효성** 컬럼 간의 관계 규칙을 만족하는 데이터 비율

일관성

**참조 무결 일관성** 참조하는 테이블 간 동일한 속성값이 일관된 데이터 비율

# 인증 심사 선택 항목

정형 데이터 내용  
인증 심사 선택 항목

데이터 구조 · 특성에 따라 적용 가능한  
**추가 품질 지표**

유효성 **데이터 값 정밀성** 명세서의 정밀도를 만족하는 데이터 값의 비율

일관성 **데이터 포맷 일관성** 데이터의 포맷(또는 구문)이 일관성 있는 데이터 비율  
**공통 어휘 일관성** 데이터 사전에 정의된 공통 어휘를 사용하는 데이터 비율

정확성 **메타 데이터 정확성** 요구명세서에 정의된 대로 정보를 제공하는 메타데이터 비율  
**데이터 값 정확성** 명세서의 업무 규칙에 따라 기대 값을 만족하는 데이터 비율

접근성 **표준 기반 데이터 접근성** 해당 표준, 협약 또는 규정에 적합한 데이터 비율

유일성 **데이터 값 유일성** 데이터 값이 중복없이 유일하게 존재하는 데이터 비율



# 향후 인증 확대 방향

---

## 인증 유형별 순차적으로 인증 시행

### 1 데이터 내용

정형 데이터      2023년 11월 시행

비정형 데이터    2024년 하반기 시행 예정

### 2 데이터 관리체계

데이터 관리체계    2024년 하반기 시행 예정

## 목적에 다양한 용도로 활용 데이터 품질 제고 및 데이터 품질 증명

데이터 품질 제고	심사 과정 중 발견된 오류 데이터 수정 및 개선 활동을 통해 품질 제고
데이터 품질 증명	거래 데이터 또는 데이터 구축 과제·용역 결과물에 대한 품질 증명
마케팅 홍보자료	거래 데이터 및 데이터 보유 기관의 마케팅 홍보
데이터 신뢰성 확보	국가공인 시험인증기관의 공정·객관적인 인증을 통한 신뢰성 확보

# 데이터 품질인증 지원 사업

---

사업관리

 **Kdata** 한국데이터산업진흥원

지원규모

총 67개 사

지원내용

데이터 품질인증 비용 최대 1,150만원(VAT 별도) 지원

지원대상

데이터를 생산·유통·활용하는 중소기업

신청방법

이메일을 통한 온라인 접수(dq@kdata.or.kr)

제출서류

신청정보, 기업정보, 납세 확인을 위한 서류 제출

# 감사합니다

